AN ELECTROENCEPHALOGRAM INVESTIGATION OF TWO MODES OF REASONING

Chaille B. Maddox

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

COLUMBIA UNIVERSITY

2012

ABSTRACT

AN ELECTROENCEPHALOGRAM INVESTIGATION OF TWO MODES OF REASONING

Chaille B. Maddox

The use of electroencephalography (EEG) to exam the electrical brain activity associated with reasoning provides an opportunity to quantify the functional and temporal aspects of this uniquely human capability, and at the same time expand our knowledge about what a given event-related potential (ERP) might measure. The question of what form of mental representation and transformational processes underlie human reasoning has been a central theme in cognitive psychology since its inception (Chomsky, 1957; McCarthy, 1955; Miller, 1956; Newell, Shaw, Simon, 1958). Two prominent, but competing views remain at the forefront of the discussion, one positing that human inference making is principally syntactic (Braine & O'Brien, 1998; Fodor, 1975; Pylyshyn, 1984; Rips, 1994), and the other that it is, fundamentally, semantic in nature (Gentner & Stevens, 1983; Johnson-Laird, 1983). The purpose of this study is to investigate the neurophysiology of mental model (MM) and mental rule (MR) reasoning using high-density electroencephalography (EEG), with the goal of providing a characterization of the time course and a general estimate of the spatial dimensions of the brain activations correlated with these specific instances of two classic views of reasoning. The research was motivated by two questions: 1) Will violations of expectancy established by the devised MM and MR reasoning tasks evoke the N400 and P600 ERPs, respectively, and 2) Will topographical scalp distributions associated with each reasoning task suggest distinct cognitive domains and processes underwrite MM and MR modes of reasoning.

The finding of a N400 component in the MM strategy condition would suggest that reasoning about the relations between entities in the type of problems presented engages a network of cortical areas previously shown to be involved in processing violations of semantic expectancies in studies of language comprehension. By comparison, incongruent events in the MR condition are expected to evoke an anterior P600, a component previously associated with recognizing and restructuring syntactic anomalies or incongruities in sentence comprehension. If the hypothesized results are obtained they would provide potentially insightful information about the chronometry of mental processes associated with the different representations and inference making mechanisms postulated to support each mode of reasoning, and as well, broaden our understanding of the neural functionality associated with the N400 and P600.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

More people than whose names can be mentioned in this space have contributed to the work that culminated in this dissertation. The advice and input I received from Professors John Black, Karen Froud, and Joy Hirsch shaped the study from the beginning. Professor John Black entrusted me with his intellectual capital, which is foundational to the study. Were it not for the generosity of Professor Karen Froud the study would not have succeeded or come to fruition. She cheerfully and courageously encouraged me to pursue the research, shared her expertise, let me fail and learn, nudged me, and finally, sprinted with me to the finish line. Professor Froud's way of mentoring stands as the example to all of us in the Neurocogniton of Language Lab of what she means when she says that, "Neuroscience is team science". In this regard, the assistance provided me by my friends and colleagues in the lab is further testament to her leadership. It is with the greatest appreciation that I acknowledge the contributions of Trey Avery, Felicidad Garcia, Paula Garcia, Heather Green, Reem Khamis-Dakwar, Etoile LeBlanc, Lisa Levinson, Laura Sanchez, and Guannan Shen. They made this research possible by lending both their hearts and brains to the work. I also thank my classmates from the Human Cognition of Learning Lab, Greg Hallman, Seokmin Kang, and Jared Katz who contributed so much to the development of the experimental stimuli. In addition, I would like to thank Sam Froud, Maria LaMadrid and Yvonne Wallace for their support and very unique friendships. Many participants gave generously of their time to participate in this study, and I thank them.

Special thanks go to my dissertation committee members – Professor Black, Professor Froud, Professor Gordon, Professor Hirsch, and Professor Johnson.

The support that comes from family members who stand by you, ask nothing of you, remain at the ready to lend a hand no matter the request, and forgive you always, is indeed of the most special kind. I thank my brothers and sisters Carleton Maddox and Antoinette Selvey-Maddox, Aama Nahuja and Kofi Lomotey, Marcia Maddox Northington, Chesley Maddox-Dorsey and Leon Dorsey, and Adam Knee - and their families.

If one looks closely, the names of Jonathan A. Knee and Vivienne Lael Maddox Knee can be seen on the title page of this dissertation, as well as, on the related degree. If this were not the case, it would be impossible for me to fully thank them for their years of endless sacrifices and love. I return the love and express my unbounded gratitude to both of them.

Lastly, I dedicate this dissertation to my parents John E. Maddox and Myrtle L.V. E. Maddox who instilled in me the belief that in knowledge there is salvation, and therefore its pursuit is the reward.

C.B.M.

1. INTRODUCTION

Two broad classes of theories of human thinking and reasoning, mental logic and mental models, have in one form or another been the subject of philosophical and psychological debate for centuries. These rival views remain prevalent, one positing that human reasoning is principally resultant of rule-governed transformations of abstract mental representations (Braine & O'Brien, 1998; Fodor, 1975; Pylyshyn, 1984; Rips, 1994), and the other that it fundementally relies on constructing an understanding of the semantic relations perceived between mental representations that model the entities in a given situation (Gentner & Stevens, 1983; Johnson-Laird, 1983). The notion that human thinking and reasoning is in general rational, and at its full potential is logical has origins in the philosophy of Plato (360 B. C. E. /2003) and was the central theorem underlying the work of early psychological theories of the development of knowledge (Piaget & Inhelder, 1955). For Plato (360 B. C. E. /2003) the purpose of reasoning was the pursuit of truth or knowledge (Ackrill, 2001). On his view it is by engaging in the dialectic method that we elicit awareness of truth or knowledge as represented in the pure and perfect universal *Forms* (Blackburn 1996, p. 104). Plato's *Forms* are abstract ideal entities that represent the essence of classes of world phenomena (White, 1976, p. 8). The *Forms* can be conceived as analogous to the more modern notion of propositions or concepts in cognitive frameworks or axioms in mathematical proofs in that they are invoked in our reasoning to understand instances of phenomena or validate their truth value (White, 1976, p. 37). Thus, the *Forms* can be seen as serving as the normative facts at the center of Plato's philosophical epistemology similar to the role of principles in Piaget's model of the development of knowledge (Piaget, 1953; Smith, 2003). Such normative concepts coupled with the dialectic process helped to ensure that in

Plato's philosophical program reasoning outcomes were necessarily and objectively true (White, 1976, p. 31). The twin notions that true reasoning is based on *a priori* abstract ideas and is guided or governed by rigorous processes of deduction forms the cornerstones of mental logic approaches (Sternberg, 1999). These ideas can be seen at work centuries later in the empirical epistemological investigations of Piaget & Inhelder (1955) which were based on the assumption that human behavior was underwritten by inherently logical thinking (Graeme, 1982, p. 288). In the seminal publication *Logic and Psychology* (1955) the researchers theorized that the ability to reason develops correspondent with the acquisition and use of increasingly objective and formal rules of inference making (Beilin, 1992). The culmination of this developmental progression in reasoning (i.e., arrival at the formal operation stage) is the ability to think about classes of phenomena, specifically, to have concepts and to evaluate relations between concepts in accordance with formal rules (Beilin, 1992, p. 200). The logical operations which the authors posited to characterize thought at this stage where captured in a set of sixteen Boolean algebraic expressions developed to describe both the nature of advanced reasoning and the mental "psychologic" operations that result in valid inferences (Beilin, 1992, p. 197). Piaget's psychologic makes explicit the assumption, that mature reasoning involves the use of a sort of mental logic analogous to standard logical formulae (Held, Knauff, Vosgereau, 2006, p. 10). That is to say, formal operational thought makes use of formal logico-mathematical like operations to evaluate concepts and relations between concepts, to test hypotheses and solve problems (Beilin, 1992, p. 197).

Philosophical epistemologies and subsequently psychological theories, rooted in the mental logic framework were dominant from the 17th to the mid 20th centuries, excepting the period marked by the reign of Behaviorism (1920s to 1950s). As will be discussed, the emergence of cognitive science in the late 1950s not only fostered a renewed interest in mental logic accounts of thought and knowledge creation, but also sponsored a new generation of mental logic theories. The resurgence was owed in large part to the compatibility between the mental logic approach and the invention of new cognitive formalisms (Chomsky, 1957, 1959, 1963; Newell, Shaw & Simon, 1958; Newell & Simon, 1972). Johnson-Laird (1983, p. 23), one of leading advocates of mental model theory, referred to the prominence, and indeed for a protracted period of time, the dominance of mental logic theories of cognition as the *doctrine of mental logic*.

In contrast to Plato's belief that imperfect natural phenomena could not be the source of true knowledge, Aristotle (4th century B. C. E. /1969) held that the development of knowledge was in fact dependent on making sense of real world phenomena. He argued against the notion that one could only make sense of the world by activating innate tacit universals originating from outside of the real world (Aristotle, 4th century B. C. E. /1969, Book III, 428a-b). For Aristotle the development of truth and knowledge rested on establishing universals (the essence of things) which, he believed, were instantiated in particular instances of phenomena (Ackrill, 2001). He held that people developed universals by abstracting common properties across many particular instances of a type (Ackrill, 2001). According to Aristotle the process of abstracting common properties across entities is enabled by the mind's natural ability to associate (e.g., to identify similar and opposite properties) and relies on the body's sense faculties (Aristotle, 4th century B. C. E. /1969, Book III). Sense-based perceptions are central to knowledge construction because they form the mechanism by which we recall or imagine our interactions with particular objects and events in the world (Gregoric, 2007, *De Memoria et Reminiscentia* 1 450a 10).

(Beare, 1906) interprets this view thusly, “When our senses are activated by our observations and experiences we are left with the “residue” or traces of the impressions”. Traces of sense-based impressions played a role in Aristotle’s epistemology very similar to that of the mental representation in modern theories of cognition (Barsalou, 1999). Specifically, sense-based traces formed the content of thought and were the ‘what’ acted on during rational contemplation (Beare, 1906; Gregoric, 2007). It is perhaps for this reason that Aristotle is credited with having developed a psychological system (Robinson, 1995, p. 49). Moreover, Aristotle’s conceptualization of understanding and knowledge creation, as making sense of the interrelationships among natural phenomena greatly influenced the subsequent development of perceptual-based associative frameworks and theories of cognition (Lachman, Lachman, & Butterfield, 1979; Robinson, 1995).

In the domain of psychology the roots of the idea that people form internal models of the world in order to represent and anticipate possible events is typically attributed to Craik (1943) who wrote:

> If the organism carries a ‘small-scale model’ of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilise the knowledge of past events in dealing with the present and future, and in every was to react in a much fuller, safer, and more competent manner to the emergencies which faces it (p. 61).

It was, however, Johnson-Laird (1980) who developed the theory and offered it as an alternative to that of mental logic. In a groundbreaking article, Johnson-Laird (1980, p. 79) proposed that reasoners made syllogistic inferences by manipulating mental models. This was accomplished by constructing and evaluating possible and variant mental models of the premises in propositional problems (Johnson-Laird, 1980). The case for model representations as the basic structure of cognition was further expounded by Johnson-Laird (1983) and Gentner & Stevens (1983), establishing mental model theory as another important construct of human thinking and reasoning. As will be discussed in future sections there are many points of view on the form of mental models and how they support thinking and inference making (Held et al., 2006). However, a general statement that may capture the central idea of the mental model framework of reasoning is to say that mental models are psychological representations that preserve the structure and internal relationships among objects, events, situations or systems and that aid the reasoner in understanding and drawing inferences about such phenomena, and even predicting the future behavior of such phenomena (Craik, 1943). This general statement highlights the point that the emphasis in most investigations of mental models reasoning is on the reasoner's understanding as an outcome of the process of examining the relations between entities and their properties as represented in the model representations (Held et al., 2006).

Despite having stimulated extensive research in the cognitive sciences it is still unresolved whether a primary representational system (i.e., a given type of representation and its associated processes of inference) underlies reasoning, and if so, if such a system is best accounted for by mental logic or mental model theories accounts. In addition, other researchers have broadened the debate by proposing models of cognition that posit the coexistence of a combination of these two classic views, and even those who have proposed that multiple levels/systems of processing are required to account for richness of human reasoning (Evans, 2003; Knauff, 2006; Sloman, 1996; Smith, Langston, & Nisbett, 1992).

Neuroscientific studies conducted over the past fifteen years are beginning to make a significant contribution to our understanding of the neural basis of reasoning (Goel, 2008; Kraft, Balazs, & Pöppel, 2009). Such investigations aim to understand how high-order cognition like reasoning is implemented by the neural system, specifically, what neural structures and neuronal processes underlie reasoning. Adopting the notion of different levels of description introduced by Marr (1982), data from these studies offer another level of explanation of how human cognition may be realized in the human brain (Ilmberger, 2009; Kraft et al., 2009; Öllinger, 2009). This seems crucially important, especially in light of the fact that the functional or computation level descriptions resulting from research in the cognitive sciences have for the most part evolved independent of probable biological constraints (Milner, Squire, Kandel, 1998; Phelps, 1999). Findings from neuroscientific studies have begun to spur interest in a reassessment of psychological models of reasoning (Barsalou, 1999; Evans, 2003; Gratton, Low, & Fabiani, 2008; Phelps, 1999). There are hints that human reasoning and inference making as reflected at the level of neural implementation, may better align with cognitive theories positing multiple and mixed-knowledge structures (Goel, 2009; Gratton et al, 2008; Kraft et al., 2009).

Significantly, to date most neuroscientific studies on the domain have used Positron Emission Tomography (PET) or functional Magnetic Resonance Imaging (fMRI) in their investigations. Data from these studies are helping reasoning researchers to gain insight into which cortical areas are most consistently engaged by different types of reasoning and problem-solving tasks.

Very few studies have made use of Electroencephalography (EEG), which has the capability to capture the kind of fast neurocognitive processing that is known to support thinking, reasoning and decision-making (Ollinger, 2009; Rugg & Coles, 1995). EEG data has the potential to inform about the temporal dimension of these cognitive processes in a way that may contribute to a richer understanding of how functional specialization within cortical areas may be instantiated and coordinated (Ollinger, 2009; Gratton et al., 2008). The proposed study seeks to address the lack of such information by operationalizing, examining and comparing the time course, degree of activation, and scalp topography of an instance of mental model reasoning with an instance of mental logic reasoning. The semantic-based processes postulated to be a component of reasoning with mental models (Gentner & Stevens, 1983; Johnson-Laird, 1983) versus the syntactic processes said to operate over mental logic systems (Fodor, 1975; Plylshyn, 1984; Rips, 1994; Braine & O'Brien, 1998) provide the theoretical framework for the event-related potentials (ERPs) hypothesized to correlate with these respective modes of reasoning.

Specifically, ERPs previously associated with semantic (N400) and syntactic (P600) processing of sentences in studies on language comprehension will be used to investigate a mental model versus mental logic experimental paradigm. The association of the N400 ERP with violations of expected outcomes set up by problem solving using a mental model (MM) reasoning strategy, and the P600 with violations of expectations set up by syntactic structure-building processes thought to be related to use of a mental rule (MR) strategy may provide relevant information about the similarity and differences of how these respective modes of reasoning may be instantiated in the brain. Of equal importance, such an outcome may provide an index of similarity between brain responses associated with linguistic processing and those associated with more domain-general cognitive processes, such as reasoning. The association of different ERP components and scalp distributions with each way of reasoning is consistent with the proposition that separable neuronal populations subserve the psychological processes assumed to underlie each mode of reasoning (Osterhout & Holcomb, 1992; Rugg & Coles, 1995), thus suggesting that the two modes of reasoning may involve distinct representational domains and not merely reflect different behavioral strategies. The addition of neurophysiological information to the debate on the nature and form of reasoning may help in refining current theories or in developing new theories of cognition. To the extent that psychological theories of cognition inform educational policy, the development of school curricula, and even therapies to address cognitive disorders a better understanding of the neural basis of cognition may be helpful in discovering ways in which to enrich and/or rebuild competencies (Gardner, 1983; Feuer, 2006; Nersessian, 1995). Lastly, indexing physiological properties of mental model and mental logic reasoning may expand our understanding of the functional characteristics of the two standard ERPs (i.e., N400 and P600) hypothesized to be correlated with MM and MR reasoning.

2. REVIEW OF THE LITERATURE

2.1 A PHILOSOPHY OF MIND

Two major theories of human thinking and reasoning mental logic and mental model have in one form or another been the subject of philosophical and psychological debate for centuries. These contending views remain at the forefront of the discussion, one positing a symbol and syntactic account of human reasoning (Fodor, 1975; Plylshyn, 1984; Rips, 1994; Braine & O'Brien, 1998), and the other a model-based semantic account of human inference making (Gentner & Stevens, 1983; Johnson-Laird, 1983). The basic notions upon which mental logic and mental model theory rest, as well as the centuries long debate on whether there is a dominant method by which humans reason, dates back to Plato and Aristotle (Ackrill, 2001; Robinson, 1995). Plato (360 B. C. E. /2003) posited that knowledge was innate and conceptual, and that *reasoning* was a structured and rigorous process of dialogue leading to awareness of the *Forms*. Plato's *Forms* are abstract ideal entities (paradigmatic examples) that represent the essence of classes of world phenomena (White, 1976, p. 8). The *Forms* can be conceived of as analogous to the more modern notion of propositions or concepts in cognitive frameworks or axioms in mathematical proofs in that they are invoked in our reasoning to understand instances of phenomena or validate their truth value (White, 1976, p. 37). Thus, *Forms* provide an objective basis for truth and understanding of the world (White, 1976, p. 7). They were said to exist independent of human perception or judgment. Phenomena in the real world, such as objects and events were considered by Plato to be only imperfect instances of the pure and perfect *Forms* (Nehamas, 1975). He therefore, surmised that neither the real world, nor our senses through which we form distorted impressions of already imperfect world phenomena could serve as a source of knowledge (Nehamas, 1975).

While the *Forms* exist independent of the human mind, they can be apprehended by the mind; in fact, Plato (360 B. C E.../2003) theorized that the *Forms* were innate structures of the mind (White, 1976). Plato (360 B. C. E. /2003) believed that although we may not be aware of it, we are born with knowledge-rich minds. As exemplified in *The Meno*, it is our awareness of the *Forms*, rather than knowledge itself that Plato held we must "recollect" or developed in order to become rational beings capable of making sense of the world around us (Vlastos, 1996). In the *Meno* Plato presents a young boy with a geometry problem, and tasks him with determining what multiple of a length of a square is required to double the area of the square. Plato leads the child through a process of knowledge recovery, which is part Socratic questioning, and part demonstration. Eventually the young boy is able to acknowledge the correct answer. The use of Socratic questioning is meant to contrast with actual teaching, and to prove to Meno (the master of the young unschooled slave boy) the correctness of Plato's theory of innate knowledge and the manner by which it is best recollected (Vlastos, 1996). White (1976, p. 7) suggests that the *Forms* are Plato's solution to epistemological questions regarding the nature and form of knowledge (i.e., what it is, how it is created, what it is to posses it, how it confers meaning).

Aristotle's (4th century B. C. /1969) epistemology reflected a seemingly opposite view by proposing that knowledge originated in the real world and that its development was a constructive process. Whereas Plato's doctrine suggests that knowledge creation is a top-down process, Aristotle's is taken to infer a bottom up process whereby *universals* (the essence of particular instances) are constructed by abstracting commonalities (types, properties or relations) across many particular instances of worldly phenomena (Ackrill, 2001; Frensch & Buchner, 1999). Thus, in Aristotle's view, universals did not exist independent of reality, that it to say they exist only in things (Ackrill, 2001).

Aristotle (Aristotle, 4th century B. C. E. /1969, Book III) held that the mind was naturally inclined to perceive relations and make associations among entities. His principles of association (i.e., similarity, contrast, frequency, contiguity) suggested the way by which the mind coheres impressions and ideas, and how thoughts become accessible for conscious contemplation (Gregoric, 2007, *De Memoria et Reminiscentia* 1 450a 10). For example, the principle of similarity states that ideas are formed on the basis of the alikeness of objects, events, and experiences. It also asserts, that ideas can be recalled by any one of many attributes with which they are associated (the color, smell, texture, taste, shape of an orange). Similarly, if ideas or experiences are perceived as opposites, are frequently encountered, or occur closely in time or space they are both naturally associated and more readily recalled (Gregoric, 2007, *De Memoria et Reminiscentia* 1 450a 10). Aristotle's commonsense, experience-based epistemology postulated that the body's perceptual faculties played a key role in the process of knowledge creation (Aristotle, 4th century B. C. E. /1995). It assumed the mental formation of perceptually derived entities to "stand for" real-world objects and events (1906). These mental formations were "residues" of sense-impressions, and because Aristotle described such residue as *phantasma* which has been translated to mean presentation or image (Beare, 1906), his epistemology has been interpreted as a "picture theory" of thought and thinking (Barsalou, 1999). Importantly, Aristotle (4th century B. C. E. /1961) described these mental structures as being structurally and relationally correspondent with the phenomena they represented (Barsalou, 1999). The capacity of mental representations to preserve the relations between real-world entities and their properties is at the core of Aristotle's epistemology of thought and contemplation (Barsalou, 1999; Beare, 1906). His perceptual-based account of knowledge is generally held to be one of the oldest theories of the content of thought (Barsalou, 1999; Robinson, 1995).

In positing his theory of knowledge, Aristotle (4th century B. C. E. /1961) accepted that knowledge thusly derived only provided a basis for explanatory hypotheses of worldly phenomena. While he held that one arrived at sense or understanding through simple processes of induction, he believed that conclusions were graded (less or more strong) depending on the weight of the empirical evidence presented in their support (Reeve, 2000). The emphasis he placed on empirical evidence or observations, to test hypotheses further underscores the importance of perceptions in the knowledge creation process (Macnamara, 1999; Reeve, 2000). Significantly, Aristotle (4th century B. C. E. /1961) made a critical distinction between the nature of our perceptual faculties and the nature of mind (Reeve, 2000). He regarded the faculties themselves to be innate, but believed the mind to be an entity that developed, albeit resultant in part of the body's perceptual faculties (Reeve, 2000). On Aristotle's view, the mind like our perceptual faculties is physical, that is, it is of the body (Reeve, 2000).

The philosophical theses developed by Plato and Aristotle on knowledge and reasoning are old representational-level accounts of the metaphysical properties of psychological phenomena (McNamara, 1994; Sternberg, 1999). The construct of the *mental representation* is assumed by almost all theorists of cognition to be the basic element of which thoughts are composed (Markman, 1999; McNamara, 1994). In the context of cognition *representation* is the term used to refer to the hypothesized internal psychological constructs that "stand for" external referents and internally generated references (Markman, 1999; McNamara, 1994). In psychological theories of cognition, a *representation* is only partially descriptive of the phenomenon of cognizing. Philosophers and cognitive scientists also specify the processes that act on representations to form concepts and thoughts, yield understanding, and create knowledge

(Markman, 1999; McNamara, 1994). Thus *representational systems* are said to underlie cognition. The twin questions of what form of mental representation and what processes of transformation underlie human reasoning have been the central theme in philosophical theses and psychological theories of cognition for centuries (Garnham & Oakhill, 1994; Holyoak & Morrison, 2005).

If the philosophical traditions of Plato and Aristotle established the broad framework within which western intellectual debate and empirical study on reasoning would develop, it was seventeenth and eighteenth century European philosophers who formalized the differences between the philosophic theses of these two ancient Greek philosophers upon founding the Rationalists and Empiricists schools of philosophy (Sternberg, 1999). Whereas, Plato and Aristotle spoke of *Forms* and *Sensibles* in a literal sense, modern philosophers viewed knowledge as a mental construct (Ackrill, 2001; Sternberg, 1999). Such a view fostered a different sort of debate, and importantly, created the desire for a theoretical and empirical philosophy (Robinson, 1995). Modern European philosophers set their sights on producing a science of mind, that moved beyond metaphysics and speculation to provide a more rigorously grounded account of mental events, mental functions, mental properties, and consciousness (Haugeland, 1985; Macnamara, 1999). The goal of developing a scientific philosophy included explicating the relationship of mind to matter, or mental phenomena to the body (Macnamara, 1999; Robinson, 1995). Their endeavors built on the ideas of Plato & Aristotle about the relation between mind and body. As is the case with the representational systems they postulated, Plato's and Aristotle's beliefs about mind-body seem to occupy opposite ends of a continuum. Plato speculated that the mind and body were qualitatively different and Aristotle that they were materially the same (Ackrill, 2001). The former held that the perfect and unchanging mind was epiphenomenal of another world and the latter that the mind arose from the senses (Ackrill, 2001).

Descartes was the leading advocate of the Rationalists school in the 17th Century (Robinson, 1995), and he shared that part of Plato's doctrine that claimed true knowledge was both innate and constrained by structural characteristics and properties inherent to the mind (Descartes, 1637/1996, First Meditation). He and other prominent rationalists (e.g., Spinoza, 1677/1996; and Liebniz, 1982/1765) asserted that such *a priori* knowledge was accessible to the intellect through logical thinking (Robinson, 1995). Like Plato, they insisted that sensation, perception, feeling, and desire played no meaningful role in a scientific explanation of cognition (Descartes, 1641/1996, First Meditation; Robinson, 1999). Descartes' unique and formidable contribution to the rationalist's thesis, and more generally, the domain of philosophy, was his formalization of thinking as symbol manipulation under rules (Chrisley, 2000; Descartes, 1637/1970, Part I). A mathematician as well as philosopher, Descartes conceived of thoughts as symbolic representations of concepts, in the same way that he had employed mathematical notations in algebraic equations to represent geometric shapes (Bogdan, 1992; Chrisley, 2000). He theorized that like mathematical tokens in expressions, thoughts could be treated systematically, combined and recombined, and arranged and rearranged according to a set of allowable and unallowable rules that rendered them interpretable as true or untrue (Bogdan, 1992; Descartes, 1637/1970, Part I). Manipulated in this way, thoughts like the mathematical propositions they mirrored could be evaluated as certain or uncertain (Descartes, 1637/1970, Part 1). Mathematics was the exemplary science during his time, and any knowledge which aspired to truth had to partake of the certainty and clarity of mathematical theorems; in Descartes' own words, knowledge had to be "clear and distinct," else it was not genuine knowledge (Descartes, 1637/1970, Part 1). Based on the reformulation of thoughts as abstract but meaningful tokens, Descartes established the notion of a computational epistemology that used combinatorial and manipulable tokens to yield binary inferences, i.e., knowledge was either true or not true (Haugeland, 1985, p. 36). Thus, Descartes' insight to redefine thoughts as symbolic tokens in the mathematical sense, led him to conceive of the intellect as a device for "pure reasoning" (Bogdan, 1992; Descartes, 1637/1970, Part V; Haugeland, 1985).

Although the concept of the computational mind was introduced by the empiricist philosopher Hobbs (1651/1994), a contemporary of Descartes, Descartes did not connect his idea of the formal manipulation of thought to the notion of a computing mind, because he did not accept the empiricist notion of a mechanical mind (Bogdan, 1992; Haugeland, 1985). As a scientist, Descartes embraced Galilean physics and considered the laws of motion to provide a scientific explanation of the workings of physical live bodies, including human and other animals (Descartes1637/1970, Part V). Therefore, Descartes (Descartes1637/1970, Part V) deduced that if the human mind was mechanical then other animals might be expected to possess something similar, whereas, he declared, all empirical evidence indicated the human mind was distinct and unique from that of other animals. The primary evidence upon which Descartes based his assertion was the uniquely human possession of language and thought (Descartes, 1637/1970, Part V). Moreover, he further decreed the human facility with language to be special, citing as proof our ability to "arrange words in various different ways so as to respond to the sense of whatever is spoken to us" (Descartes, 1637/1970, Part V). The rationalists argued that this creative ability with language (and by extension thought) could not be matched by machines, because critically, such creativity necessitated the interpretation of words which in turn required knowing the *meaning* of things, knowledge available only to humans (Descartes, 1637/1996, Second and Third Meditation; Robinson, 1999). Therefore, Descartes said, any machine capable of producing creative thought would have to have a human inside and therefore could not be said to be mechanical (Haugeland, 1985, p. 36). And if, on the other hand, there was no human inside then what the machine was doing could not be said to be reasoning (Haugeland, 1985, p. 36). Descartes' approach to reconciling the conundrum of how to explain the inner workings of the mind was to uncouple a metaphysical explanation of the mind, from a Galilean physical explanation of the workings of the body (Descartes1641/1996, Sixth Meditation). Thus, he proposed two substances, *res cogitans*, or mind, in contrast to *res extensa,* body, with the former declared to be distinct from the latter and immaterial in nature (Descartes1641/1996, Sixth Meditation). In the end, like Plato, Descartes and other rationalists defaulted to a metaphysical explanation of knowledge and mind (Robinson, 1999).

The Empiricists position was represented by philosophers, such as Hobbes (1651/1994), Locke (1690/1959), and Hume (1739/1978; 1758/1965). The empiricist theory of mind shared many of the basic tenets of Aristotelian philosophy, for example the notions that mental representations originate in the environment and are established via our senses, and that concepts are formed by abstracting similarities and assessing the temporal contiguity between worldly phenomena (Arnheim, 1969; Robinson, 1995). Like Aristotle, they also held the belief that the representations underlying cognition were perceptual in nature, often imagistic (Arnheim, 1969; Robinson, 1995). According to the empiricist doctrine, thinking is the manipulation of internal representations of real or imagined events that are external to the thinker (Robinson, 1995). Eschewing metaphysics, empiricists accounted for manipulations in terms of the laws of physics, modeling their view of the world on machines (Hobbes, 1651/1994; Hume, 1739/1978). Philosophers of this school described the mind as "operating" in accordance with Galilean and/or Newtonian principles of motion or natural forces like gravity (Hobbes, 1651/1994; Robinson, 1999). Hobbes (1651/1994) declared the effects of motion (or mechanical action) to be the causal mechanism linking the external world to the human. He said, that *sense* was no more than the motion of matter external to us "pressing" against our sensory faculties (sensitive organs) starting a continuance of motion, culminating in the motion of atoms in the brain in turn giving rise to processes in the mind, e.g., perception and thinking (Hobbes, 1651/1994). Thus, the mind, according to Hobbes' epistemology, is a physical entity, which interacts through mechanical processes with other physical objects in the world to acquire an understanding of the world. The empiricists believed this physical and mechanical explanation of the mind provided a "scientific" theory to support the Aristotelian notion that the conscious mind is a function or epiphenomenon of the body (Robinson, 1999). Also, like Aristotle, the empiricist's theory of meaning was grounded in experience (Aristotle, 4th century B. C.). That is, empiricists posited that the meaning of physical objects and events in the world developed based on our interaction with them and through the development of shared meaning with others (Hobbes, 1651/1994; Hume, 1739/1978, 1758/1965; Locke, 1690/1959). However, rationalists saw a fundamental problem with the claim that the objectivity and truth value of knowledge lay in "reality," and that was, that perception was decidedly not objective (Robinson, 1999).

Locke (1690/1959) accepted that knowledge based on perceptions was subject to natural distortion by our perceptual faculties. He suggested that the best we can do is to form contingent knowledge or testable hypotheses (Locke, 1690/1959). Hume (1739/1978; 1758/1965) argued that because all knowledge is grounded in experience we could not know causal certainty by either deductive or inductive reasoning. He asserted that concepts and causality are only true by association and customary belief, and that moreover their status was inherently contingent because they are formed on the basis of our associations between entities (Hume, 1739/1978; 1758/1965). Sorell (1992) contends that Hume attempted to undermine Descartes' proof theoretic method of argument by arguing that mathematics itself was no more than mapping relations between ideas that are also rooted in perceptions.

Philosophically, the idea that some essential elements and features of the two polarized views might together offer a third approach owes much to Kant (Robinson, 1995; Sternberg, 1999). Robinson (1995) suggests that for Kant, the essential difference between the two epistemologies rested upon the question of whether and to what extent sense-based experience contributed to knowledge. His theory of mind recalls that of Aristotle's who recognized the mediating role of innate perceptual faculties in the creation of knowledge, while maintaining that our interactions with worldly objects and events was the source and content of knowledge. Kant's (1787/1965) position on the issue of an innate versus material mind was a synthesis of the opposing views. He postulated that the human mind was constrained by certain mental "structures" (location, space, time, entity, causality, thingness) inherent to our species (Kant, 1787/1965). Human's, he posited, were uniquely equipped with a reasoning faculty, underwritten by not only innate perceptual faculties, but also by an inherent sense of conceptual categories (Kant, 1787/1965).

On this view, the rationalist insistence on innateness applied to our reasoning faculty, but not to ideas and knowledge more generally as specified by the *Forms*. Instead, Kant (1787/1965) proposed that once the mind's inherent cognitive structure is taken into consideration the empirical account, that is, that aspect of it that maintains we know what we know via our sense-based experiences in the world, is largely correct (Robinson, 1995). Significantly, Kant (1787/1965) distinguished between two worlds: the world that we experience (the *phenomenal* world), and the world of "reality" or things as they really are "in themselves" (the *noumenal* world). This distinction between two worlds (two categories of understanding) is reminiscent of Descartes' separation of physical reality from metaphysical properties. It allows Kant to say that there is certain knowledge (higher truths) that we cannot "know" by either reasoning or experiencing, and other knowledge (a distortion of real "reality"), which we construct based on our worldly experiences (Robinson, 1995). Importantly, knowledge of the noumenal world is designated as *a priori* knowledge, which is said to be an outcome of logical reasoning, whereas phenomenal knowledge is derived via sense-based experiences (Robinson, 1995).

2.2 A SCIENTIFIC THEORY OF MIND

Neither the rationalists' nor the empiricists' theories of mind provided either a coherent account of the mental processes capable of yielding psychological phenomena, or, a wholly plausible description of how the processes related to the body (Haugeland, 1985). The epistemologies of seventeenth and eighteenth century philosophers made explicit the assumptions upon which experimental psychologists and cognitive scientists of the next centuries would build their theories of cognition (Robinson, 1995; Garnham & Oakhill, 1994). Critically, their "failure" to offer a plausible treatment of the philosophical puzzle of the “mind-body problem” emerged as a major force shaping how the domains of both experimental psychology and cognitive science developed (Haugeland, 1985; Robinson, 1995).

