SPATIOTEMPORAL PATTERNS OF PARIETOFRONTAL ACTIVITY AND EYE MOVEMENTS UNDERLYING THE VISUAL PERCEPTION OF COMPLEX HUMAN TOOL USE

A Thesis
Presented to
The Academic Faculty

by

Nikhilesh Natraj

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Applied Physiology

Georgia Institute of Technology
December 2015

SPATIOTEMPORAL PATTERNS OF PARIETOFRONTAL ACTIVITY AND EYE MOVEMENTS UNDERLYING THE VISUAL PERCEPTION OF COMPLEX HUMAN TOOL USE

Approved by:

Dr. Lewis A Wheaton, Advisor
School of Applied Physiology
Georgia Institute of Technology

Dr. Minoru Shinohara
School of Applied Physiology
Georgia Institute of Technology

Dr. Richard T. Nichols
School of Applied Physiology
Georgia Institute of Technology

Dr. Eric Schumacher
School of Psychology
Georgia Institute of Technology

Dr. Boris I Prilutsky
School of Applied Physiology
Georgia Institute of Technology

Date Approved: 5th November 2015

To my parents Natraj and Nirmala and brother Nithyaesh.

ACKNOWLEDGEMENTS

It is often said that no work is ever done in isolation; without a doubt, this statement applies strongly to the work done in this dissertation. I would like to thank my advisor and director of the Cognitive Motor Control Lab, Dr. Lewis Wheaton for giving me an opportunity to do research in his lab. His broad scientific and technical knowledge and approachability on a range of other issues, personal and professional has made the PhD experience pleasurable. In particular, I would like to thank him for introducing me to and mentoring me in the burgeoning fields of motor control, affordances and neuroimaging.

During my time at GT, I have been fortunate to have interacted with many faculty who have enriched my student experience. I would like to thank Dr. Boris Prilutsky, director of the Center for Human Movement Studies at GT, who is also on my committee and with whom I have had the pleasure with working on various biomechanics and arm reaching movement projects. In addition, I would like to also thank Dr. Tom Burkholder, Dr. Eric Schumacher and Dr. Richard Nichols for the many scientific conversations and being inspiring mentors. Together, their knowledge and feedback via courses and seminars have been invaluable to my overall development as a scientist and I am deeply indebted to their willingness to take time off their busy schedules and have long meetings and talk about motor control and cognitive neuroscience. I would also like to thank Dr. Minoru Shinohara who is on my committee and with whom I have had the pleasure of interacting with over many years in courses and research seminars. The knowledge and feedback I have gained from these interactions with various faculty have tremendously enhanced my thought process, development and contributed in no small way to the work done in this dissertation. I would also like to thank Dr. Steve Potter at BME and Dr. Raj Pulugurtha at ECE for mentoring me in my initial semesters doing research at GT.

The lab members at the cognitive motor control lab (CMCL) have been incredibly helpful with experimental setups, advice, feedback, technical help and all the good times. It

was a pleasure being colleagues with the Dr. Bill Cusack, Dr. Rachel Kelly, Regan, John, Kristel, Matt and Dr. Chris Mizelle. The many experiments and data preprocessing in this thesis could not have been performed without the aid of the incredibly smart and talented undergraduate students: Bennett Alterman, Sumia Basunia, Daniel deWitz, Lauren Levinson and Yvonne Pella, in alphabetical order. I also enjoyed interacting with my fellow graduate students over the many years: Vasiliy, Brian, Ricky, Elma, George, Ted, Megan, Tracy, Ray, and Namrita, to name a few. I would like to thank my friends at GT who have made my life enjoyable; thanks to Vishwa, Raj, Rohan, Raghavan, Nikhil, Dharma, Sunil, Vaibhav, Pacino, Sanchit and Sampath to name a few. Special thanks to my roommates these past many months, Rachel Davis and her canine friend Cammie for their warmth, kindness and patience especially during the thesis writing months. Thanks also to Dennis, Sahil, Purav, Dixie, Harsha and Jayesh for more good times.

This thesis would not have been possible without my family especially my parents Natraj and Nirmala and brother Nithyaesh back in Chennai, India. Their constant prayers, words of encouragement, thoughts, support, and phone conversations have given me tremendous strength and comfort. I would like to thank my uncle Dr. Kishore Ramachandran who encouraged me to pursue a graduate degree and guided me every step of the way as I took the initial baby steps in a foreign country. His support along with his wife, my aunt Mrs. Vasanthi and cousins Hemant and Shalini provided me a family away from home and I had a total blast enjoying time with them over the many years. I will always be grateful for their love and support. Many thanks to all the cousins and relatives spread across the country and who have been my family in more ways than one: Dr. Ravi Venkatraman and Nandini, Aparna and Arijit, Rohini and Madhav. Last but not least, I would like to thank Urmila Venkat, whom I am going to be joining in San Francisco shortly. Thanks for being there through all the initial tough times in our young lives in a country far away home. Your love, kindness, incredibly large heart and strength improved my own self confidence especially in the initial and uncertain years of graduate school and I will always be grateful to you for that. You were very patient all those long nights at the library and at lab and were highly entertaining and fun to be with! I am sure spending Saturday nights in the library was not

exactly the dream date you had in mind, but hopefully we can make it up with that trip to Vegas. We have shared many wonderful memories over the years and I will always cherish Atlanta and Georgia Tech as not only did I spend my 20s here and grow as a person, but I also met and fell in love with you. I look forward to the future and continuing where we left off with many good times ahead.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

When watching a child learning to use a spoon, a mother is immediately able to recognize the error when the child grabs the bowl rather than the stem, or when the child uses the spoon to try and scoop paper. Recognizing proper tool grasp-postures and use-contexts is an ability vital for daily life and can be lost due to brain injury. A better understanding of how the brain encodes contextual and grasp-specific tool-use not only furthers basic neuroscience, but also has strong relevance to deficits arising from neural pathologies. However, the majority of research till date has studied the neural response to viewing tools in isolation or viewing simple tool-grasps. These studies have shown that the recognition of tools to be a complex visuomotor process, as not only was the visual cortex engaged but also parietal and frontal regions that underlie actual tool-use. The recognition of tools therefore involves automatically recalling their motor information (graspability and manipulability) via activation of parietofrontal motor regions, a property called action affordances. Yet, it is still unclear how parietofrontal regions encode the combination of contextual and grasp-specific tool-use scenes. In addition, parietofrontal regions are multifaceted and also underlie visuospatial attention and eye movements. It is possible a relationship might exist between eye movements, attention and tool-use understanding over parietofrontal regions. Therefore the overall goal of this thesis was to understand the spatiotemporal patterns of parietofrontal activity and eye movements underlying the perceptual of contextual and grasp-specific static tool use images. Electroencephalography (EEG) was used to measure neural activity, combined with eye tracking to measure fixation and saccades. Overall, results from this thesis present evidence that the affordances of non-functional grasp-postures perturbed an observer from understanding the contextual uses of tools, with corresponding unique patterns of parietofrontal activity and eye movements. This effect was most robust when the tool was placed in contexts that afforded a certain degree of tool-use. Results also revealed a relationship between attention, eye movements and action perception over

parietofrontal regions. Specifically, saccades perturbed activity over frontal regions during the perception of non-functional grasp postures and in addition, there was greater engagement of the left precuneus in the superior parietal lobe if the observer had to quickly parse the scene information using peripheral vision and rely on short term memory. In contrast, there was greater engagement of the left middle temporal gyrus if the observer had the ability to parse scene information continuously using foveal attention. Results in this thesis shed light on the neural and visual mechanisms in understanding the affordances of non-functional grasp postures, and the relation between the two mechanisms. The automatic sensitivity in understanding the intent of non-functional grasp-postures may correspond to a lifetime of learning the affordances of grasp-specific action outcomes with tools. Such cognitive motor knowledge may be vital in navigating a human environment almost entirely constructed on advanced tool-use knowledge and findings from this thesis have many potential applications in the field of neuro-rehabilitation.

CHAPTER I

INTRODUCTION

The ability to recognize the advantage of tools, their associated gestures and predicting their action outcomes is vital for daily living (Johnson-Frey, 2004). For example, when we want to brush our teeth, we are able to immediately differentiate the utility of a tooth-brush from the comb and razor. Indeed, while most animals use tools, humans as a species are unique in using arguably the most sophisticated tool-use knowledge to achieve daily goals and effect change in the environment (Johnson-Frey, 2004). The unique human ability of learning, deploying and transmitting complex tool-use knowledge effectively sets us apart from our near primate cousins and is even thought to be linked with the development of language and culture (Stout and Chaminade, 2012). Unfortunately, this ability can be lost due to stroke or neurodegenerative diseases in a condition called apraxia (Wheaton and Hallett, 2007), a cognitive motor deficit wherein the patient loses tool-specific knowledge without any lower level visual or motor deficit (Goldenberg, 2003, 2009). Therefore, a better understanding of how the brain encodes tools and tool-use gestures not only furthers basic neuroscience, but also has strong relevance to understanding deficits arising from neural pathologies.

1.1 Historical and clinical background on the neural substrates underlying tool-use knowledge

The first insights on the neurophysiology underlying tool-use knowledge had its roots at the turn of the previous century. In this era, it was commonly thought that complex brain outputs such as the motor act of using a tool was the result of activity over the entire cortex and could not be functionally localized to any one brain region (Goldenberg, 2003, 2009). However, pioneering clinical work of Hugo Leipmann showed that a distinction existed between pure sensorimotor functions of the cortex and higher order cognitive motor

knowledge. Based on his study of clinical patients, Leipmann hypothesized that the idea of a tool-use movement, a motor program ("movement formula" (Liepmann, 1900, 1980, Rothi and Heilman, 1996, Goldenberg, 2003, 2009)), could be functionally localized to a distinct area of the brain and that disconnections between this localized area and the motor cortex could result in deficits in imitating and performing skilled and meaningful tool-use gestures. Post-mortem operations validated Leipmann's conjectures wherein lesions to the left posterior parietal cortex and corpus callosum resulted in an inability of the patient to pantomime or imitate skillful tool-use gestures with their dominant right hand. Damages to the corpus callosum in particular disrupted the crossing of projections from left parietal regions to right motor and premotor regions. Thus the left parietal cortex was thought to be the foci of generating tool-use action information as damage to the ipsilateral left parietal cortex caused tool-use deficits in left-handers as well. It is important to note that these tool-use specific motor deficits (apraxia) were not associated with low-level motoric issues such as tremors, bradykinesia or due to any neuromuscular dystrophies. The motoric deficit was purely due to an inability to formulate higher order motor planning, or motor cognition, such as understanding how a hammer might be grasped and/or used. Leipmann thought that apraxia may arise from left parietal lesions that disrupt the flow of higher-order motoric information to premotor regions that transform the information into an actual motor act. As a result, the patient would not be able to perform successful tool-use movements (Liepmann, 1900, 1980, Rothi and Heilman, 1996, Goldenberg, 2003, 2009).

Leipmann's pioneering research technique became the basis for further work by Norman Geschwind in the 1960s (Geschwind, 1965), Kenneth Heilmann in the 1980s and 1990s (Heilman et al., 1997) and others (Buxbaum, 2001, Wheaton and Hallett, 2007), who broadly expanded upon Leipmann's hypotheses, further developing the neuroscience underlying tool-use knowledge via lesion analyses and behavioral studies in apraxic patients. In particular, Geschwind reinterpreted Leipmann's disconnection hypothesis and argued for a role of Wernicke's area in the superior temporal gyrus. In the clinical setting at that time, apraxia was typically diagnosed based on the ability of a patient to successfully respond to the clinicians instructions. An example of such an instruction would be the command "show

me how to use a hammer" (Ochipa et al., 1989, Rothi and Heilman, 1996). Geschwind hypothesized that the inability of an apraxic patient to reproduce a tool-use gesture on verbal command may arise from parietal lesions that disconnected the flow of information from Wernickes area to premotor regions via the arcuate fasciculus. The "verbalization" of a tool-use gesture's meaning was considered vital for the reproduction of that gesture itself (Geschwind, 1965, Goldenberg, 2003, 2009). However, Heilman later reformulated the disconnection hypothesis of both Leipmann and Geschwind and argued that the left parietal damage corresponded to an irrevocable loss to a stored motor program rather than disruptions in the transfer of action information to premotor regions. In this framework, parietal regions were considered to be storehouses of acquired tool-use knowledge rather than areas that purely served to relay motor information on every external visual or verbal cue (Heilman et al., 1982, Goldenberg, 2009). Functionally, Heilman implicated the left inferior parietal lobule to be the repository of tool-use knowledge. It is possible that novel motor skills might be learned via engagement of right parietal regions before consolidation into the left hemisphere (Halsband and Lange, 2006). Overall, from a historical and clinical perspective, evidence suggested a distinct role of left parietal regions in storing the motoric properties of tools, with projections to premotor regions that served to formulate this motor information in terms of the sensorimotor joint transformations that allow actual tool-use (Haaland et al., 2000).

At the same time, clinical studies also identified the role of temporal regions in differentiating conceptual rather than motoric tool-use information. In fact, it was Leipmann again who first discovered that conceptual tool-use knowledge might be functionally dissociable from its motoric aspects, as patients with parietal lesions were able to conceptually describe the uses of a tool though they were unable to actually pantomime using the tool (Liepmann, 1900, Liepmann and Psychology, 1980, Goldenberg, 2003, 2009). Based on his experiments with clinical patients, he postulated different neural substrates for conceptual and motoric tool-use knowledge. Subsequent clinical research supported Leipmann's hypothesis. Patients with damage to the posterior middle temporal gyrus exhibited deficits

in associative or semantic tool-use knowledge (Heilman et al., 1997), for example, understanding that a hammer is used to pound objects (Kalnine et al., 2010). Focal lesions to the left occipito-temporal-parietal junction resulted in a similar inability of patients to recognize the utility of tools (recognizing the advantage of a hammer's shape and weight for pounding movements (Tranel et al., 1997)). Damage to the left occipito-temporal-parietal junction (Ochipa et al., 1989, Johnson-Frey, 2004), while preserving proper tool-use action kinematics, resulted in inappropriate and faulty tool-use wherein patients used a comb to brush their teeth. Interestingly, the ability to name the tools or recall its kinematics was preserved but the conceptual knowledge to associate a tool with its proper usage was lost (Johnson-Frey, 2004). Based on the above results from the lesion data, left temporal regions were therefore thought to store conceptual and semantic knowledge of tools rather than purely its motoric properties.

One of the difficulties with lesion studies was that rarely did a patient present with only one type of deficit or with a very focal damage to the brain. The nature of the brain injury typically manifested itself with a broad range of symptoms and a wide area of damage to the cortex involving temporal and prefrontal regions, in addition to the left parietal and premotor cortices (Johnson-Frey, 2004). As a result, it was generally difficult to derive focused hypotheses and refine Heilman's/Geschwind or Leipmann's theories with respect to the visual observation of tools and tool-use gestures. Indeed, it was only with the advent of neuroimaging that further insights could be gleaned on the nature of the cortical processes underlying the visual recognition of tools and tool-use.

1.2 Neural correlates of tool-use knowledge as revealed by neuroimaging

The modern understanding of how the brain visually perceives tools and tool-use gestures stems from advances in neuroimaging technology. As a methodology, neuroimaging allows studying the brain with much finer resolution and specificity than by correlating lesion sites to loss in functional behavior. Many neuroimaging studies since the mid-1990s have explored the neural responses to viewing, naming, using and pantomiming tools across

many different imaging modalities such as functional Magnetic Resonance Imaging (fMRI), Electroencephalography (EEG) and Positron Emission Tomography (PET). For instance, fMRI studies in the past two decades have evaluated cerebral blood flow in right handed subjects during the pantomiming of tool-use movements or during the manipulation of tools with the dominant right hand. These studies have consistently reported increased blood flow or activity over superior and inferior parietal lobules, the primary motor and sensory cortices, ventral and dorsolateral premotor cortex, middle and inferior frontal gyri, and posterior middle temporal gyrus (Lewis, 2006). These neural activations were primarily left hemispheric dominant (Johnson-Frey, 2004, Johnson-Frey et al., 2005). EEG studies investigating the temporal nature of parietofrontal activity have showed that the generation of complex tool-use movements (such as pantomiming a screwdriver) were preceded by a slowly evolving negative cortical potential over left posterior parietal regions beginning about 3 seconds before movement onset (Wheaton et al., 2005b). Similar EEG studies have also shown significant changes in beta band (18-22Hz) oscillations over left posterior parietal cortex and increased synchronizations between left parietal and left frontal regions in the beta band during the execution of tool-use gestures (Wheaton et al., 2005a, Wheaton et al., 2009). Interestingly, while parietal, frontal, motor and temporal regions were active during the execution of complex tool-use movements, the aforementioned regions were also found to be active in studies wherein participants were instructed to either passively imagine using tools or viewed static images of tools (Johnson-Frey et al., 2003, Johnson-Frey, 2004, Lewis, 2006), shown in Figure 1. It was found that viewing tools also increased activity along left temporo-occipital regions (Johnson-Frey, 2004, Lewis, 2006). Therefore, as a class of objects, tools were found to be unique as they automatically potentiated or engaged action knowledge centers in the brain apart from the primary visual cortex (Grzes et al., 2003, Johnson-Frey, 2004, Creem-Regehr and Lee, 2005). For example, the mere observation of a stand-alone hammer engages parietal, temporal, premotor and frontal regions that are active when actually using the hammer (Johnson-Frey et al., 2003, Johnson-Frey, 2004). This unique property of tools gleaned from neuroimaging methods could not be predicted by Leipmann, Geschwind or others based on their models from clinical data and sets tools

apart from other non-manipulable objects such as a tree, house etc. (Chao and Martin, 2000, Lewis, 2006).

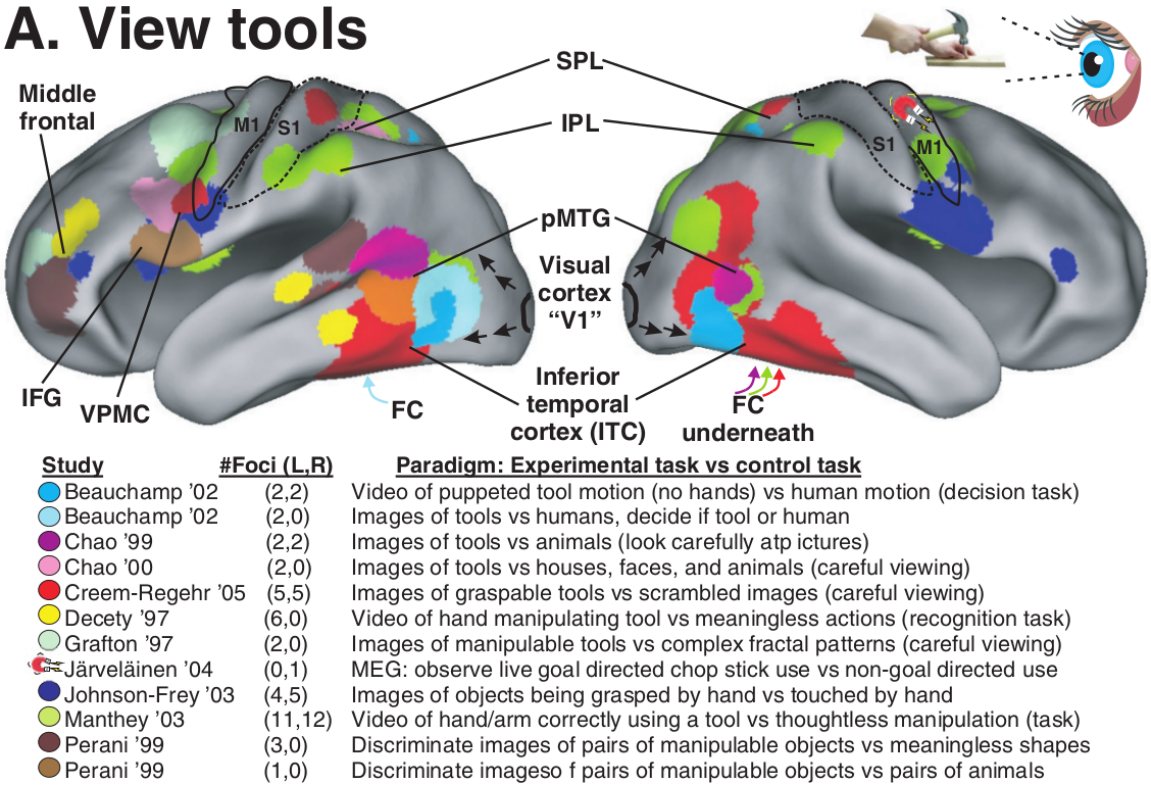

Figure 1: An overview of the role of temporal, parietal and premotor, frontal regions during the visual perception of tools (Lewis, 2006).

The main advantage of neuroimaging methods lies in the ability to test focused hypotheses given the freedom in experimental design and the ability to measure neural activations with a high degree of specificity. Although it was known that tool observation activated such diffuse regions in the brain, neuroimaging experiments were able to uncover specializations in the type of visual information processed at each of those regions. Activity over the visual cortex is to be expected as the early visual pathways aid in detecting edges, orientation and color of tools (Kandel et al., 2000). Anatomically, from the visual cortex, visual information propagates along two pathways, one to the temporal regions (the ventral stream) and another that propagates information along parietal regions (dorsal stream). With respect to the visual recognition of tools, it was initially thought that the ventral stream served to answer the 'what' question while the dorsal stream served to answer the

'where' question (Mishkin et al., 1983). However, subsequent research was suggestive of more sophisticated visual processing mechanisms over the two streams (Goodale and Milner, 1992). Specifically, rather than being only involved in answering the 'what question, the left middle temporal gyrus, fusiform gyrus in the ventral stream, along with the left inferior parietal lobule, have been shown to process semantic properties of tools such as recognizing that a hammer belongs to a class of objects used in pounding movements (Mahon et al., 2007)). It was also thought that given its location next to area V5 or the motion processing centers in the visual cortex, temporal regions may even store information on the movement related conceptual properties of tools, such as understanding the trajectory of a hammer that enable pounding movements (Johnson-Frey, 2004). With respect to the parietal cortex along the dorsal stream, research has shown parietal regions as a whole to be active when subjects viewed tools (Chao and Martin, 2000), imagined using tools (Johnson-Frey et al., 2005, Fridman et al., 2006, Lewis, 2006), evaluated how tools could be grasped (Lewis, 2006, Almeida et al., 2010) and when required to recall explicit information on the action mechanics associated with tools (Johnson-Frey, 2004). Within the parietal cortex, regions along the anterior intraparietal sulcus (aIPS) was found to encode more of the grasp-posture associated with tools based on a tools form and shape (Culham and Valyear, 2006). Areas near the intraparietal sulcus along with the angular gyrus, both located in the inferior parietal lobule, were also found to be engaged to a greater extent when participants imagined grasping tools compared to shapes or scrambled images of tools (Creem-Regehr and Lee, 2005). Alternatively, regions along the superior parietal lobule was thought to be more involved in action kinematics associated with a tool (Culham and Valyear, 2006), such as encoding the joint transformations corresponding to a tools motor properties (Glover et al., 2005, Lewis, 2006). While the successful recognition of tools activates the parietal cortex as a whole (in addition to the temporal cortex), there does appears to be a distinction within the parietal cortex itself. Inferior regions (angular gyrus, aIPS) may encode more of the cognitive motor aspects of tools, such as how it can be grasped and used (Culham and Kanwisher, 2001, Culham and Valyear, 2006, Lewis, 2006), whereas the superior parietal regions may encode more of the kinematic aspects associated with the sensorimotor joint

transformations of tool-use (Culham and Kanwisher, 2001, Culham and Valyear, 2006). It has been proposed that a potential reason for this differentiation could be the proximity of inferior parietal regions to areas along the ventral stream (Ramayya et al., 2010) given the extensive white matter tracts between the middle temporal gyrus and anterior regions of the supramarginal gyrus and angular gyrus in the inferior parietal lobule (Ramayya et al., 2010). Given this extensive connectivity, the anterior supramarginal gyrus may act as a integrative node, merging semantic tool-use information from ventral areas such as middle temporal gyrus with the motoric tool-use knowledge emanating from the angular gyrus, aIPS and areas along the superior parietal lobule (Ramayya et al., 2010). In general, this extensive integration may serve to recall the stored conceptual, cognitive motor and sensorimotor knowledge of tools, even during passive observation (Ramayya et al., 2010).

With respect to frontal regions, the dorsal visual stream continues from the angular gyrus and anterior supramarginal gyrus and terminates at the ventral premotor cortex in the frontal lobe (Johnson-Frey, 2004, Ramayya et al., 2010). The premotor cortex subsequently passes its output to putative motor neurons in the primary motor cortex that generate the descending motor signals that control the actual physical manipulation of tools (Fridman et al., 2006). During actual tool-use, the premotor and motor cortices perform the necessary sensorimotor joint transformations that enable actual tool-manipulations based on the incoming, integrated motor program from the parietal cortex. However, premotor and motor areas were also found to be active during passive tool-use observation (Lewis, 2006). Activations over premotor and motor regions during the observation of tools and tool-use may represent an ongoing motor simulation of actual tool-use (Johnson-Frey, 2004, Lewis, 2006) by encoding the grasp-specific and sensorimotor aspects of tools. These motor simulation mechanisms over motor and premotor regions may facilitate higher level motor cognition such as understanding the goal or intent of a seen tool-use gesture. For instance, the primary motor cortex was found to be engaged more fully when participants viewed chopsticks being used in a goal-directed rather than aimless manner (Jrvelinen et al., 2004). Inferior frontal regions in particular have been well implicated in encoding goal or intent of tool-use gestures. in a seminal study, researchers used fMRI to show that neural activity

over inferior frontal regions was modulated when subjects viewed a hand grasping a tool vs. a hand touching an object in a non-manipulable fashion (Johnson-Frey et al., 2003). Electrophysiology studies over analogous regions in the monkey brain revealed that neurons in the left inferior frontal regions were active during both the observation and execution of similar, goal directed hand-object gestures (Iacoboni et al., 2005). For example, a specific set of neurons in the monkey ventral premotor cortex F5 were found to be active both when the monkey grasped a rake and observed an experimenter grasp a rake, but not when the rake was merely touched (Johnson-Frey et al., 2003, Iacoboni and Dapretto, 2006). Rather than encoding the prehensile properties of tools, these specific neurons in area F5 (inferior frontal gyrus in humans) may serve to encode the goal or intent of the tool-use gesture via an internal motor simulation of the observed gesture itself (Iacoboni and Dapretto, 2006).

The fact that the mere observation of tools and tool-gestures elicits an automatic understanding of their very complex motoric, semantic and goal-oriented properties offers one potential explanation as to why humans are unique in recognizing the utility of tools when interacting with the environment and achieving complex action outcomes (Johnson-Frey, 2004).