In the view of the next generations of philosophers and scientists the fundamental challenge was to explain ephemeral mental activities in scientific terms (Bogdan, 1992; Lachman, et al., 1979; Haugeland, 1985). In the United States, the early years of experimental psychology came to be dominated by the Behaviorist School, and that school’s doctrine made the term ‘scientific’ synonymous with ’observable’ evidence (Hull, 1943; Skinner, 1950; Tolman, 1955; Watson, 1913). Given research methodologies available at the time, the imposition of this dictate by behaviorists severely constrained if not halted psychological research on mentalistic phenomena, such as those about knowledge systems, and the mind-body puzzle (Macnamara, 1999; Miller, 2003). Mental states and the mind, which had been the focus of inquiry and investigation for so many centuries, were considered to be outside of the boundaries of “real science,” (Miller, 2003). Laboratory-based experiments (largely using animals) designed to induce associations between tangible stimuli and observable and measurable behavioral responses became the new research paradigm (Watson, 1929). The “new psychologists” wanted to avoid becoming entrapped in explanations of infinite regress about mentalism, as had done the philosophers and psychologists before them (Lachman et al., 1979).

However, despite the seemingly insurmountable threshold set by the behaviorists, theorists and scientists across many disciplines resumed investigations into internal cognitive processes. The transition was aided by the research of experimental psychologists themselves (Broadbent, 1954;

Cherry, 1953; Miller, 1956) whose experimental outcomes provided compelling evidence for the influence of mental processes on observable and measurable behavior. For these scientists and many others, progressing in their research made adhering to the doctrinaire position of even neobehaviorism increasingly untenable, and their inventive work contributed to the reinstatement of the study of mentalistic activity in psychology. Ultimately, the development of certain formalisms and conceptual frameworks in other research domains provided cognitive theorists with a means of recasting what were largely abstract and intangible notions (i.e., *Ideas*) in more "legitimate" ways. One such key development came from the field of mathematical logic, and was introduced in a paper published by Turing (1936). In his seminal paper, Turing (1936) described a mathematical formalization that became known as the *Turing Machine* or *Universal Machine.*

The *Turing Machine* (Turing, 1936) was highly influential in helping cognitive theorists to recast their abstract and intangible notions of *ideas* and *knowledge* in more concrete ways. A mathematical formalism, it allowed mathematical logicians to demonstrate how abstract symbols coupled with the operations performed on them could be described in terms of explicit, concrete processes rather than intuitive abstractions (Hoare, 2004). Essentially, a *Turing Machine* with only a few properties and functions can perform any logical or mathematical procedure that can be fully specified. This new formalism created a fully rationalized method for showing that psychological processes could be represented symbolically, and that these symbolic representations could be *meaningfully* altered by precisely defined symbol-manipulating processes (Lachman, et al., 1979, p. 97). Although, Descartes (1637/1970) had advanced the idea of the computational mind much earlier, the *Turing Machine* suggested how the computing mind could be operationalized in a non-mechanical manner by demonstrating that the abstract symbols of formal logic or mathematics could be copied, transformed, rearranged, and concatenated in much the same way as physical things (Bogdan, 1992; Haugeland, 1985). The methodological tool that implemented this series of procedures was called an *effective procedure,* which is similar to the notion of the algorithm. Relevant to the emergence of rule-based theories of human cognition, "algorithms" can be thought of as "a set of rules that precisely defines a sequence of operations, or importantly, a set of axioms that, if performed, will inevitably lead to the solution of a problem" (Lachman et al., 1979, p. 94). The mathematico-logical formalisms discovered by Turing (1936) and others (Church, 1936; Post, 1936) were foundational to the development of the new discipline of cognitive science because they provided the contributing disciplines with a crucial theoretical framework and scientifically acceptable tools with which to explicate their theories of cognition (Haugeland, 1985; Lachman et al. 1979; McCarthy, 2000). Moreover, the coupling of formalism with the operations of computation, as specified by the *Turing Machine* provided a template for the development of physical computing machines (Haugeland, 1985).

2.2.1 FORMALISM AND COGNITION

The generative grammars of Chomsky (1957, 1959, 1963) and the computational production systems of Newell, Shaw & Simon, (1958) and Simon & Newell (1972) were the first widely recognized formal models of cognitive activity. The syntactic-based theory of natural language introduced by Chomsky (1957) effected a sea change in how linguists approached the study of language and how psychologists conceived of cognition (Miller, 2003). Chomsky's research program (1957, 1959b, 1965) focused on uncovering the basic structures and processes that could account for the creativity of human linguistic competency and commonalities that seemed inherent across languages. He proposed that syntax or the grammatical rules underlying the derivation of sentences should form the core of any such theory (1957). Chomsky's (1957, 1963) introduction of Standard Theory illustrated how a grammar operates at the level of the *deep structure* of a language, and how grammars provide a structure that reflects an inherent human competence (implicit knowledge) with language. The theoretical framework presented in Chomsky's seminal publication, *Syntactic Structures* (1957), illustrated how a set of formal grammatical rules could account for the *systematicity, and thus, the productivity and compositionality of language.* In other words, the creative properties of language that Descartes (1637/1970, Part V) had pondered. Searle (1972) has commented that Chomsky's persistence in maintaining the centrality and autonomy of syntax in his linguistic theories reflects his belief that the basic principles of all languages (as well as the basic range of concepts they are used to express) are innately represented in the human mind. In effect, Searle (1972) suggests Chomsky's development of a grammar is an attempt to mirror the workings of these inborn principles (the inner working of the mind) with a set of abstract, quasi-mathematical rules intended to generate the range of possible sentences in a given language. A transformational generative grammar layered on top of a propositional representational structure provided a concrete model of one kind of high-ordered mental rule cognitive system (Boden, 2008; Lachman et al., 1979).

Newell and Simon (1958) are widely recognized as having had the critical insight that computers are capable of not only performing computational operations on symbols, but also of transforming patterns of symbols according to a set of rules (Chrisley, 2000; Lachman et al., 1979). They introduced a view of computer and mind that cast both as general symbol-manipulating systems and pattern transformers (Lachman et al., 1979; Newell, 1980). They developed the first computer simulations or artificial intelligence programs, the Logic Theory Machine (LTM) and the General Problem Solver (GPS). The General Problem Solver (GPS) was not only a computer program simulating cognitive activity it was also a theory of human problem solving (Hunt, 1999; Newell, 1980). Their *production* system model of cognition provided an account of the control mechanisms constraining human problem solving, and more generally human thought (Garnham & Oakhill, 1994; Newell, 1973). Central to the development of their model was the observation that when navigating problem spaces, although participants could conceivably become ensnared in endless cycles of actions, most did not, and moreover, most completed the problem-solving tasks successfully and relatively quickly. Newell et al. (1958) and Newell & Simon (1972) accounted for this by proposing people employed natural common sense rules to limit the search and guide their behaviors. They referred to these natural rules as *heuristics*, and postulated that human reasoning and problem solving was governed by a combination of informal heuristics and formal rules (Newell et al., 1958; Newell & Simon, 1972). Newell's and Simon's production system model of cognition provided another sort of account of how human thought proceeds, as well as a functional description of how the mind is organized (Hunt, 1999, p. 6). Although, like most cognitive scientists, they were materialist, they insisted the work of cognitive scientists was to develop an explanation of the mental activity that rendered thinking and reasoning without necessarily being obligated to account for how such processes might be neurally instantiated (Hunt, 1999).

Chomsky (1959) showed how mathematical formalisms could be used to describe the mechanisms of a cognitive system such as natural language, and Newell et al. (1958) and Newell & Simon (1972) demonstrated how cognitive activity could be prescribed by a formal language (i.e., a computer programming language) that could in turn be run on computers to simulate cognitive functioning. One established a formal way of thinking and talking about how natural phenomena are enabled, and the other developed a way of using formal systems to simulate natural phenomena in physical material. While both worked within a computational framework and used symbols and rules as the explanatory entities, neither researcher claimed their theories provided a model of the mind in the working brain (Boden, 2008; Hunt, 1999). However, the ideas central to their theses were further developed by other computational cognitive scientists whose models and theories moved closer to such a claim (Fodor, 1975; Pylyshyn, 1980a, b). Putnam (1961) and Fodor (1975), respectively, introduced and developed the *Computational Theory of Mind* (CTM). A psychological theory of mind, CTM proposed the human mind be regarded as an information processing system, and thinking and intelligent behavior as an outcome of "mental computations" (Fodor, 1975, 1981). Drawing heavily on the natural language grammars of Chomsky (1959) and the physical symbol system hypothesis (PSSH) of Newell & Simon (1976), he proposed the *Language of Thought* (LOT) hypothesis to describe the representational software and operations that allow the mind to function as a "processor" and yield the sort of behavioral regularities that reflect rational intelligence rather than instinct, learned routines, or intuition. Fodor (1975) theorized that because language is the expression of thought, and language as shown by Chomsky (1957) is systematic, compositional and productive thought necessarily has the same properties. Fodor (1975) extended this basic argument theorizing that the compositional structure of thought like that of language had a combinatorial semantics. A combinatorial semantics supported Fodor's (1975) assertion that mental states are expressed in a symbolic representational system, which he called *mentalese*, the language of thought. The LOT hypothesis asserts not only that thought and thinking is enabled by a particular representational system that accounts for its systematicity, compositionality and productivity, but also that the system's symbolic representations and combinatorial semantics requires specific computational processes of inference (Fodor, 1975). According to Fodor (1975) such a system can be defined by certain characteristics: 1) A finite set of irreducible, discrete, amodal symbols; and 2) A syntax that is specified in terms of well-formed rules and inferential operators, and further, that is combinatorial with a semantic aspect that maps meaning from symbols to referents (Fodor, 1975; Fodor & Pylyshyn, 1988). Thus, the inferential computations of mentalese are syntactical in nature, and provide a working model of CTM.

Whereas Newell & Simon (1972) made a case for developing functional descriptions of the mind separate from the biological medium that gives rise to the mind, Fodor (1983) argued for a specific relation between mind and body. He asserted that properties of mental representations corresponded with and were constrained by specific functional capacities of the brain (Fodor, 1983). Fodor (1983) claims that much of the mind's functional capability is organized into fundamentally discrete modules (correspondent with mental content) that do not interact with one another or with the higher-level central processing entity that operates over the logical relations between the mind's content. For Fodor (1975) mental states are not relational, and nor do their representations depend on shared beliefs for meaning. Indeed, many of the postulated modules relate to the mind's perceptual or linguistic capabilities are defined as domain specific (Fodor, 1983, p. 103). Further, these modules are said to be *informationally encapsulated* (i.e., do not interact with one another or other entities of the mind) with respect to any other unit of the mind, but importantly not with respect to the external world (Fodor, 1983, p. 69). The former feature is important because it designates perceptual-based cognition as low-level and limited (Barsalou, 1999; Fodor, 1983), and the latter feature is important because in Fodor's (1983) theory interaction with the external world, and not interrelations in the mind, is what allows the meaning of content to be established. (Fodor, 1983) posits that the symbols of mentalese get their meaning because they exist in a certain causal relation with what they represent.

The CTM was also advocated for by Pylyshyn (1981, 1983, 1984). Pylyshyn (1984, p. 142) hypothesized that data from the natural world is encoded and tranduced at the level of the "functional architecture" of the mind into symbol strings, and that certain computations are then performed on this symbolic data to produce outputs in the form of further mental or physical states (p. 215). Like Fodor (1975) he asserts that perceptions notably, imagistic mental representations play no role in such "intelligent cognitive systems". He differs, however in his justification for rejecting a "picture theory" of cognition (Pylyshyn, 1973, 1979b). First, Pylyshyn asserts that mental images are epiphenomenal of thought, that is, people do not reason by acting on mental images, they reason by simulating what they believe would happen if they were looking at and manipulating the actual situation being visualized (Pylyshyn, 1973, 1979b). In other words, people reason based on existing tacit knowledge about the objects or events at hand (Pylyshyn, 1979b, 1981). Such tacit knowledge of the natural world can be accounted for by laws of physical substances rather than rules, laws, that explain how various physical properties of perceived entities are causally connected (Pylyshyn, 1981). Defined in this way, this is knowledge that is neither abstract, nor dependent on mediating representations and combinatorial processes. Therefore, and second, even if, as other theorists claim, images do serve as explanatory entities in inference making they do so in an utterly different kind of cognitive system from that of language-like symbolic tokens (Pylyshyn, 1981). Images (if they exist) are like percepts in that they are products of the *functional architecture* of cognitive system, and as such they are limited to depicting phenomena, they cannot refer to phenomena (Pylyshyn, 1980a, b, 1981). Consequently, Pylyshyn (1984, p. 19) contends that imagistic representations and the processes that act on them are constituents of analog cognitive systems. He has argued that models of cognition using "sentential predicate-argument" structures provide constructive proofs of *The Representational Theory of Mind,* which proposes that rational thought and thinking entails encoding of semantic-facts and use of rules to effect their transformation and yield rational inferences (Pylyshyn, 1981, 1984). In contrast, he maintains that analogue cognition is only "inferential-like," riding as it does on fixed stimulus-bound representations acted on by non-articulated, continuous, and holistic processes instantiated in the *functional architecture of the mind* (Pylyshyn, 1980, p. 126; 1981, p. 17).

2.2.2 SYMBOL AND RULE ACCOUNTS OF REASONING

The doctrine of mental logic: The assumption that human's are innately rational beings and that human intelligence is at least in part a result of an innate competence knowledge of deduction as a formal rule governed process has long been resonant both in the philosophy of mind and in empirical psychological research (Johnson-Laird, 2006) This view was the basis of the investigations undertaken in the early twentieth century by Piaget (1953) and Inhelder & Piaget (1955). Their research program aimed to explain how knowledge develops by describing the “psychological origins of the notions and operations upon which it is based” (Beilin, 1992, p. 197). The theory of cognitive development proposed by Piaget & Inhelder (1955) described the staged progression of the development of thought, from its early dependence on external objects and inference-making based on similarity and temporal contiguity to its more powerful ability to manipulate abstract concepts and make inferences based on rules (Piaget, 1955). In the seminal work, Logic and Psychology, Piaget (1955) adapted formalisms from logic and mathematics to describe the psychological aspects of thinking and inference (Beilin, 1992).

The emergence of a cognitive science unequivocally reinstated the study of mental activity in psychological investigations, and along with it a renewed interest in mental logic theories. The significant number of symbol and rule-based frameworks of cognition to emerge following the cognitive revolution was to be expected given the compatibility of the mental logic view with the new cognitive formalisms and theoretical frameworks, and the advent of rule-based computer architectures (Johnson-Laird, 1983; Lachman et al., 1979). Cognitive scientists developed descriptive and computer implementable models of reasoning and inference making aimed at explicating the mental mechanisms underlying these cognitive domains. Two paradigmatic accounts of human reasoning aligned with the mental logic view emerged at the height of this period. One was advanced by Braine (1978) and Braine & O’Brien (1998), and the second by Rips (1983, 1994). Both accounts are based on the assumption that human reasoning depends on implicit knowledge of a natural logic (Braine, 1978; Rips, 1983). Consistent with most research on reasoning up to that time, each account was developed with the goal of describing the mental processes underlying the deductive form of reasoning (Johnson-Laird, 1980). Briefly, for most of the twentieth century the study of human reasoning was nearly synonymous with the study of drawing inferences from classic problems of deductive logic (Henle, 1962; Johnson-Laird, 1980). Deductive reasoning tasks can be specified in two forms, the classical predicate calculus [If P – then Q] conditional reasoning tasks and syllogisms [All A are B; Some A are B; No A are B or Some A are not B]. The former is also referred to as propositional logic arguments, and researchers tend to use the two terms, conditional reasoning and propositional arguments, interchangeably when discussing deductive reasoning problems. This type of deductive reasoning task requires that a person reason about the relationship between conditions described in the problem statements or propositions. In the classic presentation of such reasoning problems participants are given true if then statements, and asked to reason about the validity of a concluding statement. Syllogisms, the second form of deductive reasoning tasks, are composed of two statements or premises, that are assumed to be true, and that lead to a conclusion that is either valid/ invalid or indeterminate. Statements in categorical syllogistic arguments specify quantities, denoted by words such as, all, some, none, and so on.

According to Braine (1978), Braine & O'Brien (1998), and Rips (1983, 1994) humans are endowed with a set of mental inference rules (although we may lack explicit knowledge of them) used to derive conclusions about normally occurring problems and puzzles. Essentially, this is a claim that we have innate competence knowledge (competence in the Chomskyan sense) of the syntactic and semantic properties of English language connectives, as well as the process steps of deductive inference making. This knowledge is similar in structure (i.e., the same theorems) to the formalized rules of classical logic but is built on a different foundation (Braine, 1978, p.2). The theories proposed by these researchers purport to provide psychologically plausible rules of inference making and to elucidate the mechanisms used to construct mental proofs (Chater & Oaksford, 1993).

Braine (1978) and Braine and O'Brien (1998) studied the type of propositional knowledge and reasoning underlying the ordinary assertions made by people in navigating the events of daily life. Similar to the verbal protocol methodology used by Newell and Simon (1972), these researchers captured the propositional relations reflected by people's use of natural language particles (e.g., and, or, if), and concluded that human thinking and reasoning is prescribed by a natural logic system emergent on structures of natural languages. Underlying the theoretical framework introduced by Braine and O'Brien (1998) are certain notions about how human knowledge is mentally represented and organized: 1) Stored knowledge exists in a form that explicates the relations between entities as a whole, as well as their features and properties. 2) The relational structure onto which these interrelations are mapped allows them to be traced and recognized. 3) People have a means of tracking and judging associations (sameness, similarity and differences) between entities and their properties and features. 4) Natural mental logic consists of inference rules represented in a language-like format. These inference rules are seen as analogous to the connective role served by English-language words, like not, and, or, if then, if_and_only_if, and as such instantiate the mental capability to represent alternatives among properties or the entities that have those properties, as well as conjunctions, suppositions, and negations. In other words, the key assertion of the theory is that humans hold a repertoire of inference rules derived from general knowledge and that operate in the same manner as sentential connectives such as 'if' and 'then', and quantifiers like 'all' and 'some' (Braine & O'Brien, 1998).

Rips (1983) proposed that human thinking and reasoning is governed by natural deductive logic-like principles that constrain how our propositions about the world are combined, arranged and rearranged to make sense of the world. His deduction-system hypothesis asserts the human mind reasons by applying natural algorithmic procedures to operate over abstract propositions, similar to how programming languages process symbolic binary number code (Rips, 1983). More recently in his *Unified Theory* account of reasoning Rips (1994) posits that thoughts are composed of sentence-like variables that are operated on by natural logic procedures such as truth, negation, contradiction, conjunction, and/or conditionals to yield true, untrue or provisionally true inferences (Rips, 1994). At its core, Rip's (1994) theory is a logicism view of reasoning which states that cognitive processes are proof-theoretic operations over symbolic logic-like tokens that we interpret in terms of our experience-based understanding of everyday phenomena.

Although popular for their flexibility and broad explanatory value, symbol and rule inference accounts have been criticized for a number of inherent problems. Foremost, is what Searle (1980) referred to as the symbol grounding problem. Fodor's (1975) Language of Thought (LOT) hypothesis, which has served as a motivating theoretical framework for many of the cognitive paradigms developed during the last few decades, is used to exemplify the problem. According to Goldstone & Barsalou (1998) and Barsalou (1999), LOT claims that thoughts have language-like syntactic structure and combinatorial semantics allowing propositions to be transposed and combined while retaining their meaning and creating larger still meaningful structures, but does not provide an account of how the basic propositional units have meaning in the first instance. Searle (1980) has asserted that symbolic representations cannot by themselves generate meaning. His famous thought problem the “Chinese Room” is customarily cited to illustrate this conundrum (Searle, 1980). The premise of the problem is established by describing the scenario whereby a human is placed in the role of a computer. Picture a human locked alone in a room such that communication with other persons and interaction with the world is largely restricted. The minimum communication that does occur is one-way, and that is via written messages that are passed into the room through a slot in the door. The messages are “meaningful’ and the individual is given the task of producing “meaningful” responses. However, the problem hinges on the fact that the messages are written in a language the individual does not understand (i.e., Chinese). As is the case with a computer the individual has an aid, a rulebook, specifying rules for what symbols to write down in response to particular conditioned input. In effect, like a computer the human receives meaningful communication, encrypted in symbols which are not understood, and is tasked with manipulating it to produce appropriately meaningful responses encrypted in the same symbolic code. This is to be done assisted only by rules that apply to the non-semantic properties of the symbols. The “Chinese Room” problem has been very successful at making the point that meaning is not inherent to symbols, and that some cognitive agent must instead confer it upon them. While the lack of an adequate response to the question of how symbols get their meaning does not negate the value of symbol and rule-based systems, it does in the view of some cognitive researchers leave missing an essential structural element in symbol-based theories (Goldstone et al., 1998).

Similarly, Barsalou (1999) identified another fundamental weakness with symbolic reasoning systems. Namely, that they operate under the basic assumption that perceptual states are *transduced* into a wholly different (symbolic) representational language, but the mechanisms by which transduction occurs has never been addressed (p. 579). Beyond these issues with the mechanistic properties of such reasoning systems, abstract symbols and rule-governed knowledge structures are further criticized for their inability to account for the influences of context and content on people's reasoning (Cheng & Holyoak, 1985; Cosmides, 1989; Wason & Johnson-Laird, 1972). This criticism is in fact, a general one leveled against all such domain-general cognitive processing systems. As a largely content independent formulation of human cognition, the formal inference rule approach is seen as significantly limited because it only allows for the application of a particular set of rules to solve a particular group of well-specified problems (Johnson-Laird, 1983). The question takes on importance in light of research demonstrating the extent to which content influences human thinking. The effect has been well demonstrated for several decades in many applications of the classic Wason Selection Task experiment (Wason, 1966). In the classic version of the task participants are presented with four cards whose face values are E, K, 4, 7. Participants are told that all cards have a number on one side and a letter on the other. They are then given the following proposition, "If a card has a vowel on one side, then it has an even number on the other side." The reasoner's task is to select the cards that must be turned over in order to find out whether the generalization is true or false. To succeed at the task, reasoners need to consider each card and evaluate whether each is relevant to determining the truth-value of the generalization. Most reasoners understand the need to turn over the vowel (to determine if it has an even or odd number on the other side), and also that there is no need to turn over the card bearing the consonant (the card has no implications for the proposition). Some reasoners turn over the card showing the even number, however, whether it has a vowel or consonant the proposition is not disproved. Very few turn over the most salient card, that bearing the odd number. If this card has a consonant, then the generalization holds, but if it has a vowel on the other side then the generalization is disproved. Turning over this card is just as important as turning over the card with the vowel, because it is the combinations of vowels and numbers on these cards that can refute the generalization. Johnson-Laird (1983, p. 30) has offered several reasons for the failure of participants to solve the Wason task using only formal rules: 1) uncertainty about the converse expression of the generalization 2) the tendency to take explicitly stated information as more salient, and 3) the tendency to confirm rather than disconfirm assertions. Significantly, any of these assessments suggest that reasoners integrate information beyond that which is contained in the situation into the problem solving process.

The sensitivity of human reasoning to content and context specific semantic effects is only one limitation of this kind associated with symbol and rule accounts of reasoning (Markman & Gentner, 1993; Nersessian, 1999). Other reasoning experts found that prior knowledge; particularly beliefs or false assumptions led people to make errors in logical reasoning (Evans & Barston, 1983; Johnson-laird, 1983). In a typical case, people alter or add premises, to the problem statement or even reject conventions of the deductive form and not engage in the task (Johnson-Laird, 1983; Oakhill, Johnson-Laird, & Garnham, 1989). Still other researches contended that the functional role of formal mental logic approach, knowledge explication and truth validation within a closed system, rendered it incapable of accounting for such cognitive phenomena as hypothesis generation, insight and creativity, or the deep understanding of interrelations embedded in systems (Chater & Oaksford, 1993; Gentner & Stevens, 1983; Nersessian, 1999).

Finally, cognitive modeling theorists claimed the slow serial processing implied by the classic symbol and rules form of reasoning is inconsistent with the “expert phenomenon” (Chase &

Simon, 1973). To explain, experts become faster at making inferences as their experience increases, whereas attempts to instantiate this phenomenon in formal computational models have been found to require ever more symbols, with exponentially more connections between these symbols causing slower performance. Indeed, it was by such insights that cognitive scientists realized the brain could not be a serial information processing device, and must instead function more along the lines of a distributed-parallel processing system (Cottrell and Metcalfe, 1991; Rummelhart & Mclleland, 1986; Seidenberg and McClleland, 1989). The insight was a key inspiration for new approaches to computer programming and cognitive modeling techniques (e.g., connectionist modeling).

In summary, the type of rule inference approaches reviewed above became increasingly regarded as idealized and/or specialized and unable to account for the full inferential competence exhibited by humans in reasoning. These approaches, grounded in rational philosophical thought, were viewed as disembodied from the biology of reasoning (Cottrell and Metcalfe, 1991; Rummellhart & McClelland, 1986; Searle, 1980). New theories and frameworks emerged with the aim of explicating the viability of perceptually based systems of knowledge or of accounting for the many instances of psychological processing not supported by formal mental rule theories of cognition. For example, domain rich approaches incorporating long-term knowledge stores, such as causal mental models, were introduced in Gentner & Gentner (1983). Other cognitive theorists proposed that human reasoning rested on spatial forms of knowledge representation, such as the spatial mental logic models of Johnson-Laird (1983), and the spatially mapped social networks of Cosmides (1989). New paradigms were developed based on different formulation of more general and flexible rule structures that reflect the routines of life events. The pragmatic reasoning schemas of Cheng & Holyoak (1985), and social contract theory of Cosmides (1989), Cosmides & Tooby (1994) serve as examples. There were also new computational models based on the connectionist framework, for example, neural network models (Cottrell & Metcalfe, 1991; Rummelhart and McClelland, 1986; Seidenberg & McClelland, 1989). Central to this thesis, model theoretic accounts of reasoning, the most prominent of which are reviewed in the next section, attempted to address some of these issues.

2.2.3 MODEL THEORETIC ACCOUNTS OF REASONING

Model theoretic views of reasoning introduced in the 1970s and 1980s departed from the amodal, abstract, and arbitrary symbols and rule frameworks that reflected how many researchers thought about the mind and its operations (Barsalou, 1999; Johnson-Laird, 1980). Theories associated with this view are developed on the postulate that people form mental representations that represent the structure and internal relationships among events or objects described in the situations, tasks or problems at hand (Johnson-Laird, 1983; Gentner & Stevens, 1983). There are two prominent approaches to describing reasoning with mental models. Both advance the use of models as working memory constructs to support thinking and reasoning, and draw inferences. The first, Mental Model Theory (MMT) was introduced by Johnson-Laird (1980) to describe how people reason about the same type of sentential deductive logic arguments addressed by the theories proposed by Braine (1978) and Braine & O'Brien (1998), and Rips (1983, 1994). In later publications (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991) MMT was critically expounded. This formulation of mental model reasoning is discussed in the next section.

2.2.3.1 Logical Mental Models

MMT postulates that the mind constructs mental models of the world to reason and make sense of the world (Craik, 1943; Johnson-Laird, 1980, p. 73). Mental models are real-time conceptual analog representations used to reason about the state of affairs of actual, imagined, or hypothetical situations (Johnson-Laird, 1983). They differ from the symbolic representations that are central to mental logic theories in that mental models preserve the structure, properties and relations (both implicit and explicit) of what they represent (Johnson-Laird, 1980, 1983). For example, in the case of sentential reasoning, the primary application for which MMT has been researched, Johnson-Laird (1999) asserts that modeling affords preservation of all of the implicit as well as explicit information inherent in the arguments, thus retaining the *meaning* of the relations between the entities.

A model representation takes the form of a structural spatial layout of the state of affairs that is correspondent with the structures, properties and relations they represent (Johnson-Laird 1983; Johnson-Laird & Byrne, 1991). In the MMT formulation, mental models are not generally defined as visual representations (Johnson-Laird, 1983; Knauff & Johnson-Laird, 2002). Instead, Johnson-Laird (1983) and Johnson-Laird & Byrne (1991) propose that in constructing mental models people use abstract symbols akin to tokens as stand-ins for entities and properties, and their interrelations. Although the researchers do not reject the possibility of visual, imagistic or kinematic representations, they advance the more economical view that abstract icon-like tokens are used in forming the models (Johnson-Laird 1983; Johnson-Laird & Byrne, 1991; Knauff, 2006). Tokens are economical because they do not represent every feature and attribute of the referent, but only that, which is relevant and necessary to the reasoning task (Knauff, 2006; Knauff & Johnson-Laird, 2002).

Johnson-Laird (2006, p. 29) draws a fundamental distinction between sentential reasoning by formal rule manipulation of abstract symbols and model construction, inspection and comparison. Specifically, the model-theoretic approach to reasoning posits that we have prior knowledge that allows us to comprehend the semantic content of the premises in sentential arguments (Johnson-Laird, 1983, 1999, 2006). Sentences are context-bound and subject to interpretation of the connectives in natural language (Johnson-Laird, 1983, 1999, 2006). This is in contrast to the formal rule approach which involves evaluating patterns of symbols to determine the validity of a conclusion, but not interpreting the meaning of the symbols because meaning is encoded in the syntax (Johnson-Laird, 2006, p.29). Model-based inference is concerned with drawing valid inferences based on the semantic information contained in the arguments (Johnson-Laird, 1999, 2006). Johnson-Laird (2006, p. 30) provides the following example of semantic-based inference:

> Given the assertions:
>
> > There is a triangle on the board or there is a circle, or both.
> > There is not a circle on the board.
>
> One determines that its meaning is compatible with three possibilities:
>
> > There is a triangle on the board and there is not a circle.
> > There is not a triangle on the board and there is a circle.
> > There is a triangle on the board and there is a circle.
>
> One makes the following inference:

> Knowing from the first premise that there is not a circle on the board,
> All but the first possibility can be eliminated. It follows that there is a triangle.

The example is meant to illustrate that although a reasoner could make the same inference by constructing a truth table based on a formal inclusive disjunction rule the conclusion is readily inferred by constructing and "reading off" of model-based representations that capture what is in common to all of the different ways in which the premises could be interpreted (Johnson-Laird; 1983; Johnson-Laird & Byrne, 1991). Consequently, MMT defines reasoning and inference making as an iterative process of model construction, comparison, and reconciliation with the premises and drawing successive provisional or contingent conclusions (Johnson-Laird & Byrne, 1991). However, an important note is that constructing correct models and successfully navigating the problem space parallels the use of logical rules in making inferences. For this reason the Johnson-Laird (1983) model-theoretic account is regarded by some to essentially be a proof-theoretic account of reasoning that differs from the language-like logicism only in employing abstract model representations (Chater & Oaksford, 1993). Johnson-Laird (1983, 2006) and Johnson-Laird & Byrne (1991) assert that the inference processes required to interpret the semantic relations preserved in the structural resemblance between the mental representations and the represented entities not only distinguish mental model from mental logic reasoning, but also address a weakness inherent to language-like symbol and rule-inference systems of reasoning. The researchers argue that semantics cannot be wholly reduced to syntax, specifically meaning between representations cannot be wholly determined by the logical relations between tokens (Johnson-Laird, 1999). Semantics, he asserts, "has to do with both truth and validity, whereas syntax has to do with form (and derivability)" (Johnson-Laird, 1999, p. 595).

Again, Johnson-Laird (1983) and Johnson-Laird & Byrne (1991) underscore that while it is possible for people to develop rules to describe the state of affairs between the entities described in an argument's premises, it is more likely they will imagine a spatial layout of the entities based on the combination of given information and their stored general knowledge. It is the demonstrated incorporation of world knowledge during the construction of mental models that lends support to Johnson-Laird's (1999) claim, that when using mental models to solve problems reasoners are not restricted to information contained in the argument premises or, wholly dependent on inferences drawn from linguistic knowledge of the roles of connectives. Johnson-Laird & Byrne (1991) acknowledge that the incorporation of world knowledge during reasoning can be a source of error, but describe the resulting representations as the basis for effective inductive reasoning. Conclusions based on such representations are best assessed as valid, probable or possible depending on whether they can be evaluated as true when compared to the meaning expressed in all, most or at least a single premise (Bara & Bucciarelli, 2000). The flexibility of MMT to account for both deductive and inductive processes of reasoning has Bara & Bucciarelli (2000) and Johnson-Laird & Byrne (1991) to describe it as a unified account of reasoning.

2.2.3.2 Imagistic Representations And Cognition

A further contrast between MMT and the work of Braine (1978) and Braine & O'Brien (1998), and Rips (1978, 1994) is the de-emphasis of the role of language in reasoning. The frameworks posited by the latter researchers highlight the correspondence between logical connectives and linguistic connectives in natural language, establishing an explicit, if controversial interface between logic, language, and thought. This interface has stimulated a significant body of research and an expansive debate, which is well beyond the scope of this writing to review (for a perspective see, Falmagne & Gonsalves, 1995). Similarly, there is a substantial amount of research that has investigated the relationship between language and our ability to form and use high-level concepts (Fodor, 1975; Gardner, 1983; Lennberg, 1962; Mandler, 2004b; Vygotsky, 1986). Philosophers and psychologists have long considered language to be an important and enabling component of the human capability to think, even suggesting that language underwrites true or conceptual thought (Descartes, 1637/1970; Pylyshyn, 1981). The prevalence of such thinking, particularly at the dawn of cognitive science and cognitive psychology, must be taken into consideration to appreciate the difference between MMT and imagistic dynamic model-theoretic accounts of reasoning (Gentner & Stevens, 1983) which will be reviewed in the next section. Empirical grounding for the use of images in the latter account of model-based reasoning was provided by researchers investigating the use of visual and spatial representations in reasoning (Cooper & Shepard, 1973; Kosslyn, Ball & Reiser, 1978; Shepard & Metzler, 1971). The researchers posed the question of whether thinking was possible using other than linguistic representations (Cooper & Shepard, 1973; Finke, 1989, 1980; Kosslyn, 1980, 1973; Kosslyn, Ball, & Reiser, 1978; Kosslyn, Pinker, Smith & Schwartz, 1979; Kosslyn & Pomerantz, 1977; Shepard & Metzler, 1971). Their research was instrumental in the reinstatement of imagistic representations in theories of mind (Barsalou, 1999).

The mental image theorists cited above conducted studies to investigate whether higher-order thinking was possible with non-linguistic representations. They hypothesized that people could use visual mental images as referents to make inferences about situations. Mental rotation research conducted by Shepard and Metzler (1971), demonstrated that visual images could be manipulated in real-time as working memory constructs to reason about the properties and function of real objects. In a series of experiments conducted by Shepard and colleagues (Cooper, 1975; Cooper & Shepard, 1973; Shepard, 1975; Shepard & Cooper, 1982; Shepard & Metzler, 1971), participants were shown 3-D images of pairs of objects, and ask if the items within each pair matched. One of the objects in the pair had been rotated so that it was not superficially obvious that it was the same as its member. The researchers found that the time to reason and respond to the question was linearly correlated with the degree of angle at which the one member of the pair had been rotated. This linear function (correspondence between response times and rotational requirements) has become a signature index of mental rotation and mental modeling in laboratory tasks. Finke (1989) claimed that the apparent *transformational equivalence* between visual mental images and real objects indicated the laws of physics governing manipulation of physical objects was transferred to the mental images of those objects. Because reasoning by mental rotation required operating on a whole spatially continuous image the transformational processes involved were seen to be different from processes that were said to underpin reasoning using abstract amodal language-like symbols (Finke, 1989). Evidence of mental activity based on analogue (non-formally symbolic) processes, was at the time both notable and fervently debated (Anderson, 1978; Kosslyn, 1980, 1976; Pylyshyn, 1981, 1979, 1973). On one side of the debate were cognitive scientists such as Fodor (1968) and Pylyshyn (1981) who maintained that imagistic representations were concrete, holistic, content-dependent, and non-discursive, and by virtue of these attributes could not be manipulated and analyzed by the computational processes that necessarily rendered rational thought (Fodor, 1975). As discussed, Fodor (1968) and Pylyshyn (1983, 1984) claimed that the kind of cognition that makes use of perceptual representations (especially images) is readily explained by a person's tacit knowledge, that is, knowledge that reflects an implicit understanding of the affordances of properties of the functional systems involved in any given occurrence. On the other side were researchers such as Cooper & Shepard (1973), Kosslyn (1980, 1976) and Shepard & Metzler (1971) who claimed that imagery encodes spatial, as well as modality specific properties about all of the entities in a situation, both, that which is explicit, and importantly that which is implicit (and necessary to a robust interpretation of events). Kosslyn (1980) and Johnson-Laird (1983) insisted that information encoded in the representations they described was not pictorial but rather a spatial array of the entities and only the most salient features and properties of the entities. This defense is also reflected in Barsalou's (1999) *Perceptual Symbol System Theory,* in which he asserts that such an array of information can serve as a foundation for forming abstract concepts. Imagery is further defended as serving a computationally useful function in that it informs about the whole system (Kosslyn, 1980; Stevens & Gentner, 1983). Computations on imagery are said to be efficient in that they allow information to be read by inspecting the system, instead of developing a set of definitions, conditions and rules to describe the physical attributes of entities and to specify the allowable and relevant behaviors (Johnson-Laird, (1983). The debate has continued even into this decade (Barsalou, 1999; Kosslyn, 1994; Kosslyn & Thompson, & Ganis, 2006; Pylyshyn, 2003, 2002; Knauff & Johnson-Laird, 2002; Knauff, 2006).

Another, early and influential study, that provided additional and convergent evidence of the use of images in reasoning was conducted by Kosslyn, Ball and Reiser (1978). The authors reported findings from their 'imaginary island' experiment that suggested people consulted analogue spatial imagistic representations, not sentential propositions, to reason about the spatial relations of object location on the “island”. The high correspondence found between participants’ estimation of distances based on their imagined representations and the actual distances between target objects and landmarks on the island led Kosslyn and colleagues (Kosslyn, 1973; Kosslyn et al. 1978) to conclude that visual-spatial mental images preserved the true metric information embodied in the spatial relations among real objects. Finke (1989) has referred to this intrinsic aspect of mental images as the principle of *spatial equivalence,* to designate the correspondence between the spatial arrangement of elements of a mental image of an object and the real object and the spatial arrangement of its elements.

The extent to which visual-spatial mental images are “pictures” and therefore are or are not viable as mental representations led Kosslyn (1980) to formulate a set of principles to explicate the use of visual-spatial mental representations and their concomitant transformational processes in reasoning. Three principles were specified: 1) Visual-spatial mental representations are image-like “copies,” that are spatially and structurally equivalent to the physical objects, features and properties of the physical objects and events they simulated; 2) The meaning of a visual-spatial image or scene is assigned by processes that work over the whole representation, and is not merely inherent in the mental “copy”; and 3) Imagistic referents are used only in short-term memory processes (i.e., working memory), while representations stored in long-term memory are a sub-set of percepts in the form of abstract representations (Kosslyn, 1980). Kosslyn intended these principles to clarify the difference between ‘pictures in the mind’ and the type of analogue visual-spatial representations he believed supported the real-time use of visual-spatial mental images to reason about the interrelations between entities and their features and properties in a given situation. Kosslyn's (1980) principles elucidate the means, for example, by which visual mental images could be used to make predictions about the effects of changes in one part of the scene upon other parts of the scene or the scene as a whole.

Finally, Finke (1989) has argued that the *structural equivalence* preserved between visual mental images and the real objects they represent is what allows images to be manipulated (rotated, merged, and displaced) in ways analogous to the real objects, so that new information can be inferred about the properties of the represented objects. Thus, while there is evidence visual-spatial mental images do not necessarily preserve all visual details of perceived objects or scenes, that which is preserved (e.g., shape) along with information about the properties of entities (e.g., size) can lead to important qualitative insights (Finke, 1989). The considerable evidence for the pervasive use of visual mental images in reasoning and the robustness of the effects of their manipulation led Shepard & Cooper (1982) to posit the notion of a "mental imagery faculty", similar to Chomsky's "language faculty".

2.2.3.3 Imagistic Dynamic Mental Models

The second broad formulation of the use of mental models in reasoning and inference making seeks to characterize the knowledge and processes that support understanding and reasoning in knowledge-rich domains (Markman & Gentner, 2001). The view was expounded on in an edited collection of papers published by Gentner and Stevens (1983). The research presented in the book examines the use of imagistic and dynamic *causal mental models* to reason about physical systems and mechanisms. In addition to the different nature of representation used to construct the mental models, these imagistic dynamic models differ from those described in MMT both with respect to the inference processes they postulate and the type of inferences that result (Johnson-Laird, 1983; Gentner & Stevens, 1983).

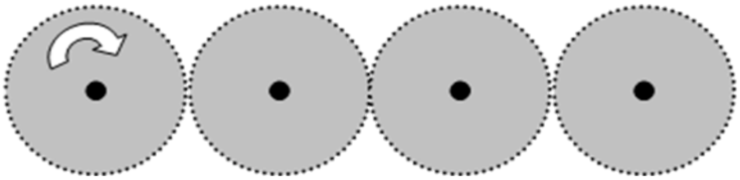
Causal mental model theories are formulated on perceptually isomorphic representations that preserve spatial and transformational equivalence with the referents in the situation modeled (Forbus, 1983; Schwartz & Black, 1996a, b). Notably, this formulation differs from the picture view of Aristotle (4th century BC/1961) & Wittgenstein (1922), or the postulate of "copies" by philosophers of the empiricist school (Hume, 1748/1965, book 2), or even the analogue token models of mental logic theory (Johnson-Laird, 1983). Causal mental models are said to exhibit *system-isomorphism* (Collins, 2010; Schwartz & Black, 1996a) – that is, there is correspondence between the system as a whole not just individual entities or parts. As such, model-theoretic paradigms posit more than mere structural similarity or resemblance between a representation and the represented entity, because causal mental models preserve the structural similarity and salient features, as well as the properties and interrelations among a system of representations and a system of the state of affairs (Finke, 1989; McNamara, 1994; Shepard & Cooper, 1982). Thus, model theorists expound on the advantages of mental models in representing real world physical systems (de Kleer & Brown, 1983; Forbus, 1983; Gentner & Gentner, 1983; Schwartz & Black, 1996a, b). First, mental models serve as imagistic and kinesthetic facsimiles of real physical systems, capturing and preserving core aspects and salient attributes of external objects (Giere, 1999). An important consequence of the system isomorphism and dynamic property of causal mental models is that they are able to simulate relational changes between entities in a system, initiated by changes in any one part of the system (Giere, 1999; Nersessian, 2002;

Schwartz & Black, 1996a, b). Second, perceptually-based reasoning systems are often more efficient for understanding and inference-making because they do not require the explicit, attention demanding and time consuming serial computing of symbol systems (Barsalou, 1999). In contrast, one is able to simply "read-off" such representational systems (Johnson-Laird, 1983). Causal model theorists (Gentner & Stevens, 1983; Giere, 1999; Nersessian, 2002) hold that this affordance, in conjunction with the imagistic and dynamic qualities of causal models, is what makes them essential aids for learners seeking to understand how complex systems work. It is this combination of affordance and other properties of imagistic dynamic models that researchers contend enables learners to obtain a relatively efficient and deep understanding of causally connected functioning of a system (Bailer-Jones, 1999; Lehrer & Schauble, 2003).

Like other resemblance-based paradigms, this type of model-based reasoning system is well suited to reasoning in real-time about real-world situations. However, in contrast to the logical model accounts, causal model systems invite deeper exploration of the subject at hand by explicitly allowing for an interface with long-term memory. This serves to import and integrate deep prior knowledge to support development of hypotheses, thus going beyond the information specified in the problem space (Bailer-Jones, 1999; Nersessian, 2002; Schwartz & Black, 1996b). In knowledge rich domains, such as biology, physics, psychology, and astronomy, the incorporation of prior deep knowledge to reason about open-ended questions is requisite to imagining and evaluating possible states of a system (Bailer-Jones, 1999). Some researchers contend that it is the evaluative aspect of model simulation that is so valuable because it is during this phase that misconceptions or errors are detected, resulting in insight and correct understanding of phenomena (Craig, Nersessian, Catrambone, 2002).

For example, in Schwartz & Black (1996a), participants were asked to think and reason about the behavior of a series of connected gears that constituted a simple mechanical system. The researchers verbally instructed adult participants to imagine several adjacent and touching gears arranged horizontally in a chain. Participants were then told that they should consider the left-most gear in the chain to be the lead gear. Then, for each problem, the direction of turn of this lead gear was specified as either clockwise or counterclockwise. Finally, participants were tasked with determining the direction of turn of the last gear in the chain (the right-most gear). A rendering of one such problem is provided in Figure 1 to help with envisioning the experimental task.

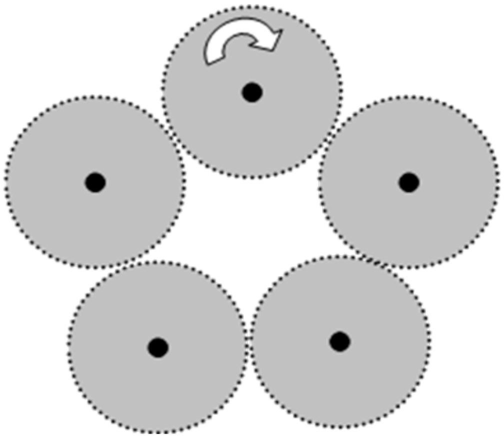
Figure 1: Open Chain Gear System



Task: *Given the direction of turn of the lead gear in which direction will the last gear turn?*

At the outset, almost all participants mentally rotated gears in the chain to simulate the expected behavior of the system. Presumably participants used prior knowledge regarding constraints imposed by the interlocking function of the gears to generate correct responses (most participants performed successfully and above chance). Importantly, having worked several problems of varying length, and in which the direction of turn of the lead gear varied between clockwise and counter-clockwise, participants invariably developed a *rule* that allowed them to quickly, respond to the posed question. For example, many participants developed the following local rule: *Each gear in the chain turns opposite to the one it touches, so, if the lead gear turns clockwise the next one turns counterclockwise, and the next clockwise, and so on.* Use of the rule became particularly helpful when participants were confronted with a problem containing a large number of gears in the chain (e.g., twelve). However, soon after participants formulated a rule, the researchers introduced a novel manipulation. They asked participants to imagine that the gears were configured in a closed chain format. Again, a rendering of one such problem is provided below in Figure 2.

Figure 2: Closed Chain Gear System



Task: *Imagine a closed system of gears, now in which direction will the last gear turn?*

If the closed system of gears imagined by participants contained an even number of gears, the rule they had developed continued to work (i.e., generate correct responses). However, if the system contained an odd number of gears the rule failed because the gears would "lock," unable to turn. The researchers found that participants who continued to "blindly" apply the rule did not detect the system malfunction. In contrast, those participants who for whatever reason, defaulted to simulating the turning of the gears (i.e., reverted to a mental model, rather than a rule-based system of reasoning) immediately discovered the problem and amended the rule so that it would apply to a variable number of gears in a closed system. This finding provides support for the assertion that qualitative reasoning is of value even if it initially leads to incorrect answers (de Kleer & Brown, 1983; Black & Schwartz, 1996a; Lehrer & Schauble, 2003).

The use of this type of mental modeling to think about theoretical problems and scientific tasks has had wide application, because once given an initial set of parameters people can construct and ‘run’ simulations of how a system might actually operate (Craig et al., 2002). This form of qualitative reasoning enables people to establish a functional and causal action path for many different types of systems, infer how or why a complex system works, and accurately predict general outcomes (Lehrer & Chasuble, 2003; Nersessian, 2002). Applied in this way, causal mental models have been used to study electrical circuitry (Gentner & Gentner, 1983), imagine astronomical events (Forbus, 1983), and induce the wave theory of sound (Holyoak & Thagard, 1995). Thus, imagistic causal mental model theory can (more easily than mental logic accounts of reasoning) account for the type of gestalt insight that is a hallmark of scientific and mathematical discovery (Holyoak & Thagard, 1995; Meuheus, 1999).

Some psychologists posit that reasoning by knowledge systems underwritten by percepts is a biologically adaptive capability that emerged based on our ability to form concepts by detecting similarities and analyzing spatial and temporal continuities between perceptual stimuli (Barsalou, 1999; Cosmides & Tooby, 1994; Galaburda Kosslyn, & Christen, 2002; Glenberg & Robertson, 2000; Mandler, 2004a). Others have posited that such perceptually based reasoning rides on an inherent grammar of graphical structures (Holyoak, 2008; Kemp & Tenenbaum, 2008). The *Perceptual Symbol Systems Theory* proposed by (Prinz & Barsalou, 1997; Barsalou, 1999) posits that perceptual representations are a neurally instantiated distributed code. Salient features and properties of objects and events are postulated to be extracted, encoded and stored in different populations of neurons, in contrast to the notion of stored holistic "pictures" or images (Barsalou, 1999; Prinz & Barsalou, 1997). Barsalou (1999) refers to the process of building a sensory percept from distributed representations as "schematic extraction", and suggest that the attributes extracted are an outcome of selective attention processes. It is these schematic perceptual representations that Barsalou (1999) and Prinz & Barsalou (1997) theorized are stored in long-term memory as symbols. These stored symbols are then processed by the cognitive system to build a fully productive formal conceptual symbol system of reasoning. Thus, by redefining the form that perceptual representations take (i.e., neural code) Barsalou (1999) interjects both a different kind and different level of mental representation into the debate on the nature of the representational systems underlying reasoning. In essence, Barsalou's theory posits that human reasoning rides on a conceptual representation system configured from neural representations derived from transformed perceptual stimuli, and that concepts thusly, derived can meet the standard of being compositional and productive.

2.2.4 HYBRID ACCOUNTS OF REASONING

More than half a century of debate has not resulted in agreement as to whether human reasoning is best accounted for by a symbol-based representational system or an associative system that probably works on perceptually-based representations, or even whether reasoning can be described by a single system. The varied nature of human thought has been recognized for well over a century (James, 1890/1950; Kant1787/1965; Neisser, 1963). Kant (1787/1965), as discussed earlier, proposed a philosophical theory of mind that was neither a wholly rationalist's, nor empiricist's view, but instead recognized the viability of elements of both. James (1890/1950) distinguished two modes of human thinking, that which is more experiential (empirical) in nature and dependent on an examination of associative relations, and analytical (reasoned) thought that engaged methodical and deliberate processes. The possibility that the robustness and sophistication of human cognition rest on more than one knowledge system was raised by Neisser (1963) in his article, *The Mulitplicity of Thought*. In the article, he famously lamented the tendency within psychological research to focus on dichotomies, e.g., deductive vs. inductive, explicit vs. implicit, rational vs. irrational, procedural vs. declarative, rules vs. models, experience-based vs. analytic (Neisser, 1963, p. 1).

Smith, Langston, & Nisbett (1992) conducted a review of research claiming evidence for reasoning by abstract rules on the one hand, and reasoning based on instance models (models that instantiate possible outcomes of a broader set of conceptual possibilities) on the other, and concluded by arguing that the literature reflects the need for at least two qualitatively different sets of mechanisms to fully describe human reasoning. The authors stated, "In addition to *pure-rule* and *pure-instance* mechanisms, hybrid mechanisms may be needed as well" (Smith et. al., 1992, p. 31). Case-based reasoning and learning theories (Hammond, Siefert & Gray, 1991; Shank, 1982; Kolodner, 1983, as cited in Smith et al., 1992) are identified as examples of theories that are better explained by hybrid mechanisms. Significantly, Smith et al. (1992) note that to posit hybrid mechanisms in such situations merely raises the more challenging issue of explaining how the two representational types and their associated mechanisms interface. On this question, the authors theorized that retrieval of an instance might trigger access to a related rule, or the converse, an abstract representation accesses an abstract rule that in turn accesses typical instances (Smith et al., 1992, p. 32). In total, the authors posit four mechanisms: pure-rule, pure-instance, instance-rule, and rule-instance that operate on abstract and analog types of representations (Smith et al., 1992).

Two of the more robust and prominent dual system reasoning frameworks were proposed by Sloman (1996) and Evans (2003). In different ways, both theories are based on the assumption that human reasoning evolved in response to an increasing need to function in a world becoming steadily more socially complex. Correspondingly, their frameworks are situated in the context of biologically adaptive evolution, and therefore, features of the models reflect a graded capability. Sloman's (1996) associative and rule-based systems are primarily distinguished by the computational processes each uses to draw inferences. The connectionist framework developed by Smolensky (1988) provides the computational architecture and processes adopted by Sloman (1996). Consequently, Sloman assumes the two systems he posits operate in one mind, specifically, they "run" on the same "hardware," each demanding a qualitatively different type and level of processing (Sloman, 1996, p. 7). The associative system encodes relations between elements of world phenomena based on principles of similarity and temporal contiguity, and yields inferences based on the underlying statistical structures (i.e., regularities, frequencies and correlations) estimated between current stimulus and stored knowledge (Sloman, 1996, p. 4). The other system is rule-based and encodes and manipulates abstract symbols that stand for any and all propositions of a specified type, thus demonstrating productivity, systematicity and compositionality. The rules of this system have both a logical structure and a set of variables, the meaning of which is derived from knowledge that exists as part of a community (Sloman, 1996, p. 5). The associative system encodes modality-specific representations of the structural relations of concrete entities such as images, stereotypes and feature sets. It operates reflexively at an implicit level and responds at an explicit level using routines, which afford fast inferential processing to yield predictions about the world. The rule-based system operates at a conscious level on symbolic representations. Its operations are effortful, deliberate, sequential, and slow but provide coherent explanations about natural phenomena. Sloman (1996) describes the two systems as functionally complementary. The associative one is automatic, less effortful, flexible, and allows quick construction of insight or hypotheses about the world. In contrast, the other is attention demanding and follows procedures or uses analytical processes that lead to the validation of ideas. The former system is unbiased and may be the source of creativity to the other's constrained analysis (Sloman, 1996, p. 7). The two systems can operate concurrently, each affording its particular type of inference. While one approach to inference making can dominate over the other, the extent to which either does or that the two work in concert depends on an individual's knowledge, ability and metacognitive functioning (Sloman, 1996).

The dual-process model proposed by Evans (2003) also draws a distinction between preconscious and conscious cognitive activities. Evans (2003) draws from the work of several theorists and researchers in formulating his theory. Preconscious activities are the product of System 1, a set of pre-programmed innate, semi-autonomous, input modules whose functional role is similar in both humans and other animals (Fodor, 1983). Although cognitive abilities are domain-specific learning occurs via a domain-general associative mechanisms, with inference making dependent on heuristics developed on this mechanism. System 1 is phylogenetically older, its theorized purpose being to enable pragmatic behaviors. Its processes are automatic, parallel and fast and inference making is intuitive in nature. While System 1 underwrites implicit instinctive behaviors it is governed by System 2, which extends human reasoning by providing an additional set of explicit well considered abilities afforded by its particular functional properties. System 2 is said to be evolutionarily more recent and unique to man (Evans, 2003, p. 454). Its processes occur at a conscious level and are volitional. Thinking is governed by rule-governed processes operating on a symbolic logic knowledge structure (Evans, 2003, p. 456). System 2 makes use of linguistic input and yields both deductive and plausible inferences (Evans, 2003). In contrast to System1, System 2 is highly demanding of limited attentional and working memory resources (Evans, 2003). Owing to the large amount of input System 2 receives from System 1, Evans (2003, p. 456) posits System 1 is the source of bias input to an otherwise rational System 2. The two systems are described as analogous to having "two minds in one brain" and competing for control of our inferences, decisions, and actions, with System 2 having the potential to suppress or overwrite the responses directed by System 1 (Evans, 2003).

2.3 NEUROSCIENTIFIC INVESTIGATIONS OF REASONING

Rather than the mental representation systems posited by the philosophical and psychological theories reviewed in previous sections, neuroscientific investigations into various cognitive capacities are designed to examine the neural activity or "brain states" correlated with specific cognitive tasks and processes. In neuroscientific studies, brain states are indexed by the spatial (anatomical) and temporal (time) involvement of neuron populations responding to specific stimuli or conditions, as measured by changes in some physiological state (i.e., patterns in blood flow, metabolic rates, neurotransmitter dynamics, and/or electrical brain activity). The view that such studies can provide insights about the nature of cognitive activities is based on two strong assumptions. First, that there is correspondence between neural activity and mental representations and processes (Kraft, Balazs, & Pöppel, 2009; Phelps, 1999; Rugg & Coles, 1995). Second, the changes in the physiological states we are able to measure are assumed to be associated with aspects of sensory encoding, recall, and recoding and other cognitive operations that underlie psychological phenomena, in accordance with the theoretical frameworks previously discussed (Rugg & Coles, 1995). More than thirty years ago in his introduction to MMT, Johnson-Laird called for the embrace of new methodologies to help change the way in which we think about the mind and investigate its operations (Johnson-Laird, 1980, p. 73). Neuroscientific research employs methodologies that allow us to think about and investigate the mind in ways that complement the philosophical views and psychological frameworks reviewed in earlier sections of this writing.

Despite extensive behavioral and computational modeling research motivated by the two classic theories of human reasoning, mental logic and mental model it remains unresolved whether a central representational system supports thinking and reasoning and whether it is best accounted for by one that is symbol and rule-based or one that uses mental models and comparative or associative processes to draw inferences. These two theoretical approaches have different implications for the neural instantiation of cognitive activity supporting reasoning. Mental logic theorists have tended to conceptualize the human brain as a computational device that uses a “mental language program” to operate on recoded perceptual stimuli using a combinatorial syntax and semantics to render new abstract amodal mental symbols interpretable as true or untrue (Braine & O'Brien, 1998; Fodor, 1975; Plylshyn, 1984; Rips, 1994). In contrast, mental model theorists speak of structure-preserving mental analogues (amodal tokens or perceptual-based models) that are manipulated and inspected to induce meaning through use of associative principles and statistically derived heuristics (Gentner & Stevens, 1983; Johnson-Laird, 1983). Although the study of the living brain has been absent from the study of cognition for most of the life of the philosophical and theoretical debate on human reasoning reviewed earlier, the gap has been increasingly addressed since the early 1990s. There now exist a quite extensive body of research attempting to delineate the neural instantiation of a wide range of diverse aspects of high-level human cognition. However, it remains unclear whether something like mental logic, or mental models could be instantiated in the brain when people perform thinking and reasoning tasks; and if so whether the neural instantiation of thinking and reasoning reflects interactions within the same knowledge domain or between distinct knowledge domains (Kraft et al., 2009).

Most neuroimaging data on human reasoning comes from studies using Positron Emission Tomography (PET) and functional Magnetic Resonance Imaging (fMRI) techniques. Neuroimaging with PET and fMRI affords quantitative mapping of certain physiological and biochemical parameters of the living brain (Goel, Gold, Kapur & Houle, 1997) and yields high-resolution spatial information about where in the living brain specific mental activity might be occurring. However, the resolution regarding timing of neurocognitive processes is much poorer (Rugg & Coles, 1995). Mapping the brain circuitry that is presumed to be associated with some specific cognitive task tells us something about the neuroanatomical structures demonstrably involved in implementing certain functions, and this is a not insignificant part of the story (Coles

& Rugg, 1995). If however, we want to know something about the time course of neurocognitive processes, electroencephalography (EEG) can be more helpful (Luck, 2005). Although EEG has poor spatial resolution (under most circumstances), it has excellent temporal resolution (in milliseconds) making it well suited to investigate the very fast neurocognitive processes underlying mental phenomena such as thinking and reasoning (Gratton, Low, Fabiani, 2008; Ollinger, 2009).

2.3.1 PET AND fMRI STUDIES ON DEDUCTIVE REASONING

The majority of neuroscientific studies on reasoning conducted over the past fifteen years investigated the neural substrates underlying deductive reasoning (i.e., classic conditional deductive logic or classic categorical syllogisms). As will be discussed most of the studies to be reviewed in this chapter were designed to contrast the natural language mental logic account of deductive reasoning (Brain & O'Brien, 1998; Rips, 1994) with that of the model-theoretic account (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991). In a few of the studies the model-theoretic account of reasoning (Bara & Bucciarelli, 2000; Johnson-Laird & Byrne, 1991) is extended to the investigation of the neural underpinnings of inductive reasoning. In general, most of the studies to be discussed can be characterized as having investigated the neural instantiation of one type of mental logic system of reasoning in contrast to one type of mental model system (Acuna et al., 2002; Goel, 2009, 2003; Goel, Buchel, Frith, & Dolan, 2000; Goel & Dolan, 2004, 2003, 2001, 2000; Goel et. al., 1997; Goel, Kapur & Houle, 1998; Goel, Makale, & Grafman, 2004; Goel, Tierney, Sheesley, Bartolo, Vartanian, & Grafman, 2007; Knauff, Fangmeier, Ruff, & Johnson-Laird, 2003; Knauff, Mulack, Kassubek, Salih, & Greenlee, 2002; Noveck, Goel, &

Smith, 2004; Osherson, Perani, Cappa, Schnur, Grassi, & Fazio, 1998; Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997; Rodriquez-Moreno & Hirsch, 2009; Wharton & Grafman, 1998). Relatively fewer studies have examined the neural instantiation of inductive and/or perceptual-based associative systems of reasoning (Christoff et al, 2001; Goel & Dolan, 2004, 2000; Goel et al., 1997; Kroger et al., 2002). Neuroscientific studies of deductive reasoning are reviewed first.

As is evident by the studies cited above, the research program on human reasoning undertaken by Goel and colleagues (Goel, 2009, 2003; Goel et al., 2000; Goel & Dolan, 2004, 2003, 2001, 2000; Goel et. al., 1997; Goel, Kapur & Houle, 1998; Goel et al., 2004; Goel et al., 2007) has figured prominently in the cognitive neuroscience literature, and consequently, comprises the largest proportion of the research reviewed in this chapter. The overarching goal of the program is to investigate the hypothesis of a central seat of reasoning (specifically, of the deductive form), and to examine to what extent the representations and operations underlying deductive reasoning align with a mental logic formal rule governed account of reasoning versus that of a model-theoretic. Solving deductive problems using a mental logic approach is theorized to implicate a neural system built on linguistic representations and syntactic processes whereas using a model-theoretic approach is thought to implicate a neural system built on spatial representations and processes and that is reliant on their inspection and comparison for drawing conclusions (Goel, et al., 1997). Consequently, the researchers hypothesized a left hemisphere (primarily frontal and temporal regions) neural system would be involved in symbolic rule-inference reasoning, whereas a model construction and comparison approach would be supported by a right hemisphere (principally parietal and occipital) system. Overall, results obtained in the studies reviewed below do not lend definitive support to these hypotheses. Instead, taken as a whole, results suggest human reasoning is underpinned by multiple brain and cognitive systems that may be dissociable under certain circumstances, but which are largely integrated. Cortical areas reported in the review that follows will be identified using Broadman classification (BA).

In one of the first PET studies on reasoning, Goel et al. (1997) asked participants to read and evaluate conditional (sentential) and categorical (quantified) arguments of a deductive and inductive kind. In effect, participants were asked to make either a valid/invalid judgment or plausible/implausible one. Conditional deductive arguments were of the form: If I am your father then you are my daughter; I am your father; therefore, you are my daughter. Categorical deductive arguments took the form: No mammoths are huge; All elephants are mammoths; No elephants are huge. Conditional inductive arguments were of the form: If Anita studies she will get an A on the exam; Anita goes skiing instead of studying; Anita does not get an A on the exam. Finally, categorical inductive arguments were of the form: Socrates was a great man; Socrates had a mother; All great men have mothers. The cognitive psychological theories upon which the experimental paradigm was based were mental logic theory (Braine & O'Brien, 1998; Rips, 1994), mental model theory (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991), and schemas theory (Holyoak & Spellman, 1993). Results of the analysis for the deductive condition revealed left hemisphere activations in the inferior frontal gyrus (BA 45, 47) and an area of the superior occipital gyrus (BA 19). Left lateralized activations were also found for the inductive condition, specifically, the medial frontal, the cingulated, and the superior frontal gyri (BA 8, 9, 24, 32). The study results suggest that both deductive and inductive reasoning are supported by the left hemisphere, differing with respect to just one cortical area, the superior frontal gyrus (BA

8, 9). Other frontal cortex activations included the orbital aspect of the left inferior frontal gyrus (BA 45, 47) in the deductive condition, and the left middle frontal gyrus (BA 10) in the inductive condition. Prefrontal cortex (PFC) activations were interpreted as suggesting that the dorsolateral region supports retrieval of rule-based knowledge, while the medial region supports world knowledge (Goel, et al., 1997). The absence of any right hemisphere involvement, as well as the lack of parietal cortex activations in either hemisphere, failed to provide support for the prediction that a Johnson-Laird mental model account of reasoning would require involvement from association cortices and/or right hemisphere structures. Instead, the results of the study suggested that many of the algorithms underpinning human reasoning could be accounted for within the language-like formal reasoning paradigms (Braine & O'Brien, 1998; Rips, 1994).

While the PET study by Goel and colleagues employed only linguistic stimuli, Prabhakaran et al. (1997) used the Raven's Advanced Progressive Matrices (RAPM) in an fMRI study to investigate the neural substrates of "fluid" reasoning (i.e., online reasoning). This study continues to be of interest to researchers because it was one of the first to discuss and confirm the nature of involvement of the frontal cortex in reasoning and to differentiate cortical areas involved in working memory relative to cognitive domain. Participants viewed and solved three types of problems: analytic (abstract/analytic processes), figural (visual-spatial processes), and pattern matching (which served as a control task for perceptual-motor processes) while their brain activations were recorded. The analytic and figural problems on the RAPM are said to require the use of two different problem reasoning algorithms, the former applying an "analytic algorithm" to features contained in the problems and the latter a "Gestalt algorithm" using visual perception processes, i.e., continuation or superimposition (Hunt, 1975). The match problems required that participants visually inspect and match figures and were used in the analyses as a control for perceptual-motor activities. The researchers were interested in examining three questions: the degree of left hemisphere involvement when solving analytic problems compared to figural, the amount of right hemisphere involvement when solving figural compared to pattern matching, and the amount of frontal activation for analytic compared to figural. Findings indicated broad but distinct networks associated with analytic and figural reasoning. Solving the analytic problems was hypothesized to engage all cortical areas involved in abstract/analytical reasoning beyond perceptual motor demands. Results revealed bilateral middle and inferior frontal gyri, premotor areas (BA 6, 9, 44, 45, and 46), and in the right rostrolateral prefrontal cortex (BA 10). Extensive activations were found in the parietal cortex in the superior and inferior parietal regions, and in supramarginal and angular gyri (BA 7, 39, and 40). Also, major foci of activity were seen bilaterally in the inferior and middle temporal gyri (BA 19, 21, and 37). The figural problems were predicted to engage cortical regions involved selectively in figural reasoning because no analytical reasoning was required to solve the problems. Solving these problems was presumed to require mostly visuospatial processing. The observed cortical activity was predominantly right lateralized revealing significant, frontal, temporal, and occipital lobes activations. Left-lateralized parietal activations were also preserved. Other left hemisphere activation was observed in the inferior and superior parietal regions (BA 7 and 40). Greatest activity was seen in the right middle frontal gyrus (BA 9 and 46). Other right hemisphere activations included anterior cingulated (BA 32), superior and inferior parietal gyri (BA 7 and 40), the inferior temporal gyrus (BA 37), and the precuneus (BA 7 and 19).