1.3 Insights on tool-use understanding from cognitive science: the theory of action affordances

The diffuse activation of parietal, premotor, frontal and temporal regions even during mere tool-use observation is suggestive of the recall of extensive motoric and semantic information necessary to recognize tools, their associated gestures and even the intent of a tool-use gesture itself. In particular, as parietofrontal regions process the motoric aspects of tools, this pathway along the dorsal stream constitutes a direct visual route to action knowledge (Humphreys et al., 2010, Yoon et al., 2010), tightly linking the visual recognition of tools with the recall of their action or motor knowledge. In the behavioral science or experimental psychology literature, this unique property of tools in automatically activating its motor knowledge is called action affordances, a theory whose general form was first postulated by Gibson (Gibson, 1977). Gibson postulated that all objects in the environment

have certain properties that uniquely afford the capability for an organism to achieve any goal based behavior. For example, a mouse may afford moving a pointer on screen for a human, but may not afford this behavior to say a dog (Borghi et al., 2012). The theory of action affordances is well suited to explain how the mere observation of tools is sufficient to activate its motoric or action properties (way it is grasped and used) especially given the diffuse activations over action knowledge regions in the brain. Affordances with respect to tools has been investigated in depth in the cognitive science literature via behavioral experiments. These experiments typically evaluated whether a tool's affordances (its spatial location, orientation etc.) automatically primed motor behavior even though affordance information was unrelated to the experimental task (akin to the Simon effect (Simon and Berbaum, 1990). A famous example of such a behavioral experiment is the seminal study by Tucker and Ellis who asked participants to push a button with either their left or right hand to evaluate if a tool was upright or inverted. They found that participants made quicker responses with either their left/right hand when the handle of the tool was slightly oriented to left/right, though the orientation of the tool was irrelevant to the task (Tucker and Ellis, 1998). In another experiment, the same authors (Tucker and Ellis, 2001) showed that when categorizing manipulable artifacts (hammer, nail) from natural objects (apple, cherry) with a power or precision grip, participants' responses were influenced by the size of the artifact even though it was not essential to the actual categorization task. As a result, smaller artifacts (nail) generated faster responses with a precision grip while larger artifacts (hammer) generated faster responses a with power grip. These two seminal papers provided behavioral evidence that simple observation automatically activated a tool's action affordances and resulted in the priming of motor behavior. Similar studies investigating the influence on affordance recognition on motor behavior have corroborated Tucker and Ellis's findings (an overview of these results is summarized in (Thill et al., 2013)). Indeed, evaluating the effect of affordances on perceptual judgments has been a recurring theme in behavioral experiments with many variations on the experimental design. For example, it has been shown that perceptual judgements of tool-use are enhanced when viewing tools that are oriented for action (a spoon positioned to scoop ice cream), when right-handers

view tools from a right handed egocentric perspective and when tools are appropriately grasped for functional use (Riddoch et al., 2003, Riddoch et al., 2006, Humphreys et al., 2010, Yoon et al., 2010, Kelly and Wheaton, 2013, Kelly et al., 2015). Results from these studies suggest that functional and appropriate tool-use scenes could optimally engage familiar and learned motoric representations of the tool over parietofrontal regions, thereby priming motor behavior leading to faster and more accurate behavioral responses.

1.4 The neural encoding of contextual tool-use affordances

The majority of neuroimaging work investigating the neural substrates underlying tool-use knowledge have either considered a tool in isolation or a simplistic tool-grasp. However, extensive behavioral experiments have shown that varying a tool's affordances correlates with motor behavior, suggestive of a differential activation of parietofrontal action encoding regions. In addition, a tool is usually seen in different contexts of use that may potentially alter its affording properties and how it is neurally encoded. Therefore, rather than considering a tool in isolation, recent neuroimaging work in our lab have aimed to understand the neural encoding of tool-use scenes where affordances were dependent on action context. In particular, these studies focused on how canonical tool-use action encoding regions (parietal, premotor and temporal regions) differentiated between incorrect and correct action contexts of tool-use. Here, a tool's context of use was defined by the type of object the tool was associated with. For example, a hammer when paired with a nail has higher action affordances as this is the hammer's typical context of use. Conversely, a hammer when paired with paper has lower affordances as this is an atypical or incorrect context of use given the hammer's motoric properties. In the former case, the nail reinforces the motoric knowledge associated with the hammer whereas in the latter case, the paper does not support a functional use of the hammer. EEG results showed that regions along the superior temporal gyrus, insula, posterior cingulate and precuneus served to differentiate between correct and incorrect tool-object pairings (Mizelle and Wheaton, 2010a). In a related experiment, subjects viewed images depicting incorrect and correct tool-object usage wherein the tool was

also grasped to interact with the object. An example of a correct tool-object usage would be viewing a hammer grasped by its handle and positioned to drive a nail. An example of an incorrect tool-object usage would be viewing a hammer grasped by its handle and positioned to hit a coffee cup. Results showed that the incorrect context of tool-use elicited greater activations along the posterior cingulate, superior temporal gyrus and insula whereas the correct context of tool-use elicited greater activations along canonical parietofrontal action encoding regions (Mizelle and Wheaton, 2010b). Engagement of ventral regions may serve in identifying the semantic errors in the incorrect tool-use context over the correct tool-use context whereas the correct tool-use context may greater engage parietofrontal regions given its higher action affordances. Together, these two studies provide evidence that the neural encoding of a tool's affordances can be modulated by its action context.

1.5 Overall goal

Broadly, behavioral, clinical and neuroimaging results show considerable and converging evidence that the successful recognition of tools is dependent on understanding its motoric properties (affordances) via the activation of a direct visual route to action knowledge regions along the dorsal stream. Further research has shown how activity over action knowledge regions (parietofrontal regions) can be modulated by the affordances of action context. However, what is unclear is how parietofrontal regions encode for more complex and real-life tool-use scenarios wherein a tool's affordances rely not just on action context, but also on the combination of grasp-specific tool-use knowledge. Indeed, tool-use is inherently adaptive and recognizing the advantage of tools must involve not only understanding its proper context of use but also the proper grasping postures that allow using tools. For example, to drive a nail into a board, we must be able to recognize the advantage of using a hammer instead of a pen (context-use) and further understand that the hammer is to be grasped by its handle rather than its head (grasp-use). Such type of knowledge may have been shaped over many years of experience and exposure to tools right from development. As outlined earlier, while prior studies have individually evaluated parietofrontal encoding of tool-object

relationships (Mizelle and Wheaton, 2010a, b) and stand-alone tool-grasps (Johnson-Frey et al., 2003), it is unclear how parietofrontal regions encode the combination of both tool-use context and tool-grasp. Therefore, the first specific aim of the thesis **(Aim 1)** focused on understanding the temporal dynamics of parietofrontal activations as healthy right handed human subjects evaluated still images depicting right-handed, egocentric, contextual and grasp specific tool-object scenes. In particular, Aim 1 focused on the sensitivity of parietofrontal responses to the affordances elicited by the type of grasp-posture within specific tool-use contexts.

As a neural network, parietofrontal regions are multifaceted and are involved in many other functions (Culham and Kanwisher, 2001, Miller and Cohen, 2001). They might also be involved in other cognitive motor processes that could potentially be coupled with action understanding. In particular, given that tool-use recognition is inherently a visual perception task, it is important to note that parietofrontal regions also underlie the control of saccades and visuospatial attention especially over inferior and posterior parietal lobules and dorsal and ventral premotor areas (Figure 2). Studies has shown these regions to be active during the control of fixations and saccades (Anderson et al., 1994, Gaymard et al., 1998, Mort et al., 2003), spatial attention (Corbetta et al., 1998) and saccadic inhibition (Chikazoe et al., 2007). Given this anatomical overlap between action encoding and visual attention over parietofrontal regions, it is possible that gaze patterns (visual encoding), could also be sensitive to complex tool-use affordances. Understanding the visual encoding of affordances can potentially shed light on the type of information an observer gathers to drive parietofrontal action encoding regions, thereby providing a valuable window into ongoing action encoding processes. Therefore, the second aim of this thesis **Aim 2** focused on understanding whether the affordances arising from contextual and grasp specific tool-use scenes could influence eye movement patterns during observation.

Aim 1 and Aim 2 focused on individually evaluating the effect of contextual and grasp specific tool-use images on parietofrontal activations and visuospatial attention mechanisms (Natraj et al., 2013, Natraj et al., 2015). In both Aims 1 and 2, the experimental design allowed the observer adequate time to visually parse the stimulus as the images were on

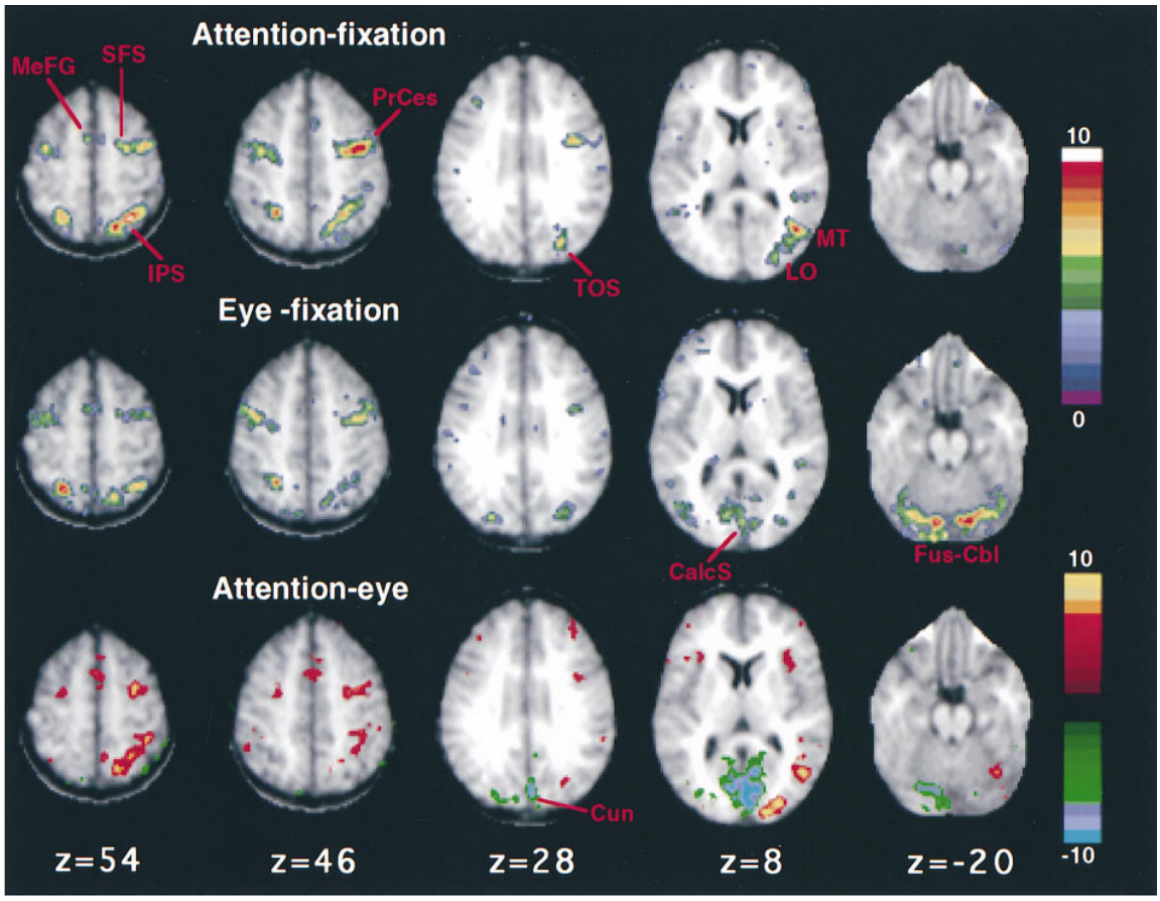

Figure 2: Example of parietofrontal regions involved in the control of eye movements and attention (Corbetta et al., 1998).

screen for at least 2 seconds. Given the anatomical overlap between tool-use understanding and visuospatial attention, it is possible that a relationship might exist between these two processes that might be reflected in measured parietofrontal activity. For instance, it is unclear if parietofrontal encoding of affordances might be attenuated, lateralized or modulated in any way by restricting an observer's ability to direct foveal attention over scene features. Such a result would suggest that parietofrontal processing of affordances could be coupled with the ability of the observer to allocate visuospatial attention towards distinct scene features. Therefore the third aim of this thesis Aim 3 focuses on contrasting the presence/absence of saccades on the spatiotemporal dynamics of parietofrontal processing of affordances. To achieve this goal, a flash experiment was designed wherein stimuli

durations were reduced to 100 milliseconds, thereby naturally forcing participants to restrict saccades and process the image extra-foveally. Subsequently, stimuli durations were increased to 500ms, thereby allowing for the reemergence of saccades.

Summarizing the three aims together, **the overall goal of the dissertation is to understand the spatiotemporal patterns of parietofrontal activity and eye movements during the visual perception of contextual and grasp specific tool-use scenes.** In the following section, a brief overview is presented of a related behavioral experiment in our lab that forms the conceptual basis for all three specific aims. The hypotheses in each of the three specific aims are then subsequently detailed that address the overall goal of the dissertation.

1.6 Related work

In order to address the goal of the dissertation, we extended previous behavioral work in our lab that evaluated how the combination of tool-use context and tool-grasp influenced response accuracies and response latencies when participants evaluated static tool-use images. Similar to previous neuroimaging work in our lab, a tool's context of use was defined based on its pairing with another object (Mizelle and Wheaton, 2010a, b). We had focused on three levels of tool-object relationships or tool-use contexts: correct (e.g., hammer-nail), incorrect (e.g., hammer-paper) and spatial (e.g., hammer-wood). The spatial context coupled a tool and object that are usually part of the same scene, but do not afford a functional tool-object action. These three levels were then orthogonally combined with four different types of grasp-postures: no hand (control condition), static hand (at the bottom of the picture roughly equidistant from tool and object), a functional grasp-posture (e.g. grasp the handle of the hammer), and a manipulative grasp-posture (e.g. grasp the head of the hammer). While the manipulative grasp-posture serves to manipulate the tool (e.g. move it), it does not afford a functional engagement of the tool on the object. Together the combination of context and grasp gave rise to 12 conditions, shown in Figure 3 for a representative tool.

The images were always presented from a right-handed egocentric perspective. The tool

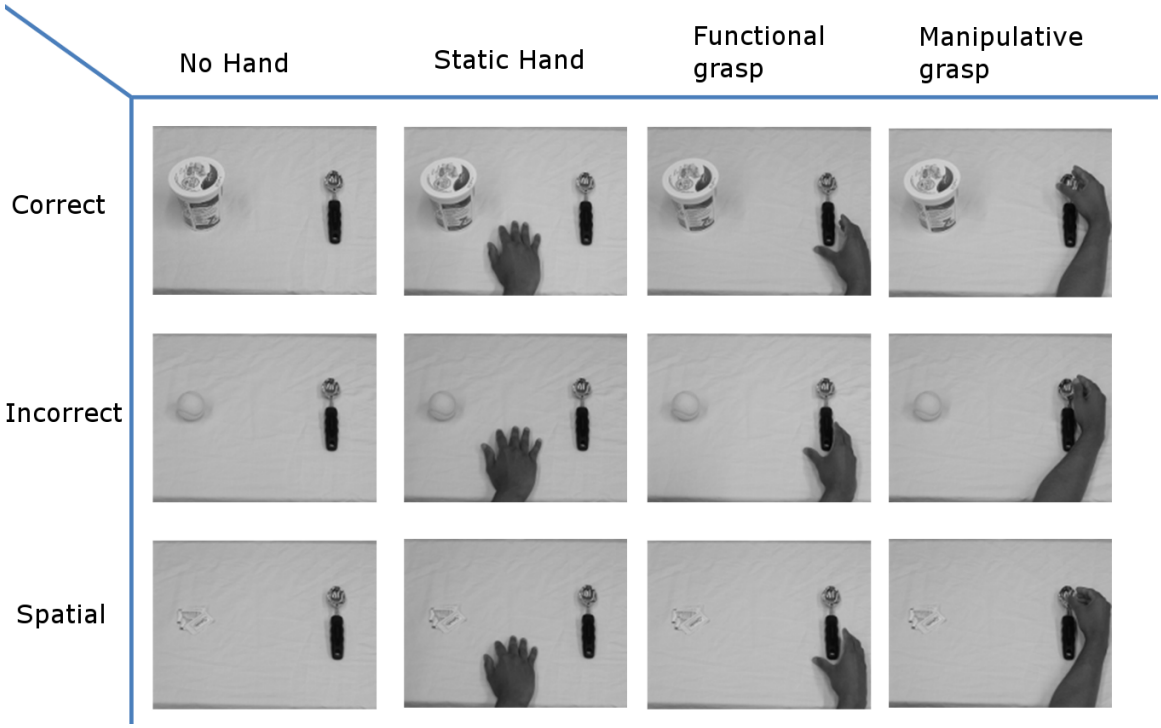

Figure 3: The combination of 3 tool-object contexts along each row and four type of grasp-postures along each column is shown here for an exemplar tool, together giving rise to 12 conditions with varying affording properties. For the purposes of this dissertation, a "tool" was defined as an object in the right hemifield that could be used by an actor to interact with an "object" in the left hemifield. Images were normed for visual complexity and for the contextual relationships between the tool-object pair. When presented with these images, participants were asked to evaluate the motoric relationship between the tool and object pairs i.e. if the tool-object pair was functionally correct or incorrect.

and object were only placed next to each other and were not interacting or positioned to interact with each other. Images were normed for visual complexity and on the contextual relationships between tool-object pairs. Participants were explicitly told beforehand that they would be seeing images of tool-object pairs wherein the tool was defined as an entity that could interact with the object and was always in the right hemifield while the object was always in the left hemifield. Participants were also told that they would see a right hand in some of the images interacting with the tool, but that their task was always to evaluate whether the tool-object relationship was functionally correct or incorrect. Participants were prepared in advance that they would be viewing right handed, egocentric tool-object images. Participants used their right hand to record a correct relationship and used their left hand to record an incorrect relationship. All participants used their dominant right

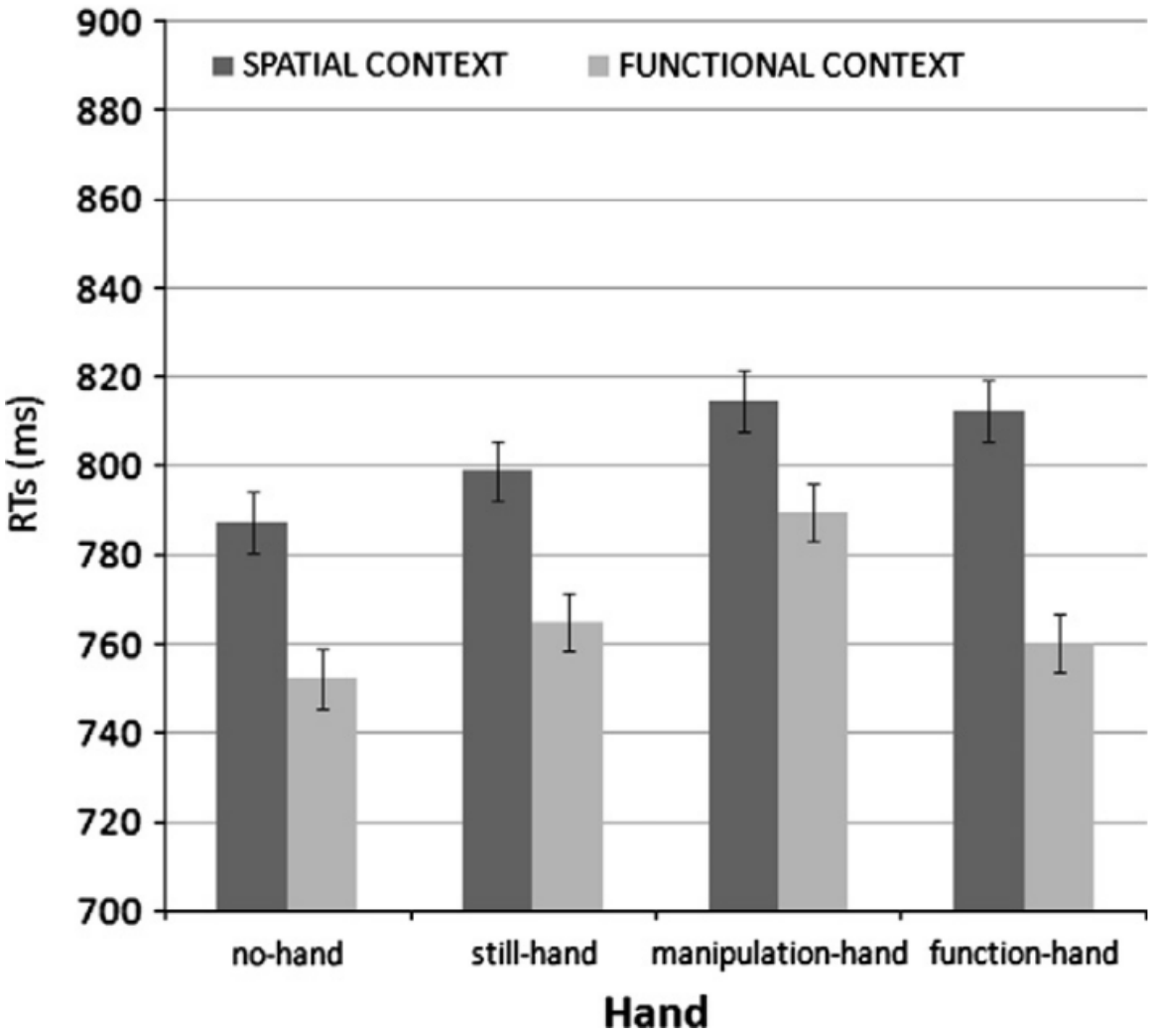

Figure 4: With respect to the main effects of Context, the spatial context elicited the longest decision times. With respect to the main effects of Hand, the manipulative grasp elicited the longest decision times, especially within the correct context and the no hand conditions elicited the shortest decision times. Please note that the terminology for the conditions in this plot is slightly different from this dissertation. Specifically, 'still hand' refers to the static hand condition and 'functional context' refers to the correct context (Borghi et al., 2012).

hand to categorize a correct tool-object relationship as we aimed to evaluate the effect of an egocentric right-handed grasp-posture on evaluating appropriate tool-object content. Subsequently, the correct and spatial contexts that afforded a certain degree of tool-object action were analyzed together and the incorrect context was analyzed separately. Behavioral data revealed that the spatial context elicited the longest response times. Amongst the four hand conditions, the no hand conditions elicited the fastest decision times whereas the

manipulative grasp condition elicited the slowest decision times. Importantly, data showed an interaction effect wherein the manipulative grasp condition elicited slower decisions times when compared to the no hand control condition with the strongest effects within the correct context (Figure 4). A similar effect was not observed for the static hand grasp or functional grasp. Thus even though the grasp per se was not essential to the task, a non-functional or manipulative grasp-posture that does not support tool-object action interfered in evaluating tool-object content especially when the tool-object pair itself clearly afforded action (Borghi et al., 2012). It should be noted that the interference effect of the manipulative grasp-posture was also observed within the incorrect context, though this effect was outside the core hypothesis that pertained to the correct context. We had suggested that the observer may be decoding tool-object content via an ongoing motor simulation process over parietofrontal regions and as such, the affordances of the manipulative grasp-posture may have delayed a behavioral response on tool-object content as they do not support tool-object action i.e., interfering with the tool-object motor simulation process (Borghi et al., 2012). This interpretation was strengthened by results in a separate experiment where participants used their feet rather than their hands to record their decisions. In this latter experiment, there was no interference effect of the manipulative grasp-posture. This finding reinforces the hypothesis of automatic parietofrontal engagement to encode the intent of the manipulative grasp-posture; the cooption of parietofrontal circuits to encode the manipulative grasp could have effectively delayed a behavioral response when subjects use their hands (a matching end effector) but not when they recorded their decision using their feet (Borghi et al., 2012).

Data from this behavioral study provided evidence that the perception of a tool's affordances could be modulated based on a combination of the type of tool-object context and the manner by which the tool was grasped. Findings from this behavioral study were used as conceptual substrates for the hypotheses in the specific aims, outlined in the following section.

1.7 Specific Aims

1.7.1 Aim 1

Purpose: The purpose of Aim 1 was to understand the temporal dynamics of parietofrontal activations underlying the perceptual judgment of the same contextual and grasp specific tool-object stimuli from the prior behavioral study (Borghi et al., 2012). The task required participants to passively evaluate whether tool-object pairs were functionally related or not. Behavioral responses were not collected to avoid any potential confounding effects of a motor response on parietofrontal activity. Specifically, we focused on the interference effects of the manipulative grasp-posture within the correct tool-object context and sought to correlate the temporal dynamics of parietofrontal activity with the time scales of the delayed response times observed in our prior data (Borghi et al., 2012).

Hypothesis: Within the correct tool-object context, the manipulative grasp-posture would elicit temporally extended activity over left hemispheric parietofrontal regions when compared to the no hand control condition, with corresponding spectral differences in the beta band. Left parietofrontal regions are well known to be involved in storing tool-use knowledge and in motoric beta band processing of tool-use (Chao and Martin, 2000, Krliczak and Frey, 2009, Mizelle et al., 2011). The temporally extended left hemispheric parietofrontal activity would correspond to the observer trying to resolve the conflict due to the atypical grasp posture and would support our hypothesis of an automatic activation of parietofrontal action regions to encode the manipulative grasp-posture. We also hypothesized that the time scales of these activations would be in line with our prior behavioral data (Borghi et al., 2012). Conversely, we hypothesized that there would be no such extended differences between the other grasp-postures (functional, static hand) and the no hand control condition. Finally, we hypothesized that the differential parietofrontal activations between the manipulative grasp and the no hand control condition would be maximal in the correct tool-object context wherein action affordances are already at a maximum and support tool-object engagement.

1.7.2 Aim 2

Purpose: Results from Aim 1 (Natraj et al., 2013) revealed that the manipulative grasp-posture did uniquely elicit temporally extended parietofrontal activity when compared to the no hand control condition. However, contrary to our hypothesis, the manipulative grasp-posture effect was observed across all three contexts and over bilateral parietofrontal networks and was not confined to the left hemisphere alone. Within each context, there were unique differences in the manner the manipulative grasp-posture was processed. In the correct tool-object context, the manipulative grasp-posture elicited early left parietofrontal activity differences (100-200ms) with the no hand condition, followed by an exclusively late appearing right parietofrontal difference (400-600ms) along with primarily theta band spectral differences. In the spatial tool-object context, the only difference between the manipulative grasp-posture and the no hand condition was a late appearing bilateral parietofrontal activity along with primarily alpha and beta band spectral differences. The time scales of these differences were in line with our prior behavioral work (Borghi et al., 2012). Together, these results suggest the engagement of a right parietofrontal network to understand the intent of the manipulative grasp-posture and a concurrent activity over left parietofrontal regions to process tool-object context. The temporally extended (400-600ms post image onset) right parietofrontal activations may correspond to the observer trying to understand the intent of the atypical manipulative grasp in relation to the tool-object relationship. This additional cognitive demand may underlie the previously observed delays on a final decision on tool-object content (Borghi et al., 2012).

Given the anatomical overlap between action understanding and attention over the same parietofrontal networks, it is possible that there might also be a corresponding difference in spatiotemporal gaze pattern when an observer evaluates tool-object scenes with and without the manipulative grasp-posture. Such a differential gaze pattern could potentially shed light on the type of action information that may be driving the right parietofrontal activity observed in Aim 1 (Natraj et al., 2013), underlying the interference effect observed in our prior behavioral study (Borghi et al., 2012). Therefore the purpose of Aim 2 was to understand spatiotemporal patterns of eye movement as participants viewed the same

stimuli as Aim 1, with the same task of passively evaluating whether tool-object pairs were functionally correct or incorrect. Similar to Aim 1, no behavioral responses were collected to avoid any potential interference of a motor task on gaze data. While all 12 conditions were included in the experimental design, the static hand conditions (across all three tool-object contexts) were dropped from the hypotheses and the analyses as they largely evoked the same behavioral and neural responses as the no hand conditions in our prior study and in Aim 1 (Borghi et al., 2012, Natraj et al., 2103). This reduced the number of conditions down in the study to 9. We primarily focused on gaze scanpaths over the areas of interest (AOI) in the stimuli, and the weighting of the AOI by the observer.

Hypothesis: It was hypothesized that hierarchical clustering of gaze scanpaths and AOI weightings would group conditions primarily by context (correct, incorrect and spatial) since the task required participants to evaluate tool-object content. Within each of the 3 context clusters it was hypothesized that the manipulative grasp-posture would be distinct from the no hand and functional grasp -posture with the strongest distinctions in the correct and spatial contexts at time scales in line with results from Aim 1 (Natraj et al., 2013). In essence, Aim 2 posited that contrasts of gaze data would mimic the results of Aim 1 and our prior data (Borghi et al., 2012, Natraj et al., 2013).