When solving both the analytic and figural problems, participants were obliged to process the same figural (imagistic) stimuli. For this reason the researchers used subtraction analysis (i.e., subtracting activations evoked by figural reasoning from those related to analytic reasoning) to isolate activations uniquely attributable to analytic. The analytic/figural analysis revealed activations in bilateral middle & inferior frontal (BA 6, 9, 44, 45, and 46), posterior cortex predominantly left inferior and superior parietal, angular, and supramarginal gyri (BA 7, 39, and 40); inferior and middle temporal (BA 37, 21, and 19); and medial occipital regions (BA 18, 19, and 37). To summarize, areas activated in the analytic task were very nearly a composite of the activations found in response to performing the figural and analytic-figural tasks, indicating that given the stimuli employed processes involved during figural and analytic reasoning were recruited to solve abstract analytic problems.

Prabhakaran et al. (1997) contended that many of the activations found in their study were consistent with previous reports of cortical areas involved during performance of working memory tasks. Specifically, right hemisphere middle (BA 9, 46) and inferior (BA 44, 45) frontal gyri, and the premotor regions found in figural reasoning were said to reflect visual-spatial working memory. By comparison, the left middle (BA 8, 9, 46) and inferior (BA 44, 45) frontal gyri and premotor (BA 6, 44) cortex activations evoked in response to analytic reasoning were associated with semantic memory and working memory for digits and phonological information. Right frontal activations found for analytic reasoning were greater and more expansive than those found in figural reasoning, and were attributed to resources needed to support the more abstract processes predicted to be necessary for analytic reasoning.

In addition to the frontal cortices, figural reasoning selectively recruited from posterior regions (parietal, occipital and temporal gyri) identified with rehearsal and retention of visual-spatial information in working memory (Prabhakaran et al., 1997). Significant activations were observed in bilateral inferior and superior parietal regions (BA 40, 7). Superior parietal areas were respectively suggested to indicate object working memory (left side activations) and working memory for spatial location (right side activations). The bilateral inferior parietal responses were thought to reflect the directing of visual-spatial attention and/or visual-spatial manipulation of stimuli. Temporal activations occurred largely in the right inferior and middle gyri (BA 19, 21, and 37), and were attributed to object transformations necessary to generate a representation of the missing form. In the occipital cortex, bilateral activations were found in the precuneus (BA 7), an area implicated in visual imagery and shape and spatial location. Occipital areas (BA 18, 19) in the right hemisphere also showed activation, attributed to working memory for spatial location. Although activations in the figural condition were predominantly right lateralized, the presence of some left hemisphere activations were thought to possibly reflect transformational demands of visual-spatial reasoning such as rotation.

Activations observed beyond the frontal lobes during analytic reasoning, were attributed to the recruitment of varying and different verbal working memory resources (Prabhakaran et al., 1997). Complementary to posterior activations found in figural reasoning, those evoked in analytic reasoning showed asymmetrically greater activity in left than right inferior and superior parietal, angular and supramarginal gyri (BA 7, 39, and 40), inferior and middle temporal gyri (BA 37, 21, and 19), and in precuneus and medial occipital gyri (BA 18, 19, and 37). Given the involvement of the middle temporal region, the authors suggested temporal activations may have been activated by working memory demands for complex non-verbal objects such as would be required to perform the type of feature mapping between elements required by the analytical task (Prabhakaran et al., 1997).

The finding of a lateralized trend in activations for analytic and figural reasoning respectively, supported the authors' first hypothesis, namely, that analytic reasoning would be underwritten by a primarily left hemisphere cortical network and figural reasoning by a predominantly right hemisphere neural complex (Prabhakaran et al., 1997). The study's second hypothesis concerned the comparative involvement of the prefrontal cortex (PFC) in analytic versus figural reasoning. The authors found that both tasks engaged the PFC, but that there was greater involvement of anterior regions relative to the posterior cortex in analytic reasoning compared to figural reasoning. Given association of the rostra PFC (BA 10) with non-routine selection of cognitive strategies and bilateral activation of the dorsolateral PFC (BA 9, 46) with performance of two tasks at once (dual tasks management) the authors suggest that the greater and bilateral involvement of dorsolateral PFC in analytic reasoning may be related to its dual management of figural processing and analytic analysis.

The quite different results obtained from the Goel et al. (1997) and Prabhakaran et al. (1997) studies serve to illustrate at least one important issue with respect to neuroscientific investigations of cognitive activity: the reasoning paradigm, task and materials can significantly impact on findings. Follow-on studies attempted to clarify some of these effects. Osherson et al. (1998) investigated the neural substrates of deductive and probabilistic reasoning basing their experimental manipulations on mental logic (Braine & O'Brien, 1998; Rips, 1994) and mental model (Johnson-Laird & Byrne, 1991) frameworks of reasoning. Participants were scanned in three conditions: logic, probabilistic, and meaning, during which they read and reasoned about formally equivalent deductive arguments (*sentential and quantified*). In the logic condition participants reached a valid/invalid conclusion, in the probabilistic condition they were asked to decide if the conclusion was more likely to be true than false assuming the premises to be true, and in the meaning condition they responded as to whether the sentences contained semantic anomalies. Surprisingly, given the similarity to the Goel et al. (1997) study, these authors reported a predominantly bilateral posterior distribution of activations, with right hemisphere dominance for deductive reasoning. Probabilistic reasoning evoked a prevalently frontal distribution of activations, with significant activation in two left dorsolateral prefrontal regions (BA 8, 10). Convergent with the thesis postulated by Goel et al. (1997), these researchers concluded that different neuronal populations supported deductive and inductive reasoning, but in contrast to their view suggested that the right hemisphere was responsible for deductive reasoning, whereas probabilistic reasoning seemed to require prefrontal involvement, possibly left lateralized (Osherson, et al., 1998).

Goel (1998) examined the use of categorical syllogisms in deductive reasoning using PET. Participants were scanned in four conditions while reasoning about three kinds of deductive reasoning problems differentiated by content (categorical syllogisms, spatial relational, and non-spatial relational problems), and a baseline condition in which they were instructed to read and comprehend the semantic content of a set of three sentences. Spatial relational arguments specified the relative spatial location of entities, such as, “Officers are standing next to generals”. Nonspatial arguments described a non-spatial attribute of the entities, such as, “Officers are heavier than generals". Findings from this study replicated those on deductive reasoning from Goel et al. (1997). Importantly, the comparison of activation patterns for the spatial and non-spatial conditions revealed the two relational conditions to be identical (Goel et al., 1998), and localized to the left hemisphere. Thus, Goel and colleagues (Goel et al., 1998) claimed that human reasoning is supported by left hemisphere brain structures, and that deductive reasoning was underpinned by a distributed computational network involving the inferior dorsolateral and medial prefrontal cortices, and the medial temporal lobe. The researchers offered a preliminary functional interpretation of the data based on reports from other fMRI and neuropsychological studies, namely, that the dorsolateral (BA 9, 46) involvement reflected management of working memory processes, the inferior frontal gyrus (BA 45, 47) supported language related processing, and the medial prefrontal cortex (BA 10) indicated executive attention processing (Goel et al., 1998). In summary, the authors proposed human reasoning was a language-based process guided by a central executive function (Goel et al., 1998).

Wharton & Grafman (1998) published a meta-review of neuroscientific and neuropsychological studies of deductive reasoning in the brain, in which they claimed support for two distinct systems of deductive reasoning differentiated by the nature of information being processed. Specifically, they proposed that a neuronal system prescribed within the left-hemisphere enabled content-independent, domain-general reasoning, whereas a right lateralized neural system in conjunction with bilateral ventromedial frontal cortex mediated content-dependent reasoning. They speculated that these two systems must interact in inference making.

The hypothesis of two neural systems underwriting reasoning was investigated in three further studies (Goel et al., 2000; Goel & Dolan, 2001; Goel, 2003). In all of the experiments, participants were asked to solve either three-term syllogistic or conditional arguments. Goel et al. (2000) scanned participants while reading and evaluating syllogisms composed with concrete content (e.g., All dogs are pets . . .) versus abstract content free sentences (e.g., All P are B . . .). Results of the study revealed two dissociable systems, running contrary to implications based on mental logic and mental model theoretical accounts which predict that either a syntactic or visual-spatial based system was necessary and sufficient for deductive reasoning (Goel et al., 2000). One of the observed systems was defined as a left hemisphere fronto-temporal linguistic complex and interpreted to support processing of arguments composed with concrete semantically meaningful content. This system was congruent with activations observed in Goel, Gold, Kapur & Houle (1997) in which, participants also reasoned about syllogisms composed with concrete semantically meaningful content. A similar left hemisphere system of activations had also been found in Goel, Kapur & Houle (1998) during reasoning about syllogisms describing both spatial and nonspatial relations between entities. However, a second bilateral occipital-parietal-frontal system functionally defined as supporting visual-spatial processing was also identified in the Goel et al. (2000) study. This system was associated with reasoning about the abstract semantic free sentences. Identification of this system in combination with the left hemisphere system argued for two distinct neuronal systems underlying deductive reasoning.

The new finding of a visual-spatial network (Goel et al., 2000) was further examined in Goel & Dolan (2001). The follow up study was designed to examine the nature of the spatial relations between entities described in the premises of three-term deductive arguments by comparing premises describing spatial relations with those devoid of spatial information. Premises were further distinguished on the basis of containing meaningful content and abstract meaningless content. For example, a meaningful content sentence describing spatial relations is of the kind, Larry is above Michael. A meaningful content nonspatial sentence is, Larry is heavier than Michael. Meaningless content spatial sentences were of the form, L is behind K; and a meaningless content nonspatial sentence took the form, L is heavier than K. However, a similar bilateral parietal-occipital system was activated in both conditions with activations being greater in the semantic free condition. Evidence from this experiment did not support the hypothesis of two dissociable systems.

Consequently, the aim of Goel (2003) was to examine the extent to which reasoning problem type (e.g., prior knowledge, concrete, content free) differently affected brain activation. Therefore, the study conducted by Goel (2003) was specifically designed to investigate the hypothesis that the nature and type of content being reasoned about might alter the composition of evoked neuronal populations. The overall hypothesis was that reasoning about semantically meaningful concrete content (e.g., All dogs are pets . .) about which participants also held prior beliefs, may differ from reasoning about semantically meaningful content that describes spatial relations between concrete entities (e.g., Adam is ahead of Bob . . .), such as had been the case with the stimuli employed in Goel (2001). In Goel (2003) participants were scanned in one condition while they evaluated three-term sentential, semantically meaningful arguments describing situations about which they both had prior knowledge and held beliefs, and a second condition in which they evaluated formally same arguments describing meaningful situations about which they had prior knowledge, but that was belief neutral. Evidence from the study supported the hypothesis that beliefs affected reasoning. The belief-laden content activated anterior left middle temporal gyrus (BA 21) and belief neutral content activated left superior parietal cortex (BA 7). In summary, the results from Goel (2003) provided additional support for the dual system model of deductive reasoning implicated in earlier studies (Goel et al., 2000), and in addition suggested the two systems could be further distinguished on the basis of a belief factor. The main findings in Goel (2003) were corroborated in a study by Noveck et al. (2004), in which one left hemisphere occipital, temporal-parietal system (BA 6, 7, 19, 37, 44) was engaged when reasoning about *modus ponens* conditional syllogisms, and a second left hemisphere parietal-frontal (BA 6, 7, 47) system inclusive of the cingulate gyrus (BA 32) was activated when reasoning about *modus tollens* syllogisms. The additional presence in this study of activations in the cingulate gyrus and the right inferior/middle prefrontal cortex motivated subsequent investigations of the modulating effect of conflict resolution in deductive reasoning.

Together the three studies (Goel et al., 2000; Goel & Dolan, 2001; Goel, 2003) investigated the proposition of dual reasoning networks within the context of deductive reasoning using premises composed of sentences with concrete and abstract, semantically meaningful and semantically unmeaningful, and belief laden and belief neutral content. In summary, the studies found evidence for the engagement of linguistic and spatial mechanisms in reasoning. In reviewing the findings from the three studies, the authors offered a new paradigm claiming dual neural pathways underpinned deductive reasoning. Specifically, in Goel (2003) it was proposed that the frontal-temporal system supported reasoning with familiar, conceptually coherent content, and the parietal system was specialized for unfamiliar, non-conceptual or incoherent material. They suggested that the frontal-temporal network corresponded with heuristic based reasoning processes, and the parietal network with formal processes of reasoning. This division, they concluded better aligned with the heuristic and formal dual mechanisms account of problem solving described by Newell & Simon (1972) than it did either a mental logic or mental model account.

The role of visual and spatial content in deductive reasoning was examined in several of the studies reviewed above (Goel et al., 1997; Goel & Dolan, 2001 Prabhakaran et al., 1997), but it was the primary focus of studies by Knauff et al. (2002) and Knauff et al. (2003). In the first study the researchers compared relational and conditional deductive reasoning problems. Relational (i.e., spatial relational) inferences were of the kind: left of, right of, overlaps, from the left, overlaps from the right, and so on. Conditional problems were either modus ponens or modus tollens (e.g., If the man is in love, then he likes pizza; The man is in love; The man likes pizza. Participants listened via headphones to the sentences being read aloud, thus received no visual input. Two important findings of the study were that reasoning in both conditions elicited activations in a bilateral occipital-parietal pathway, (greater activation in relational than conditional reasoning) even though the participants did not view any stimuli, and a bilateral parietal-frontal network. Knauff et al. (2002) found activations in prefrontal cortex (BA 6, 9) and the cingulate gyrus (BA 32), the inferior and superior parietal cortex (BA 7, 40), the precuneus (BA 7), and the visual association cortex (BA 19). Referencing the well-studied model of working memory proposed by Baddeley & Hitch (1974) and Baddeley (1986) in which it was proposed that working memory was composed of modal specific stores and a central processor, the authors interpreted the occipital-parietal pathway activations to reflect visual-spatial storage and manipulation of spatial representations used to evaluate both forms of arguments presented in the experiment. Based on the same memory construct, parietal-frontal activations found in the study were interpreted as working in concert with the occipital-parietal pathway, and to specifically support maintenance and integration of functions. Consequently, the authors concluded that, in their study, reasoning employed spatial representations and therefore provided evidence for a Johnson-Laird model-theoretic account of deductive reasoning. Further evidence for the use of mental models in reasoning was provided by Knauff, Fangmeier, Ruff & Johnson-Laird (2003), in a follow-up study that examined the neurocognitive processes underlying visual as compared to spatial imagery in deductive reasoning. The purpose of the study was to examine the relationship between content of reasoning material and form of representation, specifically, verbal, visual and spatial imagery. The researchers introduced four types of stimuli in the new experiment: visual ("The dog is dirtier than the cat"), spatial ("New York is north of Atlanta"), visual-spatial ("The shelf is above the table"), and a baseline non visual-spatial form ("The cat is smarter than the dog"). Study results confirmed findings from Knauff (2002), but provided greater specificity regarding the type of representations used in modeling different argument content. Reasoning in general evoked activity in the left middle temporal gyrus, in the right superior parietal cortex, and bilaterally in the precuneus. Prefrontal activity was observed in all problem types in the middle and inferior frontal gyri. Only that reasoning based on visual relations evoked additional activity in visual association cortex (V2). The authors concluded, the visual relations conditions induced the use of visual models and the three other conditions induced models that were spatial and/or abstract in form (i.e., not visual). Importantly, in parsing the use of models with visual content (e.g., "A rose is prettier than a daffodil") from models using a spatial array ("New York is north of Atlanta") the researchers concluded that problem content inducing visual images automatically triggered the construction of additional and unnecessary models for problem solving (Knauff, 2003). Therefore they theorized that the added mental effort to consider and then suppress these visual images actually interferes with normal processes of inference making.

Goel, Makale, & Grafman (2004) replicated earlier studies by Goel and colleagues (Goel et al., 2003; Goel et al., 2000) in order to investigate whether the proposed dual system model of syllogistic deductive reasoning extended to conditional relational reasoning, and also to investigate the effect of reasoning about familiar and unfamiliar spatial relations. Participants were presented with three-term conditional arguments describing spatial relations between places, both familiar and unfamiliar. Familiar arguments (meaningful content) referred to landmarks and took the form: Paris is south of London; London is south of Edinburgh; Paris is south of Edinburgh. Unfamiliar arguments (meaningless content) were of the same logical form and informational content, but described fictional places: The AI lab is south of Roth Center; Roth Center is south of Cedar Hall; The AI lab is south of Cedar Hill. Asking participants to reason about the spatial relations of concrete familiar and unfamiliar situations allowed the researchers to test the hypothesis that the neural networks defined in Goel (2003) underwrite heuristic and formal inference processes, respectively. Drawing conclusions about the arguments describing familiar situations elicited activations in bilateral inferior and middle occipital cortices (BA 18, 19), right inferior temporal gyrus (BA 37), and bilateral (R > L) posterior hippocampi. Reasoning about the unfamiliar situations revealed a network consisting of bilateral superior parietal lobule (BA 7) and bilateral (L > R) superior frontal gyrus (BA 6). Left hemisphere activations found in response to reasoning about familiar landmarks were interpreted as a frontal-temporal lobe linguistic system, whereas right hemisphere activations elicited when reasoning about unfamiliar landmarks were defined as a bilateral frontal-parietal visual-spatial processing network. Thus, findings from the study provided support for the dual systems model, as well as its extension to included reasoning about conditional arguments. Further, the finding of two systems corroborated results from Goel (2003). Taking the totality of findings from this and previous studies into consideration, Goel et al. (2004) concluded that deductive reasoning is not underwritten by a neuroanatomical "logic module", but instead by reasoning systems that are dynamically configured in response to varying demands forthcoming from our interactions with the world (Goel et al., 2004).

More than a decade of cognitive neuroscience studies on deductive reasoning has produced a wide range of findings regarding its neural underpinnings. As stated earlier study outcomes vary depending on research question, task manipulations, modality and content of stimuli, methodology (baseline and timing), participant population, and no doubt other factors not discussed. Thus, differences across the reported findings possibly reflect differences in the neural encoding of different types of representations and associated binding and transformation processes (Barbey & Barsalou, 2006; Ilmberger, 2009). Overall, there is evidence in support of both mental logic and mental model accounts of reasoning, without decisive conclusions emerging in favor of one or the other (Goel, 2009; Kraft et al., 2009). Instead, there appears to be emerging evidence for more than one dynamically configured underlying neural system, suggesting that neither a mental logic nor mental model theoretical paradigm sufficiently reflects what likely occurs at the neural level (Barbey & Barsalou, 2006; Ilmberger, 2009). Currently, Goel (2009) proposes (at least) three pairs of neural systems as crucial for human reasoning, each dissociable, and responsive to different perceptual input and task demands. One proposed pair of systems underpins reasoning with familiar and unfamiliar content (Goel, 2009). The second pair of systems is postulated to support reasoning about situations in which beliefs influence performance outcomes. This pair has been labeled the Belief-biased and Conflict systems (Goel, 2009). The third pair of systems subserves reasoning with certain and uncertain information, and is distributed in the left and right hemispheres, respectively. The type of well-formed closed problems represented by classic conditional arguments (*modus pollens*) are an example of reasoning with information that is certain, whereas reasoning with uncertainty is instantiated by more open-ended indeterminate arguments whose premises are ambiguous, and/or do not contain sufficient information to reach a conclusion.

2.3.2 PET AND fMRI STUDIES ON INDUCTIVE REASONING

Most of the research summarized so far has focused on brain imaging studies of deductive reasoning. There have been far fewer studies to directly examine paradigms of the inductive class of reasoning (logically indeterminate, probabilistic, similarity-based comparison tasks). As previously discussed, inductive processes are applied to reasoning about relatively unstructured, open-ended problems, and the type of inferences yielded are estimates of plausibility or reasonableness, rather than discrete judgments. Tasks requiring an assessment of similarity between entities and their properties, hypotheses generation and selection, or estimation of causal probability based on relations of temporal and physical contiguity are typically considered inductive. Inductive inferences are considered to be derived from the semantic relations of content, rather than the formal features of material (Rescher, 1980). Processing the semantics of information is thought to require a deeper level of encoding than feature-based encoding (Craik & Lockhart, 1972; Rescher, 1980).

Three of the studies already reviewed offer insights into the neural basis of inductive reasoning processes. Goel, et al. (1997) and Osherson et al. (1998) examined the neural basis of inductive reasoning as compared to deductive reasoning, and Prabharakan (1997) examined the neural basis of gestalt reasoning in a comparison with analytic reasoning. Overall, the three studies found evidence for two dissociable systems underlying deductive and inductive reasoning. The first two studies (Goel et al., 1997; Osherson et al., 1998) reported evidence for left hemisphere neural systems underpinning both the inductive and deductive forms of reasoning. Prabhakaran et al. (1997) found evidence that left and right hemisphere networks supported reasoning, but that the more formal, verbally based, structural mapping procedures were largely underwritten by a predominantly left-lateralized network and visual-spatial reasoning by a more right-lateralized complex. However, and overall, study findings showed that both categories of reasoning tasks engaged the PFC, with greater involvement of anterior regions relative to the posterior cortex in analytic reasoning compared to figural reasoning.

Two additional studies using tasks specifically designed to target inductive reasoning were conducted by Goel & Dolan (2000, 2004). In the first study, the authors compared a rule induction versus rule application paradigm; and in the second, determinate versus plausible reasoning. Results from the first study (Goel & Dolan, 2000) revealed bilateral (but stronger on the right side) activations in the parietal and prefrontal cortices for rule induction, which also involved activations in the hippocampus bilaterally (L > R). Similar bilateral activations of the parietal and prefrontal cortices (without hemispheric differences) were found for the rule application task, but hippocampal activations were not observed. The authors theorized the greater left hippocampi involvement in the inductive rule inference task was related to semantic encoding demands required to abstract similarities across the visual stimuli in order to infer a categorization rule (Goel & Dolan, 2000). This finding provides evidence at the neural level for models of cognition proposing that associative reasoning results in deeper more robust processing of information than rule application inference making (Craig & Lockhart; 1972).

In the second study, Goel & Dolan (2004) revisited the hypotheses established in their earlier PET study (Goel et al., 1997), but this time used fMRI to examine brain activity. The researchers were specifically interested in investigating the role of prefrontal cortex (PFC) in supporting inductive versus deductive reasoning. Outcomes from this follow up fMRI study supported their earlier claim (Goel, et al., 1997) that inductive and deductive reasoning engaged different and dissociable neural systems, but activations were seen bilaterally as compared to left-lateralized. While the bilateral activations found in determinate and plausible reasoning were similarly distributed, they differed with respect to the amplitude of the cortical areas activated. Specifically, determinate (i.e., deductive) reasoning showed greater activation in left inferior frontal gyrus (Broca's area, BA 44) than plausible (also, categorized as inductive) reasoning, and plausible reasoning elicited greater activity in left dorsolateral PFC (BA 8, 9) compared with determinate. The authors posited that the involvement of Broca's area in deductive reasoning reflected the verbal working memory and syntactic processing demands of deduction in contrast to induction (Goel & Dolan, 2004). Greater involvement of the dorsolateral PFC cortex in inductive inference-making was attributed to resource demands needed to access prior or world knowledge during hypothesis generation and evaluation (Goel & Dolan, 2004). Based on outcomes from this study the authors proposed the neural systems underlying the deductive and inductive reasoning were dissociable within the prefrontal cortex (Goel & Dolan, 2004).

The convergent, if varied, findings of prefrontal cortex activations across the studies reviewed in this chapter (Goel et al., 1997, 1998, 2000; Prabhakaran, et al., 1997; Osherson et al., 1998; Knauff et al. 2002, Knauff et al., 2003) are consistent with the established view of the primary role played by this cortical region in functions of executive control (Kroger et al., 2002; Christoff, 2009). In the context of higher-order cognition like reasoning, the term *executive control* often refers to the "top-down" processing involved in activating, coordinating and guiding distributed cognitive resources in support of goals and/or plans of action (Baddeley, 1996; Miller & Cohen, 2001). In other words, whenever one must mentally and simultaneously form and hold multiple representations, and manipulate them (or their features and properties) to derive new knowledge or do something, this involves executive control systems. Consequently, one would expect to see the involvement of the PFC whenever a person is confronted with navigating a new situation, engaged in contemplating a knotty issue, making plans and/or undertaking a novel task, and when engaged in tasks such as those used in the laboratory experiments discussed in this chapter (Miller & Cohen, 2001). Prabhakaran et al. (1997) found that more complex and difficult RAPM (Raven's Advanced Parametric Matrices) problems activate more anterior dorsolateral PFC, and Goel & Dolan (2000), reported activations in the right lateral orbital PFC (BA 47, 11) relative to difficulty in hypotheses evaluation. Several fMRI studies (Christoff, 2009; Christoff & Gabrieli, 2000; Christoff, Ream, Geddes, & Gabrieli, 2001; Christoff et al., 2003; Kroger et al., 2002) have specifically explored the role of the PFC in mediating executive control in reasoning. The common purpose of these studies was to determine differential roles for different areas of the PFC – specifically, the dorsolateral prefrontal cortex (Kroger et al. 2002) and the rostrolateral prefrontal cortex (Christoff & Gabrieli, 2000; Christoff, et al., 2003; Christoff et al., 2001). The four studies posited similar working hypotheses, which were that, the lateral PFC was uniquely recruited in high levels of problem complexity, and moreover, subregions of the lateral PFC were selectively activated by factors of difficulty and complexity. All of the authors theorized that ambiguity regarding the role of the lateral PFC in reasoning resultant from the varied findings across studies might be explained by differences in the levels of difficulty and/or complexity of tasks employed in the studies (Christoff et al. 2001; Christoff et al. 2003; Kroger et al. 2002). The authors used experimental manipulations that allowed them to compare monotonic changes in task difficulty and/or complexity. Because the studies used stimuli or variants of material from the RAPM, findings do provide some additional insight into the neural substrates of a perceptual-based associative reasoning system, as well as into the functional organization of the prefrontal cortex. Findings were convergent with those from Kroger (2002), specifically, that lateral prefrontal and parietal activations accompanied increasing relational complexity. Outcomes from this series of studies suggest visuospatial relations are represented and manipulated in posterior cortical regions under the control of prefrontal cortex. Specifically, dorsolateral PFC (BA 7, 9, 46) was activated by increasing relational complexity as compared with problem difficulty. Activations in this region extended into the more anterior ventrolateral (BA 10) area and became more left lateralized at the highest complexity levels. While increasing problem difficulty also activated DLPFC, the area of activation was less left lateralized and more posterior (BA 7). These results align with the organizational paradigm for the lateral PFC: RLPFC (middle and superior frontal gyri, BA 10), dorsolateral PFC (middle frontal gyrus, BA 9, 46), and ventrolateral PFC (BA 45, 47) proposed by Christoff and colleagues (Christoff, 2009; Christoff & Gabrieli, 2000) to explain the totality of their own findings and those of others regarding preferential involvement of this large cortical region.

The work of Christoff and colleagues (Christoff, 2009; Christoff & Gabrieli, 2000; Christoff, 2001; Christoff et al., 2003) was used as the basis for development of an organizational framework describing the functional role of the lateral prefrontal cortex. The framework is based on findings from their own studies cited above, but also other reasoning and problem solving research using tasks such as formal inductive logic problems, Tower of London, Wisconsin Card Sorting, and the Raven's Progressive Matrices. The proposed functional architecture unifies findings from across the many studies by categorizing activations as associated with either goal-directed or spontaneous thought, where goal-directed thought is defined as that needed to execute plans and goals, and spontaneous thought as thought akin to daydreaming or insight. At a general level the framework specifies a topographical layout of subregions of the lateral prefrontal cortex that functionally correspond with a concrete to abstract gradient of cortical representation of information. Restated, mental representations within the lateral PFC are hypothesized to be instantiated within subregions along an anterior-to-posterior pattern of organization with the most abstract level of information represented in the rostrolateral PFC (BA 10), less abstract externally oriented representations distributed in dorsolateral PFC, and concrete object representations managed in the ventrolateral PFC. Christoff (2009) and Christoff & Gabrieli (2000) propose that such a topographical gradient of representations underlies thought in most domains of higher-level cognitive functions (i.e., reasoning, working memory, and long-term memory. In functional terms, the rostrolateral PFC (BA 10) is recruited by processes associated with self generated representations, and thus is implicated in cognitive activities like task set, planning, manipulations of goal hierarchies, hypothesis generation and evaluation, and monitoring correspondence of activities with plans and goals (Christoff & Gabrieli, 2000). By comparison the dorsolateral PFC is involved during monitoring and manipulation of online information that is externally generated, which is to say of a more concrete nature (Christoff, 2009). Importantly, all of these psychological processes are required in high-level complex reasoning tasks such as those in the studies reviewed in this chapter. Similarly, Christoff & Gabrieli (2000) and Christoff (2009) suggest that, because the reasoning tasks employed in the studies they reviewed required processing both concrete (e. g., strings or forms) and self-generated (e.g., plans and goals) representations, activations were found in both the rostrolateral and dorsolateral cortices. (Christoff, 2009) proposes the reported activations are best interpreted within the framework of a three-level hierarchical processing distinction between areas of the lateral PFC: with the rostrolateral cortex primarily engaged when evaluation of abstract, internal, self-generated representations is required, the dorsolateral cortex involved in both concrete and self-generated evaluative processes, and the ventrolateral PFC specialized for online processing of object-like representations (Christoff, 2009). Thus, the most anterior area of the lateral PFC supports internal representations and processes, the dorsolateral PFC is involved when internal self-generated representations and external representations are evaluated, or when processing of external and object representations is required. The more posterior ventrolateral PFC would be involved along with either or both of the other two levels whenever object evaluation is required. Taken as a whole, PET and fMRI research on reasoning has found little or no support at the neuroanatomical level for the existence of a singular or central reasoning system as would be implicated by either the mental logic or mental model theory of reasoning. Indeed, the emerging theme based on the literature reviewed here is that human reasoning is supported by multiple broadly distributed neural networks formed in a responsive and dynamic manner (Barbey & Barsalou, 2006; Kraft et. al., 2009). This view is congruent with theoretical perspectives that conceive of high-level cognition as riding on multiple coordinated networks comprised of embedded, distributed, functionally specialized neuron populations (Edelman, 1987; Fuster, 2003; Miller & Cohen, 2001).

It is possible that the neurons participating in a given network underlying a given task also depend on demand over the time course of the task (Badre, 2008). This is an aspect of functional connectivity that remains difficult to explore using PET and fMRI imaging techniques. The time course of brain activity in response to demands of reasoning has been addressed in one fMRI study (Rodriquez-Moreno & Hirsch, 2009) in which the stages of information processing during deductive reasoning were investigated. Time course of the signal was examined by segregating activations associated with three stages of the reasoning process, i.e., premise encoding, premise integration, and conclusion validation. Consistent with findings in other studies on reasoning, a broad frontal parietal network (FPN) of activations was observed (Rodriquez-Moreno & Hirsch, 2009). The overall functional interpretation of the findings was also generally in agreement with previous authors, that is, reasoning is subserved by modality-specific components (bi-lateral parietal) in coordination with a supramodal reasoning network comprised of frontal-parietal-caudate areas (Rodriquez-Moreno & Hirsch, 2009). The researchers analyzed the amplitude of activations in cortical areas at different time points of problem solving for each of the three stages categorical syllogisms (Ex: Every politician recycles glass bottles; People who recycle glass bottles like wildlife; Every politician likes wildlife) to determine whether specific areas of the FPN were differentially engaged as the reasoning process unfolded. Further analyses focused exclusively on the time frame from integration of the second premise through validation of the conclusion (the second and third stages), said to constitute *reasoning.* The two stages were further decomposed, that is, *reasoning* was then defined as composite of four steps: encoding of premise two, integration of premise two with premise one, generation of the participant's conclusion, and validation of the conclusion. These processes were assumed to progress in a serial manner and thus to have different temporal courses, with encoding of premise two at the beginning of the magnetic resonance signal peak, premise integration and conclusion generation occurring towards the middle of the peak, and validation of the conclusion at the end. By segregating increases and decreases of the peak over this timeline the authors were able to determine for example, that middle frontal area (BA 8/6) is primarily active during encoding of premise two and integration of premise two with one, that left superior and middle frontal cortices (BA 6 and 8), and left inferior parietal regions (BA 40, 39 and 7) underwrite the same cognitive activities but remain active throughout the conclusion generation and validation phases. Left frontal regions of the middle (BA 9 and 10), medial (BA 8), inferior (BA 47) frontal gyri as well as bilateral caudate are additionally recruited during the conclusion phase (Rodriquez-Moreno & Hirsch, 2009). Results from this novel study demonstrate the contribution to be made from data on the temporal aspects of neural information processing toward the goal of developing a richer understanding of how functional specialization within the type of broadly distributed networks found to underlie reasoning might be instantiated and coordinated (Gratton et al., 2008). At the same time, the inability to further disentangle the neural processes underlying stages of premise integration and inference making using fMRI research methodology underscores the need for a different methodology much closer to the neural signal (Rodriquez-Moreno & Hirsch, 2009).

2.3.3 EEG STUDIES

Data from PET and fMRI studies reviewed in the previous section are beginning to provide important insight about the cortical areas engaged when performing different types of reasoning tasks. If however, we are to one day understand how reasoning is instantiated in the brain it will be important to integrate this spatial data on human information processing with temporal data about the brain activations thought to underpin putative cognitive activity. The ability to examine how regional areas (inter- or intra-) interact as cognitive processes unfold has the potential to augment the ongoing discussion on the functional organization of the human brain (e.g., modularity, connectionism, dynamic interplay, etc.).