1.7.3 Aim 3

Purpose: Results from Aim 2 (Natraj et al., 2015) revealed that though the task required evaluating tool-object content, eye movements were automatically primed to grasp-affordances. Specifically, clustering of gaze scanpaths and AOI weightings grouped condition by the type of grasp-posture, rather than context, into three grasp-specific clusters. It should be noted that the distinctions between these three grasp-specific clusters were most robust within the correct and spatial tool-object contexts. Results also suggested that eye movements were influenced more by the affordances of the grasp-posture rather than the mere presence of a hand per se. Specifically, in the absence of a grasp in the scene, the object was foveally weighted the most, suggestive of an object-oriented action priming effect wherein the observer may be evaluating tool-object content in terms of simulating

an engagement of the tool on the object (Thill et al., 2013). However, the manipulative grasp-posture, unlike the functional grasp-posture, drew attention away from the object and caused the greatest disruption in the aforementioned object-oriented action priming effect. Rather than being purely driven by the presence of a grasp per-se, eye movements were automatically primed to the affordances of the grasp-posture. The enhanced foveal attention towards the manipulative grasp-posture may serve to gather the visual information necessary to understand grasp-intent (Natraj et al., 2015) and could potentially be driving the parietofrontal differences in Aim 1 (Natraj et al., 2013) and the behavioral interference effect in our prior data (Borghi et al., 2012).

In both Aims 1 and 2, the visual stimuli were on screen for a minimum of 2 seconds. The observer had sufficient time to direct foveal attention across scene features and gather the visual information (Aim 2, (Natraj et al., 2015)) to ostensibly drive parietofrontal regions that encode grasp-intent (Aim 1, (Natraj et al., 2013)), thereby delaying decisions on tool-object relationships (Borghi et al., 2012). However, as outlined earlier, overlapping parietofrontal regions are also involved in the control of fixation and saccades (Anderson et al., 1994, Gaymard et al., 1998, Mort et al., 2003), visuospatial attention (Corbetta et al., 1998) and saccadic inhibition (Chikazoe et al., 2007). Results in Aim 2 (Natraj et al., 2015) raise the possibility that an unknown proportion of the parietofrontal differences between the manipulative grasp and no hand conditions in Aim 1 (Natraj et al., 2013) may in fact correspond or correlate to differential gaze patterns between the two conditions. For instance, it is unclear if the right hemispheric parietofrontal encoding of the manipulative grasp-posture might be differentially attenuated or modulated by restricting an observer's ability to direct foveal attention over scene features. The purpose of Aim 3 was to address this ambiguity by evaluating the influence of eye movements on parietofrontal differences between the manipulative grasp-posture and the no hand condition when encoding the same stimuli as Aims 1 and 2 and the prior behavioral data (Borghi et al., 2012, Natraj et al., 2013, Natraj et al., 2015).

To address the purpose of Aim 3, the experimental design contained two sub-experiments, one wherein participants executed saccades and another wherein participants were restricted

to processing the images using only peripheral vision, without any saccades. One method of restricting saccades would be instruct participants to process the tool-object image using peripheral vision while maintain fixation at center. This method of explicitly instructing participants to restrict saccades would engage inhibition related frontal activity when viewing wide field complex visual stimuli (Chikazoe et al., 2007). It is probable that the amount of active inhibition could be influenced by the type of grasp-posture, tool-object context, or a combination of both (Castelhano et al., 2009). As a result, statistical contrasts between conditions would be confounded by differential, inhibition related parietofrontal engagement. To avoid these potential confounds in the data, a flash experiment was designed wherein stimuli durations were shortened to 100ms. Previous research has shown that the lower bound to initiate a saccade is approximately 120ms (Kirchner and Thorpe, 2006). We hypothesized that a stimulus duration of 100ms would automatically eliminate saccades without any active instruction to do so. Such a short duration stimulus would naturally force the participant to gather the entire scene information using only peripheral information, and subsequently rely on short term memory to evaluate tool-object content. In a second experiment, stimuli durations were increased to 500ms, allowing for the reemergence of saccades and the ability to parse the scene using continuous foveal information. Together, these two experiments served to evaluate the role of visuospatial attention on parietofrontal processing of the manipulative grasp-posture. To ensure that participants were accurate in the task given the rapid nature of the stimuli, behavioral responses were collected. The response hand used to record a correct/incorrect tool-object relationship was counterbalanced and all three tool-object contexts were analyzed together. This was done to rule out any potential confounds in accuracy due to the response hand especially given the shortened duration of the stimuli. Eye movements were collected via concurrent eye tracking equipment and by placing two electrooculography (EOG) electrodes by the side of the left eye. For reasons similar to Aim 2, the three static hand conditions were excluded. Similar to the all previous experiments (Borghi et al., 2012, Natraj et al., 2013, Natraj et al., 2015), the task required participants to evaluate whether tool-object relationships were correct or incorrect. A mass-univariate statistical approach was utilized to precisely

capture the spatiotemporal differences between conditions at each individual electrode and each individual time-point (channel-time pair (Pernet et al., 2011)).

Hypothesis: We hypothesized that even with the rapid image presentations, participants would be able to successfully perform the task and have accuracies significantly greater than chance. Similar to our prior behavioral data, we hypothesized that the manipulative grasp-posture and the spatial tool-object context would delay decisions on tool-object content (Borghi et al., 2012). With respect to eye movement, we hypothesized that the rapid stimuli within the 100ms experiment would negate participants from directing foveal attention towards scene features. Our major hypotheses concerned the spatiotemporal patterns of parietofrontal activity underlying the interference effect of the manipulative grasp-posture. In Aim 1, we had proposed the existence of a late appearing (400-600ms after image onset) right parietofrontal network to specifically encode the intent of the manipulative grasp-posture, along with a continual left parietofrontal network to encode the tool-object content itself. However, the results of Aim 2 had suggested that this right parietofrontal activity could be driven by enhanced attention over the manipulative tool-end. Given the inability of participants to direct foveal attention in the 100ms experiment within Aim 3, we hypothesized that the right parietofrontal activity (400-600ms post image onset) would be inhibited and instead, a dominant left parietofrontal network would underlie the interference effect of the manipulative grasp-posture. In the 500ms experiment, we hypothesized that the right parietofrontal activity (400-600ms post image onset) would reemerge as participants would have to time to execute saccades and direct attention.

1.8 Integration

Given the overlap between action understanding and gaze control over parietofrontal regions, Aim 3 integrated the conceptual underpinnings of Aims 1 and 2 and sought to evaluate the overall influence of saccades on the parietofrontal processing of the manipulative grasp-posture. In the 100ms experiment, participants executed saccades in $< 10\%$ of trials and their accuracy in evaluating tool-object content was above 80% in the task, though

the stimuli durations were only 100ms. With respect to neural activity, the manipulative grasp condition elicited a larger, negative-oriented cortical potential when compared to the no hand condition, beginning at 130ms after image onset. This ERP difference originated at centro-right parietal electrodes (N100 potential), and rapidly propagated to central and right frontal electrodes, lasting till 228ms after image onset. While the functional grasp-posture also elicited a differential N100, there was no propagation to frontal regions and in addition, the differences with the no hand condition were much less sustained and lasted only till about 200ms. The centro-right parietofrontal ERP response specific to the manipulative grasp posture refuted our hypothesis that the absence of saccades would negate right parietofrontal activity. However, rather than appear 400-600ms post-image onset (Natraj et al., 2013), the ERP difference was present only in very early time-points. Importantly, results in the 500ms experiment shed light on how saccades (in 75% of trials) altered the parietofrontal processing of the manipulative grasp-posture. Similar to the 100ms experiment, the manipulative grasp condition elicited a greater N100 potential than the no hand condition, starting earlier than the 100ms experiment by 30ms. However, the ERP differences between the manipulative grasp and no hand conditions faded when saccade initiation probability was maximum (approx. 210ms), thereby delaying the propagation to frontal electrodes to 350-400ms after image onset when first saccade probabilities were minimal. When the frontal differences did emerge, the polarity difference in the ERP was switched, with the manipulative grasp-posture eliciting a greater positivity/lesser negativity than the no hand condition. This result in part does support our hypothesis on the frontal processing of the manipulative grasp-posture, as spatiotemporal activation patterns and ERP polarity characteristics were influenced by differences in visuospatial attention mechanisms (foveal vs. non-foveal vision). These spatiotemporal ERP difference patterns were unique to the manipulative grasp-posture and were not observed for the functional grasp-posture, in line with our overall hypothesis that the manipulative grasp-posture elicits unique neural encoding mechanisms. Generally, results in the 500ms experiment also showed that a temporal coupling between parietofrontal activity and eye movements. For instance, saccade initiation latency across conditions was temporally related to a late-appearing positive peak in

the ERP over parietal electrodes at around 250ms. However, this time-window was much beyond the 100-200ms window that exhibited the N100 ERP differences between conditions. In both experiments, it is probable that the N100 related to processing the affordances of the grasp-posture. Over frontal electrodes, the time-points of the ERP differences (350-400ms) preceded a differential gaze pattern (400-450ms) wherein foveal gaze position was more likely to be fixated at the object in the no hand conditions and was more likely to be focused at the manipulative tool-end in the manipulative grasp condition. Finally, when directly contrasting how the affordances of the manipulative grasp-posture was encoded in both experiments, source localization analyses revealed greater engagement of the left middle temporal gyrus in the 500ms experiment, 100-228ms post image onset, and greater engagement of the left precuneus in the parietal lobe in the 100ms experiment, 228-300ms post image onset. It should be noted that the encoding of the tool-object context itself in both experiments happened relatively much later in time, at roughly 500-700ms after image onset. The encoding of tool-object context was constrained to left parietofrontal regions in the 100ms experiment and elicited bilateral parietal activity in the 500ms experiment. The fact that only the perception of grasp-specific affordances were perturbed by different attention-related mechanisms is expounded upon further in Chapter VI.

Overall, the goal of all three Aims and this thesis was to understand the spatiotemporal activation patterns over parietofrontal action encoding regions and eye movements during the perception of contextual and grasp specific tool-use scenes. As detailed in the previous sections, the salient result observed in this thesis was with regard to the encoding of the manipulative grasp-posture. Specifically, in all three Aims, the affordances of the manipulative grasp were automatically activated and influenced both parietofrontal activity and eye movements, though the grasp per se was not essential to the task of evaluating tool-object content. Aims 1 and 2 also showed unique neurovisual encoding mechanisms of the manipulative grasp-posture especially when the tool-object pair itself afforded a certain degree of action (correct and spatial tool-object contexts). Aim 3 showed how such grasp-specific action knowledge can be rapidly recalled with differential neural mechanisms based on the ability of the observer to direct foveal attention. This automatic sensitivity

in understanding the intent of the observed manipulative grasp-posture may correspond to a lifetime of learning the affordances of grasp-specific action outcomes. Indeed, this thesis presents evidence that tool-related grasping postures strongly influence understanding the action relationships between tools and objects. Such cognitive motor knowledge may be vital in understanding the motoric errors and outcomes of grasp-specific actions. For example, when a mother watches her child learning to use a spoon, the errors in grasping the bowl of the spoon is immediately apparent to the parent, who might then correct her child's behavior. Unique to humans, such a series of singular corrections in motor behavior through development may subsequently consolidate adaptive action knowledge onto parietofrontal regions. Such grasp-specific action information may be vital in navigating a human environment that is almost exclusively shaped on complex tool-use knowledge. The potential applications and future directions of this work are detailed further in Chapter VI.

CHAPTER II

METHODOLOGY

This thesis utilizes two means of assessing neurophysiological responses during the perception of complex tool-use: Electroencephalography or EEG and eye tracking. A brief overview of each methodology is outlined in the following sections.

2.1 Electroencephalography (EEG)

Electroencephalography or EEG is a measure of the electrical currents produced by neuronal activity in the brain (Nunez and Srinivasan, 2006). These propagating currents have to travel through bone, brain tissue and skin before reaching the scalp, where they are captured as voltage drops across a recording electrode. The neurophysiology underlying a single electrode is outlined in Figure 5.

The measured signal reflects synchronous, summed excitatory post synaptic potentials (EPSPs) at the dendrites of spatially aligned cortical neurons (such as pyramidal neurons) through the dipole effect. Consider a pyramidal neuron undergoing EPSPs at its basal dendrites due to incoming neuronal communication. This neuron experiences an influx of sodium, causing a local current sink. To conserve electrical neutrality due to this charge imbalance, active sources of current are produced at the apical regions of the neuron by the opening of channels facilitating cationic outflow. As a result of these processes, a pyramidal neuron develops an electric field over its apical and basal ends and thus by basic laws of physics, behaves as a simple current dipole illustrated in Figure 5. These local cortical pools act as extracellular current generators. Spatiotemporal summation of synchronous EPSPs results in electric fields radiating outwards perpendicular to the cortical layer till they are recorded at the scalp. Action potentials on the other hand are much more transient and generate weaker fields due to asynchronous depolarization along the axon and are not captured in EEG. The EEG thus records synchronous post synaptic summed potentials at

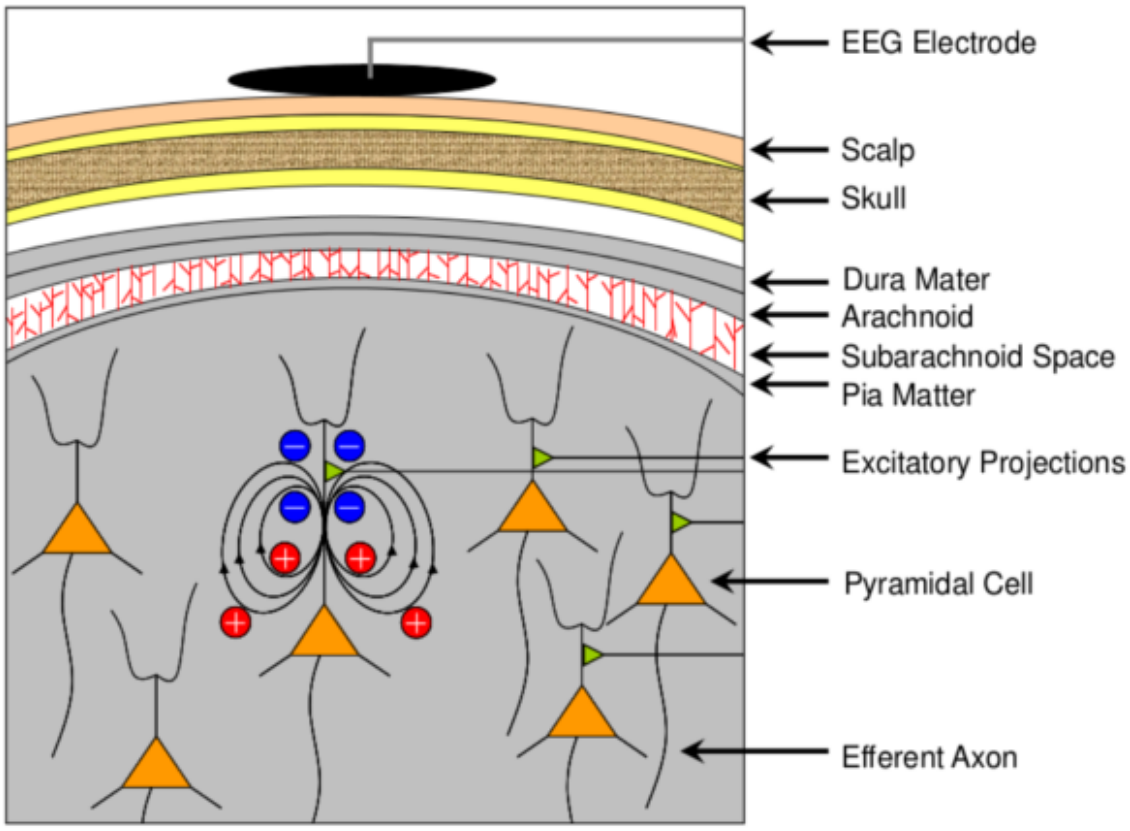

Figure 5: Snapshot of the neurophysiology underlying an EEG electrode at the scalp. The different layers of brain tissue, bone and skin separate neuronal activity from the recording electrode. Image also shows a pyramidal neuron undergoing the dipole effect. This neuron is in a cortical layer perpendicular to the surface of the scalp. Generated electric fields therefore propagate parallel to the scalp. Summation of many electric fields from individual neuron dipoles result in a large enough signal to be captured at the scalp. Here, the orientation of the dipole results in a weak signal at the recording electrode. (Figure courtesy: J.C Mizelle, PhD).

the scalp generated from layers of cortical neurons in the vicinity of a recording electrode. In this dissertation, a 58 channel EEG system is used according to the international 10-20 convention (Homan et al., 1987). A top down perspective of the 58 channel montage is shown in Figure 6. The 10-20 system signifies the ratio of separation between electrodes relative to the entire size of the head as measured from the naison (right above the nose bridge) to the inion (the bony projection at the back of the skull) along the central curvature of the head that separates the left and right hemispheres. The distance between the naison and the first electrode at the center of the forehead (FPZ in Figure 6) and the distance between the

inion and the last electrode at the center (Oz in Figure 6) is 10% of the overall head size as measured from the naison to the inion. The separation between electrodes along the center line is 20% of the overall head size. The ratio of distances between the other electrodes is 10% around the circumference, 5% between the most anterior electrodes (FP1-FPZ-FP2) and posterior electrodes (e.g. O1-Oz-O2) and 10% between the other electrodes going along the horizontal and vertical directions. The 10-20 system aims to therefore approximately place each electrode over the same anatomical position irrespective of the size of the head, which varies across the population.

As a measurement device, EEG has a few disadvantages, the chief being spatial resolution. Activity from generators that lie deeper in the brain are grossly attenuated by the time they reach the scalp and it is difficult to resolve these sources. Also, the orientation of the cortical dipoles influence the propagation of the electric fields. For example, cortical layers parallel to the scalp generate orthogonal fields that propagate towards an overhead recording electrode resulting in better quality signals. Whereas perpendicular cortical layers generate electric fields parallel to the scalp that are diminished by the time they reach the overhead electrode and may be picked up by other nearby electrodes. The EEG signal is also embedded in noise from different sources such as line noise, thermal noise, biological noise, overlapping generators etc. For these reasons, the EEG signal suffers low signal to noise ratios (SNR) and an inability to accurately identify deep neural generators. On the other hand, there are significant advantages to EEG, namely high temporal resolution and the ability to isolate task specific brain activity. For instance, EEG can be used to correlate scalp potentials to a stimulus or task (ERPs or event related potentials). This aids in understanding neural responses at fine time scales over gross anatomical regions. This is an example of time-voltage analyses and is a very commonly used tool in cognitive neuroscience. The ERPs in this dissertation correspond to visually evoked potentials or VEPs since the event itself is the onset of a visual stimulus (a tool-use scene). VEPs are typically extracted from EEG data by averaging many trials together. The task-irrelevant activity cancels out in the averaging process as noise since they do not tend to be time-locked to the stimulus. On the other hand, the VEP tends to be time locked to the stimulus onset

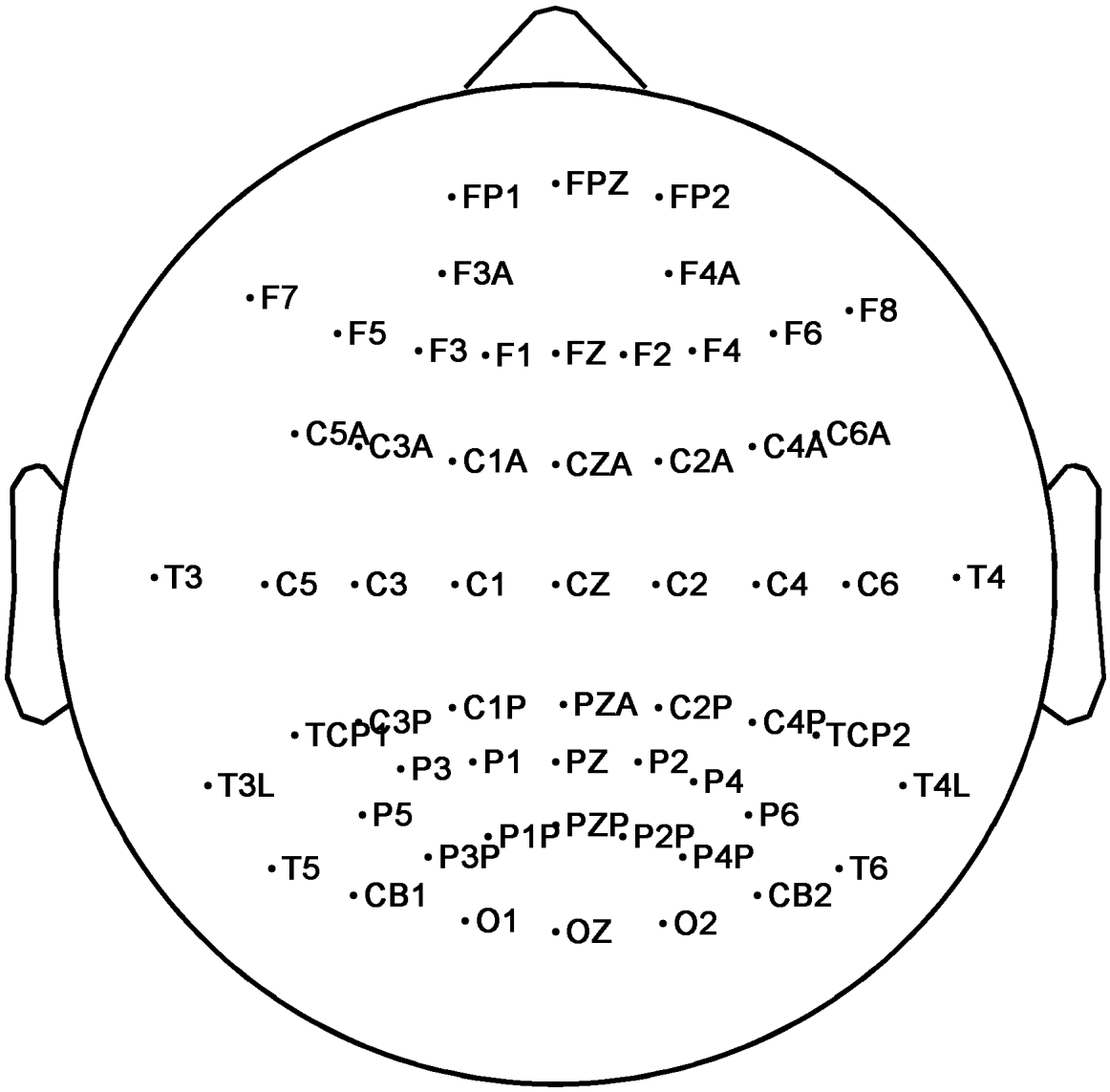

Figure 6: A top down overview of the 58 channel EEG montage used in this dissertation. The location of the electrodes over the scalp followed the international 10-20 system (see text for details). The nose is at the front and the two oblong-type shapes at the side denote the ears. The naming convention for the electrodes is as follows: FP Frontal Polar. F-Frontal. CA-central anterior. C Central. CP Central Posterior. T Temporal. TL - Temporal Lateral. P Parietal. O Occipital. CB surface above Cerebellum. Electrodes along the center dividing the head into the left and right hemispheres are along the Z-line. Odd numbered electrodes are on the left of the head and even numbered electrodes are on the right.

and while it is hard to identify it in single trials due to poor SNR, averaging many trials reveals the shape of the VEP (Luck, 2014). An example of computing a VEP is shown

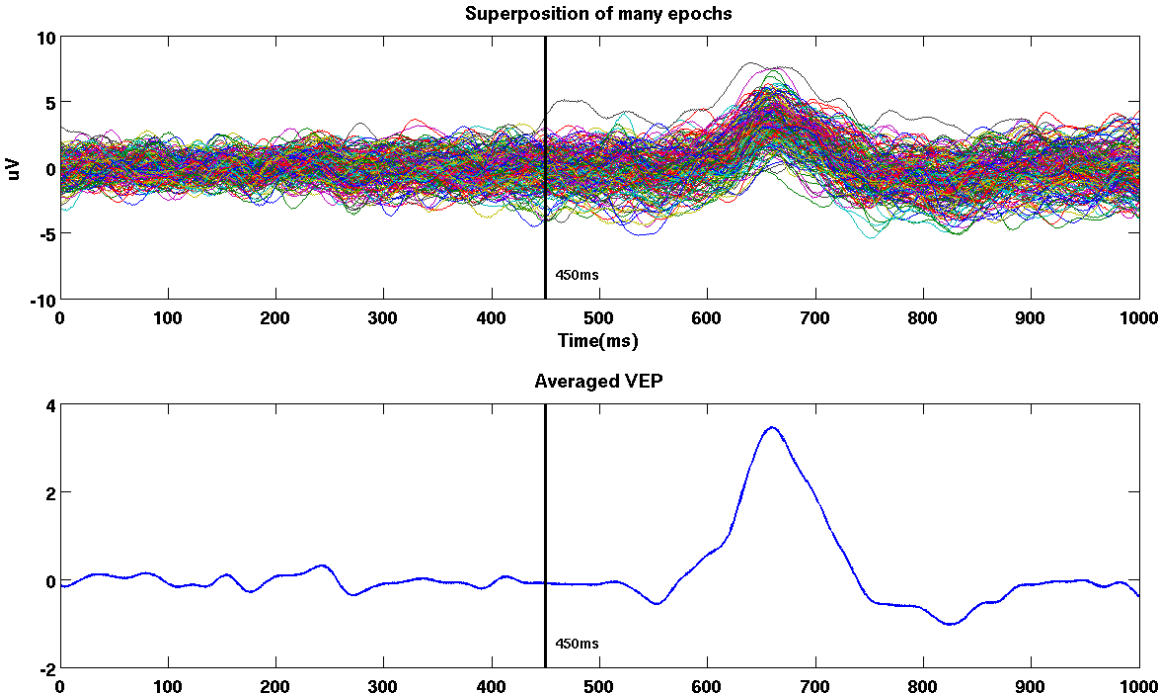

Figure 7: Example of computing a Visually Evoked Potential (VEP). The top plot showcases many single trial responses overlaid on top of each other. X-axis is time in milliseconds and Y-axis is in microvolts. Averaging the trials reveals the true shape of the VEP, in the bottom plot. The VEP is time locked to the stimulus onset (vertical solid line at 450ms) whereas the task irrelevant activity factors out as noise in the averaging process.

in Figure 7. The VEP is characterized by many distinct peaks and valleys after stimulus onset. The deflection of the peaks have typically been studied in great detail in the cognitive neuroscience literature. For example, the first peaking potential is called the C1 potential over the visual cortex that peaks within 100ms after image onset and is thought to be generated by sensory neurons in the calcarine fissure in the striate cortex. The next deflection is an early positive going deflection at around 100ms after image onset, called the P100 potential and is typically observed over occipito-parietal electrodes and is thought to reflect activity over neurons in the extrastriate cortex. The next subsequent deflection is a negative going deflection peaking around 150ms after stimulus onset (the N100) and is thought to reflect activity of neurons deep in the parietal lobe (Di Russo et al., 2002). The subsequent alternating positive and negative tending potentials are called the P200, N200, P300 and N400. The generators of these later deflections are thought to arise from

the sustained interactions between cortical neurons and deeper structures such as the hippocampus (Picton, 1992). These potentials have been broadly associated with a whole host of cognitive functions such as early sensory processing (Di Russo et al., 2002), attention related processing (Luck, 2014), memory related processing (Ruchkin et al., 1990), error identification (Picton, 1992) and can even be influenced by the gender of the subject (Steffensen et al., 2008). A complete overview of these potentials, their sources and functions can be found elsewhere (e.g. (Kutas and Federmeier, 2011).

EEG signals can also be analyzed in the frequency domain. For example, a DC-30Hz EEG signal can be decomposed (using wavelets, Fourier transforms, parametric signal modeling etc.) into linear, weighted combinations of oscillations in discrete frequency bins such as 0-4Hz (delta oscillations), 4-8Hz (theta oscillations), 8-12Hz (alpha oscillations) and 12-30Hz (beta oscillations). These frequency oscillations are thought to control the firing of neuronal assemblies. One way to visualize the role of the oscillations is to imagine the neuronal assemblies to pulsate at specific frequencies based on the location of the neuronal assemblies and the task demands. It is commonly thought that such oscillations may aid in the transfer of information across brain regions that are structurally connected and oscillating at the same rate (Engel and Fries, 2010). The density of the neuronal assembly typically correlate to the intensity of the oscillations. With respect to task specific relevance, cognitive tasks typically engage oscillations at theta cycles/frequencies, visual tasks typically engage alpha frequencies and motor tasks engage beta frequencies. These task-specific demands can either decrease/ increase oscillatory power with respect to baseline oscillatory power when the brain is not engaged in the task, a phenomenon called Event Related De/Synchronizations (ERD/S), first discovered by Pfurtscheller and colleagues (Pfurtscheller and Da Silva, 1999). It is thought that increase in oscillatory power (ERS) corresponds to a stabilization in the basal or idling state. This may serve to deactivate brain regions that are involved in inhibitory control. Conversely, decreases in oscillatory power (ERD) are thought to reflect destabilizations of the idling rhythms and as the brain is no longer in an idling state, cortical neurons may engage in task-specific demands. For example, voluntary hand movement is characterized by ERD over the hand motor knob

and ERS over the foot motor knob at the motor cortex. An example of a wideband view of ERD/S is shown in Figure 8.

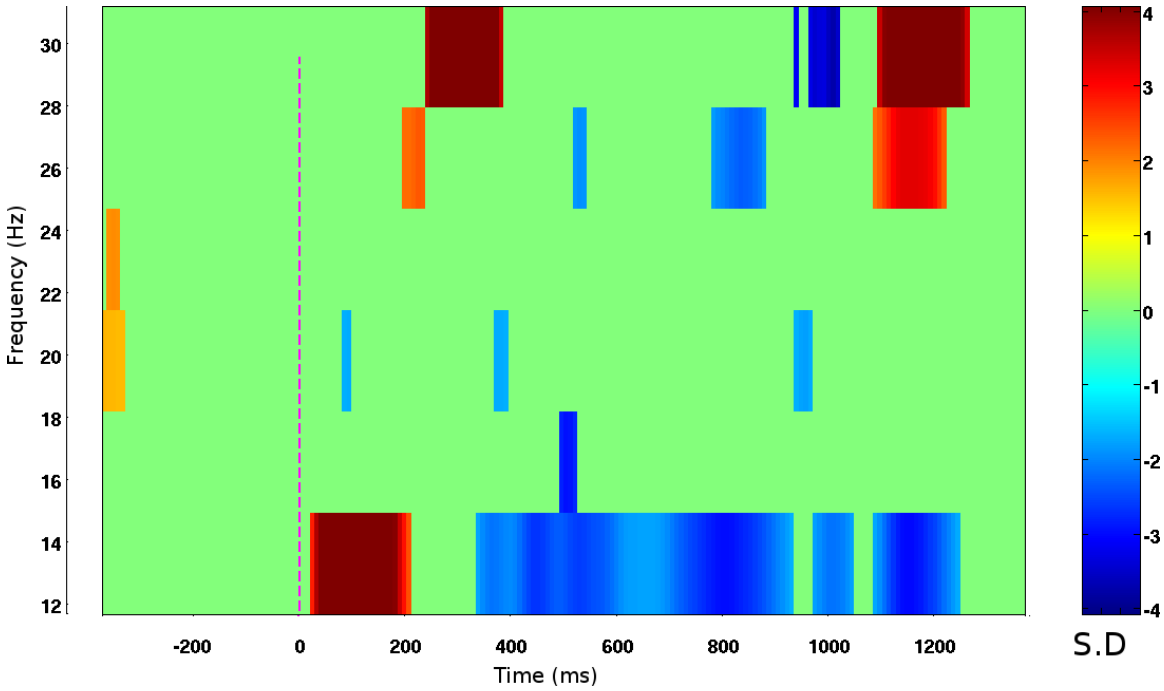

Figure 8: Example of computing Event Related Desynchronization/Synchronization at a particular channel. X-axis denotes time, Y-axis denotes frequency and the color coding represent the normalized change in oscillatory power (in S.D) with respect to the baseline power (time periods to the left of the vertical black line at 0ms) at each specific frequency. ERD/S is computed at each channel over many trials. In this particular figure, the frequency content of the EEG signal was identified by sliding a 256ms Hanning window with 50% overlap over the EEG signal and applying a Discrete Fourier Transform to the snippet within each window (Delorme and Makeig, 2004). It can be seen in this plot that there is significant ERD in the low beta band (12-14Hz) from 400 to 1200ms preceded by ERS in the same frequency bins from around 20ms to 200ms.

2.2 Eye tracking

The control of human eye movements can be characterized by two key variables: fixations and saccades (Bahill and Stark, 1979). Fixations serve to place key features of interest on the fovea, located at the center of the macula of the retina. Saccades on the other hand, transition foveal gaze from one feature to another. Information processing is limited in saccades (Bahill and LaRitz, 1984) which are high-speed ballistic eye movements while fixations capture the visual scene in fine detail. To a large degree, fixations and saccades are

dictated by the anatomy of the eye (Kandel et al., 2000). The fovea, located at the center of the retina contains the highest concentration of cones (color photoreceptors) and the highest innervation to the optic nerve that relays to the primary visual cortex. There is limited blood supply at the fovea and thus the retinal neurons at the fovea are arranged such that light directly hits the fovea. The fovea is thus specialized for maximum visual acuity and sharp vision, essential in understanding fine visual details. However, such a high resolution apparatus with its complex connections is computationally expensive to be replicated all over the retina. Hence visual acuity and photoreceptor density decays exponentially when going away from the fovea. Therefore the brain has to plan a sequence of fixations and saccades to place various features of interest on the fovea, thereby capturing information necessary for perception. Pioneering work by Alfred Yarbus in the 1960s (Yarbus, 1967, Yarbus et al., 1967) showed that fixations and saccades are very specific and do not scan the complete field of view. When studying eye movements when viewing faces, he found that participants focused on the eyes and mouth alone. When participants viewed natural scenes depicting a forest, participants focused on animals or bright canopy spots and did not scan the entire image. The brain therefore continuously weights all features of a visual scene before initiating a saccade using a process of covert or extra foveal attention. In covert attention, peripheral features of the visual scene not currently on the fovea are processed in parallel. Therefore fixations serve two purposes: first, the feature currently on the fovea is processed in finer detail. At the same time, other parts of the scene not on the fovea are attended to using peripheral vision and the brain saccades to the next feature of interest. There is naturally a trade-off when attending to an 'important' extra foveal scene feature and the distance of this feature from the center of the fovea. Thus it is currently hypothesized that eye movements may be driven not just to gather visual information to activate the visual cortex (and enable vision), but also in part by task specific cognitive demands that act as a top-down filter to discard irrelevant visual information. The resultant continuous planning of fixations and saccades forms the basis of high level visual perception. Based on these ideas, the eye has commonly been considered as a window into the mind (Van Gompel, 2007). Therefore, analyses of eye movements during tool-use observation can potentially be

a valuable tool in understanding ongoing action encoding mechanisms.

CHAPTER III

SPECIFIC AIM 1

3.1 Introduction

The purpose of Aim 1 was to understand the temporal dynamics of parietofrontal activations underlying the perceptual judgment of contextual and grasp specific tool-use. As detailed in Chapter 1, while research has shown how parietofrontal differentiate between correct and incorrect contexts of tool-use (Mizelle and Wheaton, 2010a, b) and encode a stand-alone tool grasp (Johnson-Frey et al., 2003), it is unclear how these regions process more complex and adaptive tool-use scenarios wherein a tool's action affordances are modulated by a combination of its context of use and the way it is grasped. To this end, we had performed a behavioral experiment that measured participants' reaction times when they evaluated contextual and grasp specific tool-use images. There were 3 possible tool use contexts within 4 possible hand postures, creating 12 conditions for any one tool (3 contexts by 4 hand postures). Figure 9 gives a depiction of all conditions for a particular tool. The context of tool usage was determined by its pairing with another object. For example, hammer-nail would be a correct tool-object pairing, hammer-paper would be an incorrect tool object pairing, hammer-wood would be a spatial tool object pairing. In these tool-object scenes, there were four hand variations. The no hand control condition, a static hand posture at the bottom of the scene equidistant between the tool and object, a functional tool-grasp (grasp the hammer-handle), or a manipulative tool-grasp (hold hammer-head). The key behavioral results revealed that unlike other grasp-postures, the manipulative or non-functional tool-grasp-posture delayed decisions on tool-object content when compared to the no hand control condition. It should also be noted that broadly, this interference effect of the manipulative grasp was observed in all three contexts, but with the strongest effects within the correct tool-object context. Thus even though the grasp per se was not essential to the task, a non-functional or manipulative tool-grasp

interfered in evaluating tool-object content, especially when the tool-object pair itself afforded a certain degree of action (Borghi et al., 2012). Importantly, the interference effect of the manipulative grasp was present only when participants responded with their hands and not with their feet. We had interpreted this finding as suggestive of an automatic activation of parietofrontal regions to encode the manipulative grasp that may have in turn disrupted participants motor response with their hands (a matching end effector) but not with their feet (Borghi et al., 2012). Therefore the goal of Aim 1 was to understand the parietofrontal correlates underlying this interference effect observed in this prior behavioral experiment. In the experiment here, we did not collect behavioral responses as we did not want motor control (motor execution/inhibition) to confound activity over parietofrontal and motor regions. To correlate the spatiotemporal dynamics of parietofrontal activations with the prior behavioral results, the hypotheses here included EEG data up to 600ms post image onset. This would allow for the 150ms differential in theoretically producing a motor response given that the average response time in our prior behavioral data was 750ms.

3.2 Hypotheses

Within the correct tool-object context, the manipulative grasp-posture would elicit temporally extended activity over left hemispheric parietofrontal regions when compared to the no hand control condition, with corresponding spectral differences in the beta band. Left parietofrontal regions are well known to be involved in storing tool-use knowledge and in motoric beta band processing of tool-use (Chao and Martin, 2000, Krliczak and Frey, 2009, Mizelle et al., 2011). The temporally extended left hemispheric parietofrontal activity therefore would correspond to the observer trying to resolve the conflict due to the atypical grasp posture. We also hypothesized that the time scales of these activations would be in line with our prior behavioral data (Borghi et al., 2012). Conversely, we hypothesized that there would be no such extended differences between the other grasp-postures (functional, static hand) and the no hand control condition. Finally, we hypothesized that the differential parietofrontal activations between the manipulative grasp and the no hand control

condition would be maximal in the correct tool-object context wherein action affordances are already at a maximum and support tool-object engagement.

3.3 Methods

3.3.1 Subjects

Sixteen (16) right-handed adult subjects (8 male, mean age, 21.2; SD, 1.3) were recruited for this research study. The experimental protocol was approved by the Institutional Review Board at Georgia Institute of Technology and each subject provided their written informed consent before the start of the experimental session. Subjects were healthy based on self-report, and had no history of neurological illness or injury.

3.3.2 EEG data acquisition

Subjects were seated comfortably in a chair with no restraints and fitted with a standard 58 channel tin electrode EEG cap (Electrocap, Eaton, OH) in accordance with the international 10-20 system. Neural activity was recorded using Synamps 2 (Neuroscan, Charlotte, NC). As well, two electrodes were placed above and below the left eye to record electrooculographic activity (EOG). These EOG channels were used offline to extract ocular artifacts (eye movement and eyeblink) from the EEG signal using adaptive filtering. Data acquisition was performed using a right ear reference at a sampling rate of 1000 Hz. The left ear was also recorded and was used offline in creating a linked ears reference. Visual stimuli were presented on a 20 inch widescreen display using Stim (Neuroscan, Charlotte, NC), at a distance of 6 feet from the subjects. The display was placed in the middle of their visual field and the height of the display was matched to the seated eye of the subject.

3.3.3 Experimental design

We used the same stimuli from our previous behavioral study which were images composed of tools, objects, and a hand interacting with the tools in various grasps. All images were normed according to visual complexity, tool recognition, and familiarity (Borghi et al., 2012). There were 23 tools with 12 possible variants of grasp and context within each tool

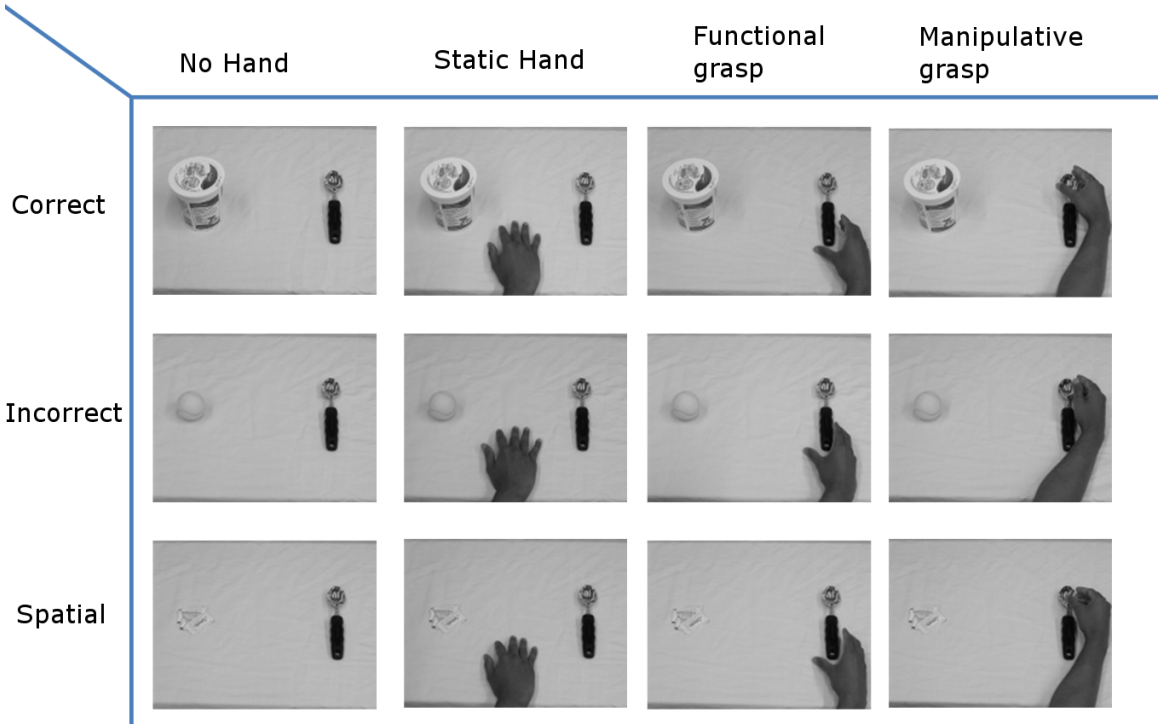

Figure 9: The combination of 3 tool-use contexts along each row and four type of hand-grasps along each column shown here for an exemplar tool, together giving rise to 12 conditions with varying affording properties. A tool is defined by an object in the right hemifield of the image that can be used by an actor to interact with an object in the left hemifield. When presented with these images, participants were asked to evaluate the motoric relationship between the tool and object pairs i.e. if the tool-object pair was functionally correct or incorrect.

(12 conditions); therefore there were a total of 23 X 12 = 276 images. Sorting all the images based on condition gave 23 images per condition. Each image was presented twice. Due to the large number of potential images, 17 (out of 23) images per condition were pseudo-randomly pre-selected for each subject. Therefore each subject viewed 408 images in total (17 pictures X 12 conditions X 2 presentations per picture). The presentation of these visual stimuli was equally distributed across 4 data acquisition blocks. Each block lasted approximately 11 minutes. The design of each block is outlined in Figure 10. Subjects viewed two sequential cues followed by the stimulus. The first cue was a circle, lasting 4000ms, to indicate a rest phase. The second cue was a cross lasting 500ms, which indicated an image was pending. Immediately following the cross, the target tool-object image was

presented on the screen for 2000ms. The circle cue then reappeared, and this cycle of rest-warning-stimulus repeated for the duration of the block. Continuous EEG was recorded for this duration. Image presentations were synched with the continuous EEG traces using Stim2 (Compumedics, Charlotte, NC). Unique codes generated by Stim2 identified the onset of each image presentation in the EEG signal to facilitate analysis. Participants were required to passively evaluate the appropriateness of tool-use i.e., whether the tool-object pair was functionally correct or incorrect.

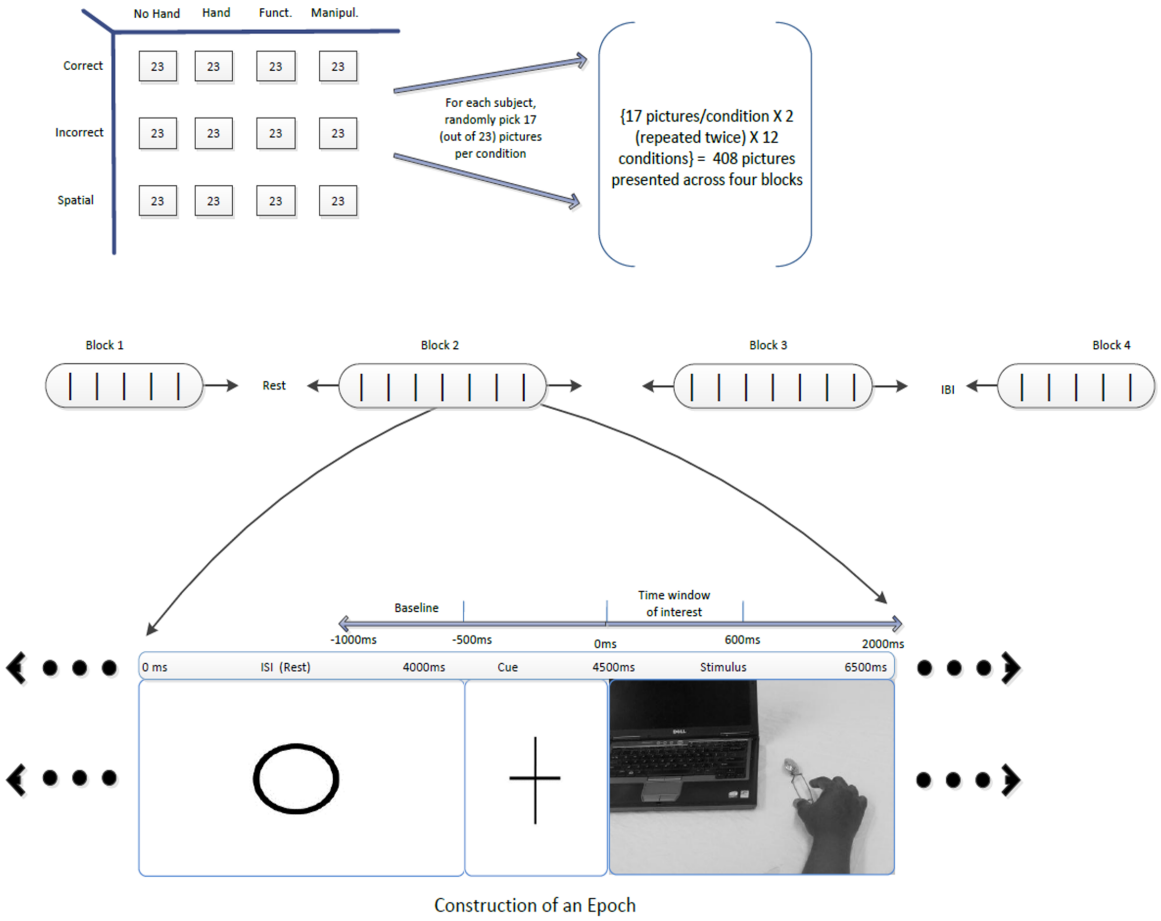

Figure 10: This figure details the design of this study and also highlights the time course of an EEG epoch.

3.3.4 Data analysis

The raw EEG data was first low pass filtered from DC-30Hz. An epoch was then constructed by extracting data from 1000ms before to 2000ms after tool-object image onset (Figure

10). Each epoch was therefore 3000 ms long (-1000 2000ms), with 0 corresponding to appearance of the tool-object image on screen and -500 corresponding to presentation of the cross indicating that an image was pending. Using custom scripts in Neuroscan and MATLAB (Mathworks, MA), the extracted epochs were then linear detrended and baseline corrected to the first 500ms (-1000 to -500ms), until onset of the warning cue. This allowed identifying ERPs in each epoch by factoring in task-irrelevant neural activity in the baseline interval. Intervals right before the image onset were not used as baseline as active processing of the cue might be ongoing. Epochs were sorted based on the 12 conditions. Epochs were then corrected for ocular artifacts (eye movement and eye blink) using the Recursive Least Squares Algorithm (Liavas and Regalia, 1999, He et al., 2004) implemented through the EEGLAB AAR Toolbox (Delorme and Makeig, 2004, Gmez-Herrero, 2007). Typically, EEG artifact correction is done by visual inspection of trials and discarding those with blinks, a processing step that results in a substantial loss of data. Alternatively, blind source separation methods which aim to decompose the EEG dataset into its principal components (PCA) or independent components (ICA) (Hyvrinen et al., 2004) typically are able to isolate the noise component very well as the noise component produces the most variance in the data. However, this method also requires subjective identification of the noise component. An alternative and automated method is to use the data from the EOG electrodes themselves which pick up the artifact caused by eye movements and eye blinks. A simple subtraction of the EOG signal from the EEG signal is not feasible as the polarity, intensity and latency of the artifact is different at every electrode as the eye artifact propagates from the anterior to the posterior electrodes. However, the EOG signal is statistically correlated with the EEG signal during eye movements and eye blinks. The RLS algorithm tries to identify this correlation between the EEG signal and EOG signal to correct for artifacts by minimizing the least squares error at every time point between the true signal and the signal with noise. The exponential weighting of the least squares error while calculating the adaptive filter coefficients makes RLS attractive for EEG artifact correction, as it follows the non-stationary trends of the EEG signal very well (Hayes, 2009) (Figures 11 and 12).

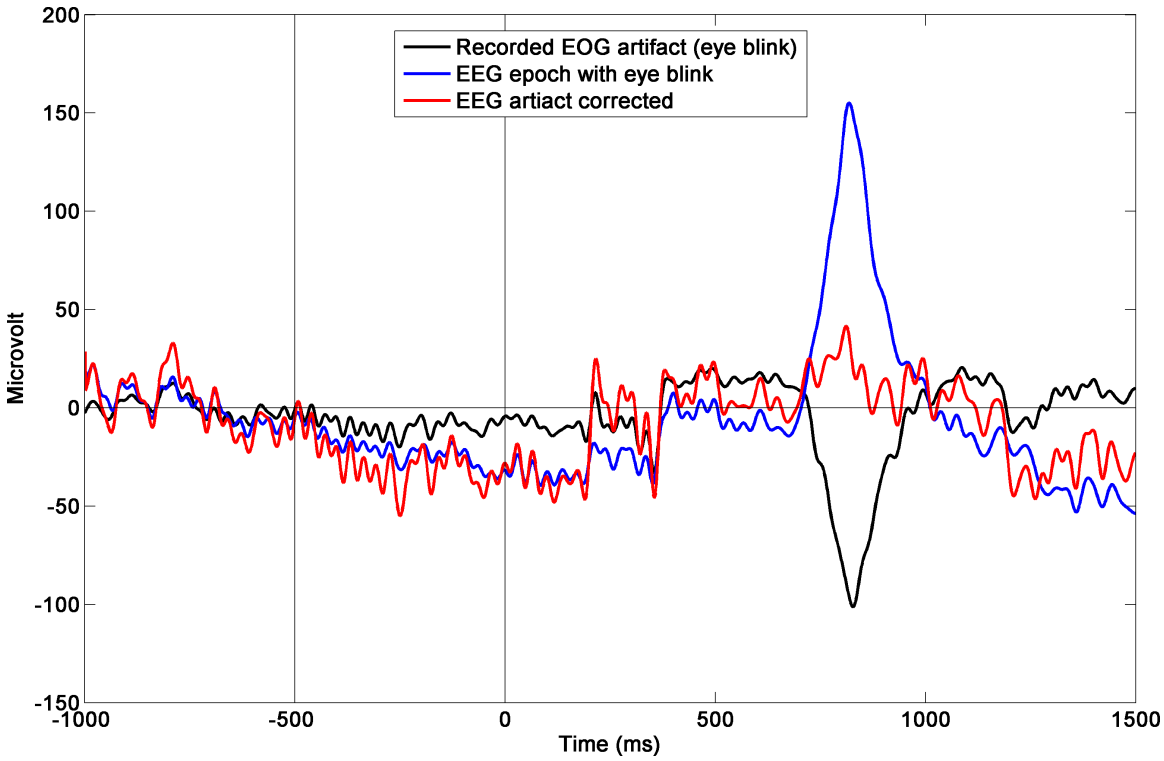

Figure 11: Example of artifact correction in the EEG signal due to an eye blink using adaptive filtering

A threshold magnitude of 75V was applied to reject epochs with residual noise. Research has shown that visually evoked response amplitudes are susceptible to non-task related differences, such as gender, societal influences (Steffensen et al., 2008) and age (Falkenstein et al., 2001). To counter this inter-subject variance, data for all 16 subjects were resampled to 1000 epochs according to the bootstrap procedure (Efron and Tibshirani, 1993, Mizelle and Wheaton, 2010a). Sixteen (16) new pseudo-subjects were created from the resampling procedure and these data were then subjected to time voltage analysis.

3.3.5 Time voltage analyses

To assess regional voltage-based effects, five Regions of Interest (ROI) were constructed based on the 10-20 convention and our hypotheses concerning parietofrontal activations: Left frontal: electrodes F7, F5, F3; Right frontal: electrodes F4, F6, F8; Left Motor: electrodes C3, C1, Cz; Left Parieto-temporal: electrodes TCP1, C3P, P5, P3; Right Parieto-temporal: electrodes TCP2, C4P, P4, P6, highlighted in Figure 13.

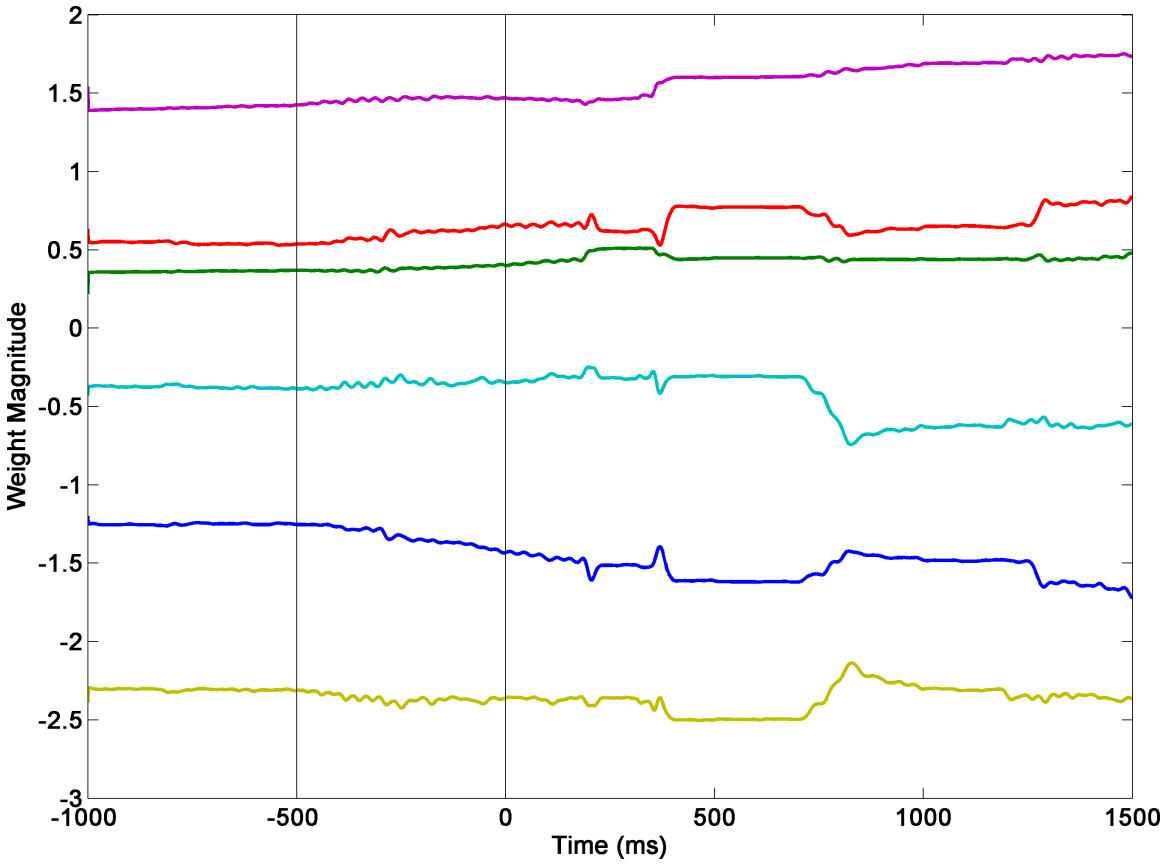

Figure 12: Evolution of the recursive least squares filter's weights with respect to the artifact correction in Figure 11.