This EEG literature review only found one electroencephalography research program on human reasoning (Qiu et al., 2007). The study examines the electrophysiology underlying conditional relational reasoning employing four different types of deductive arguments: modus ponens (MP), modus tollens (MT), denial of the antecedent (DA), and affirmation of the consequent (AC). High-density (64 scalp sites embedded in an elastic cap) EEG recordings were made while participants read and solved three-term deductive reasoning and baseline problems. The baseline manipulation required participants to engage in memory retrieval activities, but not to integrate the premises. Premise statements in all five conditions described visual-spatial content using words for color, or geometric form, or spatial layout (Ex: If the figure is a square then it is red; It is a square; So, it is red *or not red*). The researchers found that conditional reasoning tasks (MP, MT, DA, AC) elicited a more negative ERP deflection in the left fronto-central region in the 500-800ms and 1400-2500ms time windows compared to the baseline task. Dipole analysis located a generator of the early deflection, referred to as N600 by the authors, to the left dorsal anterior cingulated gyrus (ACC), suggesting that left frontal–central areas are involved in inferential processing. The earlier time window (500-800ms) was selected to correspond with cognitive activities following onset of the minor premise (2nd premise) but prior to onset of the conclusion, thus the authors suggest that ACC activity was related to memory encoding and retrieval processes enabling the application of inference rules to premise integration. This interpretation was said to be consistent with findings from fMRI studies regarding involvement of the ACC in deductive reasoning tasks (Goel et al. 2000; Ruff, Knauff, Fangmeier, & Spreer, 2003). Activations recorded over the left hemisphere were also interpreted to be consistent with fMRI study findings on reasoning in which left lateralized activity was said to indicate execution of functions of formal logic manipulation (Deglin & Kinsbourne, 1996, as cited in Qiu, 2007). Dipole source analysis for the second waveform corresponding with cognitive activities following onset of the conclusion, showed a generator in the right ACC. Right lateralized ACC activity during late time window (1400-2500 ms) was interpreted as indicating greater recruitment of cognitive resources associated with validating the conclusion, something not required in the baseline activity.

In contrast to research on reasoning, there have been several electroencephalographic studies conducted on mental rotation. The research undertaken for this dissertation employed dynamic rotating stimuli, thus a brief review of the selected rotational studies may provide relevant information about the ERPs previously associated with this cognitive function. Investigation of the electrophysiology of mental rotation was first undertaken by Wijers, Otten, Feenstra, Mulder, & Mulder (1989) who observed that the presentation of graphical characters elicited a pronounced positive component (P300) at parietal electrodes 300-700 ms after presentation. They found that the P300 (P3b) ERP was inversely related to the changing orientation of the character, becoming less positive with increasing angular disparity from an upright position. The authors suggested the gradual decrease of the positivity was unrelated to the modulation of the P300 itself, but rather caused by variations in a slow negative simultaneously occurring waveform superimposed on the P300. Increasing amplitude of the slow negativity was said to correlate with increasing rotational effort necessitated by the greater angular disparity of the target stimuli (Rösler, Heil, Bajricˇ, Pauls, and Hennighausen, 1995; Röder, Rösler, & Hennighausen, 1997; Vitouch, Bauer, Gittler, Leodolter, & Leodolter, 1997). Based on this interpretation, Wijers, Otten, Feenstra, Mulder, & Mulder (1989) proposed that this modulating slow negativity be interpreted as an index of cognitive activities supporting the mental rotation of objects, an interpretation also argued for by Rösler, Schumacher, Sojka, 1990. An additional assumption upon which this interpretation is based is that the P300, functionally interpreted as related to processes of context updating and/or participant generated expectations (Donchin & Coles, 1988; Johnson, 1988), remains constant throughout task execution.

Studies conducted by Heil (2002), Heil, Rauch, & Hennighausen (1998), and Heil & Rolke (2002a, 2002b) continued the line of research on mental rotation seeking to replicate and extend the outcomes of earlier studies (Wijers, et al., 1989; Rösler, et al., 1990) and confirm a mental rotation ERP index. In an initial study, Heil, et al. (1998) tested the hypothesis that if the amplitude modulation was an index of mental activities supporting rotation independent of the nature of the stimuli, then the modulation should be absent in tasks involving similar graphical characters (e.g., F, P, R, 2, 3, 4), but not requiring rotation. Participants viewed stimuli rotated to different angles, and were asked to classify whether the characters were presented in a normal position. In support of their hypothesis, the authors did not observe the amplitude modulation in the classification task, although reasoners were presented with rotated characters (Heil et al., 1998). However, related research (Ullsperger & Gille, 1998) provided evidence that increases in P300 amplitude corresponded directly with increases in the degree of angular disparity of task stimuli, thus the P300 was said to serve as an index for problem difficulty. Since varying the angular disparity of experimental stimuli is the key manipulation in tasks of mental rotation, and it was now demonstrated that increasing the angle of rotation of presented stimuli correlated significantly with increasing problem difficulty, it could not be ruled out that the variations previously found in the amplitude of the P300 were attributable to a problem difficulty factor and not the postulated negative deflection (Wijers et al., 1989; Rösler et al., 1990).

Bajric, Rösler, Heil, & Hennighausen (1999) conducted a study aimed at disambiguating brain activations associated with problem difficulty from those thought to support mental rotation. The experiment was designed to separate the contingent negative variation (CNV) waveform, from the hypothesized composite P300/negative deflection. Previous research had shown that as a task preceding waveform, the CNV varied in association with varying levels of effort anticipated to be required to complete assigned tasks (Lang et al., 1988; Lang, Zilch, Koska, Lindinger, & Deecke, 1989). Based on the experimental protocol, the CNV was predicted to occur in response to presentation of the cue signaling the angular orientation of the upcoming to be rotated target, thus allowing participants to evaluate difficulty of the rotational task. Analysis of the EEG recording revealed three distinct responses: a cue evoked bilateral parietal P300 whose magnitude increased correspondingly with a signal of increased angular disparity; a frontocentral anticipatory negativity occurring prior to presentation of the target, and a target response P300 over Cz and Pz scalp areas whose amplitude showed a monotonic decrease with increasing angular displacement of the target. These results allowed the authors to conclude distinct P300 effects in response to the angular manipulation, an initial P300 related to stimuli categorization and evaluation and a follow on P300 that interacted with the hypothesized rotation specific DC – like negative wave (Bajric, et al., 1999). Finally, Heil and colleagues (Heil & Rolke, 2002a; Heil & Rolke, 2002b) undertook several experiments aimed at examining the temporal relationship between onset of the mental rotation amplitude modulation and onset of the mental rotation activity. For example, in one set of experiments tasks were devised to prolong the encoding and stimuli assessment phases (e.g., varying perceptual quality of presented stimuli) proceeding onset of mental rotation. The authors predicted that the different onset times of mental rotation would correlate with the onset of amplitude modulation of the P300. Results revealed an onset latency effect, specifically, delay in the onset of cognitive processing corresponded with delay in the onset of amplitude modulation (Heil & Rolke, 2002a). Taken as a whole (for a review see Heil, 2002) these experiments provide support for the hypothesis of a mental rotation amplitude modulation as an index of mental rotation processes. A neurophysiological index of cognitive activities associated with mental rotation is regarded as important given the prominent use of mental rotation tasks in psychological experiments (Heil, 2002). The presence or absence of such a marker may be useful in determining whether or not mental rotation occurred in performing a given task.

Milivojevic, Johnson, Hamm, & Corballis (2003) investigated an interesting variant of the classic mental rotation studies. These researchers recorded EEG signals while people were engaged in mental rotation of letters, as well as a mental paper folding task. The mental rotation of letters was rated to be a simple task as compared to the more complex task of mental paper folding. A positive going ERP effect at parietal electrodes approximately 420-700ms latency was found to be associated with both tasks (Milivojevic et al., 2003). However, this ERP component was right lateralized for the simple task of rotation, and bilaterally distributed for the more complex paper-folding task. The researchers concluded that the right hemisphere might be primary for simple tasks of mental rotation, but that the left hemisphere becomes involved in tasks requiring more complex processing such as a sequence of mental transformations (Milivojevic et al., 2003).

3. PURPOSE OF THE STUDY

The study purpose was to use electroencephalography (EEG) methodology to investigate the neural level time course and spatial topography of reasoning with certain hypothesized mental representations, and their associated inference making processes. Current prominent theories of reasoning were developed largely based on the results of behavioral research and/or computational modeling. To date, most neuroscientific (fMRI, PET, EEG) studies have investigated the neural underpinnings of deductive reasoning in language based paradigms or those using other abstract symbols and rule-governed processes, with only a few having examined inference-making involving figural representations. There are no known studies reported in the neuroscientific literature that have directly compared imagistic mental model reasoning (reasoning on imagistic dynamic causal models) with mental logic reasoning (reasoning on symbolic representations), which would provide an approximate comparison of the two major representational systems that have been posited for centuries to describe human reasoning. The study reported herein tests the hypothesis that mental model-based (henceforward, MM) and mental rule-based (henceforth, MR) reasoning, as operationalized by the use of two different strategy paradigms to answer a simple binary question (clockwise or counterclockwise about the behavior of a simple mechanical system of gears), would show two different neurophysiological signatures.

The semantic cognitive processes postulated to be a property of model-based representational systems of reasoning (Gentner & Stevens, 1983; Johnson-Laird, 1983) versus the syntactic processes said to operate over symbol-based reasoning systems (Braine & O'Brien, 1998; Fodor, 1975; Plylshyn, 1984; Rips, 1994) were taken as the theoretical framework for predicting the event-related potentials (ERPs) of interest. Consequently, EEG recordings were analyzed for evidence that problem-solving using the MM strategy relies on semantic inference-making processes, whereas reasoning by the MR strategy engages syntactic inference-making processes.
Two ERP components, the N400 and P600, long associated with semantic and syntactic sense making in language processing were identified as candidates for neural indices of MM and MR reasoning, respectively. In the domain of EEG research, characteristics (i.e., amplitude, latency, scalp distribution) of such components are correlated with cognitive processes occurring immediately following specific events of interest (e.g., presentation of a stimulus or a behavioral response to an event).

Mental model (MM) and mental rule (MR) reasoning were contrasted using each participant as their own control, while they completed a simple binary-choice task that required them to determine the direction of turn of a target gear in a horizontally presented chain of gears. EEG was recorded during performance of the task, and changes in the waveforms, time courses, and scalp distributions of event-related potentials derived from the EEG recordings were subsequently identified offline. The study sought to further test the hypothesis that high-order complex cognitive tasks, such as reasoning, recruit a distributed network of neurons and that this would be evident in the EEG recorded data by: (a) specific components found to correlate with each mode of reasoning; (b) the timing or sequencing of activations; and/or (c) the field distribution of the waveforms.

The N400 event related potential (ERP) was originally described in the research of Kutas & Hillyard (1980a, b). In these classic studies, a sentence final word that deviated from the context of the preceding sentence fragment was shown to elicit a negative going deflection peaking around 400 milliseconds following onset of the incongruent terminal word as compared to sentences with congruent final words. As an example, Kutas & Hillyard (1980a) asked participants to read sentences such as the following:

1a) Congruent final word: It was his first day at work,

1b) Incongruent final word: He spread the warm bread with socks.

The negativity evoked to the incongruent sentences endings was quantified as the area in the latency zone between 300 and 600 ms post-target and given the label of N400 (Kutas &

Hillyard, 1980a, b). The N400 was found to be significant at frontal electrode Fz, central electrode Cz, and parietal electrode Pz, and was therefore described by Kutas & Hillyard (1980a, b) as having a central-parietal scalp distribution. Elicited in response to semantically incongruous final words, the N400 was functionally interpreted as reflecting participant reprocessing of the "semantic mismatches" existing between the prevailing context and the stimulus word (Kutas & Hillyard, 1980a, p. 204). It has subsequently become the standard index of semantic expectancy in studies on language processing (for a review, see Kutas & Federmeier, 2011; Osterhout & Holcomb, 1995).

More recently, the N400 has been investigated in paradigms using non-linguistic stimuli, as a general index for unexpected deviance from prior context (e.g., semantic matching of pictures, Barrett & Rugg, 1990; goal directed action, picture and sound, Cummings, Ceponiene, Koyama, Saygin, Townsend, & Dick, 2008; multiplication verification, Dohmas et al., 2007; Reid, Striano, & Koops, 2007; incongruous endings in video, c). Findings from such studies have prompted a view of the N400 as a neural correlate of cognitive activities related to the processing of semantic deviations from the established context regardless of stimulus modality (Barrett & Rugg, 1990; Dohmas et al., 2007). Federmeier & Laszlo (2009) and Kutas & Federmeier (2011) have suggested that the N400 represents a broader range of processes indexing access to semantic memory. On this view, the N400 is said to correlate with cognitive processes that involve the binding of information from stimulus input with established representations from short- and long-term memory (such as the prevailing context of a sentence stem, and accessing a word's meaning in long-term memory) that work together to create a whole conceptual representation (Federmeier & Laszlo, 2009).

The P600 ERP is also typically thought of as a language-specific component because it has largely been found in studies on language processing. The component, which is also referred to as the syntactic positive shift (SPS), was observed in several experiments conducted in the early 1990s, and was shown to be sensitive to anomalies related to syntactic analysis of sentences (Friederici, Pfeifer, & Hahne, 1993; Hagoort, Brown, & Groothusen, 1993; Neville, Nicol, Barss, Forster, Garrett, 1991; Osterhout & Holcomb, 1992). A functional interpretation of the waveform along with the P600 label was given by Osterhout & Holcomb (1992) based on the outcomes from a study in which they investigated the "garden-path effect," defined as the phenomenon whereby a reader builds-up an initial syntactic representation based on a beginning sentence fragment and then upon reading the sentence continuation must revise the meaning and structure of the just-completed sentence. The effect is illustrated by comparing the grammaticality of the continuations in the following two phrases:

a. Grammatical continuation: The mother tried <u>to</u> calm the baby.
b. Ungrammatical continuation: The nurse trusted <u>to</u> help the patients.

In the referenced study, participant responses to ungrammatical sentence continuations elicited a widely distributed positivity quantified as the mean voltage between 500 – 800 ms following presentation of the input words (Osterhout & Holcomb, 1992). The component's midpoint rested at approximately 600 ms post-stimulus, therefore it was labeled the P600 ERP. The waveform was widely distributed, but largest at fronto-central electrodes and was right-lateralized. The authors interpreted the P600 as a marker of the "syntactic garden-path effect" triggered when participants "backtracked" on a previously generated syntactic structure that proved inappropriate to input words (Osterhout & Holcomb, 1992). However, based on study outcomes the researchers claimed only that the component *co-occurred* with syntactic anomalies, and not that it was either *unique or specific* to processes associated with the syntactic analysis of phrases.

There now exists, a substantial number of studies in which the SPS or P600 waveforms have systematically been shown to index a variety of syntactic anomalies or incongruities in sentence processing. For example, the garden-path effect has been replicated (Osterhout & Holcomb, 1993; Osterhout, Holcomb, & Swinney, 1994), as have phrase structure violations (Gouvea, Phillips, Kazanina, & Poeppel, 2009; Kaan & Swaab, 2003a, b). In addition, the P600 has been elicited to violations of subject-noun verb agreement (Coulson, King, & Kutas, 1998b; Friederici et al. 1993; Gunter, Stowe, & Mobley, 1997; Hagoort et al., 1993; Kaan & Swaab, 2003a; Osterhout & Mobley, 1995), and to acceptability judgment tasks related to level of difficulty or increasing complexity of syntactic integration in sentence comprehension (biem Graben, Gerth & Vasishth, 2008; Frieiderici, Hahne, & Saddy, 2002; Kaan, Harris, Gibson, & Holcomb, 2000; Kaan & Swaab, 2003a). Based on these studies, several functional psychophysiological theories of the SPS and P600 ERP have been advanced. One proposal is that it reflects cognitive processes related to rescuing the interpretation of a sentence (repair) when a word continuation is ungrammatical relative to the preceding fragment, and/or trying to rearrange the structure of a sentence (revision) to yield a correct analysis when the continuation is either grammatically correct or even ungrammatical, but nonpreferred (Friederici, 1999, 1995; Kaan & Swaab, 2003a; Osterhout & Holcomb, 1993, 1992; Osterhout, Holcomb, & Swinney, 1994). Other views represent that these late positivities may correlated with more general cognitive processes required to build-up a coherent structure in a sentence (Hagoort, 2003), or the construction and subsequent destruction of the syntactic structure of a phrase (Coulson et al., 1998).

Importantly, the scalp distribution of the P600 component has been observed to differ depending on the nature of the cognitive process assumed to be engaged. Of particular relevance to this paper, P600 waveforms elicited to repair processes triggered by ungrammatical syntactic structures (e.g., agreement violations) have been observed to correlate with more posterior distributions (Coulson et al., 1998b; Friederici et al., 1993; Gunter, et al., 1997; Hagoort et al., 1993; Osterhout & Mobley, 1995), whereas revision processes have been associated with a frontally distributed P600 (Hagoort, Brown & Osterhout, 1999; Osterhout & Holcomb, 1992; van Berkum, Brown, & Hagoort, 1999). One interpretation of these data has been proposed by Hagoort et al. (1999) who suggest that cognitive resources required to overwrite a preferred current structural representation of a sentence (revision) evoke the frontal P600, and that the more posteriorly distributed P600 is triggered by processes associated with collapse of the structural representation (repair) owing to ungrammatical word endings. However, frontal positivities have also been elicited to increases in sentence complexity, such as in the study conducted by Friederici, et al. (2002) in which the level of complexity of unambiguous sentences was manipulated and compared. Kaan & Swaab (2003b) reviewed the issue of the relation between the P600 and sentence complexity in a study comparing sentences containing one-noun phrases (grammatical and ungrammatical) with those consisting of two noun-phrases (grammatical preferred and nonpreferred, and ungrammatical). In addition, the study incorporated a comparison of reader repair and revision of phrase structures. Specifically, the ungrammatical continuations (both one-noun phrase and two noun-phrases) constituted the repair condition, the grammatical nonpreferred continuations (two noun-phrases) the revision condition, and the two nouns-phrases sentences in comparison to the one noun-phrase sentences the complexity condition. The results revealed a significant posteriorly distributed P600 for both the ungrammatical conditions and the nonpreferred grammatical condition relative to the preferred condition (repair), and also for the grammatical nonpreferred continuation compared to the grammatical preferred ending. A significant frontally distributed P600 was evoked to the critical verbs preceded by the more complex sentences constructed with the two noun-phrases compared with the grammatical continuations in the simple one noun-phrase sentences (complexity).

While the P600 has been reliably shown to be sensitive to the syntactic aspects of linguistic input, late positivities, similar to the P600/SPS component have also been elicited to violations of arithmetic sequencing rules (Nunez-Pena & Honrubia-Serrano, 2004), mathematical rules (Lelekov, 2004), and harmonic musical structure (Patel, Gibson, Ratner, Besson, Holcomb, 1998). On the basis of these observations, there is an emerging view of the P600/SPS as an index for a general purpose mechanism related to restructuring and reintegration of constituent structure (Gouvea et al., 2009; Patel, 1998).

The EEG recordings for this dissertation study were analyzed for the ERPs hypothesized to yield information about the mental representations established during MM and MR reasoning as well as the nature of the processes posited to operate over them. Behavioral performance measures (accuracy and latency of responses) were analyzed for evidence that participants followed the strategy instructions, reliably completed the tasks, and that the level of problem solving difficulty was comparable between the manipulated modes of reasoning. In the long history of mental rotation studies, a finding of longer response times associated with MM reasoning as compared to MR reasoning has served as evidence that participants mentally rotated analogue representations of task stimuli compared to using rules to assess an abstracted representation of the task (Hachey, 2005; Kosslyn, 1976, 1980; Schwartz & Black, 1996a, b; Shepard & Metzler, 1971). However, in this dissertation study response times are not expected to differ significantly between the two conditions both because participants were trained to floor with both strategies, and responses in both conditions were time-locked to presentation of the target gear (not recorded from the start of problem reasoning as in the behavioral studies). Accuracy scores will be analyzed for evidence that the reasoning strategies established tasks of comparable difficulty for participants.

4. RESEARCH QUESTIONS/HYPOTHESES

The study focused on the following question:

In a strategy manipulation paradigm designed to simulating mental model and mental rule reasoning, will participants show that ERP effects previously associated with semantic (N400) and syntactic (P600) processing of sentences in studies on language comprehension also correlate, respectively, with these two modes of reasoning.

Hypotheses for the study were:

EEG recordings of participants' reasoning, using mental model (MM) and mental rule (MR) strategies to predict the behavior of a target gear in a simple mechanical system of gears, will show significantly different ERP signatures for each strategy as evidenced by amplitude, timing, and scalp topography of recorded brain activations.

ERP hypotheses:

(a) Unmet expectations of the direction of turn of the target gear established by using the MM strategy to reason about the behavior of the gears in the problems will elicit a N400 ERP, reflecting the semantic nature of mental model inference making; and

(b) Unmet expectations of the direction of turn of the target gear generated by use of the MR strategy to reason about the gear problems will elicit a P600/SPS ERP, reflecting the syntactic / rule-governed nature of mental logic inference making.

Behavioral hypotheses:

(a) Problem accuracy scores for reasoning with the two strategies will not be significantly different.

(b) Response times for MM and MR reasoning will not be significantly different.

5. RESEARCH DESIGN AND METHODS

A 2 X 2 within-subjects experimental design was implemented, allowing brain activity from the same participant to be recorded during problem solving using the MM and MR reasoning strategies in two different sessions. Thus, participants served as their own controls. Participants applied the strategies to solving the same series of problems about a simple univariate mechanical system modeled on those developed by Schwartz & Black (1996a). The current study contrasted two modes of reasoning (Conditions: MM and MR), used to reason about two types of problems (Problem types: congruent vs. incongruent). Reasoning was defined by the use of two reasoning strategies, which were devised to operationalize an instance of mental model (MM) and mental logic (MR) reasoning, respectively. MM reasoning was induced by asking participants to visually track the dynamically presented stimuli in order to affect a mental simulation of the system of turning gears, and to "read off" the causal relation between adjacent contiguous gears in the system to infer an answer. MR reasoning was induced by asking participants to recode the visual-spatial stimuli as symbolic tokens (transducing the visual-spatial percepts to form a symbol-based representational reasoning system), count the number of token units in the gear system, and then apply the trained rule to solve the problems. Similar to the study by Prabhakaran et al. (1997) in which the same visual-spatial stimuli from the Raven's Progressive Matrices Test were used for all conditions, both MM and MR reasoning as manipulated in this study involved viewing the same imagistic dynamic stimuli but processing them differently. However, unlike the task manipulations designed in the former study (both of which required figural or visual-spatial processing), it was hypothesized that when reasoning with the MM and MR strategies (assuming participants follow instructions) participants need not process the imagistic dynamic stimuli during MR reasoning. Therefore, the hypotheses for the current study were driven by two presumed distinctions between the representational systems underpinning MM and MR reasoning. The first is that use of the MM strategy involves reasoning with an imagistic dynamic representation, whereas MR strategy use requires only that a token entity be considered. Second, the inferential processes employed by mental model reasoning (MM) are computations over relations of temporal and spatial contiguity and implicit laws of physics, and in contrast, conclusions reached using mental rule reasoning (MR) are a product of formal rule governed computations. As devised, the two strategies are intended to operationalize psychological distinctions theorized to exist between mental model and mental logic accounts of reasoning.

Two problem types were developed to correspond with 1) meeting participant expectations; and 2) violating participant expectations as established by use of MM and MR strategies to reason about the problems.

2 X 2 Experimental Design:

- Reasoning strategies:
    - Condition 1: mental model (MM)
    - Condition 2: mental rule (MR)
- Problem types:
    - Type 1: congruent turning target gear
    - Type 2: incongruent turning target gear

Participants completed the study in two sessions, the first ran approximately one hour during which participants were trained in either the MM or the MR reasoning strategy (order was counterbalanced across participants), and asked to solve thirty-two gear-turn problems using the trained strategy while EEG was recorded. Session 2 was conducted within two-four weeks following the first session. During the second session participants were trained in the alternate of the two reasoning strategies (whichever they did not use in session 1), and were again asked to solve the same thirty-two problems while EEG was recorded. Since participants reasoned about the same thirty-two gear problems in sessions 1 and 2, the primary difference between the two sessions was the reasoning strategy they were instructed to use during problem reasoning.

5.1 PARTICIPANTS

Participants were graduate students at Teachers College, Columbia University, or friends of the Principle Investigator (PI). Graduate students were volunteers who participated out of research interest or to satisfy a course requirement. Friends of the PI were recruited by word-of-mouth and participated out of interest in the research. Volunteers were screened for right-handedness, normal or corrected-to-normal vision, normal hearing, and no reported history of neurological illness or trauma. Study participants were 6 females and 4 males (n=10) aged 28 to 43 (mean = 33 ± 5.75 years).

5.1.1 RECRUITMENT AND INFORMED CONSENT

The experiment was performed in accordance with the requirements of Institutional Review Board for the Protection of Human Subjects at Teachers College, Columbia University. On the first meeting, the consent and Handedness Inventory forms were presented and reviewed with participants (as required) by the PI and/or trained lab assistants. Participants were given a tour of the lab, including a viewing of the sound attenuated chamber where they were to be seated for the EEG recording. Further, participants were shown the EEG net and explained its functionality prior to its being fitted on their head. All participants were informed that it was within their rights and that they should feel free to withdraw from participation in the experiment at any time during the course of the two sessions. An overview of the experiment was then given, including the assurance that the experiment did not constitute a test and that they were not expected to compete for either time performance or achievement of highest score. Every step of the procedure was explained and discussed as it occurred, and ample opportunity was created for participants to ask questions, or to express concerns or anxieties. All participants were encouraged to ask questions, and in cases were they evidenced tiredness or sickness they were encouraged to withdraw temporarily with an offer made to reschedule the session as appropriate. All consents and other forms were presented in the same manner to each participant at each session. Finally, all participants were provided with a lab telephone number and email address to contact the researcher in the event they should have questions or concerns to arise at a time subsequent to their participation.

5.2 SAMPLE SIZE AND STATISTICAL POWER

Estimations of power and appropriate sample size for measuring ERPs are notoriously difficult (see, e.g., Picton et al., 2000 for an overview of some of the issues involved in statistical approaches to analyzing EEG and ERP data). Power estimation requires knowledge of the expected percent signal change between two conditions (effect size), as well as estimates of the variability in signal change, and these are usually unknown in brain imaging studies. Signal-to-noise is typically low, requiring repeated presentations of stimuli within each condition while subjects are recorded over a period of time. The experiment reported here took approximately 20 minutes of EEG recording time. The raw data consisted of continuous digital recordings (sampling 250 times per second) of voltage deflections at 128 different points on the participants' scalp. This means that, for this ERP experiment, a time series of approximately 300,000 (i.e., 250 samples per second x 60 seconds per minute x 20 minutes per session) data points for each of the 128 sensors for each condition for each participant was captured. Within this time series data, there are two sources of variability of interest: within-subject time course variability (fluctuations from one time point to another) and within-subject experimental variability (variation in the effectiveness of the experimental manipulations in producing a percentage signal change). Analyses of power and sample size for brain imaging data are therefore complex, and little work has been done on generation of power curves for ERP. Sample sizes and numbers of trials per condition were established with reference to available guidelines relative to the predicted ERPs, and the previous experimental experiences of the sponsor. Additionally, experimental design parameters to reduce variability were used where possible (e.g. within-subject variability can be minimized by ensuring trial-by-trial consistency: Handy 2005; Luck, 2005).

The research study reported here was conducted to investigate the N400 and P600 event-related potentials both standard, relatively slow and large components. Results of a pilot study showed these ERP effects to be large and identifiable (although in both the pilot and current study both components presented with a topography differing from what is typically reported – discussed in more detail below). For investigations of large components like these, 30-60 trials per condition is recommended (Luck, 2005). The EEG experiment reported here was designed with 128 trials per condition (four cycles through 32 trials), which created an acceptably tolerable length of recording time for participants while meeting recommendations. A review of the literature showed that, for the ten most comparable ERP studies on problem-solving and/or reasoning using a similar participant group, a mean sample size of 15.1 was used (SD=4.4, range = 16). The stated goal to recruit twenty participants for the study was met, and therefore the study was executed well within parameters utilized in the extant literature in this field.

Completion of the sessions for the twenty participants was greatly hindered owing to operational challenges with the EEG platform. Quality of the recordings was a related challenge, and the high number of artifacts evoked by the experimental stimuli (head movement, eye saccades and blinks) added to the overall difficulty of collecting sufficient trials from each participant. Consequently, of the 20 participants recruited, files from 6 participants were eliminated because the quality of the recordings was too poor to warrant analyses. Other participants were dropped from the analyses because at least one of their two files contained too few trials to assess. The 10 remaining pairs of participant files are included in the analyses reported in this paper.

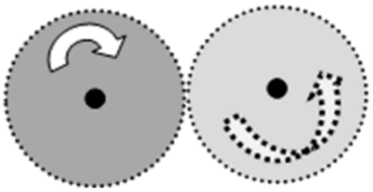
5.3 MATERIALS

Two sets of materials were developed for use in the study. Both sets consisted of problems modeled on a simple mechanical system of gears. The original images used to create the gear problems were obtained from a free open-source public library, and modified using graphical software programs (PhotoShop, Microsoft PowerPoint, AfterEffects). The first set of materials was incorporated in a self-instructional computer training program which participants used to become familiar with the experimental task, learn the assigned reasoning strategy, and practice its use in problem reasoning. The second set of materials, which differed from the first with respect to the graphical rendering of the gears, was used to construct the experimental stimuli (gear problems). Both the training and experimental problems sets consisted of from 2 to 8 gears arrayed horizontally in an "open chain" configuration (Figure 3).

Participants were able to view the training problems in both a static and dynamic format (by clicking the 'Play' button), which gave them the opportunity to first solve the problem and then watch a dynamic modeling of the problem (correct answer). Observing a correct simulation of the problem (MM or MR strategy solution) offered the participant feedback on their predicted answers, as well as a short tutorial of strategy application. The training program is designed such that problems initially appear on screen as presented in Figure 3 below. Written instructions describing the strategy (MM or MR) are provided on the first few slides. Participants can initiate a dynamic modeling of the problem for either strategy solution by clicking or pressing a designated key. The dynamic version of the problem shows the *problem stem* (the gears shaded in medium grey) rotating through one full revolution (360º), during which the *target gear* (shaded in light grey) remains stationary. As the problem stem begins a second revolution, the target gear activates and completes the next 360º revolution in conjunction with the problem stem. In effect, the target gear "joins" the problem stem at the beginning of the second revolution completing the system of gears and depicting the correct direction of turn of all gears in the problem chain. All training problems are in an *open-chain* format (Figure 3).

The training programs were designed with and are presented using Microsoft PowerPoint. The gears used to configure the problems are a uniform 1.5” in size; dimension of the image is 480 pixels high by 640 pixels wide. Problems are shown on a white background, with the lead and any other gears in the problem set colored a medium gray and the target gear shaded a light grey (Figures 3 and 4). Markings on the gears appear in white or black, as seen in Figures 3 and 4. All problems are positioned on the slide at 2.75” horizontally and 2.42 inches vertically from the top left corner.

Figure 3: MM and MR Strategy Training Problems



Mental Model Training Problem: The open chain gear problem pictured above depicts a one-gear *problem stem* shaded in a medium grey color and a target gear shaded in light grey. Both gears are shown with arrows indicating their correct direction of turn. The arrows on each gear are placed to aid in imagining that the tip of the arrow of the lead gear turns and connects with the tail of the arrow of the target gear, striking it and setting it in motion.

Mental Rule Training Problem: The open chain gear problem pictured above depicts a two-gear problem stem shaded in medium grey and a target gear shaded in light grey. Together, the three gears comprise a gear problem with an odd number of gears. The direction of turn of each gear is indicated in accordance with the odd-even rule, and the arrows of the lead and target gears are both shown in white to highlight that they turn in the same direction.

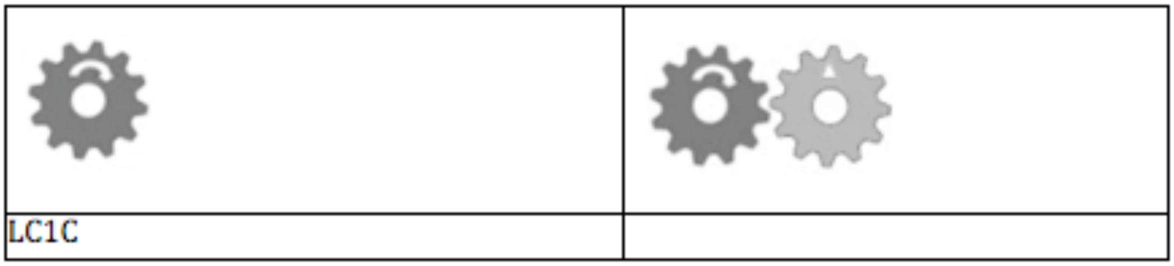
Each of the thirty-two experimental problems (stimuli) developed for the EEG recording were created as MPEG 2 video files, allowing the experimental stimuli (gear problems) to be presented in a visually dynamic format. Presentation of dynamic stimuli ensured that all participants reasoned about the problems in the same format, without having to construct a dynamic representation using gestures. The video files were created offline and imported into E-Prime, the computer software used to program, present, and acquire specific data points for the research study.

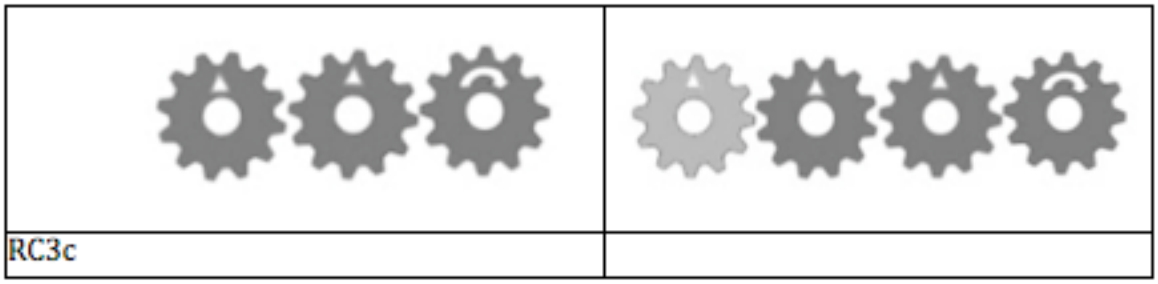
Each video presents a problem in two scenes. The first scene depicts the problem stem (from one to seven interlocked gears depicted horizontally in an open-chain configuration), that is, a simultaneously and congruently rotating open chain of gears. In the first scene, gears in the problem stem complete a full rotation (360º). In the second scene, a single rotating “target” gear appears and merges (interlocks) with the last gear in the problem stem. The target gear appears as either turning in a direction congruent (meshing) with the other gear(s) in the problem stem or incongruent (not meshing). More examples of the gear problem stimuli are shown below in Figure 4. Images of gears used to configure the experimental stimuli are approximately 1.8” high by 1.8” wide. Gears in the problem stem are shaded in medium grey, and the target gear in light grey. Problems were presented against a white background on a 15” computer screen. The thirty-two open chain problems were created from the following basic units of gears:

- Open chain problem stem consisting of either two gears, three gears, four gears, or five gears for a total of four type;
- Problem stem with lead gear turning either clockwise or counterclockwise;

- Problem stem with lead gear in far left or far right position;
- Target gear turning either congruently or incongruently relative to the problem stem.