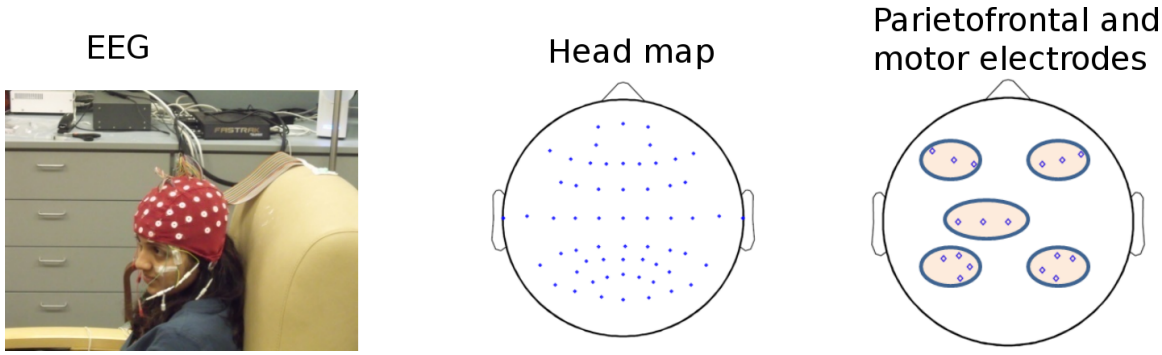

Figure 13: (Left) Visual of the EEG cap and the ocular electrodes (EOG). (Right) Selection of parietofrontal-motor electrodes from the whole brain (middle) based on the hypothesis of the study.

Time voltage waveforms were grand averaged per subject and condition across each ROI to obtain visual evoked responses at the single subject level. Given the complexity of the stimuli that could elicit multiple peaks in the visually evoked responses (VEPs)

with subtle variations from person to person, the VEPs were windowed and mean voltages within the window were analyzed (area under the evoked response curve). Based on the location and characteristics of these peaks, three windows of interest were chosen post hoc: Early phase (100-200 ms) after picture presentation; Middle phase (200-400 ms) after picture presentation; Late phase (400-600 ms) after picture presentation. Since our prior behavioral study had average subject response times of about 750 ms for the same visual stimuli, we did not analyze any EEG epoch beyond 600ms (with the 150 ms differential to account for the motoric processing underlying subjects' response required in our prior work). Analyzing binned data this way made it easier for us to relate spatiotemporal patterns of parietofrontal activations to time scales from our behavioral study. Within each time window, voltage values were averaged within subjects and conditions for each ROI to yield a mean voltage for each ROI, Subject and Condition. Statistical analyses were performed on the resultant means. For each time window, voltage means were compared using a 3-way analysis of variance (ANOVA) at the $p<0.05$ level, with the following within-subject factors: Hand (4 levels; no hand, static hand, functional, manipulative), Context (3 levels; correct, incorrect and spatial), and Region (5 levels; Right and Left Frontal, Left Motor, Left and Right Parieto-Temporal). Post-hoc t-tests were used to determine effects of Hand within each ROI ($p<0.01$), wherever allowable. As our interests were focused on the role of hand postures within a context, the no hand condition was used as control and individual t-tests were run (static hand vs. no hand, functional vs. no hand, manipulative vs. no hand) at each ROI, within each context to estimate the effect of hand posture on visually evoked responses. In line with our hypotheses, the post-hoc t-tests design did not compare all hand and context conditions with each other directly (such as manipulative vs. functional) (Nieuwenhuis et al., 2011) since we were not interested in the differences between a functional and manipulative grasp per se in this study. Our primary interests were in contrasting the manipulative posture vs. the no hand control condition within each context to infer unique spatiotemporal activations due to this hand posture that might speak to our prior behavioral results. As well, we aimed to examine if similar differences against control exist for the functional and static hand postures. To this end post-hoc

tests were carried out relative to the no hand condition and we qualitatively discuss the spatiotemporal differences in post-hoc contrasts (functional vs. no hand, static hand vs. no hand within each context).

3.3.6 Time frequency analyses

We evaluated the time varying properties of the EEG signal in the frequency domain. We carried out time frequency analyses in Matlab using EEGLAB (Delorme and Makeig, 2004), specifically evaluating Event Related Spectral Perturbations (ERSP) (Makeig, 1993, Makeig et al., 2004). ERSP are generalizations of ERD/ERS (Event Related De/Synchronization) in the power spectra as they offer a wide band view of the data in the frequency domain [DC-30Hz here] and additionally, statistically quantify the increase/loss in power with respect to baseline. First, all the epochs for each condition across all subjects were pooled together. Within this pool, epochs corresponding to specific electrodes within each ROI were then extracted and grouped together and the data were resampled. EEG power spectrum (DC-30Hz) was then computed per epoch using a 512 data point (512ms as Fs = 1000Hz) Fast Fourier Transform (FFT) tapered with a moving Hanning window, 50% overlap. This window size was chosen to obtain adequate frequency resolution while at the same time corresponding to the window length in the time-voltage analyses (200ms). The first 500ms (-1000 -500ms) was considered as baseline and used to compute significant changes in power ($p < 0.01$, False Detection Rate corrected (Genovese et al., 2002)). The baselined power spectra were averaged across all subjects resulting in grand averaged EEG ERSP for each of the 12 conditions, across all 5 ROI. We focused on the effect of the manipulative hand posture against the no hand condition within a context, as was done in the time-voltage analyses. To this end, ERSP of the no hand condition was statistically ($p < 0.001$, uncorrected) compared to the manipulative posture. Similar to the time-voltage analyses, the hand posture conditions were not directly compared to each other (such as functional vs. manipulative). Analyses were primarily restricted to the comparison of the manipulative posture with respect to the no hand control condition within the correct and spatial contexts. The ROI used in these contrasts were derived from significant findings in

the time-voltage statistics and are described in detail in the subsequent results section.

3.4 Results

3.4.1 Time-voltage analyses

For each time window (early, middle and late), a separate 3-way ANOVA was computed with the following factors: Hand (four levels: No hand, Static hand, functional and manipulative), Context (three levels: correct, incorrect and spatial) and Region (five levels: Left and Right Frontal, Left Motor, Left and Right Parieto-temporal areas). The 3-way ANOVA computed on the early phase, [100 - 200ms] after picture presentation, revealed a three way interaction of Context x Hand x Region [$F(24,900) = 2.26$, $p < 0.05$], with a main effect of Region [$F(4,900) = 551.96$, $p < 0.05$] and a two way interaction of Context x Hand [$F(6,900) = 8.15$, $p < 0.05$]. The middle phase [200 - 400ms] revealed a significant three way interaction, Context x Hand x Region [$F(24,900) = 1.54$, $p < 0.05$], with a main effect of Region [$F(4,900) = 884.88$, $p < 0.05$] and a significant two way interaction: Context x Hand [$F(6,900) = 11.8$, $p < 0.05$]. Analysis on the late phase [400 - 600ms] revealed a significant three way interaction of Context x Hand x Region [$F(24,900) = 1.85$, $p < 0.05$] with a main effects of Region [$F(4,900) = 707.92$, $p < 0.05$] and Hand [$F(3,900) = 24.05$, $p<0.05$] and the following significant two way interactions: Context x Hand [$F(6,900) = 3.86$, $p < 0.05$], Hand x Region, [$F(12,900) = 4.27$, $p < 0.05$]. Given the three way interactions with subsequent main effect of Region and two way interaction of Context x Hand, post-hoc t-tests enabled us to evaluate the contrasts of the various hand conditions (manipulative, functional and static hand) with respect to the no hand control condition, within a specific context and over individual ROIs across all three time windows. The results of these ROI analyses are presented in the following sections and Figures 14-16 depict the results. Additionally, 3-way ANOVAs on all conditions confirmed ($p<0.05$) that i) EEG traces did not differ in the baseline period, ii) EEG waveforms were not divergent at stimulus presentation and iii) convergence of EEG waveform traces of all conditions back to baseline 1.5 seconds post stimulus.

3.4.2 Left frontal area

3.4.2.1 Early phase

Since we were primarily concerned with the effect of Hand, we conducted two tailed t-tests within each context, comparing no hand to other hand conditions. Figure 14A illustrates the results of the left frontal area. When context was correct, the no hand posture differed significantly from static hand [$t(30) = 6.91$, $p < 0.01$], functional [$t(30) = 3.31$, $p < 0.01$] and manipulative [$t(30) = 3.52$, $p < 0.01$] hand postures. When context was incorrect, only manipulative differed from no hand, [$t(30) = 3.97$, $p < 0.01$]. When context was spatial, there were no statistical differences among the hand conditions.

3.4.2.2 Middle phase

In the middle phase, some of the early effects were maintained; two tailed t-tests show that when context was correct, no hand differed significantly from static hand [$t(30) = 3.11$, $p < 0.01$] and manipulative [$t(30) = 2.95$, $p < 0.01$] postures. Within the incorrect context, only the functional grasp significantly differed from no hand [$t(30) = -2.95$, $p < 0.01$]. Whereas when context was spatial, there were no statistical differences among the hand conditions.

3.4.2.3 Late phase

Finally, in the late phase, two tailed t-tests show that when context was correct, there was no statistical difference between hand conditions. When context was incorrect, no hand differed significantly from functional [$t(30) = -5.53$, $p < 0.01$] and manipulative [$t(30) = -5.68$, $p < 0.01$] postures. Uniquely, when context was spatial, no hand differed from the manipulative posture [$t(30) = -3.82$, $p < 0.01$].

3.4.3 Left Parieto-temporal Area

3.4.3.1 Early phase

Figure 14B illustrates the results of the Left Parieto-Temporal area. When context was correct, no hand differed significantly from static hand [$t(30) = 4.21$, $p < 0.01$], functional [$t(30) = 4.68$, $p < 0.01$] and manipulative [$t(30) = 5.19$, $p < 0.01$] postures. When context was incorrect, no hand differed significantly from static hand [$t(30] = 3.84$, $p < 0.01$],

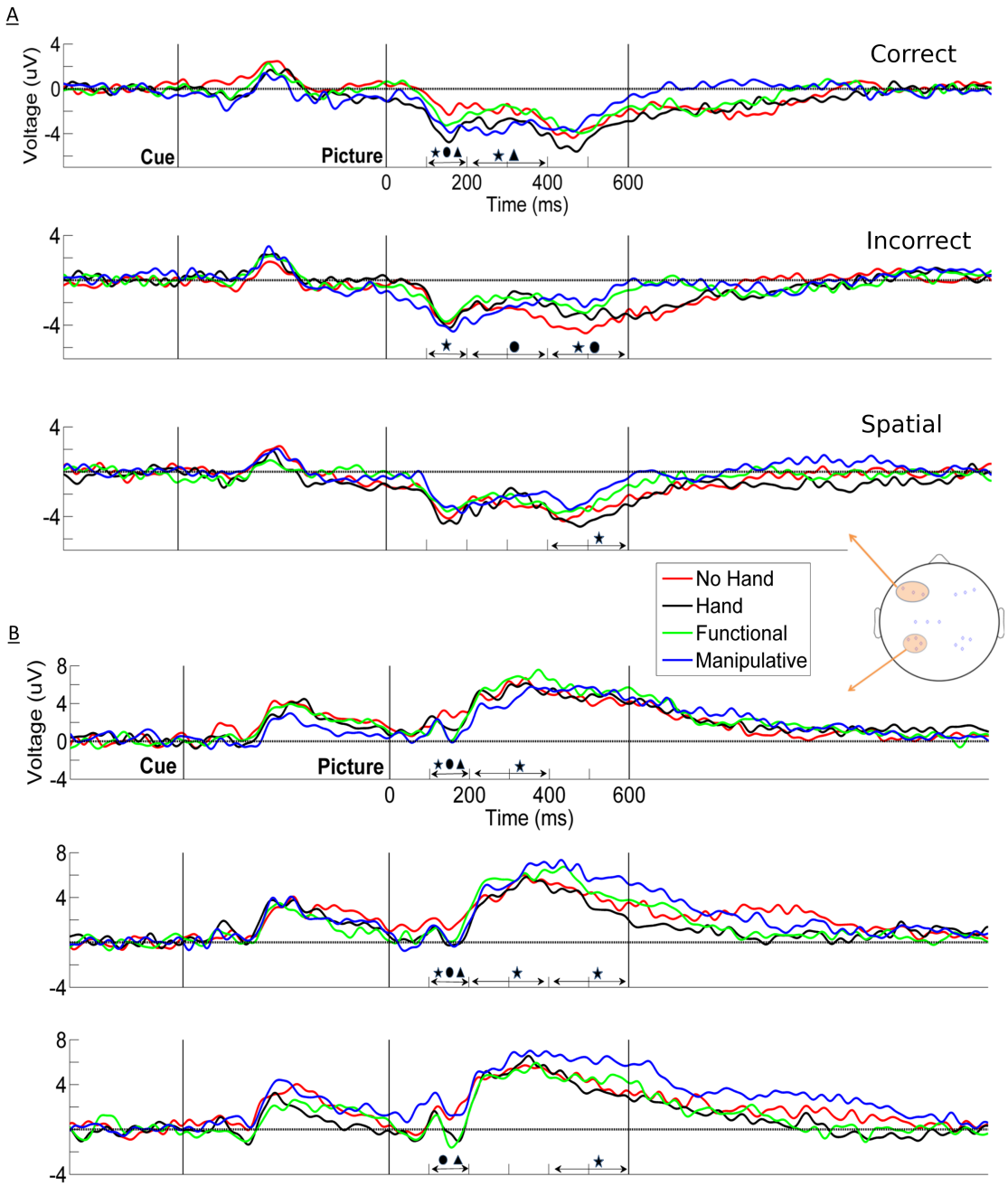

Figure 14: Time-Voltage Activations over the A) left frontal and B) left parietotemporal areas. Time windows of interest are from [100-200ms], [200-400ms], [400-600ms] post stimulus presentation. All statistical comparisons are made by comparing static hand, functional and manipulative hand postures with respect to the no hand condition at the $p<0.01$ level. These comparisons are performed for each tool object context along each row, with the following markers of significance in a particular time window of interest. ★ : manipulative significant, ● : functional significant and ▲ : static hand significant.

functional [t(30) = 2.81, p < 0.01] and manipulative [t(30) = 3.31, p < 0.01] postures. When context was spatial, no hand differed significantly from static hand [t(30) = 3.59, p < 0.01] and functional posture [t(30) = 6.19, p < 0.01].

3.4.3.2 Middle phase

When context was correct, no hand differed significantly from manipulative posture [t(30) = 3.39, p < 0.01]. When context was incorrect, no hand differed significantly from manipulative [t(30) = -2.91, p < 0.01] posture. When context was spatial, no hand did not differ from any other conditions.

3.4.3.3 Late phase

When context was correct, there were no statistical differences. When context was incorrect, no hand differed significantly from manipulative posture [t(30) = -4.99, p < 0.01]. When context was spatial, no hand differed significantly only from the manipulative posture [t(30) = -6.8, p < 0.01].

3.4.4 Right frontal area

3.4.4.1 Early phase

Figure 15A illustrates the results of the right frontal area. When context was correct, the no hand condition differed significantly from static hand posture [t(30) = 4.99, p < 0.01]. When context was incorrect, no hand differed significantly from manipulative posture [t(30) = 3.32, p < 0.01]. When context was spatial, there was no statistical difference among the hand conditions.

3.4.4.2 Middle phase

There were no statistical differences in the middle phase across all hand postures with respect to no hand condition.

3.4.4.3 Late phase

Over the right frontal area, two tailed t-tests show that when context was correct, no hand differed significantly only from the manipulative posture [t(30) = -6.06, p < 0.01], evidence

that the manipulative posture could drive extended neural processing . When context was incorrect, no hand differed significantly from functional [t(30) = -5.08, $p < 0.01$] and manipulative postures [t(30) = -9.14, $p < 0.01$]. When context was spatial, no hand differed from the manipulative posture [t(30) = -6.16, $p < 0.01$].

3.4.5 Right Parieto-temporal Area

3.4.5.1 Early phase

Figure 15B illustrates the results of the Right Parieto-Temporal area. When context was correct, no hand differed significantly from static hand [t(30] = 3.61, $p < 0.01$], functional [t(30] = 4.65, $p < 0.01$], and manipulative [t(30) = 5.08, $p < 0.01$] postures. When context was incorrect, no hand differed significantly from static hand [t(30] = 5.22, $p < 0.01$], functional [t(30) = 5.02, $p < 0.01$] and manipulative [t(30) = 3.65, $p < 0.01$] postures. When context was spatial, no hand differed significantly from static hand [t(30) = 3.06, $p < 0.01$] and functional posture [t(30) = 4.52, $p < 0.01$].

3.4.5.2 Middle phase

When context was correct, no hand differed significantly only from the manipulative posture [t(30) = 3.33, $p < 0.01$]. When context was incorrect or spatial, no hand did not differ significantly from any other hand condition.

3.4.5.3 Late phase

When context was correct, no hand differed significantly from manipulative postures [t(30) = -2.87, $p < 0.01$]. Together with the Right Frontal, the Right Parieto-temporal area is sensitive to the manipulative grasp, with extended neural activations. When context was incorrect, no hand differed significantly from manipulative posture [t(30) = -4.67, $p < 0.01$]. When context was spatial, no hand differed significantly only from the manipulative posture [t(30) = -6.35, $p < 0.01$].

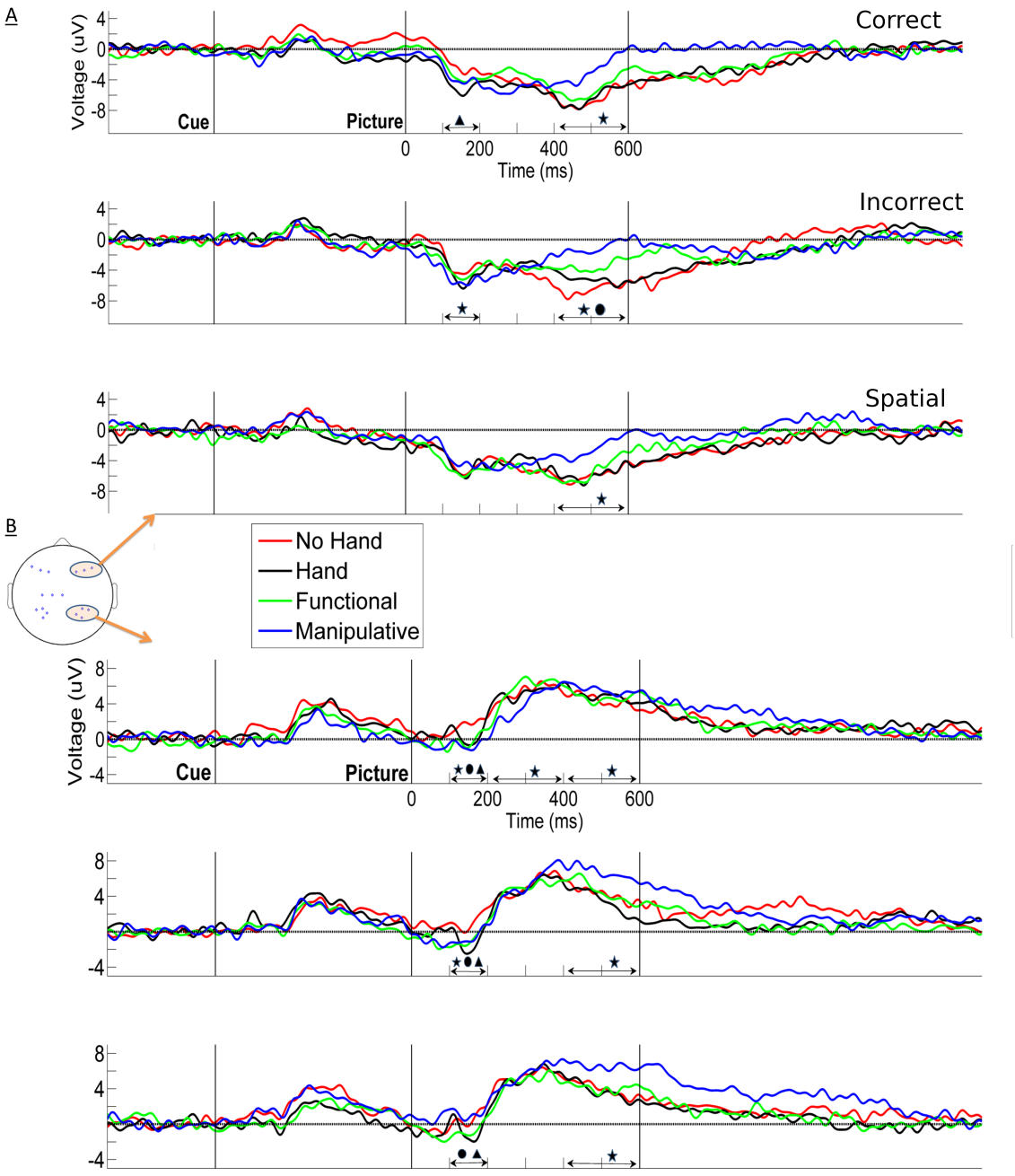

Figure 15: Time-Voltage Activations over the A) right frontal and B) right parietotemporal areas. Time windows of interest are from [100-200ms], [200-400ms], [400-600ms] post stimulus presentation. All statistical comparisons are made by comparing static hand, functional and manipulative hand postures with respect to the no hand condition at the $p<0.01$ level. These comparisons are performed for each tool object context along each row, with the following markers of significance in a particular time window of interest. ★ : manipulative significant, ● : functional significant and ▲ : static hand significant.

3.4.6 Left motor areas

3.4.6.1 Early phase

Figure 16 illustrates the results of the left motor area. When context was correct, no hand differed significantly from static hand [t(30) = 3.04, p < 0.01], functional [t(30) = 3.64, p < 0.01] and manipulative postures [t(30) = 4.06, p < 0.01]. When context was incorrect, no hand differed significantly from static hand [t(30) = 4.09, p < 0.01], functional [t(30) = 3.34, p < 0.01] and manipulative postures, [t(30) = 4.05, p < 0.01]. When context was spatial, no hand differed significantly from the functional posture alone [t(30) = 3.53, p < 0.01].

3.4.6.2 Middle phase

When context was correct, no hand differed significantly from manipulative posture [t(30) = 2.97, p < 0.01]. There were no differences when context was incorrect or spatial.

3.4.6.3 Late phase

Focusing on the late phase, when context was correct, no hand did not differ significantly from any other hand condition. When context was incorrect, no hand differed significantly from functional [t(30) = -4.18, p < 0.01] and manipulative, [t(30) = -4.84 p < 0.01] postures. When context was spatial, no hand differed significantly only from the manipulative posture [t(30) = -6.43, p < 0.01] advancing insight into the neural processing of the manipulative grasp and the spatial context.

3.4.7 Time-frequency results

The time varying ERSP of the EEG signal [DC 30Hz] was computed to investigate outcomes in frequency power. Only ROIs with significant differences in the time-voltage analyses that were relevant to our hypotheses were analyzed. Using ERSP plots, we were able to identify distinct changes in power within unique bands. Statistical differences in the ERSP for the correct and spatial contexts when manipulative was compared to the no hand condition are shown in Figures 17 and 18.

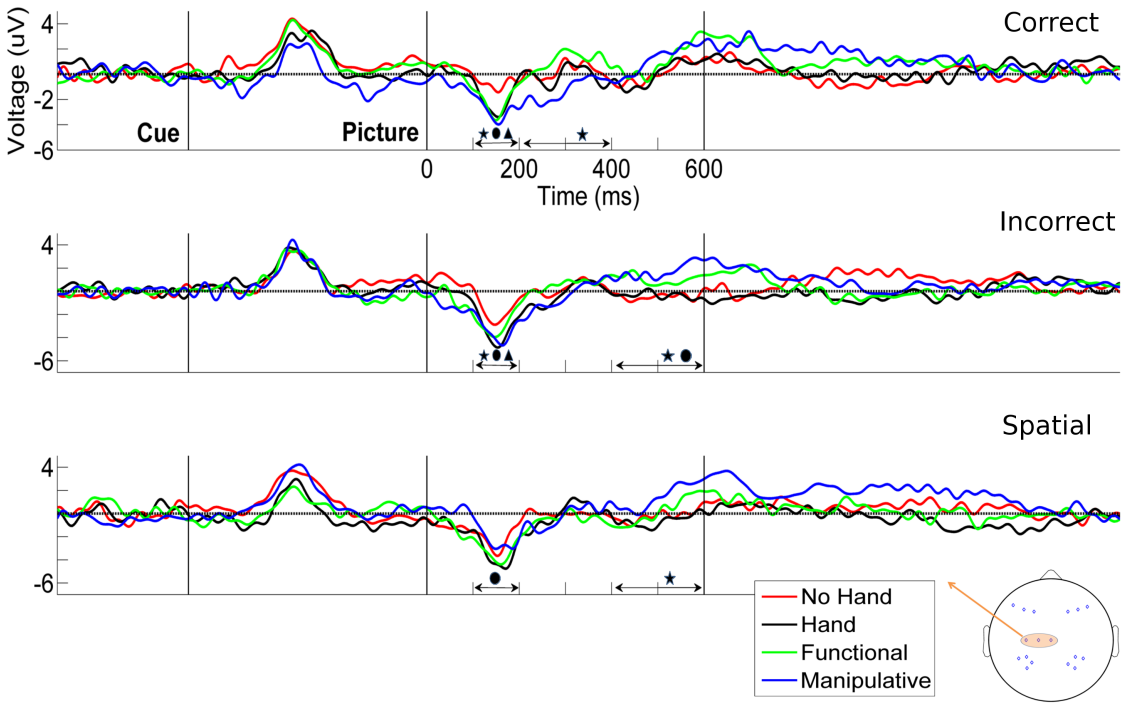

Figure 16: Left motor time-voltage results. Time-Voltage Activations over the left motor electrodes. Time windows of interest are from [100-200ms], [200-400ms], [400-600ms] post stimulus presentation. All statistical comparisons are made by comparing static hand, functional and manipulative hand postures with respect to the no hand condition at the $p<0.01$ level. These comparisons are performed for each tool object context along each row, with the following markers of significance in a particular time window of interest. ★ : manipulative significant, ● : functional significant and ▲ : static hand significant.