Figure 4: Example of 2-gear and 4-gear Open-chain Problem





To help maintain attention and mental task set the thirty-two problems were presented serially in four blocks. The four blocks were: LCW (Far left clockwise turning lead gear); LCCW (Far left counter clockwise turning lead gear); RCW (Far right clockwise turning lead gear); and RCCW (Far right counter clockwise turning lead gear). Each block consisted of eight problems, four with a congruent turning target gear and four with an incongruent turning gear. These eight problems were randomly presented (i.e., random presentations within blocks). Finally, the set of thirty-two problems was cycled four times to obtain 128 trials or approximately twenty minutes of sampling time.

5.4 EEG/ERP EXPERIMENTAL PROCEDURES

EEG is a continuous recording of voltage fluctuations that are associated with intracellular communication in the brain (mainly, synaptic potentials generated in thalamocortical pathways). By recording electrical brain activity during cognitive processing, EEG can provide a real-time measure of mental activity. Thus, data collected during EEG recordings afford a means by which the time course and/or sequence, amplitude, and topographical distribution of activations that are associated with cognitive processes constitutive of reasoning can be examined *in vivo,* providing essential information about the underlying neuronal mechanisms and by inference the nature of the representations on which they are thought to act (Ollinger, 2009). These scalp potentials are measured using small electrodes distributed across the scalp at standard locations yielding high temporal resolution, but only a rough estimate of the brain regions and lateralization of the source of electrical activations associated with cognitive activity due to smearing and distortion of the electrical potentials as they travel through brain tissue and the scalp (Handy, 2005). The electrode placement is typically done according to an accepted system (the "10-20 System" in which electrode placement occurs over the frontal, central, temporal, parietal, and occipital portions of the scalp (Handy, 2005; Luck, 2005). The 10-20 System has 19 electrodes, while the system used in the current study has 128 electrodes offering comparatively better spatial information (Tucker, 1993; Tucker, Liotti, Potts, Russell, & Posner, 1994**).**

5.4.1 ASSIGNMENT TO CONDITION AND TRAINING

Assignment to condition (MM or MR) was counterbalanced across participants. Following assignment to reasoning condition participants were first given an overview of the study and then received instructions on the training program. Training objectives were twofold: 1) Instruct participants on use of the reasoning strategies and ensure they are able to solve the gear problems; 2) Allow participants to practice application of the strategies until they obtained a 90% accuracy rate in problem performance. Instructions for the MM and MR reasoning strategies were presented as follows:

- MM strategy as mental simulation: “Imagine that the lead gear is turning in the direction indicated by the arrow. Now imagine that it causes the second gear to turn by pushing it in the opposite direction. Next, imagine that the second gear causes the third gear to turn in a direction opposite to it, and so forth. You can picture each gear making a full turn or you can also imagine a sort of “serpentine” like or “wave-like” movement which would be a bit more like a half-turn.”

- The MR strategy as an “Odd/Even” rule: “If there are an even number of gears in the chain, then the target gear will turn opposite to the lead gear; if there are an odd number of gears, then the target gear will turn in the same direction as the lead gear. In other words, even – opposite and odd – same.”

5.4.2 MEASUREMENT OF HEAD SIZE AND VERTEX LOCATION

The circumference of each participant’s head was measured to ensure the correct size sensor net was selected, and their vertex marked to ensure accurate placement of the net. The participant was fitted with an appropriate 128-channel geodesic sensor net (Electrical Geodesics, Inc., Eugene, OR – Tucker, 1993) with electrodes referred to the vertex. These nets are arrangements of electrodes, held in relative positions to each other with fine elastic. The electrodes are embedded in sponges, which are soaked in a weak electrolyte solution (potassium chloride). The geodesic sensor net is quick to apply and comfortable to wear, and does NOT require scalp abradement or the application of any electrode glue. Once participants successfully completed the training program, they were fitted with the selected pre-soaked sensor net. Following proper seating of the net, sensors were adjusted until they made good contact with the scalp.

Next, the participant was seated in a chair in front of a computer screen in a sound attenuated chamber within the lab. The amplifier was checked and calibrated before the net was connected, and following this procedure impedances (loss of signal between scalp and sensor) were measured by feeding a minute (400 microvolt) electrical field through each electrode, which was then ‘read back’ by the acquisition system so that the amount of signal loss was calculated. A response button box was provided for the participant to indicate the response choice to each trial presentation (gear problem). A visual overview of the EEG recording set-up is provided below in Figure 5.

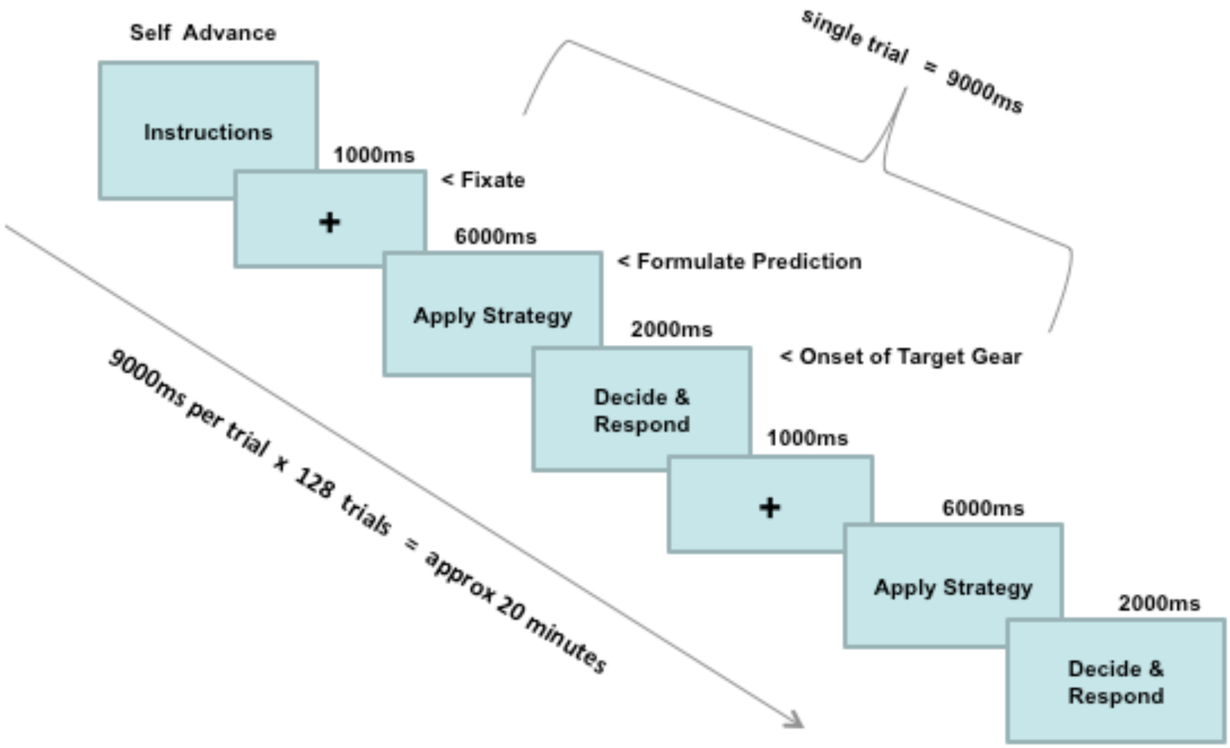
Figure 5: EEG Methodology

5.4.3 INSTRUCTIONS AND EXPERIMENTAL TASK

Before beginning the recording session participants were given one last set of instructions aimed at reducing movement artifacts. The instructions included a demonstration of the unwanted effects of various body movements (eye blinks and saccades, head turning, foot tapping, etc.) on the EEG recording, plus an explanation of when during the programmed presentation such movements would have the least impact. Lastly, they were asked to refrain from moving to the extent possible. Participants initiated the start of the experimental program via button press.

Written task instructions, a repeat of what had previously been communicated verbally to participants, appeared as the first slide in the experimental presentation. After reading these instructions, and again at their own initiation, participants advanced to the next slide, which was a crosshair (+) presented in black type on a white background centered on the computer screen. The crosshair (+) appeared for 1000 ms followed by the first gear problem, and served to fixate eye gaze at the center of the screen prior to onset of problem presentation. Thereafter, the computer program advanced via response button press or the auto-programmed feature until completion of four cycles of the set of thirty-two problems, resulting in 128 trials and approximately twenty minutes of EEG signal sampling. Each of the thirty-two separate video clips is 8000 ms in length, with the first scene presenting a 360º rotation of the problem stem gears in 6000 ms, and the second scene showing the incoming rotating target gear merging congruently or incongruently with the gears in the problem stem. The target gear, which is programmed to complete a 360º revolution in 2000 ms serves as the time-locked event of interest. Participants were able to respond at anytime during the 2000 ms timed interval during which the target gear appeared and completed its 360º revolution. Participant responses were collected during this 2000 ms interval. Participants responded by pressing one of two designated buttons on the response box. They were instructed to press button "1" if the incoming target gear turned in the predicted or expected direction and button "2" if it turned in the direction not predicted or unexpected. Either the participant's button press or a response delay of 2000 ms marked the end of a given trial and advanced the program to the subsequent trial. Figure 6 below depicts the timeline for two successive trials.

Figure 6: Stimuli Presentation and Trial Timeline



When all the trials were completed, the end of the recording was announced on screen, and the experimenter entered the recording chamber to disconnect and remove the sensor net as quickly as possible. The participant was debriefed upon leaving the recording chamber. The continuous digital recording captured during each session was saved for later processing.

5.5 DATA ANALYSIS AND INTERPRETATION

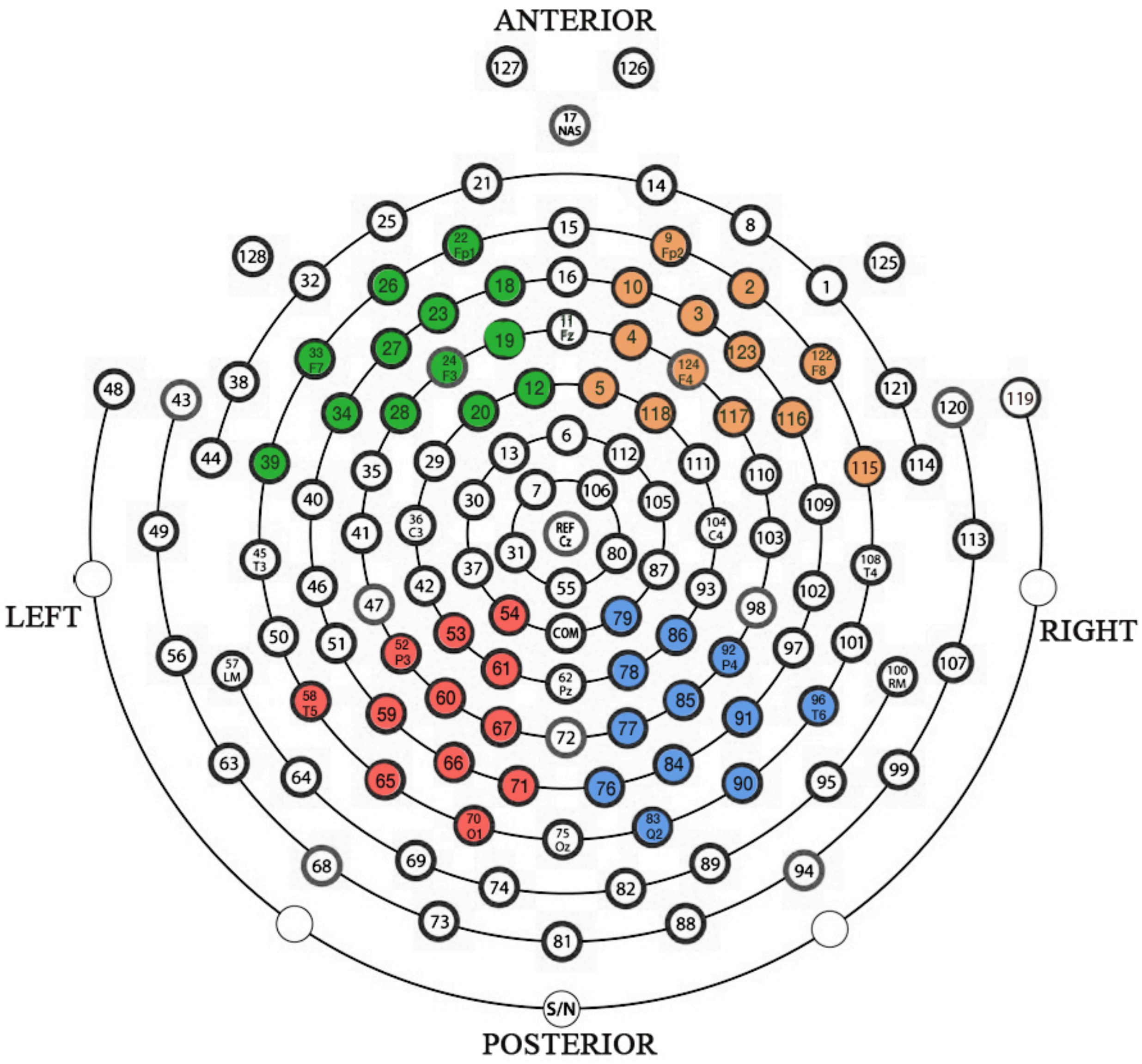
The EEG recordings resulting from participant completion of 128 trials of problem reasoning in each session (condition) were saved for later processing offline. A standard ERP analysis protocol was followed for the analysis of the EEG data (following principles described in detail in Picton, Bentin, Berg, Donchin, Hillyard, Johnson, Miller Ritter, Ruchkin, Rugg, Taylor, 2000; Luck, 2005; Handy, 2005). The following two stages of processing were performed.

5.5.1 DATA PRE-PROCESSING

The recorded raw EEG data was digitally filtered offline using a 30 Hz LowPass filter, after which a decision was made whether or not to subject the recorded data to automatic artifact rejection protocols for removal of externally induced noise, movement and physiological artifacts (EKG, EMG, EOG). Given the potential for artifacts associated with eye movements (e.g., saccades) manual artifact review was completed instead of using an automatic artifact detection routine. Subsequent to LowPass filtering noisy channels were marked as bad and interpolated using spherical spline interpolation based on recorded data from surrounding sensors. Data were vertex referenced during recording and this reference was kept for further data processing. Error trials and timeout trials were removed from the analysis process. To examine the EEG waveform for the predicted ERP components following onset of the target gear, the continuous recording was segmented into relevant millisecond epochs, including a 200 milliseconds pre-stimulus (the “baseline period”) and a 900 milliseconds post-stimulus window. The latter epoch was further segmented into three non- contiguous time windows. The first time window (T1) was defined from 0-200 ms, the second window (T2) from 300-500 ms, and the third (T3) from 600-900 ms. T1 was selected for examination of early sensory components, and to ensure examine separation of them from the first ERP of principle interest, the N400 component. The N400 ERP is known to peak around 400 milliseconds post-target stimulus onset, but can extend from 250-500 milliseconds. Thus, T2 (300-500ms) was selected for inspection for the N400 waveform. T3 was selected for examination for the P600/SPS component, which has a known onset around 500 milliseconds after the eliciting stimulus, and often peaks around 600 milliseconds following presentation of the eliciting stimulus lasting for several hundred milliseconds. A window (600-900 milliseconds) was selected given the study's implementation of dynamic visual stimuli (Dien, Michelson, & Franklin, 2010; Sitnikova, Goff, & Kuperberg, 2009).

Epoch segments were averaged together to reduce variance in the data due to random noise, and to permit identification of time-locked event-related responses associated with the onset of the target gear rotation. EEG epochs were averaged separately for congruent trials and incongruent trials for each condition, for each individual participant. Next, averaged waveforms were baseline-corrected to control for drift. Base-line correction procedures involve using the average electrical potential during the 200 millisecond baseline period to calculate a mean "zero" from which values of the positive and negative voltage deflections across the scalp following onset of target gear processing will be derived. Finally, four montages were applied to the data in order to examine the different responses by electrodes in specific areas of the scalp. The four montages applied to these data correspond with the following four quadrants of the scalp: left anterior; right anterior; left posterior; right posterior (Figure 7). The regional montages are shown as blocks of differently colored electrodes: Light Green = Left Anterior sensors, Orange = Right Anterior sensors, Red = Left Posterior sensors, Blue = Right Posterior sensors.

Figure 7: Four Quadrant Montages



The montaged data was exported in a format permitting further analyses using data analysis packages such as Excel, MATLAB and PASW. First, individual data segments (T1=0-200 ms; T2=300-500 ms; T3=600-900 ms) from the pre-processed data were averaged together for congruent and incongruent trials for both the MM and MR reasoning conditions. These individual averages were then grand-averaged (Handy, 2005; Luck, 2005; Picton et al., 2000). This enabled us to identify the predicted ERP components for the MM and MR conditions by comparing grand averaged waveforms obtained in response to the expected congruent target stimuli with those obtained to the unexpected incongruent target stimuli for both modes of reasoning. Component identification was based on distribution, topography, and latency of activations.

Repeated (participants) measures analysis of variance (ANOVAs) were used to evaluate the main effects and interactions by time window (T1, T2, T3) in the EEG recordings in a 2 (Condition: MM and MR) x 2 (Problem Type: Congruent vs. Incongruent) x 2 (Laterality: Left vs. Right) x 2 (Anteriority: Anterior vs. Posterior) comparison (Dien & Santuzzi, 2005). The dependent variable was grand-averaged voltages across relevant sensor arrays, determined following data preprocessing. The ANOVAs were followed by planned comparisons at each level of each significant variable in order to determine the sources of significant main effects and interactions (e.g., to answer the specific question of whether a N400 ERP was evoked to the unexpected endings in the MM reasoning condition and a P600 to the unexpected endings in the MR condition). The Greenhouse-Geisser correction was applied to all repeated measures with more than one degree of freedom, and Bonferroni corrections are reported for the multiple planned comparisons (Dien and Santuzzi, 2005; Luck, 2005). All statistical tests were conducted to evaluate data within *a priori* selected time windows. Analyses for each mode of reasoning are reported separately along with graphs depicting significant ERP waveforms by Montage (*a priori* selected regions of interest) in association with specific sensors as they are distributed over the scalp.

6. RESULTS

6.1 Behavioral Data

The statistical comparison of participant responses to the gear problems revealed that they were equally able to solve the problems across both the MM and MR task manipulations. For MM reasoning the accuracy rate was 88.6% with a standard deviation of 11%. The accuracy rate for MR reasoning was 88.4% with a standard deviation of 11%. The difference in accuracy for the two modes of reasoning did not reach statistical significance (F (1,14.844) = .002, p = .969). The reported degrees of freedom reflect use of the Welch correction for unequal sample sizes. The comparison of response time between the two conditions was not statistically significant (F (1,2046) = .266, p = .606). The mean response time for MM reasoning was 808.3477 milliseconds with a standard deviation of 499.7507. Mean response time for MR reasoning was 819.8955 milliseconds with standard deviation of 473.5428 milliseconds. Response time results are consistent with expectations given participant preparation and training.

6.2 Event-Related Potential Data

Statistical analyses of the observed components were conducted using 2 x 2 x 2 x 2 ANOVA for each of the three time windows. For time window 1 (T1 = 0-200ms) only the main effect for

Laterality was significant ($F (1,9) = 5.875$, $p < .05$). The two-way interactions for Congruency x Laterality was significant: ($F (1,9) = 8.900$, $p < .05$). In time window 2 (T2 = 300-500ms) the main effect of congruency was significant ($F (1,9) = 14.545$, $p < .005$. The two-way interaction of Congruency x Laterality was significant $F (1,9) = 7.142$, $p < .026$. One three-way interaction, Congruency x Laterality x Anteriority was significant ($F (1,9) = 5.544$, $p < .05$). For time window 3 (T3 = 600-900ms) only the main effect of Anteriority was significant, ($F (1,9) = 5.355$, $p < .046$). Two-way interactions for Congruency x Laterality and Congruency x Anteriority were significant ($F (1,9) = 5.671$, $p < .041$ and $F (1,9) = 9.013$, $p < .05$, respectively). The three-way interaction of Reasoning x Congruency x Anteriority was significant $F (1,9) = 6.385$, $p < .05$. A four-way interaction of Reasoning x Congruency x Laterality x Anteriority was also significant ($F (1,9) = 12.268$, $p < .007$).

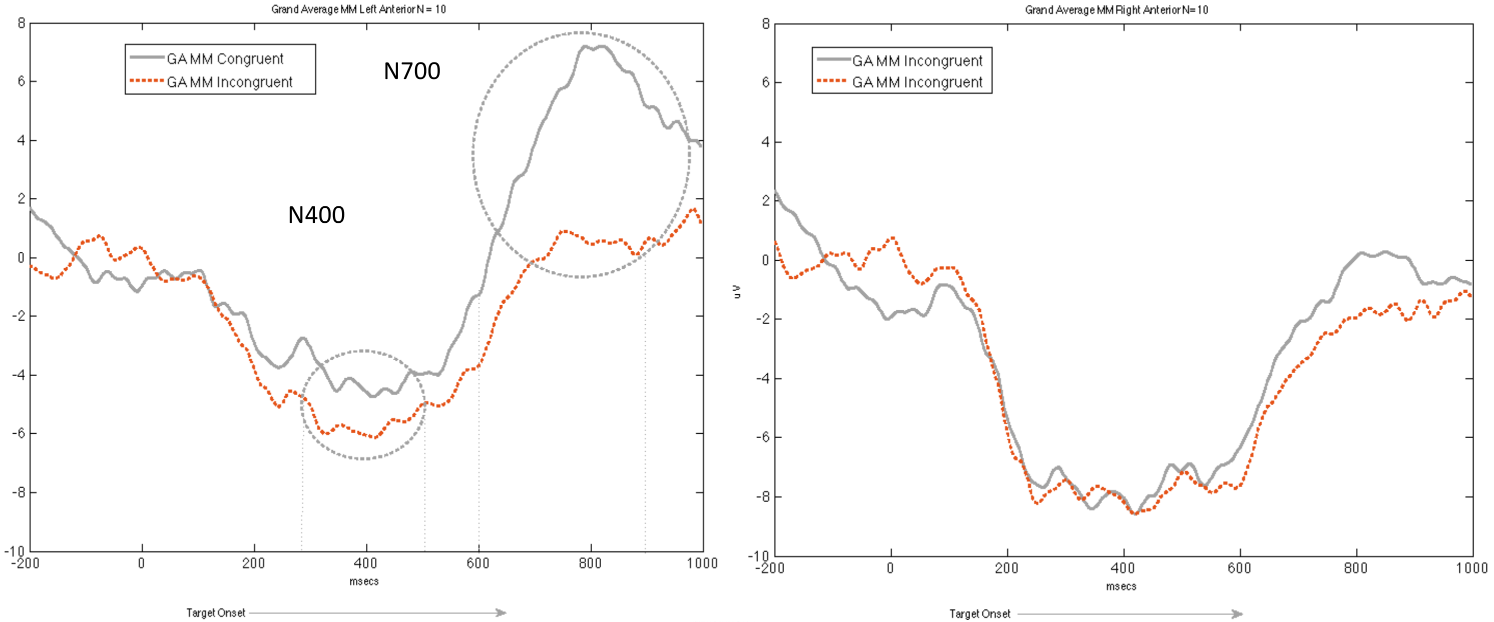
Event Related Potentials (ERPs) in the Mental Model Reasoning Condition:

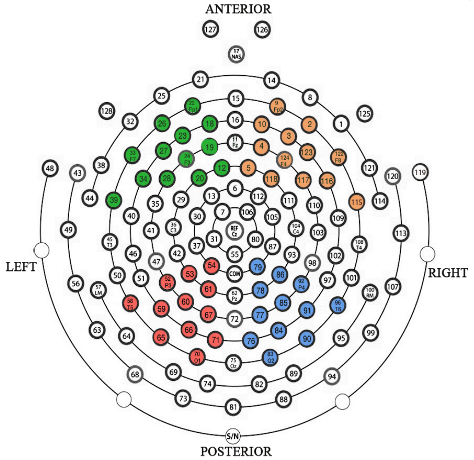
The Grand-average ERPs to the target gears in each problem ending type (congruent and incongruent) are shown below in Figures 8. The head map in the center of the page shows a view of the 128 electrodes making up the HydroCel Geodesic Sensor Net, in their positions on the scalp, as if viewed from above. The regional montages are shown as blocks of differently colored electrodes: Light Green = Left Anterior sensors, Orange = Right Anterior sensors, Red = Left Posterior sensors, Blue = Right Posterior sensors. Each of the regional waveform plots shows the regionally averaged event-related responses to all the congruent meshing target gears (grey line) versus all the responses to incongruent meshing target gears (red line) for the mental model reasoning task manipulation. Time intervals (milliseconds) are plotted on the X-axes and microvolts (μV) on the Y-axes. The effective zero time point for onset of the target stimuli is at the 200-millisecond mark as identified by the annotation under the X-axes. The approximate midpoints of all ERPs found to be associated with significant differences between responses to congruent and incongruent trials are encircled with a light grey ring and labeled according to their polarity and latency and/or functional association.

Figure 8: Mental Model Reasoning ERP Waveforms

Left Anterior

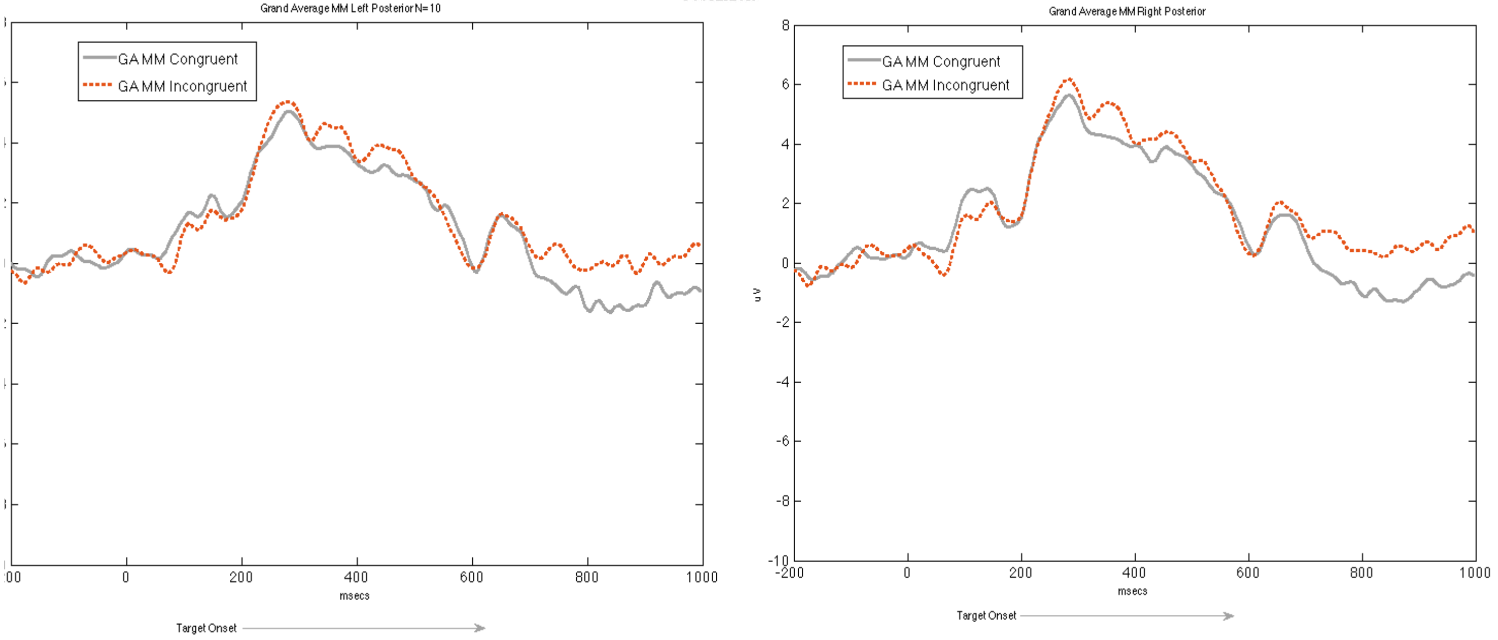
Right Anterior





Left Posterior

Right Posterior



Regional plots for Mental Model reasoning are displayed above in Fig. 8. Each of the plots shows the regionally averaged event-related responses to all congruent target gears (grey line) versus all responses to incongruent target gears (red line) for mental model reasoning. Time intervals (milliseconds) are plotted on the X-axes and microvolts (uV) on the Y-axes. The zero time point of target onset (200ms) is identified on the X-axes by a text arrow, and indicates the time of onset of the target gear (congruent meshing or incongruent meshing) for each problem. Approximate midpoints of significant ERPs have been encircled in light grey and labeled according to their polarity and latency.

It was hypothesized that a N400 ERP would be evoked in response to participants' unmet expectations when reasoning using the mental modeling strategy. Visual inspection of the waveform plots in Figure 8 shows a N400 ERP over the left anterior sensor montage, elicited in response to the incongruent target gears (dotted red line) relative to the congruent target gear (solid grey line). A secondary finding, a late negative deflection, can also be seen in the same waveform plot.

Follow on planned comparisons (dependent measures t-tests) were conducted to evaluate differences in mean amplitude between congruent and incongruent turning of the target gears in each of the four Montages for each of the three Time Windows. These tests showed that the incongruent manipulation resulted in significant ERP effects in time windows T2 (300-500 ms) and T3 (600-900 ms) in the Left Anterior scalp region: (T2: $t(9) = 2.525$, $p < .05$ and T3: $t(9) = 3.231$, $p < .05$, respectively.

Mental model reasoning (MM) mean amplitudes for grand averaged voltage responses to congruent and incongruent stimuli for each of the three Time Windows are given below in Tables 1-4.

Table 1 – MM: Left Anterior Region ERP Mean Voltages by Reasoning Condition and Time

ERP Identification	*Congruent Meshing Gear*	*Incongruent Meshing Gear*
Time Windows	Mean uV (SD)	Mean uV (SD)
Time 1 (0 - 200ms)	-1.0925 (2.7164)	-1.2819 (1.2082)
Time 2 (300 – 500ms)	-4.2244 (6.3077)	-5.7353* (6.6347)
Time 3 (600-900ms)	4.5642 (10.1149)	-.2488* 7.7917

*p < .05, two tailed

Table 2 – MM: Right Anterior Region ERPs Mean Voltage by Reasoning Condition and Time

ERP Identification	*Congruent Meshing Gear*	*Incongruent Meshing Gear*
Time Windows	Mean uV (SD)	Mean uV (SD)
Time 1 (0 – 200ms)	-2.1102 (3.4009)	-1.6014 (1.8954)
Time 2 (300 – 500ms)	-8.0354 (7.5126)	-8.373 (7.1677)
Time 3 (600-900ms)	-.7031 9.3752	-3.0701 6.5046

Table 3 – MM: Left Posterior ERPs Mean Voltage by Reasoning Condition and Time

ERP Identification	*Congruent Meshing Gear*	*Incongruent Meshing Gear*
Time Windows	Mean uV (SD)	Mean uV (SD)
Time 1 (0 – 200ms)	.8199 (1.9878)	.49502 (2.0182)
Time 2 (300 – 500ms)	2.5631 (5.9161)	3.149 (5.7172)
Time 3 (600-900ms)	-.8841 (4.3740)	.0923 (4.1470)

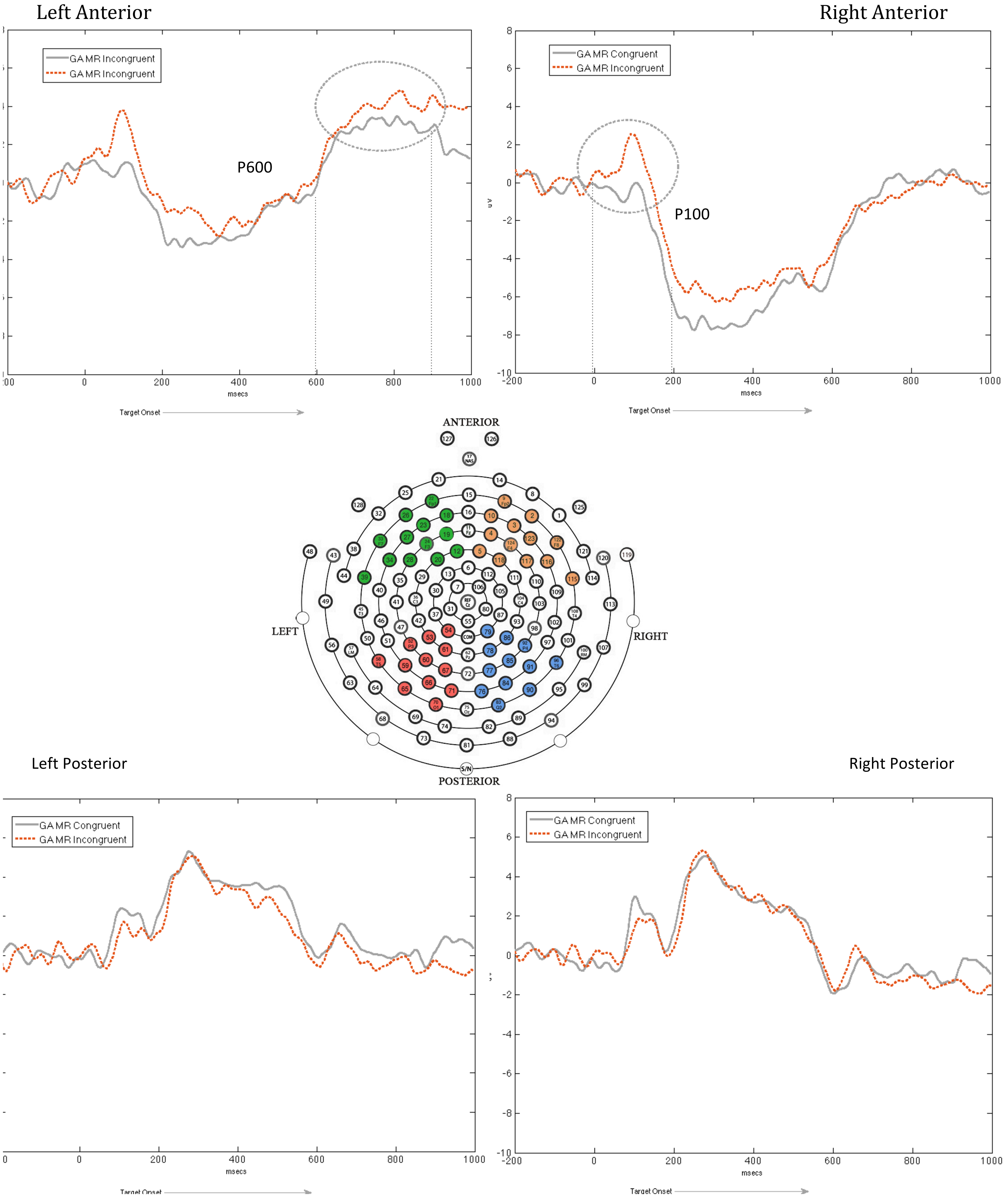
Table 4 – MM: Right Posterior ERPs Mean Voltage by Reasoning Condition and Time

ERP Identification	*Congruent Meshing Gear*	*Incongruent Meshing Gear*
Time Windows	Mean uV (SD)	Mean uV (SD)
Time 1 (0 – 200ms)	1.0742 (1.8742)	.7806 (1.9654)
Time 2 (300 – 500ms)	3.5115 (5.2554)	4.1038 (5.2258)
Time 3 (600 – 900ms)	-.1677 (4.466)	.7157 (4.0278)

Event Related Potentials (ERPs) in the Mental Rule Reasoning Condition:

Regional plots for mental rule reasoning (MR) are displayed below in Figure 9. Each of the plots shows the regionally averaged event-related responses to all congruent target gears (solid grey line) versus all responses to incongruent target gears (dotted red line) for mental rule reasoning. Time intervals (milliseconds) are plotted on the X-axes and microvolts (uV) on the Y-axes. The effective zero time point relative to target gear is at the 200-millisecond mark on the X-axes. Approximate midpoints of significant ERPs have been encircled in light grey and labeled according to their polarity and latency and/or functional association.