3.4.7.1 Correct context

Over the right frontal area, when context was correct, the manipulative posture had significantly more positive differences in power ($p < 0.001$, uncorrected) when compared to the no hand condition, specific to theta and low alpha frequency bands [4-10Hz]. This modulation emerges early and sustains throughout all time-windows of interest as can be seen in Figure 17A. Over the right parietal area, manipulative posture exhibited both positive and negative differentials. In the beta band [15-20Hz], no hand condition had more negative ERSP than manipulative; therefore (manipulative no hand) resulted in a positive differential (see figure caption for detailed explanation). In alpha band [8-10Hz], manipulative exhibited a larger negative ERSP (ERD) than no hand, giving rise to a negative differential. These modulations were present only in the late window ($p<0.001$, [400-600ms]) and the statistical

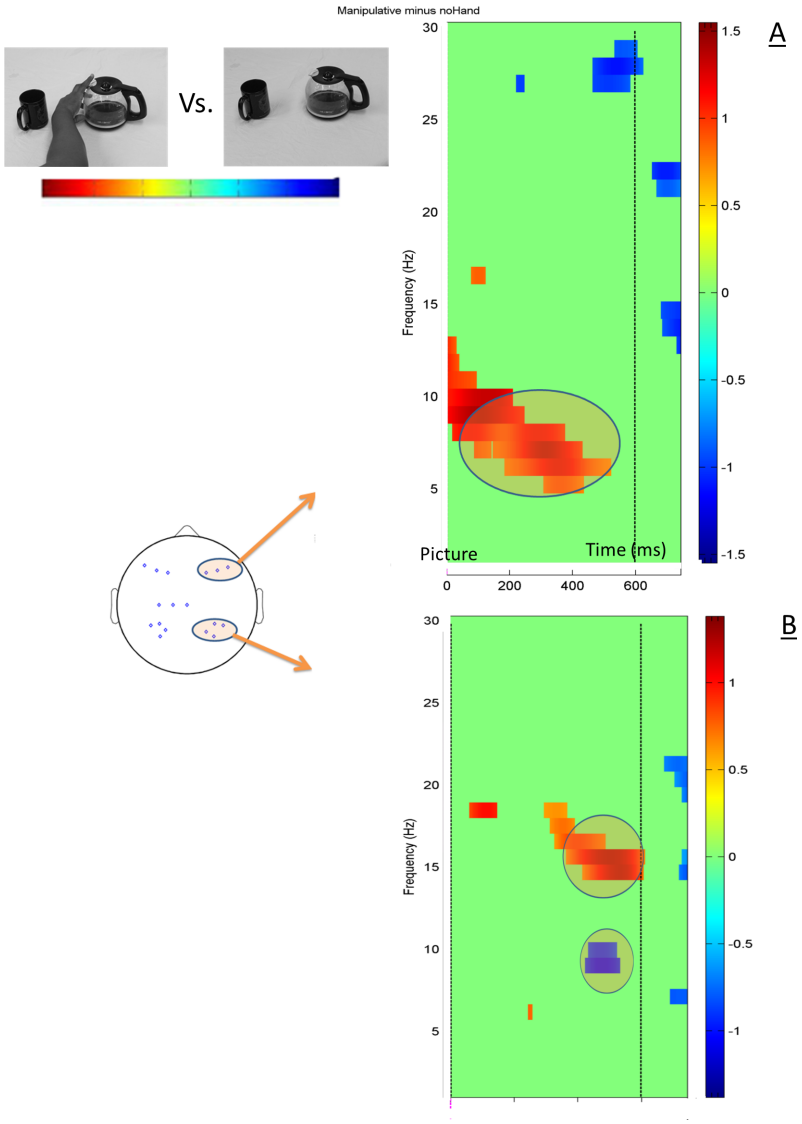

Figure 17: Differential Event related spectral perturbation statistical comparisons at the $p<0.001$ level (uncorrected) between the manipulative and no hand conditions (when context is correct) in the time-frequency domain. Red signifies that the manipulative condition has significantly more positive power. Blue signifies that no hand condition had significantly more positive power, as measured along the entire real line $[-\infty$ to $-\infty$. For example -2 is more positive than -4, while 4 is more positive than 2]. A) Spectrogram outlining the differences in frequency power over the right frontal area, from [0-600ms] post stimulus presentation. B) Spectrogram outlining the differences in frequency power over the right parietotemporal area from [0-600ms] post stimulus presentation.

differences are shown in Figure 17B.

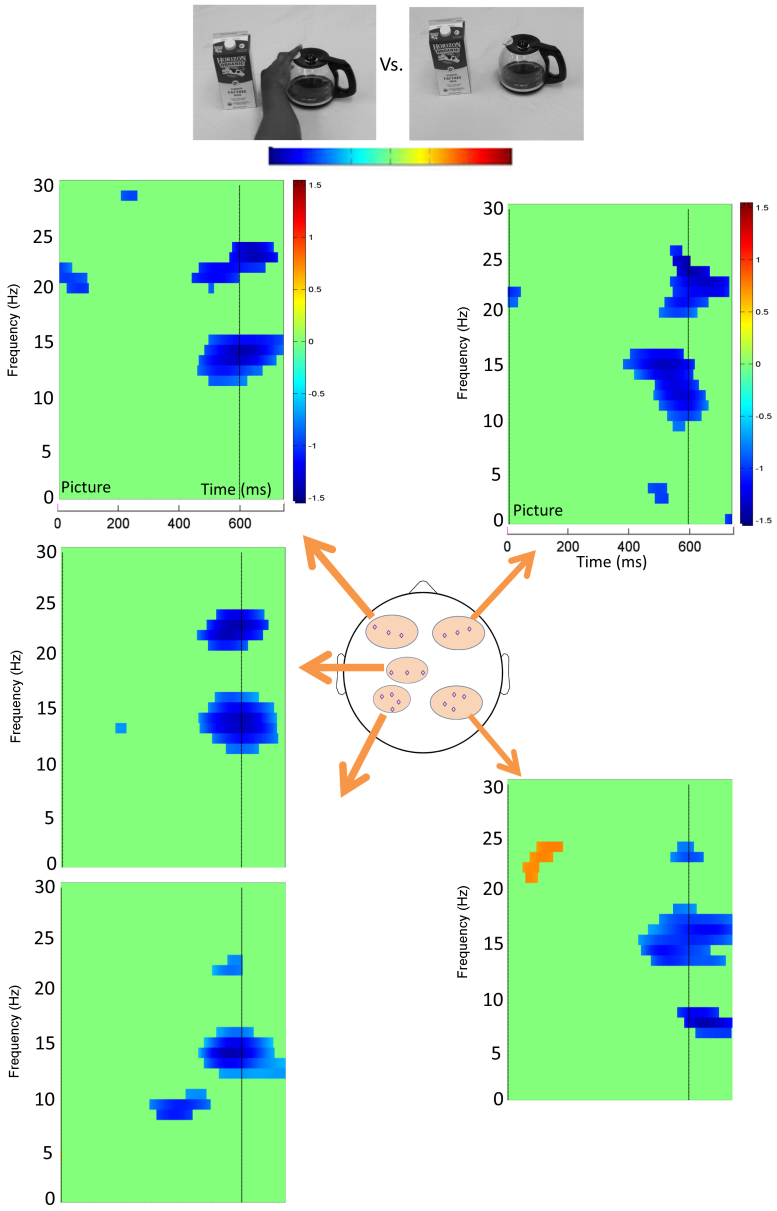

Figure 18: Differential Event related spectral perturbation statistical comparisons at the p<0.001 level (uncorrected) between the manipulative and no hand conditions (when context is spatial) in the time-frequency domain from [0-600ms] post stimulus presentation. Blue signifies that manipulative condition had significantly more negative power, as measured along the entire real line [$-\infty$ to ∞. For example, -4 is more negative than -2]. Red signifies that the no hand condition has significantly more negative power. Spectrograms outline the differences in power over bilateral parietotemporal and left motor areas.

3.4.7.2 Spatial context

The manipulative posture alone engages bilateral parietofrontal networks and the left motor area, with significantly greater beta band ERD than no hand [12-16Hz, 20-25Hz], resulting

in an overall negative differential. These power differences are present only in the late phase ($p<0.001$, [400-600ms], uncorrected) and are shown in Figure 18.

3.5 Discussion

The overarching goal of Aim 1 was to probe the neural correlates underlying results from our prior behavioral data (Borghi et al., 2012). We aimed to understand the spatiotemporal patterns of parietofrontal activations in response to viewing contextual and grasp specific stimuli. Specifically, we hypothesized that compared to the no hand control, the manipulative grasp condition would uniquely elicit prolonged activity over left parietofrontal regions known to be involved in tool-use understanding with corresponding beta band spectral differences. We hypothesized these activations to be maximal in the correct tool-object context that had maximal action affordances. Given the diffuse nature of the the EEG recordings at the scalp, temporal electrodes were also included in the parietal electrode pool. Results here did in part support our hypotheses; unlike the static hand and functional grasp conditions, the manipulative grasp uniquely elicited robust and prolonged activations when compared to the no hand control. The static hand and functional grasp conditions only elicited early differences with respect to the no hand control conditions (up to 400ms post image onset). However, the laterality of the parieto-temporal-frontal differences were not in line with our hypotheses, as we observed a late appearing right parietofrontal difference between the manipulative grasp and no hand conditions in the correct context, and a late appearing, bilateral parietofrontal difference within the spatial context. It should be noted that the difference between the two conditions was present in the incorrect tool-object context as well. These results offer a neural explanation of the behavioral outcomes seen in our prior work, and introduce potential neural mechanisms outside of the left parietofrontal network commonly engaged in deciphering tools and tool-object based action (Chao and Martin, 2000, Johnson-Frey et al., 2005, Krliczak and Frey, 2009).

With respect to the spatiotemporal patterns in the parietofrontal differences between the manipulative grasp condition and the no hand control within the correct context, ERP

results showed early left parietofrontal-motor and right parietal differences followed by an exclusively late appearing, right parietofrontal difference. In addition, time-frequency results contrasting the two conditions primarily showed significant right frontal theta ERS for the manipulative grasp condition. Therefore, it appears that there are two networks at play: an early left parietofrontal-motor and right parietal network to process the familiar, correct tool-object context and a unique, late appearing right parietofrontal network to process the manipulative grasp-posture with increased oscillatory power in the theta band over the right frontal area. It is likely that the right parietofrontal activity could correspond to the mirror neuron system (MNS), a set of specific parietofrontal circuits that are active during both action observation and action execution (Iacoboni et al., 2005). Rather than process the motoric properties of tools and the prehensile properties of tool-use gestures, research has shown these neurons to process the goal or intent of the gesture via an internal motor simulation of the gesture itself (Iacoboni et al., 2005, Iacoboni and Dapretto, 2006). While many studies have shown the MNS to be left hemispheric dominant, studies have also suggested right hemispheric involvement (Aziz-Zadeh et al., 2006). Of particular interest is the fMRI experiment by Iacoboni et al. who demonstrated that the posterior right inferior frontal gyrus was significantly active in deciphering the goal or intent of a tool-use gesture when subjects differentiated the same grasping gesture embedded in two different contexts (pick up a cup to drink from a clean table vs. pick up a cup to clean from a dirty table) (Iacoboni et al., 2005). Therefore the right frontal activity here might be indicative of the MNS in deciphering the intent of the manipulative grasp. Since the functional grasp is already familiar, it may not have elicited extended right frontal activity and only elicited early left parietofrontal differences with respect to the no hand condition. With respect to the frequency oscillations over the right frontal area, theta synchronizations have been implicated in having a role in the MNS via Hebbian learning mechanisms (Del Giudice et al., 2009) and may be in involved in motor imagery of the grasp (Caplan et al., 2003). Theta frequency oscillations may demonstrate a mechanism whereby specific frequency oscillations serve to bind action perception with action memories.

With respect to the spatial tool-object context, the manipulative posture alone was characterized by extended (around 600ms) bilateral parietofrontal and left motor differences in time-voltage and time-frequency domains, while the functional grasp and static hand elicited only early bilateral parietal, left motor and bilateral parietal activity respectively. Time-frequency results showed significant bilateral alpha and beta ERD for the manipulative grasp condition with respect to the no hand control condition. This result was contrary to out hypotheses where we had posited that parietofrontal differences would be restricted to the correct context alone where action affordance was highest. However, there were qualitative differences in the manner the manipulative grasp-posture was processed within the spatial tool-object context. Specifically, the differences between the manipulative grasp and the no hand condition surfaced only in the late phase (400-600ms post image), while being similar until 400ms. We theorize this delay to be attributed to the ambiguous nature of the tool-object relationship where action affordances are not immediately apparent, with or without the manipulative posture that ostensibly does not affect early processing of affordances. At the same time, the differences were bilateral and not just restricted to the right hemisphere. The late bilateral parietofrontal and left motor differences in both time-voltage and time-frequency domains may underlie two networks at play in evaluating the image: a right parietofrontal network to comprehend the unfamiliar hand action (presumably right mirror neuron networks) and a left parietofrontal, motor network to decode possible action in the spatial context. In the frequency domain, increased ERD was present for the manipulative posture vs. no hand over the aforementioned areas dominant within two frequency bins [12-17Hz and 20-25Hz]. Use of bilateral beta band parietofrontal networks continually in the late phase would permit for ongoing evaluation of the intent of the action scene and perhaps ongoing activation related to the content of the scene itself (Chao et al., 2002, Mizelle et al., 2011), to determine if action is possible between the tools and objects as discrete entities.

With respect to the incorrect context, parietofrontal differences did emerge between the manipulative grasp and no hand conditions over bilateral parietal and left motor areas. The late phase (400-600ms) was also characterized by left motor and bilateral frontal differences

between the no hand and functional grasp-posture. Thus even in the incorrect context, the manipulative grasp elicited significant differences with respect to the no hand condition. However, the neural encoding between the two conditions was not as clear as the correct and spatial tool-object contexts given that the functional grasp condition also differed from the no hand conditions. While this result is outside the scope of our hypotheses, they provide evidence that by and large, the interference effect of the manipulative grasp-posture is observed across all three contexts. Importantly, it should be noted that it is not the mere presence of a hand per se that elicited these differences, as similar parietofrontal differences were not observed in the static hand and functional grasp conditions.

3.6 Conclusion

The results of Aim 1 revealed that when passively evaluating tool-object content, the manipulative grasp elicited temporally extended parietofrontal activations when compared to the no hand control condition within all three contexts. In addition, the time scales of these activations were in line with the prior behavioral data. Within the correct context, the manipulative grasp elicited extended responses over right parietofrontal areas with significant power increases in the theta band [4-8Hz] over the right frontal area. Within the spatial context, the manipulative grasp elicited significant bilateral parietofrontal and left motor activations with significant power decreases in the beta band [12-16Hz, 20-25Hz]. The static hand and functional grasp conditions did not elicit such temporally extended and reliable parietofrontal activations in all three contexts. Results from Aim 1 suggest that though the grasp per se is not essential to the task, the manipulative grasp automatically primes action encoding regions to decode grasp intent by various mechanisms that appears to be sensitive to the tool-object context itself. These phenomena can therefore possibly underlie previously observed delayed responses with the same stimuli (Borghi et al., 2012). Given that action understanding and gaze control overlap over the same parietofrontal networks, in the next chapter we investigate whether gaze patterns are also sensitive to the interference effects of the manipulative grasp and follow results from Aim 1.

CHAPTER IV

SPECIFIC AIM 2

4.1 Introduction

The results from Aim 1 showed differential parietofrontal involvement when participants were processing the manipulative grasp-posture as compared to the no hand condition (Natraj et al., 2013). From a behavioral perspective, the manipulative grasp-posture also elicited longer decision times on tool-object content when compared to the no hand control (Borghi et al., 2012). This interference effect was observed in all three contexts with the most robust effect within the correct and spatial tool-object contexts. Results from Aim 1 therefore suggest that parietofrontal action encoding regions were automatically primed to process the intent of the manipulative grasp though the grasp per se was not essential to evaluating tool-object content. However, given the anatomical overlap between action encoding and gaze control over the same parietofrontal regions, it is likely that eye movements may also be influenced by the affording properties of tool-use scenes. Understanding how the observer parses or visually encodes complex tool-use scenes in real-time can shed light on the type of action information that may potentially be driving action encoding regions. This issue is especially important as the visual encoding of tool-use affordances is still poorly studied in the literature. Indeed, an important premise in action encoding theory is that affordances influence how an observer gathers the visual information to activate parietofrontal action encoding regions (Humphreys et al., 2010, Yoon et al., 2010). However, the link between tool-use affordances and visual attention has typically been studied in an indirect manner by correlating latencies in neural and behavioral responses to different types of tool-use scenes, a method we have also used in our prior work and in Aim 1 (Borghi et al., 2012, Natraj et al., 2013). While such an approach sheds light on how affordances influence the global visual processing of tool-use scenes, it does not allow making inferences on the complete relationship between tool-use affordances and visuospatial

attention, i.e., how affordances influence the manner in which an observer spatiotemporally parses the discrete features of the tool-use image. For instance, it is unclear if there could a difference in spatiotemporal gaze position when an observer evaluates tool-object scenes with and without the manipulative grasp-posture. Such a difference can help identify the type of visual information that may be driving the neural activations and interference effect observed in our previous studies with the same stimuli (Borghi et al., 2012, Natraj et al., 2013). Understanding how an observer visually parses tool-object scenes in real-time and under free viewing conditions can provide a valuable window into ongoing action encoding processes.

To address this aforementioned issue, we used eye tracking to evaluate the relationship between tool-use affordances and gaze control. Eye tracking has the advantage of directly evaluating the influence of affordances on the temporal allocation of visuospatial attention. We used the same stimuli as Aim 1 and the prior behavioral work. However, the static hand conditions across all three contexts were not included in this study as the results from the prior behavioral data and Aim 1 showed that they were processed largely similar to the no hand conditions (Borghi et al., 2012, Natraj et al., 2013). Our primary goal in this study was to assess how the combination of tool-object context and tool-grasp influenced the control of eye movements as participants passively evaluated the appropriateness of tool-object relationships, with emphasis on the visual encoding of the interference effect of the manipulative grasp. We focused on two key aspects of gaze behavior: gaze scanpaths after saccade initiation over the areas of interest in the stimuli (AOI), and the relative weighting of the AOI.

4.2 Hypotheses

We posited that clustering of gaze scanpaths and AOI weightings would group conditions primarily by context (correct, incorrect, spatial) since the task required participants to passively evaluate tool-object content. Within each of the 3 context clusters, we hypothesized that the manipulative grasp-posture would be distinct from the no hand and functional grasp-posture with the strongest distinctions or between-condition separation in the correct

and spatial contexts, in line with results from Aim 1 (Natraj et al., 2013).

4.3 Methods

4.3.1 Participants

Eight subjects (4 male, 4 female, mean age = 23.1 years, S.D. = 1.7) participated in the study after giving their informed consent as per the Georgia Tech Institutional Review Board (IRB). Subjects were right handed based on the Edinburgh Handedness Inventory (Oldfield, 1971). All participants had normal or corrected normal vision and self-reported not having any prior neurological problems.

4.3.2 Stimuli and Experimental design

We used stimuli consisting of tools and objects identical to Aim 1 and the prior behavioral data (Borghi et al., 2012, Natraj et al., 2013). While we included all 12 conditions in the experimental design for consistency with our prior neurobehavioral work, we did not include the static hand conditions in our hypotheses and analyses as they had evoked largely similar neurobehavioral responses as the no hand conditions in our prior data (Borghi et al., 2012, Natraj et al., 2013). In addition, this allowed us to contrast the main effect of an actual grasp-posture (either manipulative or functional) directly with the control no hand condition across tool-object contexts, similar to our prior work with the same stimuli (Natraj et al., 2013). This reduced the number of experimental conditions in this study down to 9. Figure 19 details the matrix of the 9 conditions for a single representative tool. Each of the 8 subjects viewed a total of 144 unique stimuli across the 12 conditions (including the 3 static hand conditions). To avoid the effect of repetition and familiarity on eye movements, subjects viewed each stimulus only once. In total, the dataset consisted of 864 experimental trials across all subjects and the 9 experimental conditions. The sequence of images was randomized and each stimulus had an on-screen duration of 3500ms. Stimuli were separated from each other by a fixation circle presented at the center of the screen that had a random duration of 2500-3500ms. The randomized stimuli were presented across 4 blocks to provide adequate rest intervals for the subject. Participants were instructed to

fixate at the fixation circle and not move their eyes prior to image presentation. The task required participants to passively evaluate (without any behavioral response (Mizelle et al., 2013, Natraj et al., 2013)) the appropriateness of tool-use, i.e., if the relationship between the tool and object was correct or incorrect.

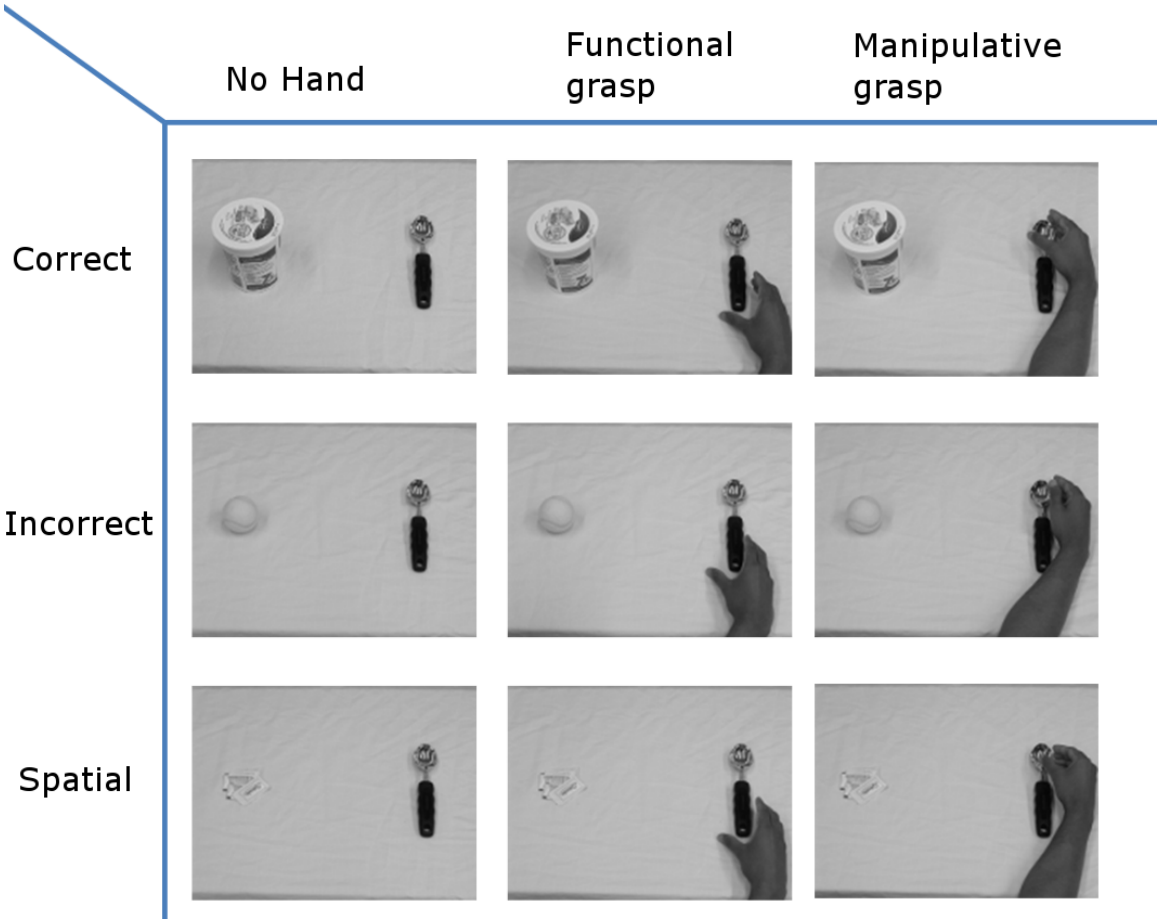

Figure 19: The stimuli were similar to Aim 1 and the prior behavioral data with the exception that the static hand conditions were not included in our hypotheses

4.3.3 Data collection

Subjects were fitted with an iViewX (SensoMotoric Instruments, Boston, MA, USA) head mounted eye tracker hat from the Center for Human Movement studies, Georgia Tech. It consisted of an infrared pupil detector and a scene camera both fixed on the lid of the hat. The scene camera captured a 2D representation of the environment with a resolution of 752X480 pixels. The system recorded gaze data at 50Hz, with an overlay of gaze position

on the scene at a rate of 40 frames per second (Figure 20A), had a tracking resolution ¡0.01 degrees, had a gaze position accuracy between 0.5 - 1 degree and generated an ascii file of the time course of gaze position mapped onto the video captured by the 752X480 resolution scene camera. The eye tracker was connected to an iViewX laptop via USB, allowing for real time visualization and data acquisition using iViewX software. Stimuli were presented from a separate laptop that was custom programmed in MATLAB (The MathWorks Inc., Natick, MA, USA) to pass digital pulses to the iViewX laptop via the serial port. This allowed identifying temporal events in recorded gaze data, such as the appearance of the fixation circle and subsequent start and end times of stimulus presentation. A comfortable chin rest was used to adjust head position and minimize head movements during data acquisition. All the stimuli in the study were presented on a 106.7cm (42") LCD screen that was approximately at the center of a participant's field of view with a resolution of 1080X1920 pixels. To enable gaze tracking, the eye tracker was calibrated by having participants foveate at the edges and center of the 42" LCD TV screen and the coordinates of these spots were passed to calibration software on the iViewX laptop. Subjects' gaze perspective at the chin rest was approximately 170cm away the LCD screen. The total viewing angle was 30.6 degrees in the horizontal direction (left-right) and 17.5 degrees in the vertical direction (up-down). The average horizontal separation between the centers of the tool-object pair was 18.9 degrees (1.3 degrees S.D) and therefore either the tool or object was on average 9.45 degrees from the screen center.

4.3.4 Preprocessing

Each condition had 3 areas of interest (AOI). These were the object and the two ends of the tool that were defined based on the type of grasp-posture associated with it. The two ends were either the functional tool-end (spoon-handle) or the manipulative tool-end (spoon-bowl). While the manipulative and functional grasp conditions also included the portion of the arm below the wrist that was not interacting with the tool, in ¿95% of trials, the arm was not attended to (a similar result was observed in (Land et al., 1999)). The boundaries of the AOI were extracted by applying Sobel's edge detection algorithm to the

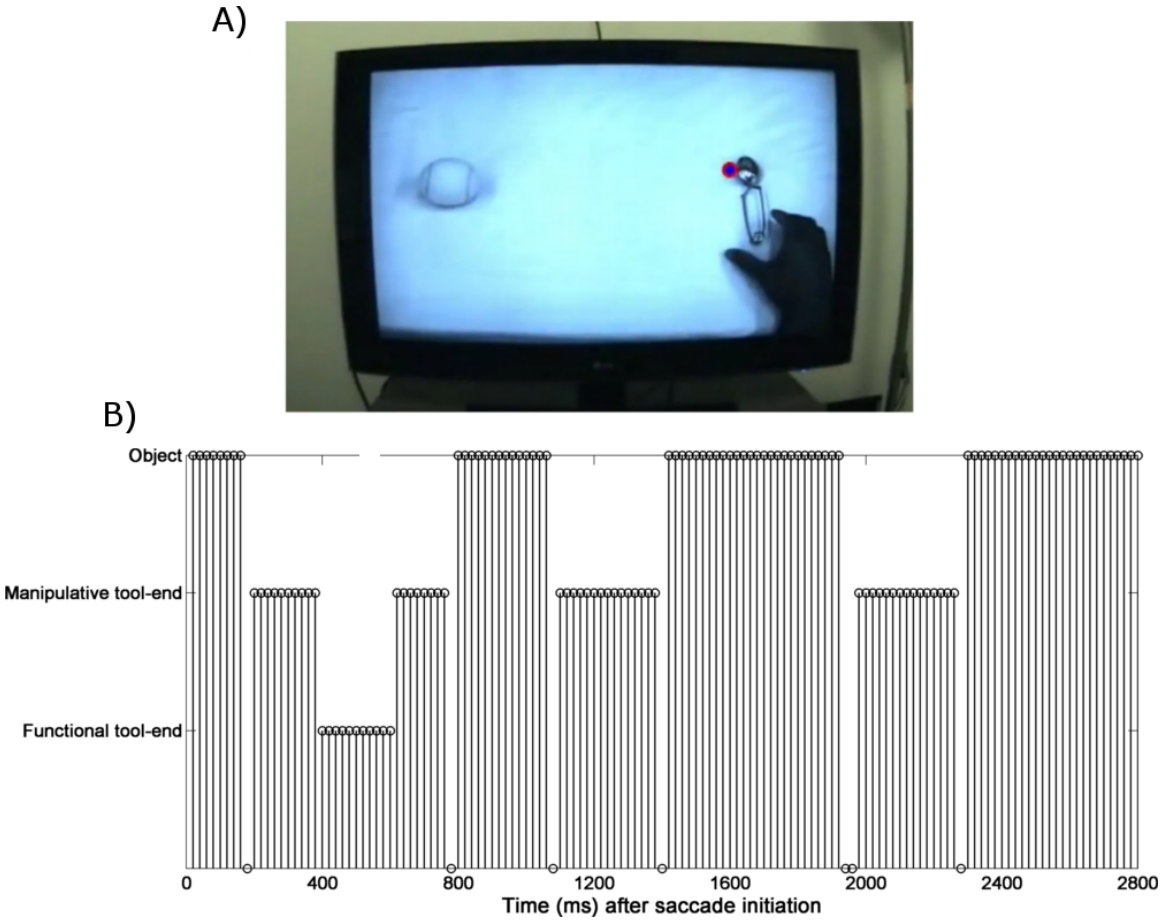

Figure 20: A) A frame from the video of scene-gaze overlay generated by the eye tracker. B) Example of foveal gaze scanpath in a trial.

scene video-gaze overlay (Gonzalez et al., 2009) and by fitting a rectangle to the detected boundary. The average width of tools and objects was 4.5 degrees (2.04 degrees S.D) and the average height of tools and objects was 5.6 degrees (2.3 degrees S.D). With respect to the demarcation of the functional and manipulative tool-end, on average 34.2% (6.1% S.D) of the length of the tool from the top constituted the length of the manipulative tool-end and the bottom 65.8% (6.1% S.D) of the tool constituted the length of the functional tool-end. The width of the functional tool-end or manipulative tool-end increased in the presence of a functional or manipulative grasp-posture to accommodate the fingers and upper-wrist that were positioned next to the tool. It should be noted that the exact position of the grasp-posture was not controlled for across images and conformed more to the tool shape and size. Slack was given to the AOI to account for jitter in gaze position over the

AOI and noise in the eye tracking system in estimating gaze position (Ambrosini et al., 2011). The distances between tool-object centers in terms of viewing angle (average of 19deg) was large enough to warrant saccades in every trial given that visual acuity at the fovea drops exponentially (Kandel et al., 2000) and that eye movements in this experiment were free-viewing in nature. The time to initiate a saccade after image presentation was determined using a rolling window method i.e., whenever gaze position exceeded the mean of the previous 100ms of data by 2 S.D. A saccade was then identified if subsequent eye movement resulted in gaze position transitioning to an AOI with a single peak in gaze velocity profile (Haith et al., 2012). Saccade initiation times were additionally confirmed by analyzing the gaze data video overlay on a frame by frame basis (Ambrosini et al., 2011). The time course of eye movement data was set relative to the time point when a saccade was first initiated and entered an AOI post image presentation. Eye movement data were then smoothed using a second order Savitzky-Golay filter (Haith et al., 2012). The time-points in a trial therefore denoted the position of foveal gaze over the AOI after saccade initiation, i.e., foveal gaze scanpath (Figure 20B). After obtaining gaze scanpaths, the following trial exclusions rules were applied: a) time to initiate a saccade after stimulus presentation either exceeded 500ms or was less than 50ms, b) trials with excessive blinks and c) trials where a saccade was initiated before stimulus presentation.

4.3.5 Pattern recognition framework

Each of the 9 conditions had 3 AOI and images were on screen for 3500ms. The resultant eye movement data is inherently multidimensional given gaze position across all subjects, conditions, time and AOI. Many approaches in the literature exist for analyzing large multivariate spatiotemporal datasets (Raudenbush and Bryk, 2002, Cressie and Wikle, 2011). However, for this study, we took a simple pattern recognition based approach to address the hypothesis on spatiotemporal gaze position distributions and AOI weightings, explained in the following sections.

4.3.5.1 Clustering of gaze scanpaths

For every condition at the single subject level, the spatiotemporal probability of foveal gaze position once a saccade had been initiated was computed by counting the number of trials when foveal gaze position was within a particular AOI at a given time and then dividing this number by the total number of trials. Each condition therefore had 8 such sets (one for each subject) of individual AOI gaze position probabilities through time. The population level empirical distribution of mean AOI gaze position probabilities was computed using the bootstrap at each individual time point (using the *bootstrp()* function with 250 samples). Each of the 9 condition's gaze scanpath statistics were therefore represented by the time course of mean AOI gaze position probability with the bootstrap distribution around the mean at each time point. Subsequently, the statistical similarity between any two conditions a and b was computed by a version of the Mahalanobis distance metric (Bishop, 2006) given by the following equation:

$$d(a,b) = \sqrt{2\sum_{k=1}^{L}\sum_{t=1}^{N}\frac{(\mu_{k_a}^t - \mu_{k_b}^t)^2}{(\sigma_{k_a}^t)^2 + (\sigma_{k_b}^t)^2}} \quad (1)$$

In Equation (1), $\mu_{k_a}^t$ and $\sigma_{k_a}^t$ are the mean and standard deviation respectively of the population level distribution denoting probable gaze position at AOI k, at time t and in condition a. The distances between conditions therefore depended on both spatial and temporal similarities in mean gaze position weighted by the precision of the mean position. The above computation resulted in the formation of a symmetric distance matrix with 9 rows and 9 columns. An entry in any row i and column j corresponded to the similarity between spatiotemporal gaze position distributions of conditions i and j. Distributions that were more spatiotemporally similar to each other had a smaller Mahalanobis distance between them. The diagonal entries of the distance matrix were all zero as each condition was perfectly similar to itself. Hierarchical clustering on the distance matrix was performed using the complete link algorithm (Theodoridis et al., 2010) that agglomerated conditions with similar gaze scanpath statistics (using the *linkage()* function). A dendrogram was utilized to visualize the clustering outcome (using the *dendrogram()* function). The y-axis of the obtained dendrogram was scaled to the largest global numerical value of Mahalanobis

distance, thereby changing the limits on the y-axis to 0 and 1. Distinct between-cluster separability in the dendrogram was evaluated by inspecting the tri-cluster link cut level in the dendrogram; lower levels indicated better between-cluster separability and by extension, better within-cluster condition similarity.

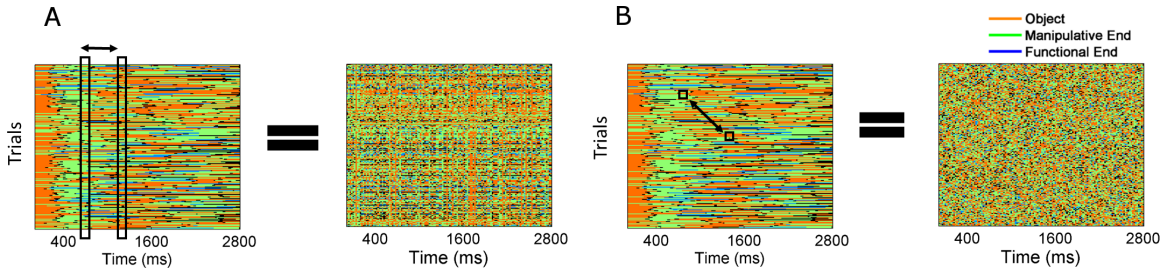

Figure 21: A) Example of temporal shuffling wherein gaze position individual time-points are randomly swapped across trials. B) Example of space-time shuffling wherein gaze positions at any instant within any trial are randomly swapped.

Any clustering algorithm is guaranteed to identify a set of clusters in a dataset. Therefore to assess the cluster significance, random data permutations or data shuffling was performed. Permutation tests are well established non-parametric tests (Maris, 2004, Maris and Oostenveld, 2007) that have been previously used in neuroscience research (e.g. (Rolston et al., 2007, Chao et al., 2010)). We employed three different type of permutation tests or shuffling schemes. In the first shuffling scheme, trials were permuted between-subjects but within-condition (trial shuffling), to identify the dendrogram structure that was an inherent property of the dataset, resistant to subject-level variability due to gender or cultural influences on perception (Steffensen et al., 2008, Goh et al., 2009). In essence 8 pseudo-subjects (Mizelle and Wheaton, 2010a, Natraj et al., 2013) were created within each condition and a roughly equal number of trials were randomly allocated (without replacement) to each of the 8 subjects. In the second shuffling scheme, the temporal order of gaze position was randomly shuffled across all trials within a particular condition (temporal shuffling) and the temporally shuffled trials were then randomly allocated to the subjects. Temporal shuffling was used to evaluate the effect of temporal perturbations on clustering gaze scanpaths. The null permutation scheme involved shuffling foveal gaze positions across any time, trial,

subject and condition (space-time shuffling) and the space-time shuffled trials were subsequently randomly allocated to the subjects. Each permutation scheme was iterated 1000 times and the permuted trials from each iteration were input to the aforementioned clustering algorithm, generating a dendrogram at each iteration. To obtain confidence intervals for mean occurrences of a resultant dendrogram, the 1000 permutations were randomly divided into 20 'sessions' of 50 'experiments'. The statistics primarily used to assess the permutation tests was the number of times a particular dendrogram structure was observed in the permutations, its associated distribution of tri-cluster link cut levels (Levenstien et al., 2003) and the distribution of linear correlation coefficients with the distances between conditions in the original dataset (using the *cophenet()* function in MATLAB). The statistics from the trial and temporal shuffling schemes were contrasted with those produced by the space-time (null) permutation scheme. A pictorial representation of the shuffling methods and clustering algorithm is shown in Figure 21 and Figure 22 respectively.

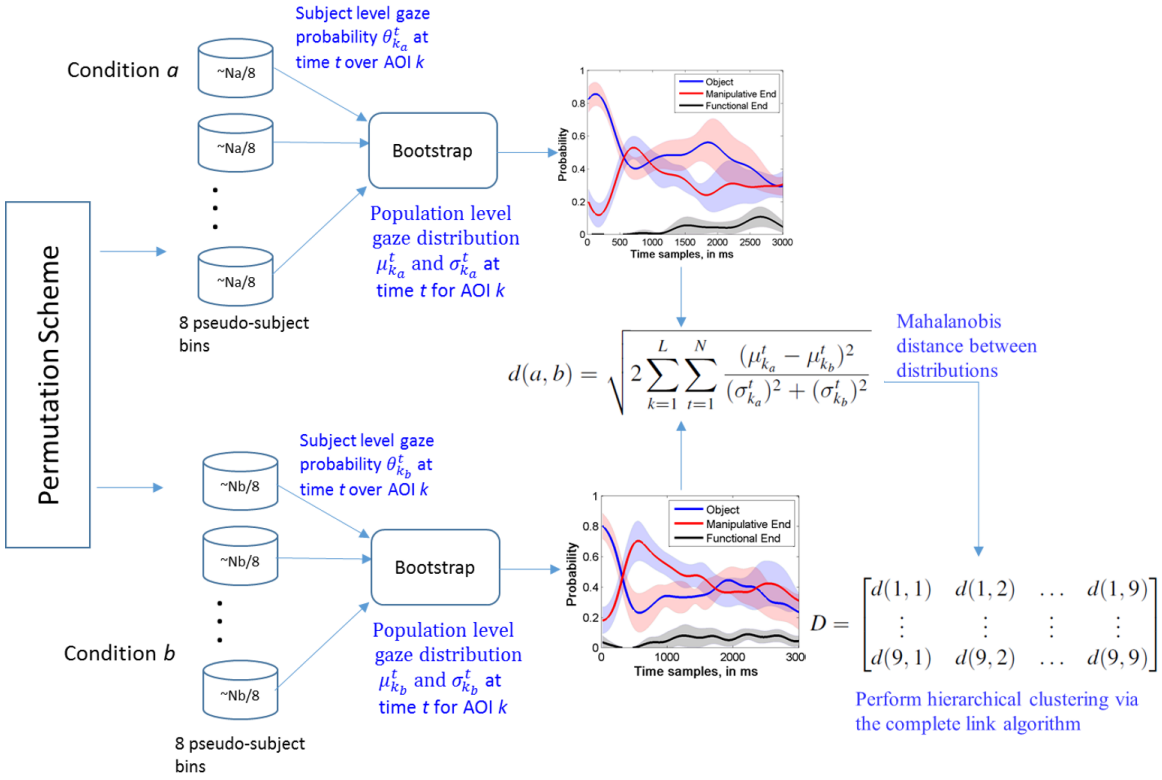

Figure 22: Framework to cluster conditions based on their similarities in spatiotemporal gaze position distributions.

Any repeatedly observed dendrogram produced with either trial or temporal shuffling was associated with a distribution of link-cut levels at which the number of distinct clusters were equal to the original dendrogram. It was also associated with a distribution of linear cophenetic correlations (Theodoridis et al., 2010) with the original dataset. The validity of the link-cut levels distribution and cophenetic correlations distribution were contrasted with those produced by space-time shuffling, the null permutation scheme.

4.3.5.2 Clustering of AOI weights

To determine the relative weighting of the AOI or the foveal bias given a gaze scanpath, the AOI in a stimulus were considered as nodes in a graph and eye movements over the nodes were modeled as a first order discrete Markov process (Hacisalihzade et al., 1992). Consecutive gaze position at a node constituted a fixation and transitions between nodes constituted a saccade. Given the low sampling resolution of the eye tracker (50Hz), the above method adequately served to differentiate fixations and saccades without the inclusion of saccade points as fixations. The categorical fixation and saccade probabilities from one node to any other node were represented by a transition matrix $\mathbf{P}$ wherein an entry in row i and column j corresponded to the probability of gaze transitioning from AOI i to AOI j in one time-step. A Bayesian approach was taken to estimate $\mathbf{P}$ at every instant of the gaze scanpath with a priori assumption that the observer was unbiased towards any AOI (flat prior) given that the stimuli order were randomized. Formally, Dirichlet distributions were used as priors (Minka, 2000, Bishop, 2006, Gupta, 2010) to estimate the categorical distribution in each row of $\mathbf{P}$. The implementation of this method involved simple increments to the elements in each row of the transition count matrix $\mathbf{A}$ (that were a priori all set to 1) at each recorded gaze transition and the corresponding row of $\mathbf{P}$ was determined by dividing the row elements of $\mathbf{A}$ by its row sum (Minka, 2000, Bishop, 2006, Gupta, 2010). To account for extrafoveal attention, elements in each row of $\mathbf{A}$ were incremented by 1 approximately every 300ms. The Bayesian method of updating $\mathbf{P}$ implicitly modeled every AOI to be reachable to the observer (irreducible graph) without any constraints on gaze transitions, allowing for gaze jitter (aperiodic transitions). This allowed

computing a unique steady state categorical distribution of the Markov chain (DeGroot and Schervish, 2002) at each update. The steady state distribution $\mathbf{v}$ was the left eigenvector of the transition matrix $\mathbf{P}$ that satisfied the equation $\mathbf{v}^\top \mathbf{P} = \mathbf{v}^\top$ (Ng et al., 2001), associated with the eigenvalue of 1 (i.e., the dominant left eigenvector of $\mathbf{P}$) . The entries of $\mathbf{v}$ denote the relative importance or weights of each AOI as all the entries of $\mathbf{v}$ sum to 1. The main advantage of this framework is that one single AOI weight vector, $\mathbf{v}$ accounts for both fixation durations and saccade frequencies across all AOI cumulatively through time, largely reducing the dimensionality of eye tracking data. A pictorial representation of computing an AOI weight vector from an individual trial along with the mathematical details are outlined in the Appendix.

AOI weight vectors were computed for every trial and were subsequently averaged to obtain 8 subject-specific mean AOI weight vectors for each of the 9 conditions. It should be noted that though the computed mean AOI weight vectors are 3 dimensional (a weight for each of the 3 AOI), the degree of freedom is 2 as knowing two of the weights is sufficient to calculate the third (as all three weights sum to 1). Thus the computed mean AOI weight vectors were multiplied with a transformation matrix to obtain its 2D representation. Subsequently, the distance between any two conditions' transformed mean AOI weight vector distributions was calculated by the multidimensional Mahalanobis distance metric (Bishop, 2006, Theodoridis et al., 2010)

$$d(a,b) = [\mu_a - \mu_b] \left[\frac{C_a + C_b}{2}\right]^{-1} [\mu_a - \mu_b]^\top \tag{2}$$

In Equation (2), μ_a and C_a are the population mean vector and population covariance matrix of the 2D-transformed AOI mean weight vectors for condition a. Equation (2) was used to construct a distance matrix as explained in the previous section and hierarchical clustering was performed on the distance matrix using the complete link algorithm. The clustering algorithm agglomerated conditions' AOI weight vector distributions that were more similar in Mahalanobis distance. A dendrogram was utilized to visualize the clustering outcome and verify that there would be three distinct context-specific clusters. The y-axis of the obtained dendrogram was scaled to the largest global numerical value of Mahalanobis

distance, thereby changing the limits on the y-axis to 0 and 1. Distinct between-cluster separability in the dendrogram was evaluated by inspecting the tri-cluster link cut level in the dendrogram.

Similar to clustering gaze scanpaths, trial shuffling was carried out to verify if there was a repeatedly occurring and consistent dendrogram that was an inherent property of the dataset resistant to subject-level variability. Temporal shuffling (of gaze position across trials within a condition) was not performed as the AOI weight vector (and the Mahalanobis distance metric here) takes into account the cumulative effects of fixation counts and saccade frequencies, without respect to order. Specifically, the AOI weight vector at any given time-point is independent of the order in which the transition count matrix is populated using gaze data from previous time-points. Similar to gaze scanpath analyses, space-time shuffling constituted the null permutation scheme wherein foveal gaze positions were swapped across conditions, trials, time and subjects. It was hypothesized the tri-cluster link cut levels produced by trial shuffling would be lower (indicating better between-cluster separability) that those produced by space-time shuffling, the null permutation scheme.

4.3.6 Statistical comparisons

Wherever appropriate, repeated measures within-subject ANOVAs ($\alpha = 0.05$) were conducted on participants' gaze characteristics. Bonferroni corrected two sided paired t-tests were used as a post-hoc procedure (significance threshold set at α=0.05). The two sided Kolmogrov-Smirnov (KS) was used to contrast distributions (significance threshold set at $\alpha = 0.05$). All statistical tests were performed in MATLAB and SPSS (The IBM Corporation, Armonk, NY, USA). Effect sizes for the ANOVA tests (SS_{effect}/SS_{total}) were computed from the Sum of Squares in the ANOVA tables (Levine and Hullett, 2002) and an open source tool was used to compute Cohen's d_z, the effect size for the paired t-tests (Lakens, 2013). The strength of the clustering results was assessed using the probability of occurrence of a dendrogram across the iterations of a particular permutation test.

4.4 Results

4.4.1 Saccade initiation times

Across all trials and participants, the average saccade initiation time was $231 \pm 55.2ms$ and the majority of saccades (¿60%) were initiated within a short time window (180-260ms after image presentation, Figure 23A-B). The time course of a gaze scanpath within each trial was then set relative to the time when gaze position first entered within an AOI post saccade-onset.

4.4.2 Clustering gaze scanpaths

The central hypothesis in this study posited that hierarchical clustering of gaze scanpaths or spatiotemporal gaze distributions would primarily group conditions into three clusters by tool-object context, since the task required evaluating tool-object relationships. The mean probable gaze position over the AOI for each condition at each time point is depicted in Figure 23C, along with the bootstrap confidence interval. The bootstrap distribution of the mean was assumed to be normally distributed at each individual time-point, verified using the Kolmogorov-Smirnov goodness-of-fit test at the 0.01 level at each time-point with FDR correction for multiple comparisons (False Discovery Rate, (Genovese et al., 2002, Delorme and Makeig, 2004)). The 95% confidence interval of mean probable gaze position was thus computed using the 2 S.D metric of the bootstrap distribution. The results of the trial and temporal shuffling schemes to address the hypothesis are presented in the following sections. It should be noted that the results in the following sections are relative to gaze data up to 1760ms after first saccade onset, or roughly 2000ms after stimulus onset (similar to stimuli duration times in (Natraj et al., 2013)) taking into account mean saccade initiation time.

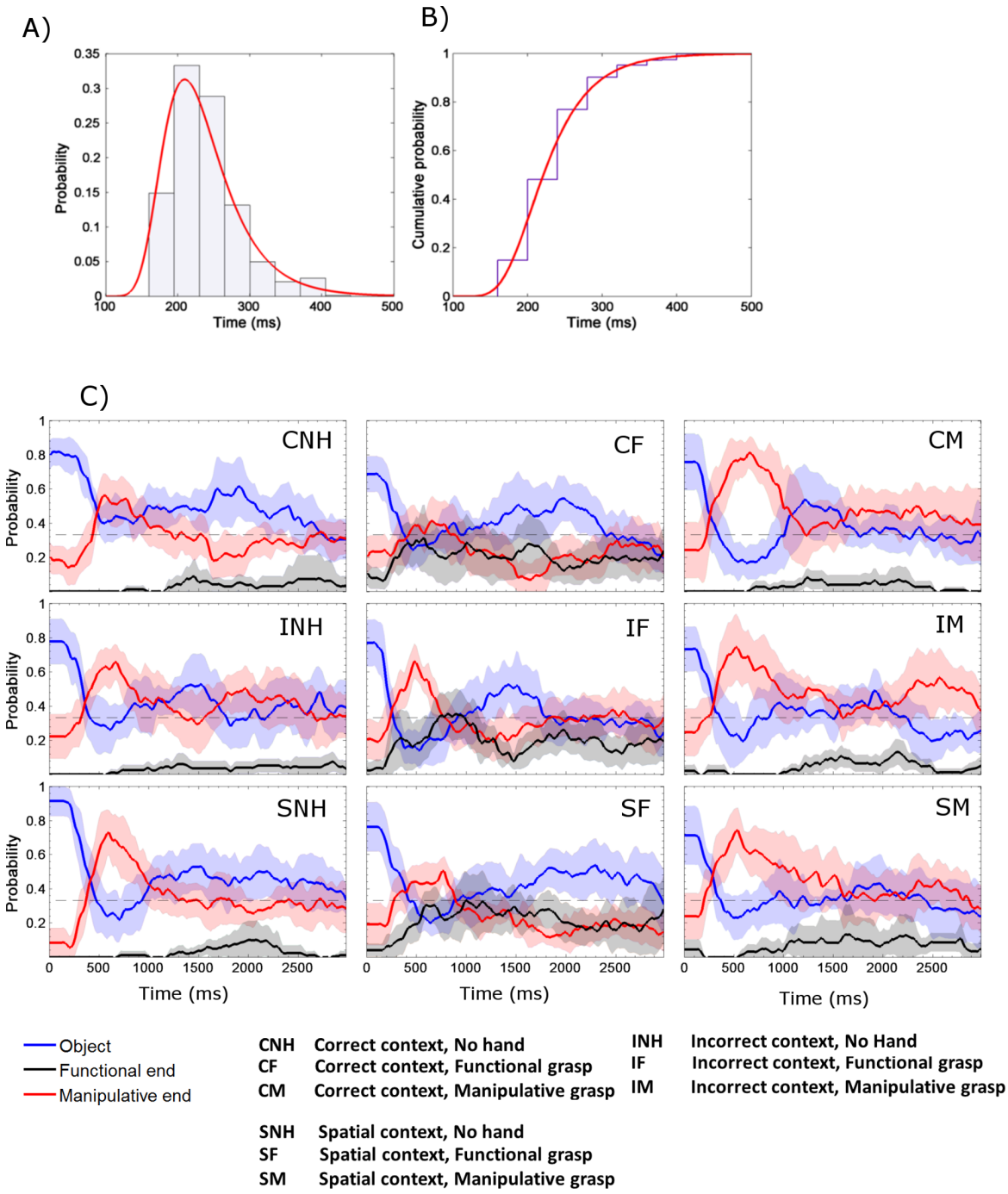

Figure 23: (A and B) The Frechet probability density function for saccade initiation times with a mean of $231 \pm 55.2ms$ (S.D.) C) The time course of mean probable foveal gaze position over each AOI (with 95% bootstrap C.I shading) for each of the 9 conditions, smoothed with a 5 sample Savitzky-Golay filter. As reference, the thin black dotted line represents the uniform or maximally random distribution if the observer were to be unbiased towards any AOI and is a constant at $\frac{1}{3}$ given the 3 AOI.

4.4.2.1 Trial shuffling

Results of the trial shuffling scheme are presented in Figure 24A. There were two stable dendrograms in the dataset, called Cluster 1 and Cluster 2 (Figure 24A) defined by the conditions that formed three distinct clusters. Cluster 1 was observed in $72.2 \pm 4.98\%$ (S.D.) of the 20 simulated sessions and Cluster 2 was observed in $25.7 \pm 4.91\%$ (S.D.) of the 20 simulated sessions. Cluster 1 was therefore the more dominant feature of the dataset ($t(38) = 29.73, p < 1 \times 10^{-5}, d = 9.647$). 98.89% of Cluster 1's dendrograms and 98.83% of Cluster 2's dendrograms had tri-cluster link cut levels less than 0.5 (50% of maximum Mahalanobis separation distance). Thus the between-cluster separability (and within-cluster condition similarity) within the dendrograms of Cluster 1 and Cluster 2 was significant with respect to at least 50% of the maximum Mahalanobis distance. On average, both Cluster 1 and Cluster 2 were highly correlated with the distances between conditions in the original dataset (Cluster 1: 0.897, Cluster 2: 0.899). The distribution of tri-cluster link-cut levels and linear correlations for Cluster 1 and 2 were then individually assessed against those produced by space-time shuffling, the null permutation scheme (Figure 24C). Results revealed that the distribution of tri-cluster link-cut levels and correlation coefficients for both Cluster 1 and Cluster 2 to be significantly different from those produced by space-time shuffling, the null permutation scheme ($p < 1 \times 10^{-5}$ in all KS tests). From the three clusters in the dominant dendrogram (Cluster 1), it can be seen that gaze distributions clustered by the type grasp-posture rather than tool-object context, contrary to the original hypothesis. Therefore, we rejected the null hypothesis of observing three context-specific clusters (null dendrogram) in a significant number of iterations. However, clustering results were in line with our secondary hypotheses as all three manipulative grasp conditions were distinct from the no hand conditions. The three clusters of Cluster 2 revealed this effect of grasp to be weakest within the incorrect tool-object context as the incorrect context-no hand condition was similar to the manipulative grasp conditions in a smaller but significant number of iterations. The effect of temporal perturbations in gaze scanpaths on Clusters 1 and 2 is detailed in the next section.