Figure 9: Mental Rule Reasoning ERP Waveforms



Left Anterior

Right Anterior

Left Posterior

Right Posterior

Regional plots for Mental Rule reasoning are displayed above Fig. 9. Each of the plots shows the regionally averaged event-related responses to all congruent target gears (grey line) versus all responses to incongruent target gears (red line) for mental rule reasoning. Time intervals (milliseconds) are plotted on the X-axes and microvolts (uV) on the Y-axes. The zero time point of target onset (200ms) is identified on the X-axes by a text arrow, and indicates the time of onset of the target gear (congruent meshing or incongruent meshing) for each problem. Approximate midpoints of significant ERPs have been encircled in light grey and labeled according to their polarity and latency.

It was predicted that when employing the mental rule strategy to reasoning about the gear problems, incongruent target gears would violate participants' expectations eliciting a P600 ERP. Visual inspection of the waveform plots (Figure 9) for time window T3 (600-900ms) shows the presence of a P600/SPS ERP over the Left Anterior groups of sensors. It can be observed in the waveform plot that responses to incongruent meshing gears yielded higher-amplitude positivities than congruent responses to turning of the target gear. Secondarily, an early (T1: 0-200ms) positive deflection can be seen over the Right Anterior sensor group. The P1 ERP seen in this waveform appears prior to onset of the eliciting target.

Tests of the simple effects comparing participant brain responses to congruent and incongruent problem endings revealed one significant ERP effect and another approaching significance in two different regional scalp montages in two different time windows. In the Right Anterior scalp montage for time window 1 (T1: 0-200ms) the P1 ERP was significant, t (9) = -2.296, $p < .05$. In time window 3 (T3: 600-900ms) in the Left Anterior scalp quadrant the P600/SPS ERP was marginally significant, t (9) = -2.248, $p < .051$.

Means and standard deviations for the grand averaged voltages during the four time windows are reported for Mental Rule reasoning below in Tables 5-8.

Table 5: MR: Left Anterior Region ERP Mean Voltage By Condition and Time

ERP Identification	*Congruent Meshing Gear*	*Incongruent Meshing Gear*
Time Windows	Mean uV (SD)	Mean uV (SD)
Time 1 (0 – 200ms)	.2370 (2.9652)	1.4750 (6.967)
Time 2 (300 – 500ms)	-2.4624 (5.2953)	-1.9747 (5.9457)
Time 3 (600 – 900ms)	2.6610 (6.7215)	3.424** (6.4756)

**p = .051, two-tailed, marginally significant

Table 6: MR: Right Anterior Region ERP Mean Voltage By Reasoning Condition and Time

ERP Identification	*Congruent Meshing Gear*	*Incongruent Meshing Gear*
Time Windows	Mean uV (SD)	Mean uV (SD)
T1 (0 – 200ms)	-1.5102 (3.9810)	.1757* 5.3801)
T2 (300 – 500ms)	-6.7635 (5.5078)	-5.4980 (4.9761)
T3 (600 – 900ms)	-1.5475 (6.8136)	-.9265 (6.6922)

*p < .05, two-tailed

Table 7: MR: Left Posterior Region ERP Mean Voltage by Reasoning Condition and Time

ERP Identification	*Congruent Meshing Gear*	*Incongruent Meshing Gear*
Time Windows	Mean uV (SD)	Mean uV (SD)
T1 (0 – 200ms)	1.0479 (1.4486)	.6368 (2.1162)
T2 (300 – 500ms)	3.6970 (3.1102)	3.3110 (3.0723)
T 3 (600 – 900ms)	.2797 (3.2090)	-.1863 (3.0323)

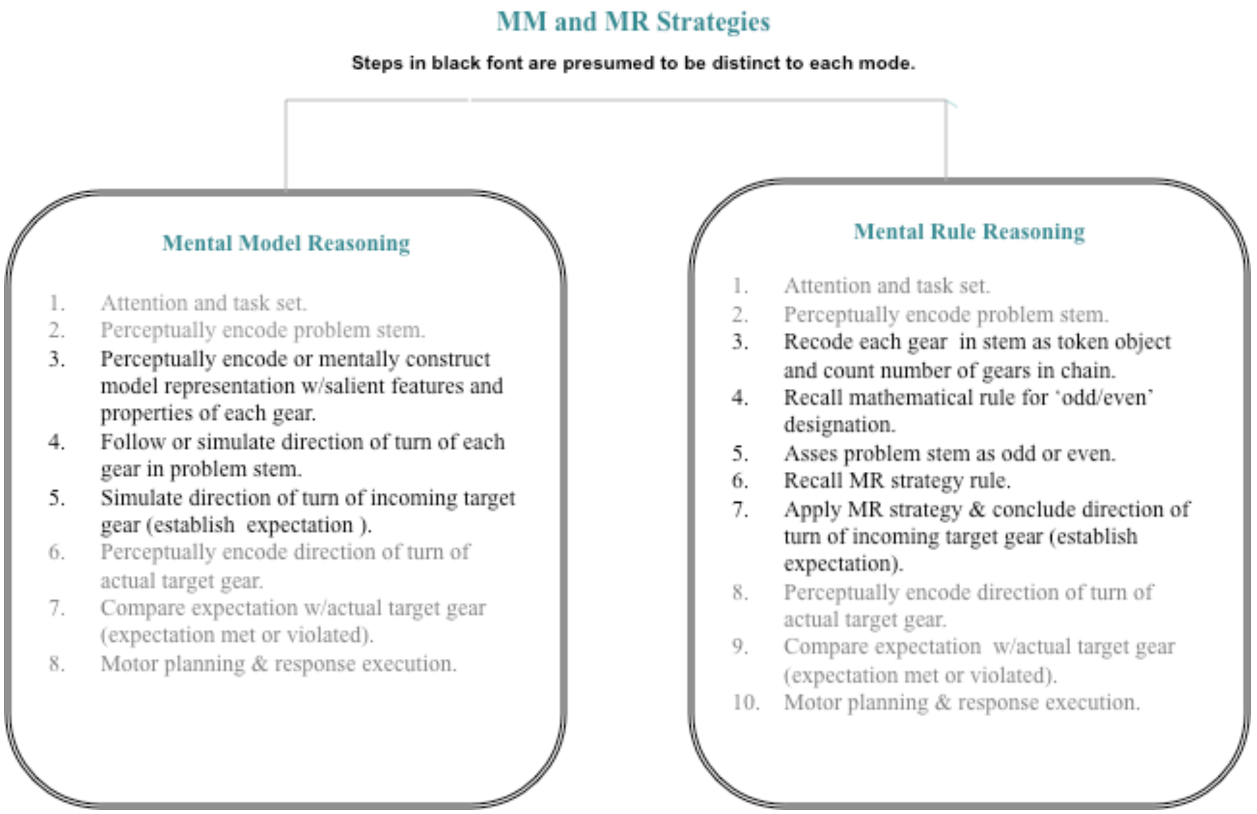
Table 8: MR: Right Posterior Region ERP Mean Voltage By Reasoning Condition and Time

ERP Identification	*Congruent Meshing Gear*	*Incongruent Meshing Gear*
Time Windows	Mean uV (SD)	Mean uV (SD)
Time 1 (0 –200ms)	.7727 (1.6844)	.6541 (2.3575)
Time 2 (300 – 500ms)	2.9944 (3.3659)	3.0962 (4.0538)
Time 3 (600-900ms)	-.96270 (4.9854)	.9673 (5.0073)

7. DISCUSSION

In the mental model reasoning task (MM) it is assumed that participants solved the gear problems by mentally simulating the rotation of the gears in the problem stem (i.e., by imagining and mentally inspecting each gear in the chain as it turns or by tracking the onscreen dynamic simulation of the gears turning), and then "reading off" a prediction based on the interdependency of adjacent gears. For mental rule reasoning (MR), it is assumed that use of the strategy induced participants to solve the problems by first counting or estimating the number of gears in the problem stem (treating each observed gear merely as a token entity), retrieving the mathematical concept of "odd" versus "even," from memory stores, and finally using the trained rule to make a prediction. An analysis of the cognitive processes thought to be involved in use of the strategies to reasoning about the gear problems follows in Figure 10.

Figure 10: Cognitive Analysis



MM and MR Strategies

Steps in black font are presumed to be distinct to each mode.

Mental Model Reasoning

1. Attention and task set.
2. Perceptually encode problem stem.
3. Perceptually encode or mentally construct model representation w/salient features and properties of each gear.
4. Follow or simulate direction of turn of each gear in problem stem.
5. Simulate direction of turn of incoming target gear (establish expectation).
6. Perceptually encode direction of turn of actual target gear.
7. Compare expectation w/actual target gear (expectation met or violated).
8. Motor planning & response execution.

Mental Rule Reasoning

1. Attention and task set.
2. Perceptually encode problem stem.
3. Recode each gear in stem as token object and count number of gears in chain.
4. Recall mathematical rule for 'odd/even' designation.
5. Asses problem stem as odd or even.
6. Recall MR strategy rule.
7. Apply MR strategy & conclude direction of turn of incoming target gear (establish expectation).
8. Perceptually encode direction of turn of actual target gear.
9. Compare expectation w/actual target gear (expectation met or violated).
10. Motor planning & response execution.

It was expected that a N400 ERP would be evoked in response to participants' unmet expectations in the mental model (MM) problem reasoning condition. The main finding in the MM reasoning tasks was the significant left lateralized anteriorly presented N400 (T2: t (9) = 2.525, p < .05; Time 3: t (9) = 3.231, p < .05. A finding of a N400 response in the MM condition suggests that reasoning about the relations between entities in the type of problems presented recruits cortical areas previously shown to be involved in processing violations of semantic expectancies in studies of language comprehension (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980a, b). In addition, an N400 elicited to MM reasoning is consistent with a more general interpretation of it as a more general index of relational processing (Barrett & Rugg, 1990; Dohmas et al., 2007; Federmeier & Laszlo, 2009; Kutas & Federmeier, 2011; West and Holcomb, 2002).

However, the topography of the N400 observed in response to using the MM strategy is more similar in its latency and distribution to that reported in (Sitnikova, Goff, & Kuperberg, 2009; Sitznikova, Holcomb, & Kuperberg, 2008a; Swaab et al., 2002; West & Holcomb, 2002). In these studies, incongruous video, pictures and concrete word (easily envisioned) endings produced an increased N400 effect that was more anteriorly distributed. The left lateralized and frontally distributed N400 ERP found in the current experiment fits with the proposition of a N400 effect that is sensitive to processing imagistic input (versus the classic more posteriorly distributed N400 that is associated with the processing of linguistic stimuli). Moreover, a finding of a frontally distributed N400 is consistent with other reports of non-typical distributions of the waveform, which has contributed to an emerging view of this component as a composite of several sub-components Dien et al. (2009). Aligned with this view, West & Holcomb (2002) have suggested that the N400 may represent the summation of coordinate neural activity recruited from varying large networks each responsive to different task demands and stimuli input. Finally, the topographical distribution of the N400 elicited in the MM condition is in line with fMRI research suggesting that complex reasoning preferentially draws upon left anterior neural resources (Christoff, 2009; Goel, 2009; Goel & Dolan, 2004; Kroger et al., 2002; Prabhakaran et al., 1997).

For the mental rule reasoning (MR), task it was predicted that incongruent target gears would violate participants' expectations eliciting a P600/SPS ERP component. Although the elicited effect (T3: 600-900ms) did not quite reach statistical significance (t (9) = -2.248, p = .051), its clear presence in the waveform coupled with the near significance suggest that the strategy manipulation may provide an approach to assessing rule-like conceptual reasoning. The neurofunctional specificity of the P600 has been related to mental operations supporting reprocessing of sentence meaning owing to unexpected violations of syntactic structures (Friederici, et al., 1993; Hagoort, Brown, & Groothusen, 1993; Neville, Nicol, Barss, Forster, Garrett, 1991; Osterhout & Holcomb, 1992). Although the P600 observed in the MR condition presented at more anterior scalp locations than those found in classic language studies, the topographic distribution of the P600s has been shown to differ depending on the nature of the assumed cognitive activities. Frontally presenting P600s have been observed in studies in which participants were required to “revise” as opposed to “repair” phrase structures (Hagoort et al., 1999; Osterhout & Holcomb, 1992; van Berkum, Brown, & Hagoort, 1999). In these studies the authors distinguish repair processes as those defined by phrase restructuring following collapse of meaning introduced by ungrammatical syntactic continuations, and revision as processes supporting the overwriting of a preferred current structural representation (Hagoort et al., 1999). More frontal P600s have also been noted in studies of sentence processing in which complex (two-noun) fragments preceding the critical verb were compared with simple (one-noun) fragments, even when the phrase completing words formed an unambiguous sentence (Friederici et al., 2002; Hagoort et al., 1999). An interpretation of the P600 evoked in the MR condition to incongruent turning target gears, as indexing cognitive processes related to overwriting a preferred solution is quite plausible. It is also, equally plausible that the anterior topography of the P600 seen in MR reasoning can be explained as reflecting greater neural resources associated with more complex cognition, such as reasoning (Goel et al., 1997, 1998, 2000; Prabhakaran, et al., 1997; Osherson et al., 1998). Lastly, the presence of a P600/SPS ERP effect in a reasoning paradigm lends support to the proposition that the P600/SPS component is a more general index of rule governing processing (Gouvea et al, 2009; Patel, 1998).

The predicted ERP components were observed to be left lateralized in both the MM and MR conditions. Whereas, it might have been expected that reasoning with the MM strategy would result in more right lateralized activations, and MR uniquely in left lateralized responses, the left frontal findings in the current study are in fact, consistent with results from over a decade of neuroscientific research on reasoning concluding that the left pre-frontal cortex (PFC) plays a dominant role during tasks of thinking and reasoning (Goel, et al., 1997; Prabhakaran, et al., 1997; Goel, 1998; Kroger et al., 2002; Goel, 2003; Noveck et al., 2004; Goel & Dolan, 2004; Christoff, 2009).

While only the statistical test for mental model reasoning (N400) reached significance, the two different neurophysiological signatures (i.e., polarity and latency) observed to be associated with the two modes of reasoning as operationalized by the use of two different strategy paradigms to solve the same simple binary question (clockwise or counterclockwise turning of a target gear in a simple mechanical system of gears), provide evidence for dual or multiple-code theories of cognition. Specifically, the hypothesis that MM reasoning depends more on direct sensory representations and relational processing, and that MR reasoning makes use of symbolic representations operated over by syntactic-like processes, is supported by these research findings. However, as discussed above, the N400 and P600 waveforms observed in these experimental results were both distributed in the left anterior scalp region. Although there is evidence from several fMRI studies that inductive and deductive reasoning may recruit from different cell assemblies within the left prefrontal cortex (Christoff, 2009; Goel & Dolan, 2004; Kroger, et al., 2002), further spatial differentiation within the left PFC cannot be determined at the present time for the MM and MR task outcomes. Nonetheless, the N400 component effect was elicited to incongruous target gear endings in the MM task, but not a P600; and in the MR task condition a P600 waveform was observed, but not a N400 ERP. This dissociation within the left anterior scalp quadrant suggests that different neuronal sources underwrite the representations and mechanisms giving rise to the two modes of reasoning, and that mental model (MM) and mental rule (MR) reasoning represent at least partially different domains of reasoning (Dien, Michelson, & Franklin, 2010; Christoff, 2009; Luck, 2005; Rugg & Coles, 1995).

The secondary findings of other significant waveforms in both modes of reasoning may provide further support for the idea of different reasoning networks. The significant late negativity (tentatively labeled N700) found in the MM task application, and the P100 found in the MR task execution may reflect certain cognitive capacities that are thought to support reasoning, such as sensory processing, attention and memory, or as well, activity related to the transformation and integration of mental representations (Bender, Oelkers-Ax, Hellwig, Resch, & Weisbord, 2008; Gratton et al., 2008). A frontal N700 has previously been reported in the literature in a study contrasting concrete (imageable) and abstract words (West & Holcomb, 2000). In the study, the N400 was elicited to both incongruent concrete and abstract word endings, whereas the N700 was seen only in response to easily imageable final words that introduced unexpected semantic input. The distinct N400 and N700 responses seen in this experiment, led the authors to propose that the N400 may reflect a unified semantic processing system supporting evaluation of both categories of words, but that the N700 may uniquely index image-mediated representations. Such an interpretation of the N700 fits well in the context of the MM task condition implemented in the current study.

Recent research (Bender et al., 2008) provides another possible interpretation of the late negativity found in the MM task condition. The authors reported evidence of a late visual processing component, which they referred to as the “visual N700” based on results from two experiments. Methodologies allowing for the quantification and spatial-temporal analysis of EEG data were used in both experiments. The aim of the research was to localize the neural generators of a previously noted late negative wave, separate from those of well known early visual potentials (e.g., C1, P1), the sustained potential (SP) or even endogenous components such as the P300 complex. In the first experiment participants passively viewed rapidly presented simple short duration pairs of visual stimuli, and in the second experiment participants viewed a long-lasting visual stimulus (7000 ms). Results from the two experiments provided compelling evidence that the early exogenous sensory components were more localized to mid-occipital and occipital extrastriate cortices (C1 and P1, respectively), and that the late negative wave was generated by sources beginning in posterior bi-lateral extrastriate occipital cortex but extending up into secondary visual cortex (bilateral occipito-temporal junctions). Results of the spatial-temporal analysis located the posterior bi-lateral sources of the N700 near to the dipole sources of the early P1. Based on these findings the authors proposed the existence of a “visual N700” complex that is phasic in nature, consisting of an early occipital part and a later occipital-temporal phase (Bender et al., 2008). This phasic “visual N700” was hypothesized to represent a late perceptual post-processing stage in the ventral visual stream (i.e., the visual pathway concerned with object identification) concerned with assessing the visual features of stimuli as would be necessary for integration of sequential stimulation (Bender et al., 2008).

There are parallels between the late negative deflection observed in the current research and that found in the Bender et al. (2008) study that make this interpretation of the N700 feasible. First, the T3 (600-900ms) time window specified in the current experiment aligns with the 600-1000 millisecond window reported by Bender et al., (2008). Second, stimuli presentation was similar in the two experiments, that is, participants viewed long lasting visual stimuli (7000ms and 8000ms in Bender et al., 2008 and the current research, respectively). Finally, the cognitive processes proposed to support implementation of the MM task are consistent with Bender et al.'s (2008) functional interpretation of a "visual N700" component. However, as was the case with the predicted N400 and P600 ERPs, the late negativity reported in the current research presented with a more anterior scalp topography than the N700 observed by Bender et al., 2008. It is worth noting, however, that the stimuli presented in the latter's research were simple symbols (an exclamation mark, a circle, a cross, a star) and that the task only required passive viewing. In contrast, stimuli and task demands in the current experiment were considerably more robust and this may also offer an explanation for the different topographical distributions of the observed "visual N700" found by Bender et al. (2008) and the late negative deflection reported here (Christoff, 2009; Ishai, Ungerleider, & Haxby, 2000; Shimamura, 1995; West & Holcomb, 2000). The late negative wave found in these results was not predicted, and of course, the experiment was not designed to either elicit it or correlate it with a specific cognitive event.

Thus, while a "visual N700" is a reasonable interpretation of the finding in the current research it remains an untested but ready hypothesis for future research.

The significant P100 waveform observed during MR reasoning would seem to be an occurrence of the well documented visually evoked positive deflection known to peak around 100 milliseconds in response to visual stimuli (Jeffreys & Axford, 1972). The P100 is regarded to be an index of low-level perceptual analysis (Heinze et al., 1994; Martinez et l., 1999, 2001). It has been elicited to both words and nonwords, with evidence pointing to a later presentation to words (toward 158 ms), and an earlier presentation (closer to 100 ms) to nonwords (Segalowitz & Zheng, in press). Its presence at 100 milliseconds following presentation of the target may reflect early perceptual processing of this incoming last gear. Assessed in the context of the cognitive processes thought to support mental rule (MR) reasoning, a low-level evaluation of the target, [that is, making a nominal/binary comparison of its direction of turn] merely comparing its direction of turn (same vs. different), may have been sufficient for responding in the MR reasoning condition.

This interpretation is also consistent with the research of Bender et al. (2008) reviewed above, in which a P100 component was also observed and described as an automatic response to exogenous stimuli, by comparison to the endogenous "visual N700, " thought to be related to more internally generated representations.

Conclusions

The primary objective of the current experiment was to investigate the neurophysiological correlates of two modes of reasoning in order to examine the question of whether the human brain executes conceptual processing on different types of mental representations, and also if the inferential computations associated with each of the respective representations differ. Outcomes of the study suggest that MM and MR reasoning represent two different domains of reasoning both psychologically and neurophysiologically. West & Holcomb (2002, p. 363) have pointed out that the human brain is uniquely equipped to store and process conceptual representations formed on direct sensory stimuli as well as symbolic stimuli. The results of the current study revealed a N400 ERP effect thought to be in response to manipulation of the relatedness of visual-spatial stimuli based on an inspection of their visual-spatial properties and temporal contiguity. This finding also contributes evidence to the hypothesis that the N400 ERP component may be an index of more general cognitive processes concerned with assessing relatedness between entities regardless of modality. The marginally significant P600 ERP effect (in response to symbolic-based representations) provides weaker, but suggestive evidence for a reasoning system that makes use of abstract (digit) stimuli and rule-governed computations. This characterization of the P600 component, also lends support to a more general interpretation of its cognitive functionality as indexing cognitive activities that underwrite the construction of a mental proposition based on rules.

8. STUDY LIMITATIONS AND DELIMITATIONS

EEG data collection protocols are well established, and one should expect that experimental procedures would be implemented as planned. In fact, technological advancements with the computer platforms used to collect such data have rendered the process less stable and subject to unpredictable disruptions. The impact of equipment failure experienced during the data collection was to introduce an undesirable level of noise in many recordings, resulting in relatively high attrition of the sample size. These effects on the research were minimized to the extent possible by recording a sufficient number of trials, applying appropriate data pre-processing methodologies and recruiting more than the required number of participants needed to obtain statistical power. Data were closely examined after acquisition. Where movement artifacts were too extreme for data to be salvaged for a given trial, that trial was exclude from the ongoing analysis. The same standard was applied to whole participant files. If data were contaminated by disruption to the recording, unexplained noise, or movement artifacts for the whole recording period, data from that subject were not included in any of the group analysis. This resulted in the inclusion of 10 out of 20 participant files, allowing only reasonable statistical power to be obtained.

The use of peak and mean amplitude quantification and measurement procedures are well developed, but are not regarded as wholly satisfactory for either fully interrogating EEG data or satisfying the assumptions upon which statistical tests, such as ANOVA and Student T, are implemented. The analytical procedures used in this research included establishing *a priori* hypotheses and time windows based on prevailing research in the domain, and the reporting of conservative adjusted test statistics. Questions of how best to quantify and measure ERPs define an important area for further study within the domain of ERP research.

Item analysis was not conducted on the stimuli used in the research. The gear problems used as stimuli in the experiment were not analyzed for their effect on reasoning relative to differences in length, nor the placement or direction of turn of the lead gear. Therefore, it cannot be said that some of the effects measured were not owing to these factors.

The conclusions in this study might be reserved due to the chronological age and years of schooling used as the classificatory parameter for defining the participant group. Frequently, physiological age is found to be different from chronological age and presents with significant variation between and among individuals. Years of schooling may mean that individuals in the targeted population were simply adept at following task instructions that elicit ways of reasoning that are neither natural nor spontaneous.

9. FUTURE DIRECTIONS

An examination of the neurophysiological correlates of an instance of mental model (MM) and mental rule (MR) reasoning may produce brain-based evidence of the human capacity for reasoning in different ways about the same presenting situation. If human biology affords reasoning in dual or multiple ways, expanding our knowledge about the neural implementation of the representational systems that support understanding and learning may provide opportunities to rebuild and enrich competencies (the term is used in the Chomskyan sense) in ways that have not as yet been explored in any systematically meaningful manner (Gardner, 1983).

A study has been developed and conducted, and data collected and analyzed, which may contribute to existing literature on the physiological correlates underlying two classic modes of reasoning and the functional interpretation of two standard ERPs. To my knowledge the current study is the first to directly compare the brain activity evoked during model-based versus rule-based reasoning using EEG.

The study has compared the temporal and spatial properties of MM and MR reasoning, with clear hypotheses about expected differences related to the psychological processes and the postulated representational systems with which they are associated. The work has the potential to add hitherto unavailable information about the temporal dimension of reasoning that may contribute to a richer understanding of how functional specialization within cortical areas may be instantiated and coordinated across brain regions. ERP research has largely followed a serial information processing model of human cognition, but recent technical advances in both imaging methodologies and data analysis are beginning to suggest that multi-stage (early and late) dynamic semantic and visual-spatial processing models may support high-order cognition (see Dien, et al., 2009 for a discussion of an evolving view of more phasic processing model). Toward this end, further work to interrogate these data in ways that will allow for the event-related potentials to be examined as a function of time by condition for the whole scalp could provide crucial insight about how reasoning processes unfold, both intra-regionally and inter-regionally.

10. REFERENCES

Ackrill, J. L. (2001). *Essays on Plato and Aristotle.* USA: Oxford University Press.

Acuna, B. D., Eliassen, J. C., Donoghe, J. P., &Sanes, J. (2002). Frontal and parietal lobe activation during transitive inference in humans.*Cerebral Cortex, 12,* 1312-1321.

Anderson, J. (1978). Arguments concerning representations for mental imagery. *Psychological Review, 85(4),* 249-277.

Arnheim, R. (1969). *Visual thinking.* London: Faber and Faber.

Aristotle (1995). Sense and sensibilia. (J. I. Beare, Tran.). In Complete Works of Aristotle: Volume I. In J. Barnes (Ed.). Princeton: Princeton University Press. (Original work published in 4th century B. C.).

Aristotle (1961). *De anima, Books II and III,* (D. W. Hamlyn, Tran.). Oxford: Oxford University Press. (Original work published in 4th century B. C.).

Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology, 49A(1)*, 5-28.

Baddeley, A. D. (1986). *Working memory.* Oxford: Oxford University Press.

Baddeley, A. D., & Hitch, G. (1974).Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-90). New York: Academic Press.

Badre, D. (2008). Ventrolateral prefrontal cortex and controlling memory to inform action. In S. A. Bunge & J. D. Wallis (Eds.), *Neuroscience of rule-guided behavior* (pp. 365-389). Oxford: Oxford University Press.

Bailer-Jones, D. M. (1999).Tracing the development of models in the philosophy of science.In L. Magnani, N. J. Nersessian, & P. Thagard (Eds.), *Model-based reasoning in scientific discovery* (pp. 23-39). New York: Kluwer Academic/Plenum Publishers.

Bajric, J., Rösler, F., Heil, M., & Henninghausen, E. (1999). On separating process of event categorization, task preparation, and mental rotation proper in a handedness recognition task. *Psychophysiology, 36*, 399-408.

Bara, B. G., & Bucciarelli, M. (2000). Deduction and induction: Reasoning through mental models. *Mind & Society, 1(1),* 95-107.

Barbey, A. & Barsalou, L. W. (2006). Intelligence: models of reasoning. In L. Squire, T. Albright, F. Bloom, F. Gage, & N. Spitzer (Eds.), *New Encyclopedia of Neuroscience* (pp. 35-43). Oxford: Elsevier.

Barrett, S. E., & Rugg, D. (1990). Event-related potentials and the semantic matching of pictures.*Brain and Cognition, 14(2),* 201-12.

Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22(4),* 577-660.

Beare. J. I. (1906).Greek theories of elementary cognition from Alcemeaon to Aristotle. Oxford: The Claredon Press.

Beilin, H. (1992). Piaget's enduring contribution to developmental psychology. *Developmental Psychology,*28 (2), 191–204.

Bender, S., Oelkers_Ax, S., Hellwig, S., Resch, F., Weisbord,. M., (2008). The topography of the scalp-recorded visual N700. *Clinical Neurophysiology, 119,* 587-604.

beim Graben, P., Gerth, S., Vasishth, S. (2008). Toward dynamical system models of language-related brain potentials. *Cognitive Neurodynamics, 2*, 229-255. doi 10.1007/s11571-008-9041-5.

Blackburn, S. (1996). *The oxford dictionary of philosophy.* Oxford: Oxford University Press.

Boden, M. A. (2008). *Mind as machine*: *A history of cognitive science.* Oxford: Oxford University Press.

Bogdan, R. J. (1992). Cognitive science. In H. Burckhardt & B. Smith (Eds.), *Handbook of Metaphysics and Ontology: Philosophia* (pp. 69-73). Munich: Springer Verlag.

Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, *85,* 1-21.

Braine, M.D.S., & O'Brien, D.P. (1998). *Mental logic*. Mahwah, NJ: Lawrence Erlbaum Associates.

Broadbent, D. E. (1958). *Perception and communication.* London: Pergamon Press.

Chase, W. G., & Simon, H. A. (1973).The mind’s eye in chess. In Chase, W. G. (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.

Chater, N., & Oaksford, M. (1993). Logicism, mental models and everyday reasoning: Reply to Garnham. *Mind & Language, 8,* 73-89.

Cheng P. W., & Holyoak, K. J. (1985). Pragmantic reasoning schemas. *Cognitive Psychology. 17,* 391-416.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and two ears. *Journal of Acoustic Society of America 25(5)* 975–979. doi:10.1121/1.1907229.

Chrisley, R. (2000). *Artificial intelligence: Critical concepts*. London: Routledge.

Christoff, K. (2009). Human thought and the lateral prefrontal cortex. In E. Kraft, B. Gulyas, & E. Poppel (Eds). *Neural correlates of thinking* (pp. 219-252). Berlin: Springer-Verlag.

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., & Gabrieli, J. D. E. (2001). Rostrolateral prefrontal cortex involvement in relational intergration during reasoning. *Neuroimage,14*, 1136-1149.

Christoff, K., Ream, J. M., Geddes, L. P. T., &Gabrieli, J. D. E. (2003). Evaluating self-generated information: Anterior prefrontal contributions to human cognition. *Behavioral Neuroscience, 117(6)*, 1161-1168.

Christoff, K., &Gabrieli, J. D. E. (2000). The frontopolar cortex and human cognitive cognition: evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychology 28(2)*, 168-186.

Chomsky, N. (1972). *Language and Mind.* New York: Harcourt Brace Jovanovich.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge: MIT Press.

Chomsky, N. (1963). Formal properties of grammars. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 3, pp. 323-418). New York: Wiley.

Chomsky, N. (1959a). A review of B. F. Skinner's verbal behavior [Review of the book verbalbehavior, by B. F. Skinner]. *Language, 35,* 26-58.

Chomsky, N. (1959b). On certain formal properties of grammar. *Information and Control, 2,* 137-167.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton. [Reprint. Berlin and New York].

Church, A. (1936). An unsolvable problem of elementary number theory. *The American Journal of Mathematics 58(2)*, 345-363. doi: 10.2307/2371045.

Coles, M. G. H., Rugg, M. D. (1995). Event-related potentials: An introduction. In M. D. Rugg, and M. G. Coles, (Eds.), *Electrophysiology of mind: Event-related potentials and cognition.* New York: Oxford University Press.

Collins, M. (2010). *The nature and implementation of representation in biological systems.* (Unpublished doctoral dissertation). Graduate Center, New York, NY.

Cooper, L. A. (1975). Mental rotation of random two-dimensional shapes. *Cognitive Psychology, 7,* 20-43. doi:10.1016/0010-0285(75)90003-1.

Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visusal information processing* (pp. 135-142). New York: Academic Press.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? studies with the Wason selection task. *Cognition 31,* 187-276.

Cosmides, L. Tooby J. (1994). Origins of domain specificity: The evolution of functional organization. In L. A. Hirschfield & S. A. Gelman (Eds.), *Mapping the mind (pp. 84-116).* New York: Cambridge University Press.

Cottrell, G., & Metcalfe, J. (1991). Face, gender and emotion recognition using Holons. In R. Lippman, J. Moody, & D. Touretzky (Eds.), *Advances in Neural Information Processing Systems 3,* 564 – 571. San Mateo, CA: Morgan Kaufman.

Coulson, S., King, J. W., & Kutas, M. (1998b). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13,* 21–58.

Craig, D. L., Nersessian, N. J., & Catrambone, R. (2002). Perceptual simulation in analogical problem solving. In L. Magnani, & N. J. Nersessian (Eds.), *Model-based reasoning: cience, technology, & values* (pp. 167-191). New York: Kluwer Academic/Plenum Publishers.

Craik, F. I. M., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Craik, K. (1943). *The nature of explanantion*. Cambridge: Cambridge University Press.

Cummings, A., Ceponiene, R., Koyama, A., Saygin, A. P., Townsend, J., & Dick, F. (2008). Auditory semantic networks for words and natural sounds. *Brain Research*, *1115,* 92-107.

Daugman, J. G. (1990). An information-theoretic view of analog representation in striate cortex. In E. Schwartz (Ed.), *Computational Neuroscience* (403-423). Cambridge: MIT Press.

Descartes, R. (1996). *Meditations on First Philosophy*, (J. Cottingham, Trans.). Cambridge: Cambridge University Press, 1996. (Original work published 1641.)

Descartes, R. (1970). Discourse on Method. In E. S. Haldane, & G. R. Brown (Eds.). *The philosophical works of Descartes* (Vol. 1, pp. 178-291).Cambridge: Cambridge University Press. (Original work published 1637).

Dien, J., Michelson, C., Franklin, M. (2010). Separating the visual sentence N400 effect from the P400 sequential expectancy effect: Cognitive and neuroanatomical implications. *Brain Research, 1355,* 126-140.

Donchin, E., & Coles, M. G. H. (1988a). Is the P300 component a mnifestation of context updating? *Behavioral and Brain Sciences, 11*, 355-72.

deKleer, J. D., & Brown, J. S. (1983). Assumptions and ambiguities in mechanistic mental models. In D. Gentner &A. L. Stevens (Eds.), *Mental models* (pp. 155-190). Hillsdale, NJ: Erlbaum.

Dien, J., Michelson, C. A., Franklin, M. S. (2010). Separating the visual sentence N400 effect from the P400 sequential expectancy effect: Cognitive and neuroanatomical implications. *Brain Research,* 1355 (2010) 126-140.

Dien, J., 2009. The neurocognitive basis of reading single words as seen through early latency ERPs: a model of converging pathways. Biol. Psychol. 80 (1), 10–22.

Dien, J., & Santuzzi, A. M. (2005). Application of repeated measures ANOVA to high-density ERP datasets: A review and tutorial. In T. C. Handy (Ed.), *Event-related potentials: A methods handbook*. Cambridge: MIT Press.

Dohmas, F., Dohmas, U., Schlesewsky, M., Ratinckx, E., Verguts, T., Willmes, K., & Nuerk, H. (2007). Neighborhood consistency in mental arithmetic: Behavioral and ERP evidence. *Behavioral and Brain Functions, 3(1),* 66.doi:10.1186/1744-9081-3-66.

Edelman, G. M. (1987). *Neural Darwinism: The theory of neuronal group selection.* New York: Basic Books.

Evans, J. St. B. T. (2003). In two minds: dual process accounts of reasoning. *Trends in Cognitive Sciences, 7(10)*, 454-459.

Evans, J. S. B. T., Barston, J., Pollard, P. (1983). On the conflict bewteen logic and belief in syllogistic reasoning. *Memory Cognition, 11*, 295-306.

Falmagne, R. J., & Gonsalves, J. (1995). Deductive inference. *Annual Review of Psychology, 46,* 525-559.

Federmeier, K. D., & Laszlo, S. (2009). Time for meaning: electrophysiology provides insights into the dynamics of representation and processing in semantic memory. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 1-44). San Diego, CA: Elsevier.

Feuer, M. J. (2006). *Moderating the debate: Rationality and the promise of American education.* Cambridge: Harvard Education Press.

Finke, R. A. (1989). *Principles of mental imagery.* Cambridge: MIT Press.

Finke, R. A. (1980). Levels of equivalence in imagery and perception. *Psychological Review, 87(2),* 113-132.

Fodor, J. (1983). *The modularity of mind*. Cambridge: MIT Press.

Fodor, J. (1981). *Representations: Philosophical essays on the foundations of cognitive science.* Cambridge: Bradford Books/MIT Press.

Fodor, J. (1975). *The language of thought.* Cambridge: Harvard University Press.

Fodor, J., & Pylyshyn, Z., (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28,* 3-71.

Forbus, K. D. (1983). Qualitative reasoning about space and motion. In D. Gentner & A. Stevens (Eds.), *Mental models* (pp. 53-72). Hillsdale, NJ: Lawrence Erlbaum Associates.

Friederici, A., D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research, Vol. 31(1),* 45-63.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 1,* 183–192.

Fuster, J. M. (2006). The cognit: a network model of cortical representation. *International Journal of Psychophysiology, 60,* 125-132.

Galaburda, A. M., Kosslyn, S. M., & Christen Y. (2002). Introduction. In A. M. Galaburda, S. M. Kosslyn, & Y. Christen (Eds.), *The languages of the brain* (pp. 1-14). Cambridge: Harvard University Press.

Gardner, H. (1983). *Frames of mind. The theory of multiple intelligences*. New York: Basic Books.

Garnham, A., & Oakhill J. (1994). *Thinking and Reasoning.* Malden, MA: Blackwell Publishers, Inc.

Gentner, D. 1983. Structure-mapping: a theoretical framework for analogy. *Cognitive Science, 23,* 155-170.

Giere, R. N. (1999). Using models to represent reality. In L. Magnani, N. J. Nersessian, & P. Thagard (Eds.), *Model-based reasoning in scientific discovery* (pp. 41-57). New York: Kluwer Academic/Plenum Publishers.

Glenberg, A. M. & Robertson, D., A. (2000). Symbol Grounding and Meaning: A Comparison of High-Dimensional and Embodied Theories of Meaning. *Journal of Memory and Language, 43*, 379 – 401.

Goel, V. (2009). Fractionating the system of deductive reasoning. In E. Kraft, B. Gulyas, & E. Pöppel (Eds.), *Neural correlates of thinking* (pp. 203-218). Berlin: Springer-Verlag.

Goel, V. (2003). Evidence for dual neural pathway for syllogistic reasoning. *Psychologica, 32,* 301-309.

Goel, V., Buchel, C., Frith, C., & Dolan, R. J. (2000). Dissociation of mechanisms underlying syllogistic reasoning. *Neuroimage 12,* 504-514.

Goel, V., & Dolan, R. J. (2004). Differential involvement of the left prefrontal cortex in inductive and deductive reasoning. *Cognition*, 93*(3),* B109-B121.

Goel, V., Dolan, R., J. (2003). Explaining modulation of reasoning by belief. *Cognition, 87,* B11-B22.

Goel, V., & Dolan, R. J. (2001). Functional neuroanatomy of three-term realtional reasoning. *Neuropsychologia, 39(9),* 901-909.

Goel, V., & Dolan, R. J. (2000).Anatomical segregation of component processes in an inductive inference task. *Journal of Cognitive Neuroscience*. 12:1, 110-119.

Goel, V., Gold, B., Kapur, S, &Houle, S. (1997). The seats of reason: A localization study of deductive and inductive reasoning using PET (O15) blood flow technique. *NeuroReport, 8(5),* 1305-1310.

Goel, V., Kapur, S., &Houle, S. (1998). Neuroanatomical correlates of human reasoning. *Journal of Cognitive Neuroscience*, *10(3),* 293-302.

Goel, V., Makale, M., & Grafman, J. (2004). The hipppocampal system mediates logical reasoning about familiar spatial environments. *Journal of cognitive neuroscience, 16(4),* 654-664.

Goel, V., Tierney, M., Sheesley, L., Bartolo, A., Vartanian, O., & Grafman, J. (2007). Hemeispheric specialization in human prefrontal cortex for resolving certain and uncertain inferences. *Cerebral Cortex, 17,* 2245-2250.

Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition,65*, 231-262.

Gouvea, A. C., Phillips, C., Kazanina, N., Poeppel, D. (2009). The linguistic processes underlying the P600. *Language and Cognitive Processes, 25(2)*, 149-188. doi:10.1080/01690960902965951.

Gratton, G., Low, K. A., & Fabiani, M. (2008). Time course of executive processes: Data from the event-related optical signal. In S. A. Bunge & J. D. Wallis (Eds.), *Neuroscience of rule-guided behavior* (pp. 197-223). Oxford: Oxford University Press.

Gregoric, P. (2007). *Aristotle on the common sense* (*De Memoria et Reminiscentia* 1 450 ^{a} 10). Oxford: Oxford University Press.

Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology, 34,* 660–676.

Hachey, A. (2005). *An Inquiry into the ontogeny of mental models and the etiology of phenomenological inferencing* (Unpublished doctoral dissertation). Columbia University. New York, NY.

Haagort, P., Brown, C. M., & Groothusen, J. (1993).The syntactic positive shift as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8,* 439-84.

Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. *The Neurocognition of Language*. New York: Oxford University Press.

Halford, G. S. (1982). The development of thought. Hillsdale, NJ: Lawrence Erlbaum.

Handy, T. C. (2005). *Event-related potentials: A methods handbook.* Cambridge: Bradford/MIT Press.

Haugeland, J. (1985). *Artificial intelligence: The very idea.* Cambridge: MIT Press.

Heil, M. (2002). The functional significance of ERP effects during mental rotation. *Psychophysiology, 39,* 535-545.

Heil, M., & Rolke, B. (2002a). Towards a chronopsychophysiology of mental rotation. *Psychophysiology, 39*, 414-422.

Heil, M., & Rolke, B. (2002b). ERP effects in a dual-task mental rotation paradigm. Manuscript in Preparation.

Heil, M., Rauch, M., & Hennighausen, E. (1998). Response preparation begins before mental rotation is finished: Evidence from event-related brain potentials. *Acta Psychologica, 99*, 217-232.

Henle, M. (1962).The relationship between logic and thinking. *Psychological Review, 69, 366-378.* doi: 10.1037/h0042043.

Hoare, G. T. Q. (2004). 1936: Post, Turning and ‘ a kind of miracle’. *Mathematical Gazette, 88(511)*, 2-15.

Hobbes, T. (1994). *Leviathan* (Latin edition), In E. Curley (Ed.), *Leviathan, with selected variants from the Latin edition of 1668,* Indianapolis, IN: Hackett. (Original work published in 1668).

Holyoak, K., J. (2008). Relations in semantic memory. In M. A. Gluck, J. R. Anderson, & S. K. Kosslyn (Eds.), *Memory and Mind: A Festschrift for Gordon H. Bower* (pp. 141-158). New York: Erlbaum.

Holyoak, K. J., & Morrison, R., G. (2005). Thinking and reasoning: A reader's guide. *The Cambridge handbook of thinking and reasoning* (pp.1-12). Cambridge: Cambridge University Press.

Holyoak, K. J., & Spellman, B. A. (1993). Thinking. *Annual Review of Psychology*, *44*, 265-315.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought.* Cambridge, MA: MIT Press.

Hume, D. (1978). *A treatise on human nature* (2^{nd} edition). Oxford: Oxford University Press. (Original work published 1739.)

Hume, D. (1965). *An enquiry concerning human understanding.* In Cohen, R. (Ed.), *The essential works of David Hume*. New York: Bantam. (Original work published 1748.)

Hunt, E. (1999). What is a theory of thought? In Sternberg, R., J. (Ed.), *The nature of cognition,*(pp. 3-49). Cambridge: The MIT Press.

Hunt, E. (1975). Quote the raven? Nevermore! In L. W. Gregg (Ed.), Knowledge and cognition (pp. 129-158). Hillsdale, NJ: Erlbaum Associates.

Ilmberger, J. (2009). Knowledge systems of the brain. In E. Kraft, B. Gulyas, & E. Pöppel (Eds.), *Neural correlates of thinking* (pp. 175-186). Berlin, Germany: Springer-Verlag.

Ishai, A., Ungerleider, L. G., J.V. Haxby, (2000). Distributed neural systems for the generation of visual images. *Neuron* 28, 979–990.

James, W. (1890). *The Principles of Psychology*. New York: Dover.

Jeffreys, D. A. & Axford, J. G. (1972b). Source locations of pattern-specific components of human visual evoked potentials. II. Component of extrastriate cortical origin. *Experimental Brain Research 16*, 22-40.

Johnson-Laird, (2006). Mental models, sentential reasoning, and illusory inferences. In C. Held, M. Knauff, & G. Vosgerau (Eds.), *Mental models and the mind* (pp. 127-152).

Johnson-Laird, P. N. (1999). Formal rules versus mental models in reasoning. In R. J. Sternberg (Ed.), *The Nature of Cognition* (pp. 587-624). Cambridge, MA: The MIT Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness.* Cambridge: Harvard University Press.

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science, 4,* 71-115.

Johnson, R. (1988). The amplitude of the P300 component of the event- related potential: Review and synthesis. In P. K. Ackles, R. Jennings, & M. G. H. Coles (Eds.) *Advances in psychophysiology, Vol. III* (pp. 69-137). Greenwich, CT: JAI Press.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. J. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15,* 159–201.

Kaan, E., Swaab, T. Y. (2003a). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience. 15(1),* 98-110.

Kaan, E., & Swaab, T. Y. (2003b). Electrophysiological evidence for serial sentence processing: a comparison between non-preferred and ungrammatical continuations. *Cognitive Brain Research*, *17(3),* 621-635.

Kant, I. (1965). The critique of pure reason (N. K. Smith, Trans.). New York: St. Martin's Press. (Original work published in 1787.)

Kemp, C., & Tenenbaum, J. B. (2008).The discovery of structural form. *Proceedings of the National Academy of Science, USA,105,* 10687-10692.

Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate.* Cambridge: MIT Press.

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

Kosslyn, S. M. (1973). Scanning visual images: Some structural implications. *Perception and Psychophysics, 14,* 90–94.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance, 4(1),* 47-60.

Kosslyn, S. M., Pinker, S., Smith, G. E., & Schwartz, S. P. (1979). On the demystification of mental imagery. *Behavioral and Brain Sciences2*, 535–581.

Kosslyn, S. M., & Pomerantz, J. R, (1977). Images, propositions, and the form of internal representations. *Cognitive Psychology, 9,* 52-76.

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery.* New York: Oxford University Press.

Knauff, M. (2006). A neuro-cognitive theory of relational reasoning with mental models and visual images. In C. Held, M. Knauff, & G. Vosgerau (Eds.), *Mental models and the mind* (pp. 127-152).

Knauff, M., Fangmeier, T., Ruff, C. C., & Johnson-Laird, P. N. (2003). Reasoning, models, and images: Behavioral measures and cortical activity. *Journal of Cognitive Neuroscience, 14:4,* 559-573.

Knauff, M., & Johnson-Laird, P. N. (2002). Visual images can impede reasoning. *Memory and Cognition, 30(3),* 363-371.

Knauff, M., Mulack, T., Kassubek, J., Salih, H., & Greenlee, M. W. (2002). Spatial imagery in deductive reasoning: A functional MRI study. *Cognitive Brain Research*, *13*, 203-212.

Kraft, E., Balazs, G., & Pöppel, E. (2009). Introduction. In E. Kraft, B. Gulyas, & E. Pöppel (Eds.), *Neural correlates of thinking* (pp. 3-11). Berlin: Springer-Verlag.

Kroger, J. K., Saab, F. W., Fales, C. L., Bookheimer, S. Y., Cohen, M. S., & Holyoak, K. J. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of realational complexity. *Cerebral Cortex, 12,* 477-485.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621-647.

Kutas, M., & Hillyard, S. A. (1980a). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology, 11,* 99-116.

Kutas M., & Hillyard, S. A. (1980b). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science,207,* 203-205.

Lachman, R., Lachman, J. L., & Butterfield (1979). *Cognitive psychology and information processing: An introduction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Lang, M., Lang, W., Uhl, F., Koska, C., Kornhuber, A., & Deecke, L. (1988). Negative cortical DC shifts preceding and accompanying simultaneous and sequential finger movements. *Experimental Brain Re- search*, *41*, 1–9.

Lang, W., Zilch, O., Koska, C., Lindinger, G., & Deecke, L. (1989). Negative cortical DC shifts preceding and accompanying simple and complex sequential movements. *Experimental Brain Research*, *74*, 99–104.

Lehrer, R. & Schauble, L. (2003). Modeling in mathematics and science. In R. Glaser (Ed.), *Advances instructional psychology* (Vol. 5, pp. 161-238). Mahwah, NJ: Lawrence Erlbaum Associates.

Liebniz, G. W. (1982). *New Essays on Human Understanding*. P. Remnant and J. Bennett (Eds., Tran.). Cambridge: Cambridge University Press. (Original work published in 1765).

Lennberg, E. H. (1962). The relationship of language to the formation of concepts. *Synthese, 14(2/3)*, 103-109.

Lelekov, T., Dominey, P. F., & Garcia-Larrea, L. (2000). Dissociable ERP profiles for processing rules vs. instances in a cognitive sequencing task. *Neuroreport: For Rapid Communication of Neuroscience Research 11,* 1129–32.

Locke, J. (1959). *An essay concerning human understanding*.Vols.I and II. (1st Edition). New York: Dover Publications, Inc. (Original work published in 1690.)

Luck, S. J. (2005). *An Introduction to the event-related potential technique.* Cambridge: MIT Press.

Macnamara, J. (1999). *Through the rearview mirror*: *Historical reflection on psychology*. Cambridge: MIT Press.

Mandler, J. M. (2004a). *The foundations of mind*. Oxford: Oxford University Press.

Mandler, J. M. (2004b). Thoughts before language. *Trends in Cognitive Science, 8*, 508-513.

Markman, A. B. (1999). *Knowledge representation.* Mahwah, NJ: Lawrence Earlbaum Associates.

McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. (1955). *A proposal for the Dartmouth summer research project on artificial Intelligence.*

Markman, A. B., & Gentner, D. (2001).Thinking. *Annual Review of Psychology, 52,* 223-247.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* New York: Freeman.

McClelland, J. L., & Rummelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 144*, 159-188.

McLeod, P., Plunkett, K., & Rolls, E. T. (1998). *Introduction to connectionist modeling of cognitive processes*. Oxford: Oxford University Press.

McNamara, T. P. (1994). Knowledge representation. In E. C. Caterette & M. P. Fiedman (Series Eds.) & R. J. Sternberg (Vol. Ed.), *Handbook of perceptual cognition*: Vol. 12, *Thinking* (pp. 81-117). Orlando, FL: Academic Press.

McPherson, W. B. & Holcomb, P. J. 1999 An electrophysiological investigation of semantic priming with pictures of real objects. Psychophysiology 36, 53–65. doi.10.1017/S0048577299971196.

Meuheus, J. (1999). Model-based reasoning in creative processes. In L. Magnani, N. J. Nersessian, & P. Thagard (Eds.), *Model-based reasoning in scientific discovery* (pp. 199-217). New York, NY: Kluwer Academic/Plenum Publishers.

Milivojevic, B., Johnson, B. W., Hamm, J. P., & Corballis, M. C. (2003). Non-identical neural mechanisms for two types of mental transformation: event-related potentials during mental rotation and mental paper folding. *Neuropsychologia, 41,* 1345-1356.

Miller, G. A. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Neuroscience, 7(3)*, 141-144.

Miller, G. A. (1956). The magical number seven, plus or minus two – bottleneck of limited short term memory. *Psychological Review, 63,* 81–97.

Miller, E. K., & Cohen, J. D. (2001).An integrative theory of prefrontal cortex function. Annual *Review of Neuroscience, 24,* 167-202.

Milner, B., Squire, L. R., Kandel, E. R. (1998). Cognitive neuroscience and the study of memory. *Neuron 20(3)*, 445--468.

Nehamas, A. (1975). Plato on the imperfection of the sensible world. *American Philosophical Quarterly, 12,* pp.105-117. [Reprinted in his *Virtues of Authenticity*, Princeton, NJ: Princeton University Press, 1999, 138-158.]

Neisser, U. (1963).The multiplicity of thought*. British Journal of Psychology, 54*, 1-14.

Nersessian, N. (2002). The cognitive basis of model-based reasoning in science. In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The cognitive basis of science* (pp. 133-153). London: Cambridge University Press.

Nersessian, N. J. (1999). Model-based reasoning in conceptual change. In L. Magroni, N. Nersessian, & P. Thagard (Eds.), *Model-based reasoning in scientific discovery* (pp. 5-22). Dordrecht, The Netherlands: Kluwer.

Nersessian, N. J. (1995). Should physicists preach what they practice? Constructive modeling in doing and learning physics. *Science & Education, 4,* 203-226.

Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of cognitive Neuroscience*, 3, 151-165. doi: 10.1162/jocn.1991.3.2.151.

Newell, A. (1980). Physical symbol systems. *Cognitive Science, 4,* 135-183. doi.10.1207/sl5516709cog0402_2.

Newell, A. (1973). Production systems: Models of control structures. In W. G. Chase (Ed.), *Visual information processing* (pp. 463-526). San Diego, CA: Academic Press.

Newell, A., Shaw, J. C., & Simon H. (2000). Report on a general problem-solving program. In R. Chrisley (Ed.), *Artificial intelligence: critical concepts* (Vol. 2, pp. 69-181). London: Routledge.

Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological Review, 65*, 151-166.

Newell, A., & Simon, H.A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM, 19,* 113-126.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice Hall.

Nirenberg, R. (1996). The Birth of modern science: Galileo and Descartes. Paper presented at Project Renaissance: The University at Albany, Albany, NY

Nunez_Pena, I. M., & Honrubia-Serrano, M. L. (2003). P600 related to rule violation in an arithmetic task. *Cognitive Brain Research, 18(2),* 130-141.

Noveck, I. A., Goel, V., & Smith, K. W. (2004).The neural basis of conditional reasoning with arbitrary content. *Cortex*, *40,* 613-622.

Oakhill, J., Johnson-Laird, P., & Garnham, A. (1989). Believability and syllogistic reasoning. *Cognition, (31),* 117-140.

Ollinger, M. (2009). EEG and thinking. In E. Kraft, B. Gulyas, & E. Pöppel (Eds)., *Neural correlates of thinking* (65-82). Berlin, Germany: Springer-Verlag.

Osherson, D., Perani, D., Cappa, S., Schur, T., Grassi, F., & Fazio, F.(1998). Distinct brain loci in deductive versus probabilistic reasoning. *Neuropschologia, 36,* 369-376.

Osterhout, L., & Holcomb. P. J. (1995). Event-related potentials and language comprehension. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind* (pp. 171-215). New York: Oxford University press.

Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and CognitiveProcesses, 8,* 413-438.

Osterhout, L. & Holcomb, P. (1992), Event-related brain potentials elicited by syntactic anomaly, *Language and Cognitive Processes 8, 413-438.*

Osterhout, L., Holcomb, P. J., & Swinney, D. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 786–803.

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34,* 739–773.

Patel, A., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10 (6),* 717-733.

Phelps, E. (1999). Brain versus Behavioral Studies of Cognition. In R. J. Sternberg (Ed.),*The Nature of Cognition* (pp. 51-78). Cambridge: MIT Press.

Piaget, J. & Inhelder, B. (1955). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books, 1958. [*De la logique de l'enfant à la logique de l'adolescent* (1955)].

Piaget, J. (1953). *Logic and psychology.* Manchester, England: University of Manchester Press.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37,* 127-152.

Plato (2003). *The republic* (2nd Edition). (D. Lee, Trans.). New York: Penguin Books. (Original work published 360 B. C.).

Post, E. L. (1936). Finite combinatory processes – Formulation I. *Journal of Symbolic Logic, 1,* 103-105.

Putnam, H. (1980). Brains and behavior. In N. Block (Ed.), *Readings in the philosophy of psychology (*Vol.1, pp. 24-36). Cambridge, MA: Harvard University Press. [Reprint December 27, 1961.]

Prabhakaran, V., Smith, J. A., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices. *Cognitive Psychology, 33(1)*, 43-63.

Prinz, J., J. & Barsalou, L., W. (1997). Acquisition and productivity in perceptual symbol systems: An account of mundane creativity. In T. Dartnell (Ed.), *Creativity, cognition and knowledge: An interaction* (pp. 231-251). Westport, CN: Praeger Publishers.

Pylyshyn, Z. W. (2003). Return of the mental image: Are there really pictures in the head? *Trends in Cognitive Science, 7*, 113-118.

Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences, 25,* 157-237.

Pylyshyn, Z. W. (1984). *Computation and cognition.* Cambridge: The MIT Press/Bradford Books.

Pylyshyn, Z. W. (1983). Representation, computation, and cognition. In F. Machlup & U. Mansfield (Eds.), *The study of information: Interdisciplinary messages* (pp. 115-118). New York: John Wiley & Sons.

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge, *Psychological Review, 88(1)*, 16-45.

Pylyshyn, Z., W. (1980a). Cognitive representation and the process-architecture distinction. *The Behavioral and Brain Sciences, 3(1)*, 154-169.

Pylyshyn, Z. W. (1980b). Computation and cognition: Issues in the foundation of cognitive science. *The Behavioral and Brain Sciences, 3(1)*, 111-132.

Pylyshyn, Z. W. (1979). The rate of "mental rotation" of images: A test of a holistic analogue hypothesis. *Memory and Cognition, 7*, 19–28.

Pylyshyn, Z. W. (1973). What the mind's eye tells the minds brain. *Psychology Bullentin, 80,* 1-24.

Qiu, J., Li, H., Huang, X., Zhang, F., Chen, A., Luo, Y., Zhang, Q., & Yuan, H. (2007). The neural basis of conditional reasoning: An event-related potential study. *Neuropsychologia*, *45*(7), 1533–1539.

Reeve, C., D., C. (2000). *Substantial knowledge: Aristotle's metaphysics*. Indianapolis, IN: Hackett.

Reid, V. M., Striano, T., & Koops, W. [Eds.] (2007). Special Issue: Social cognition during infancy. *European Journal of Developmental Psychology.* Hove: Psychology Press.

Rescher, N. (1980). *Induction.* Oxford: Basil Blackwell.

Rips, J. L. (1994).*The psychology of proof: Deductive reasoning in human thinking*, Cambridge, MA: MIT Press.

Rips, J. L. (1983). Cognitive processes in propositional reasoning. *Psychological Review, 90*, 38–71.

Robinson, D. N. (1999). Rationalism versus empiricism in cognition. In R. J. Sternberg (Ed.), *The nature of cognition* (pp. 79-110). Cambridge: The MIT Press.

Robinson, D. N. (1995). *An intellectual history of psychology* (3rd ed.). Madison, WI: University of Wisconsin Press.

Röder, B., Rösler, F., & Hennighausen, E. (1997). Different cortical activation patterns in blind and sighted human subjects during encoding and transformation of haptic images. *Psychophysiology*, *34*, 292–307.

Rodriquez-Moreno, D., & Hirsch, J. (2009). The dynamics of deductive reasoning: An fMRI investigation. *Neuropsycholgia*. *47,* 949-961.

Rösler, F., Heil, M., Bajricˇ, J., Pauls, A. C., & Hennighausen, E. (1995). Patterns of cerebral activation while mental images are rotated and changed in size. *Psychophysiology*, *32*, 135–150.

Rösler, F., Schumacher, G. & Sojka, B. (1990). What the brain reveals when it thinks. Event-related potentials during mental rotation and mental arithmetic. *The German Journal of Psychology, 14*, 185-203.

Ruff, C. C., Knauff, M., Fangmeier, T., & Spreer, J. (2003). Reasoning and working memory: Common and distinct neuronal processes. *Neuropsychologia, 41*, 1241–1253.

Rugg, M., & Coles, M. G. H. (1995). The ERP and cognitive psychology: Conceptual issues. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and Cognition.* Oxford: Oxford University Press.

Rummellhart, D. E., McClelland, J. L., & the PDP Research Group (1986). Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations. Cambridge: MIT Press.

Schwartz, D. L., & Black, T. (1999). Inferences through imagined actions: Knowing by simulated doing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25 (1),* 116-136.

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

Schwartz, D. L. & Black, J. B. (1996b). Analog imagery in mental model reasoning. *Cognitive Psychology***,** *30,* 154-219.

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

Searle, J. (1972, June 29). Chomsky’s revolution in linguistics. The New York Review of Books, 18(12).

Seidenberg, M. S. & McClelland, J. L. (1989). A Distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523-568.

Shepard, R. N. (1975). Form, formation and transformation of internal representations. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium.* Hillsdale, NJ: Erlbaum.

Shepard, R., & Cooper, L. (1982). *Mental images and their transformation*. Cambridge: MIT Press.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,* 701-703.

Shimamura, A. P. (1995). Memory and frontal lobe function. In: M.S. Gazzaniga (Ed.), *The Cognitive Neurosciences*, (pp. 803–813). MIT Press: Cambridge, MA.

Sitnikova, T., Goff, S., & Kuperberg, G. R. (2009). Neurocognitive abnormalities during comprehension of real-world goal-directed behaviors in schizophrenia. *Journal of Abnormal Psychology. 118(2)*, 256-277.

Sitnikova T., Holcomb P.J., Kiyonaga K.A., Kuperberg G.R. (2008a). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience 20*, 11-22.

Sitnikova, T., West, W., C., Kuperberg, G., R., & Holcomb, P., J. (2006). The neural organization of semantic memory: Electrophysiological activity suggests feature-based segregation. *Biological Psychology, 71,* 326-340.

Sitnikova, T., Kuperberg, G., & Holcomb, P. J., (2003). Semantic integration in videos of real world events: an electrophysiological investigation. *Psychophysiology, 40,* 160–164.

Sloman, S. (1996).T he empirical case for two systems of reasoning. *Psychological Bullentin, 119*, 3-22.

Smith, L. (2003). From epistemology to psychology in the development of knowledge. In Brown, T., & Smith, L. (Eds.), *Reductionism and the development of knowledge.* Mahwah, NJ: Erlbaum Associates.

Smith, E. E., Langston, C., & Nisbett, R. E. (1992). The case for rules in reasoning. *Cognitive Science, 16,* 1-40.

Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences II,* l-23.

Spinoza, B. (1996). *Ethics*. In E. Curley (Ed., Tran.). London: Penguin books. (Original work published in 1677).

Sorell, T. (1992). *Hobbs: The arguments of the philosophers*. London: Routledge Publications

Stevens, A., & Gentner, D. (1983). Introduction. In D. Gentner, & A. Stevens, *Mental Models* (pp.1-6), Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (1999). A dialectical basis for understanding the study of cognition. In R. J. Sternberg (Ed.). *The Nature of cognition* (pp. 51-78). Cambridge: The MIT Press.

Swaab, T.Y., Baynes, K., Knight, R.T., 2002. Separable effects of priming and imageability on word processing: an ERP study. *Brain Research, Cognition Brain Researh,15 (1),* 99–103.

Tucker, D. M. (1993). Spatial sampling of head electrical fields: The geodesic sensor net. *Electroencephalography and Clinical Neurophysiology, 87*, 154-163.

Tucker, D. M., Liotti, M., Potts, G. F., Russell, G. S., & Posner, M. I. (1994). Spatiotemporal analysis of brain electrical fields. *Human Brain Mapping, 1*, 134-152.

Turing, A. M. (1936–7). On computable numbers, with an application to the Entscheidungs Problem. Proceedings of the London mathematical Society Series 2 (42) 230–265. doi: 10.1112/plms/s2-42.1.230.

Ullsperger, P. & Gille, H. G. (1998). The late positive component of the erp and adaptation-level theory. *Biological Psychology, 26*, 299-306.

van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language. 41*, 147-182.

Van Voorhis, S. & Hillyard, S. A. (1977). Visual evoked potentials and selective attention to points in space. *Perception & Psychophysics, 22(1),* 54-62.

Vitouch, O., Bauer, H., Gittler, G., Leodolter, M., & Leodolter, U. (1997). Cortical activity of good and poor spatial test performers during spatial and verbal processing studied with slow potential topography. *International Journal of Psychophysiology*, *27*, 183–199.

Vlastos, G. (1996). *Studies in Greek Philosophy: Socrates, Plato, and their tradition, Vol., 2.* Princeton: Princeton University Press.

Vygotsky, L. (1986). *Thought and language*. Cambridge: MIT Press.

Wason, P. C. (1966). Reasoning. In Foss, B. M. *New horizons in psychology*. Harmondsworth: Penguin.

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning structure and content.* London: Routledge.

West, C. W., & Holcomb, P., J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Brain Research, 13(3),* 363-375.

West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *Journal of Cognition Neuroscience, 12,* 1024–1037.

Wharton, C., & Grafman, J. (1998). Deductive reasoning and the brain. *Trends in Cognitive Sciences, 2(2)*, 54-59.

White, N. (1976). *Plato on knowledge and reality*. Indianapolis, IN: Hacket.

Wijers, A. A., Otten, L. J., Feenstra, S., Mulder, G., & Mulder, L. J. M. (1989). Brain potentials during selective attention, memory search, and mental rotation. *Psychophysiology, 26,* 452–467.