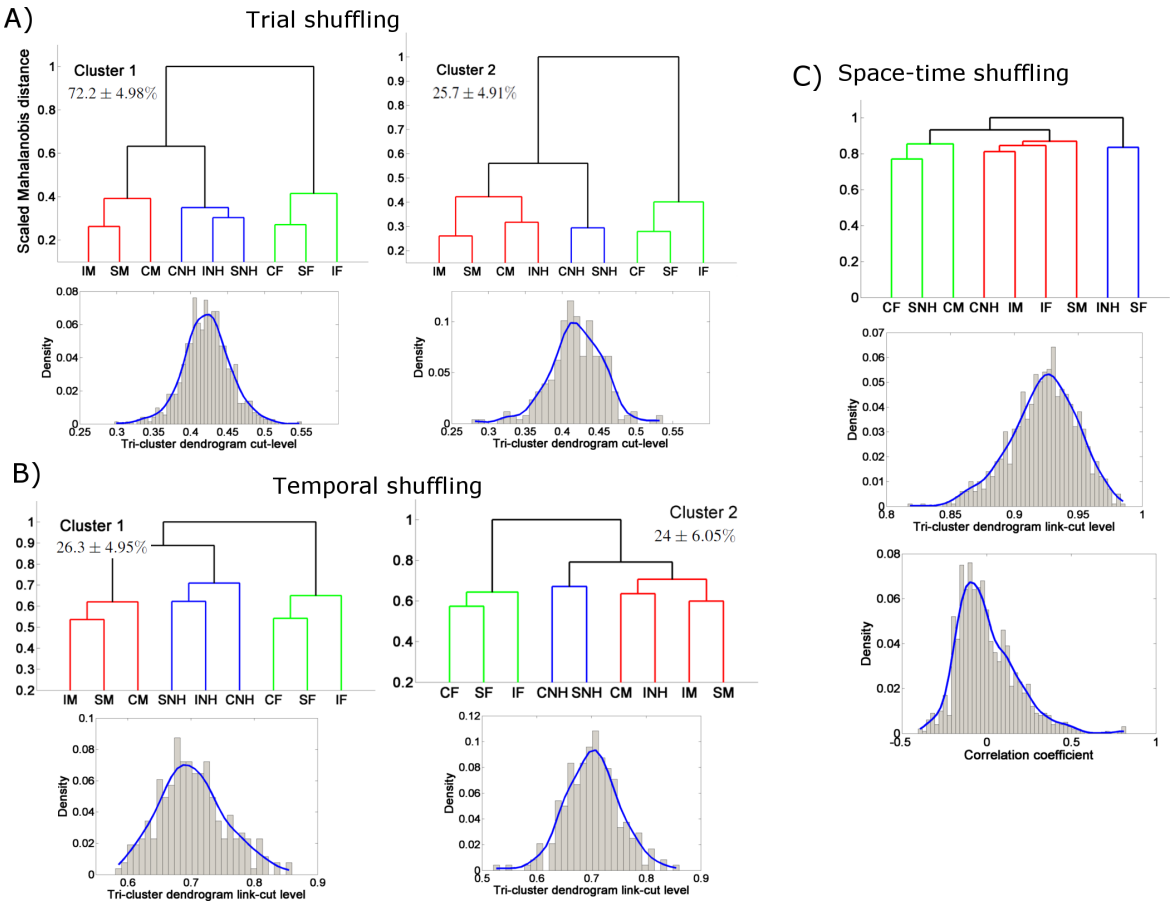

Figure 24: A) Results from the trial shuffling permutation test. The two dominant dendrograms (Cluster 1 and 2, defined by the conditions grouping into three distinct clusters) from trial shuffling are shown in the top row with their mean occurrences (and S.D) from the 20 sessions of 50 simulated experiments. The distributions of each dendrogram's tri-cluster link-cut levels are shown in the second row. B) Results from the temporal shuffling permutation test. The two dominant dendrograms (defined by the conditions grouping into three distinct clusters) are shown along with their mean occurrences (and S.D) from the 20 sessions of 50 simulated experiments. C) Results from the space-time (null) shuffling permutation test. There were no significant, repeatedly occurring, three cluster dendrograms in the dataset. For illustration, an exemplar dendrogram with three clusters is shown. The null distribution of tri-cluster dendrogram link cut-levels and correlation coefficients is also depicted. For illustration purposes, a non-parametric curve has been approximately fitted to each of the distributions.

4.4.2.2 Temporal shuffling

The temporal shuffling scheme involved shuffling the temporal order of gaze position across all trials within each condition. We temporally shuffled data up to 1760ms after saccade-onset to associate with results from trial shuffling. Results revealed that together, Clusters 1 and 2 were reasonably robust to temporal perturbations in gaze scanpaths (Figure 24B)

as together they were observed in aprox. 50% of the 1000 permutations. On average, both Cluster 1 (0.85) and Cluster 2 (0.86) were well correlated with the distances between conditions in the original dataset. Comparisons with the tri-cluster link-cut distribution and correlation distribution of the null permutation scheme in Figure 24C showed them to be significant features of the dataset ($p < 1 \times 10^{-5}$ in all KS tests). However, Cluster 1 was no longer the dominant dendrogram and in fact, both Clusters 1 and 2 had an equal likelihood of occurrence across the 20 simulated sessions (likelihood of occurrence in a session: Cluster 1: $26.3\% \pm 4.95\%$ S.D, Cluster 2: $24\% \pm 6.05\%, t(38) = 1.3152, p = 0.196$). Thus with respect to Cluster 1, temporal shuffling caused significant loss in the probability of occurrence (power). The drop in the number of occurrences for Cluster 1 is suggestive of its dependence on the temporal pattern of the gaze scanpaths, whereas data suggests that Cluster 2 may not be as sensitive to the temporal pattern of gaze scanpaths as Cluster 1. In addition, the loss in the temporal order of gaze scanpaths also resulted in significantly poorer between-cluster separability (and poorer within-cluster condition similarity). For instance, the three clusters in both Clusters 1 and 2 formed at significantly higher levels in the dendrogram when compared to their counterparts that were generated without temporal shuffling (Cluster 1 KS statistic = 0.993, $p < 1 \times 10^{-5}$, Cluster 2 KS statistic = 0.992, $p < 1 \times 10^{-5}$). In addition, all of Cluster 1's and Cluster 2's tri-cluster link-cut levels were greater than 50% of the max Mahalanobis distance.

4.4.2.3 Summary

While the above permutation tests are specific to gaze scanpaths up till approx. 2000ms post image-onset it should be noted that similar permutation tests can be theoretically performed at every time instant. However, to avoid the very large computational burden and statistical complexity, we did not perform the permutation tests at each time-point. Therefore, we only qualitatively inspected the temporal evolution of the dendrogram produced by applying the clustering algorithm on the original dataset. We observed that Cluster 1 was the more dominant dendrogram from 720ms to 1920ms. From 1940ms onwards, we observed that Cluster 2 started to emerge as the dominant dendrogram (Video 1). However, the mean

tri-cluster link-cut level tended to be greater than 0.5 and in addition, this time period was well beyond the stimulus duration times in our previous studies (Borghi et al., 2012, Natraj et al., 2013). Summarizing, clustering results revealed that the grasp-specific gaze scanpaths, with a secondary effect of the incorrect context, was a consistent feature of our dataset. This result was contrary to our hypotheses as we had hypothesized gaze scanpaths to cluster by tool-object context.

4.4.3 Clustering of AOI weights

The AOI weighting algorithm generated a time-varying estimate of each subject's mean AOI weight vector for every condition. While hierarchical clustering on similarities in conditions' population level mean AOI weight vector distributions could be performed at any time instant, we chose 1760ms (Figure 25A) to associate with results from the gaze scanpath clustering analyses. In particular, based on the results from clustering gaze scanpaths, we were primarily concerned with Clusters 1 and 2. Results from trial shuffling showed that Cluster 1 was an inherent property of the dataset, and was observed in $76.2 \pm 4.72\%$ (S.D) of the experiments across the 20 sessions (Figure 25B). Comparison of the tri-cluster link cut level distribution showed between-cluster separability of Cluster 1 to be significantly different from those produced by the null permutation scheme (Figure 25B, KS statistic $= 0.6015$, $p < 1 \times 10^{-5}$). Similar to gaze scanpath analyses, we rejected the null hypothesis of observing three context-specific clusters (null dendrogram) in a significant number of iterations. However, clustering results satisfied our secondary hypothesis as the manipulative grasp conditions were distinct from the other grasp conditions. With respect to Cluster 2, results showed that it was observed only in 6 of the 1000 permutations. One possible reason for this low occurrence for Cluster 2 here could be due to manner in which the 2D Mahalanobis distance equation evaluated the similarity between conditions' AOI weight vector distributions. In line with gaze scanpath analyses, clustering results here as well were contrary to our hypothesis that conditions would cluster by tool-object context. However, results did confirm our secondary hypothesis as the manipulative grasp conditions were distinct from the no hand and functional grasp conditions.

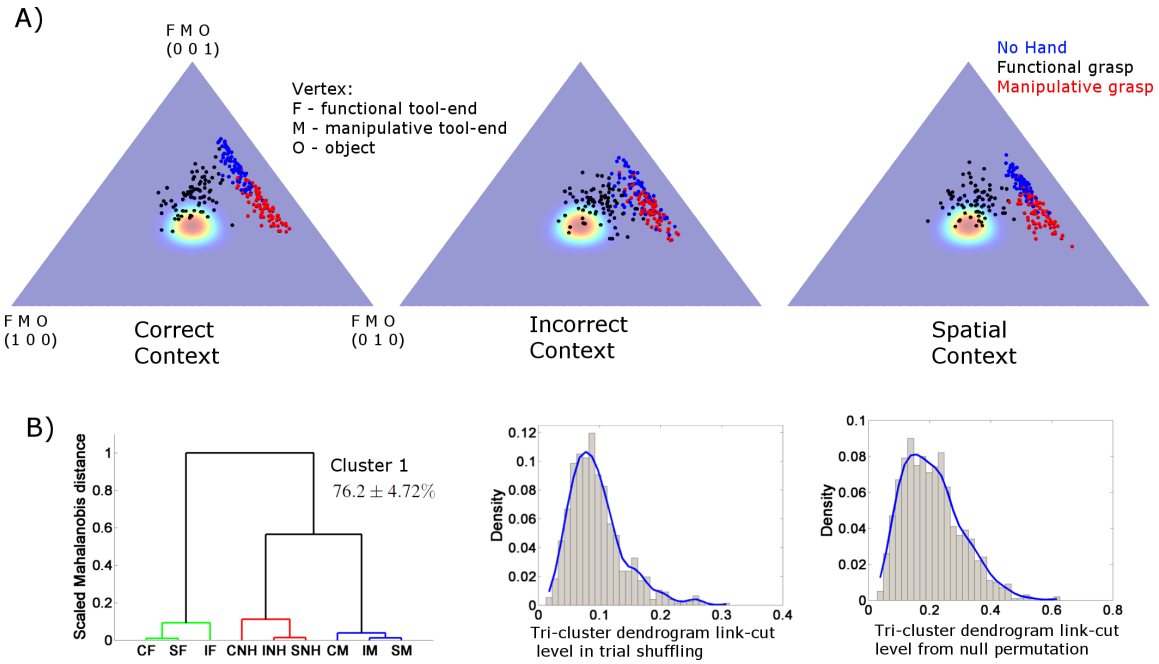

Figure 25: A) AOI weight vectors from 10 permutations of the trial shuffling scheme that produced the grasp-specific clustering result, Cluster 1 in Figure 24A. It should be noted that each permutation of the trial shuffling scheme produced 8 mean AOI weight vectors, one for each subject, across all 9 conditions. The grasp-specific AOI weight vectors are plotted on the simplex or the probability plane by each tool-object context. The edge or vertex of each simplex corresponds to the bias towards an AOI. The sum of the coordinates on any point on the simplex sum to 1. The halo at the center of each simplex corresponds to the likelihood that the three AOI have been weighted equally by the observer (Dirichlet distribution around uniform foveal bias (Minka, 2000, Gupta, 2010)). B) Representative dendrogram depicting Cluster 1, along with its mean and S.D of occurrence across the 20 sessions of 50 experiments from the trial shuffling scheme. Cluster 1's distribution of the tri-cluster dendrogram link-cut level is shown along with the distribution produced by the null permutation scheme. The two distributions were significantly different.

With respect to the visualization of the AOI weight vectors, we plotted data from 10 randomly chosen permutations of the trial shuffling scheme that produced Cluster 1. It should be noted that each permutation produced 8 mean AOI weight vectors, one for each pseudo-subject, across all 9 conditions. The AOI weight vectors were plotted on the simplex or the probability plane, a commonly used method to represent probability vectors (Minka, 2000, Bishop, 2006, Gupta, 2010), shown in Figure 25A. The simplex is well-suited to represent the AOI weight vector as any point on the simplex sums to 1, similar to the entries of the AOI weight vector. Given that the clustering results here showed the grasp-posture to be the principal driver of how the AOI were weighted by the observer, in the

following section we detail the differences between the three grasp-postures on the mean bias towards each AOI derived from the original dataset. Specifically, we collapsed subjects' data across Context and evaluated the mean bias towards each AOI separately using a one-way RM-ANOVA with factor Hand (no hand, functional grasp, manipulative grasp), detailed in Figure 26.

4.4.3.1 Bias towards the object

Results of the RM-ANOVA showed a main effect of Hand ($F(2,46) = 15.656$, $p = 1.3 \times 10^{-5}, \eta^2 = 0.405$). Post-hoc tests showed that the no hand conditions elicited greater bias towards the object than the functional grasp conditions (by $7.6 \pm 2.11\%$ S.E., $t(23) = 3.61$, $p = 1.474 \times 10^{-3}, d_z = 0.737$) and manipulative grasp conditions (by $12.67 \pm 2.62\%$ S.E., $t(23) = 4.828, p = 7.15 \times 10^{-5}, d_z = 0.98$). The functional grasp conditions tended to have greater bias towards the object than the manipulative grasp conditions, but the difference between the two did not reach Bonferroni adjusted significance threshold ($5.045 \pm 2.06\%$ S.E., $t(23) = 2.449, p = 0.022, d_z = 0.5$).

4.4.3.2 Bias towards the manipulative tool-end

Results showed a main effect of Hand ($F(2,46) = 39.79$, $p < 1 \times 10^{-5}, \eta^2 = 0.634$), with an adjustment of the degrees of freedom using the Greenhouse-Geisser correction. Post-hoc tests on the main effect of Hand showed that the manipulative grasp conditions had greater bias towards the manipulative tool-end than the no hand conditions (by $11.64 \pm 2.5\%$ (S.E.), $t(23) = 4.66, p = 1.08 \times 10^{-4}, d_z = 0.95$) and functional grasp conditions (by $19.2 \pm 2.4\%$ (S.E.), $t(23) = 8.011, p < 1 \times 10^{-5}, d_z = 1.635$). In addition, the no hand conditions had greater bias towards the manipulative tool-end than the functional grasp conditions (by $7.56 \pm 1.46\%$ S.E. $t(23) = 5.19, p = 2.92 \times 10^{-5}, d_z = 1.06$).

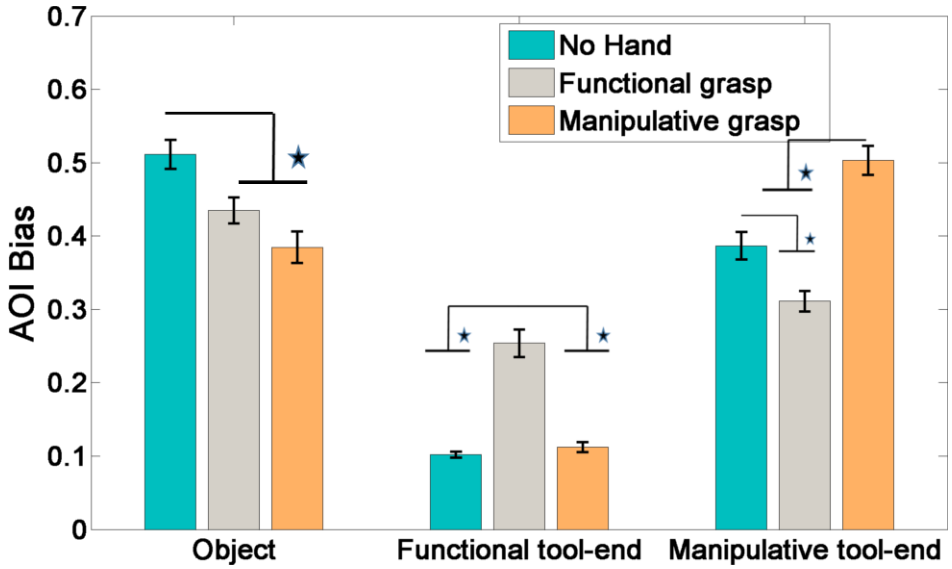

Figure 26: A-C) Mean foveal bias (with S.E bars) towards each AOI across the three Hand conditions (no hand, functional grasp, manipulative grasp). Data were collapsed across Context in line with the Clustering results of Figure 25B. Significant differences are marked with the ★ sign.

4.4.3.3 Bias towards the functional tool-end

Results showed a main effect of Hand ($F(2,46) = 51.89$, $p < 1 \times 10^{-5}, \eta^2 = 0.693$), with an adjustment of the degrees of freedom using the Greenhouse-Geisser correction. Post-hoc analyses showed that functional grasp conditions elicited higher bias towards the functional tool-end than the no hand conditions (by $15.18 \pm 1.85\%$(S.E.), $t(23) = 8.2, p < 1 \times 10^{-5}, d_z = 1.67$) and manipulative grasp conditions (by $14.15 \pm 2.06\%$ (S.E.), $t(23) = 6.86, p < 1 \times 10^{-5}, d_z = 1.4$). The no hand conditions and the manipulative grasp conditions did not differ from each other ($t(23) = 1.294, p = 0.209$). However, as it can be seen in Figure 26, it is important to note that within the functional grasp conditions, the functional tool-end was not the dominant AOI. A RM-ANOVA on the mean bias towards the three AOI across the functional grasp conditions revealed a main effect of AOI (object, functional tool-end, manipulative tool-end, $F(2,46) = 20.05$, $p < 1 \times 10^{-5}, \eta^2 = 0.466$) and post-hoc Bonferroni corrected paired t-tests revealed that the functional tool-end was weighted significantly lesser than the object (by $18.14 \pm 3.35\%$ S.E, $t(23) = 5.412, p = 1.69 \times 10^{-5}, d_z = 1.1$), while being weighted statistically similar to the manipulative tool-end (difference of $5.724 \pm 2.78\%$

S.E, $t(23) = 2.063, p = 0.051, d_z = 0.42$). In fact, on average, across conditions, a RM-ANOVA on the 8 subjects' mean bias towards the three AOI revealed a main effect of AOI (object bias: $44.36 \pm 2.04\%$ S.E, functional tool-end bias: $15.6 \pm 1.06\%$ S.E, manipulative tool-end bias: $40.04 \pm 1.46\%$ S.E, $F(2,14) = 65.12$, $p < 1 \times 10^{-5}, \eta^2 = 0.903$) and post-hoc Bonferroni corrected paired t-tests showed that the functional tool-end was weighted lesser than the manipulative tool-end (by $24.432\% \pm 1.532\%$ S.E, $t(7) = 15.951, p < 1 \times 10^{-5}, d_z = 5.64$) and the object (by $28.76\% \pm 2.898\%$ S.E, $t(7) = 9.927, p = 2.24 \times 10^{-5}, d_z = 3.51$) while the latter two were weighted statistically similar to each other ($t(7) = 1.281, p = 0.241$).

4.4.3.4 Summary

Broadly, results showed that the AOI weighting patterns were primarily driven by the grasp-posture as the no hand conditions elicited the greatest bias towards the object, the functional grasp conditions elicited the greatest bias toward the functional tool-end and the manipulative grasp conditions elicited the greatest bias towards the manipulative tool-end. These results can be interpreted of the grasp-posture being a gaze attractor. However, across conditions, participants primarily focused on the object and manipulative tool-end and sparsely attended to the functional tool-end, even within the functional grasp-conditions. It was only the manipulative grasp-posture that caused the manipulative tool-end to become a gaze attractor. In the absence of any grasp-posture, the object was weighted the most.

4.5 Discussion

The overall goal of Aim 2 was to understand whether gaze data followed the results from Aim 1 (Natraj et al., 2013) and the prior behavioral result (Borghi et al., 2012) given the overlap between action understanding and gaze control over parietofrontal networks. We had hypothesized that hierarchical clustering of spatiotemporal gaze distribution patterns and AOI weighting patterns would group conditions by tool-object context since the task involved evaluating tool-object content. However, results showed that both spatiotemporal gaze distributions and AOI weightings up to 2000ms after image-onset clustered primarily by the type of grasp with unique effects of tool-object context within each grasp cluster.

Clustering results in part did validate our secondary hypotheses as the manipulative grasp conditions were distinct from the no hand and functional grasp conditions in all three contexts. This result is in line with the EEG results in Aim 1 and our prior behavioral data wherein the manipulative grasp elicited significantly extended parietofrontal activations with temporally coinciding response delays when compared to the no hand conditions in all three contexts. In addition, permutation tests showed that the grasp-specific clusters were disrupted by the incorrect tool-object context as the incorrect context-no hand condition elicited similar gaze scanpaths as the manipulative grasp conditions in a small but significant number of iterations. Thus the statistical difference in spatiotemporal gaze distributions between the manipulative grasp condition and no hand condition was robust in the correct and spatial tool-object contexts but weakest within the incorrect tool-object context that clearly does not afford tool-object action. One potential reason underlying this phenomenon is that gaze scanpaths could correspond to the observer evaluating the action error in the scene (Mizelle et al., 2013), given that both the manipulative grasp conditions and the incorrect context does not afford a functional engagement of the tool on the object. Indeed, a qualitative inspection of the gaze scanpaths in Figure 23B show that unlike the correct and spatial context-no hand conditions, gaze tended to oscillate more between the object and the manipulative tool-end, similar to the manipulative grasp conditions. These grasp specific gaze scanpath clusters were also reasonably robust to the temporal order of gaze position, though the strength of gaze scanpath clustering results depended on both the spatial and temporal similarities in gaze position, derived from the Mahalanobis distance metric.

It should be noted that it was not the mere presence of a grasp-per se that influenced gaze patterns as the tool-end associated with the grasp was a gaze attractor only in the manipulative grasp conditions and not in the functional grasp conditions. In fact, in the absence of a grasp in the scene, the object was foveally weighted the most. Assuming that the observer is evaluating the dynamic properties of the static tool-object scene (Kourtzi and Kanwisher, 2000, Proverbio et al., 2009), the enhanced attention towards the object may correspond to an object-oriented action priming effect (Thill et al., 2013), wherein the

observer may 'map' how the tool could be used on the object via the activation of his/her own motor knowledge of the tool-object pair (Flanagan and Johansson, 2003, Ambrosini et al., 2011). Indeed, previous research has shown that when making tea or a sandwich, people's saccades tended to gravitate towards objects that were targets of potential tool-use (Land et al., 1999, Land and Hayhoe, 2001). As a result, the greater bias towards the object in the no hand conditions (especially within the correct and spatial tool-object contexts, qualitatively seen in Figure 25A) may facilitate the evaluation of the motoric relationship between the tool-object pair and could underlie results from our previous behavioral data wherein the no hand conditions elicited the fastest reaction times compared to the other grasp-postures when participants evaluated tool-object content (Borghi et al., 2012). Conversely, observing an egocentric hand-posture grasping the tool in the right hemifield could have automatically primed or engaged grasp-specific neurons in the action encoding system to encode grasp-intent via motor resonance (Iacoboni et al., 2005), thereby drawing gaze to the grasp-posture and lessening the bias towards the object. However, the disruptive effect of the grasp-posture on the object-oriented priming effect was most pronounced within the manipulative grasp conditions. Unlike the functional grasp-posture, the manipulative grasp-posture caused to tool-end associated with it to become a gaze attractor and the dominant AOI. There are two synergistic reasons that could underlie this attractor effect of the manipulative grasp-posture. First, the manipulative grasp-posture may have perturbed an ongoing motor simulation to evaluate the tool-object content (Borghi et al., 2012, Natraj et al., 2013) as this grasp-posture does not afford tool-object action. Indeed, in our prior data, the manipulative grasp-posture elicited the most delayed behavioral response times (Borghi et al., 2012) and elicited greater neural activity along action-encoding regions (Natraj et al., 2013). Second, the effect of the manipulative grasp-posture may also be driven by its non-functional interaction with the manipulative tool-end, an aspect we had not previously considered. Specifically, across conditions, eye tracking data revealed that participants primarily focused on the object and the manipulative tool-end and sparsely attended to the functional tool-end, even within the functional grasp conditions. A potential reason for participants to focus more on the manipulative tool-end than the functional tool-end could

be due the former's affordances with respect to the object. While we had classified the manipulative tool-end with respect to the grasp-posture associated with it, it is in fact the operant tool-end that interacts with the object. Thus for example, the bowl of the spoon may be more relevant than the stem in evaluating the spoon's relationship with a cup of yogurt, as it is the bowl that actually affords the act of scooping. Given the importance of the manipulative or operant tool-end in evaluating the relationship between the tool-object pair, a manipulative or a non-functional grasp maybe be a gaze attractor as it interferes with that part of the tool that ostensibly has higher affordances with respect to the object. Therefore, the gaze attractor effect of the manipulative grasp-posture could synergistically be due to its non-functional interaction with the operant tool-end that actually engages with the object (e.g. hammer-head to nail), thereby requiring more attention from the observer to process its affordances with respect to the tool-object pair.

The question then remains, why focus foveal attention at the functional tool-end at all in the functional grasp conditions when this tool-end seems to be less important to the task? As outlined earlier, it is likely that the presence of functional grasp-posture could automatically prime or engage grasp-specific neurons in the action encoding system via motor resonance (Iacoboni et al., 2005), thereby drawing a proportion of gaze towards the functional tool-end ostensibly to encode the functional grasp (Iacoboni and Dapretto, 2006). However, the amount of bias or attention drawn to the functional tool-end was never dominant when compared to the object and the manipulative tool-end, highlighting the importance of the latter two AOI in evaluating tool-object action relationships. The fact that the functional grasp-posture (and the portion of the arm below the wrist) was sparsely attended to follows the results reported in (Land et al., 1999), wherein the authors found that when performing day-to-day functional tasks such as making tea, people rarely fixated on their hands and predominantly directed gaze towards the object that was the target of upcoming action.

4.6 Conclusions

Results here broadly support our prior interpretations from Aim 1 and the behavioral experiment that the observer may automatically try to encode the affordances of the manipulative grasp though it is not essential to evaluating tool-object content. As a consequence, this interference effect of the manipulative grasp-posture may result in delayed decision times (Borghi et al., 2012), extended parietofrontal activity (Natraj et al., 2013) and differential gaze patterns (Natraj et al., 2015) when compared to the no hand control condition. Aim 2 also provide an attentional mechanism underlying results in Aim 1 and the prior data (Borghi et al., 2012, Natraj et al., 2013). Given that parietofrontal circuits also underlie the control of saccades and visuospatial attention (Anderson et al., 1994, Mort et al., 2003), it is unclear as to how much of previously observed parietofrontal activity in Aim 1 may correspond to action encoding alone. Specifically, an unknown proportion of the parietofrontal differences between the manipulative grasp and no hand conditions may in fact correspond to the planning of differential eye movements between the two conditions. The fact that we did find differential gaze patterns between the two grasp conditions therefore gives us an empirical justification to probe the coupling of action encoding and eye movements over parietofrontal regions, and form the basis for Aim 3.

CHAPTER V

SPECIFIC AIM 3

5.1 Introduction

Results from both Aims 1 and 2 showed that the manipulative grasp-posture automatically elicited differential parietofrontal and gaze patterns when compared to the no hand control condition as participants passively evaluated tool-object content (Natraj et al., 2013, Natraj et al., 2015). It should be noted that the grasp per se was irrelevant to the task of evaluating the relationship between the tool-object pair and a similar neurovisual effect was not observed for the functional grasp-posture. Together, these results provide a neural and attentional mechanism underlying our prior behavioral data wherein manipulative grasp-postures exclusively interfered and delayed decisions on tool-object content. The goal of Aim 3 was to further probe the neural and attentional mechanisms underlying this interference effect of the manipulative grasp-posture. Specifically, in both Aims 1 and 2, the observer had sufficient time to parse the features of the scene as the stimuli durations were on the order of 2000ms. Given that parietofrontal networks are involved in both evaluating tool-object images (Mizelle and Wheaton, 2010b, Natraj et al., 2013) and attentional mechanisms (Nobre et al., 1997, Corbetta et al., 1998), it is possible that a relationship might exist between attention and affordance understanding over parietofrontal networks. Such a relationship might create ambiguity in the interpretations of neural activations. For instance, it is likely that an unknown proportion of the parietofrontal difference between the manipulative grasp condition and no hand condition might actually be due differential visuospatial attention. More specifically, it is possible that the greater right parietofrontal involvement to process the manipulative grasp-posture could be driven by the observer allocating greater foveal attention towards the manipulative tool-end. The right hemispheric parietofrontal encoding of the manipulative grasp-posture might be differentially attenuated or modulated by restricting an observer's ability to direct foveal attention over scene

features. Therefore, the purpose of Aim 3 was to understand the influence of eye movements on parietofrontal differences between the manipulative grasp-posture and the no hand condition when encoding the same stimuli as Aims 1 and 2 and the prior behavioral data (Borghi et al., 2012, Natraj et al., 2013, Natraj et al., 2015).

To isolate the role of foveal attention and eye movements on parietofrontal activations, it was necessary to design an experiment wherein participants processed the images extrafoveally or using peripheral vision, without any saccades. It is well known that instructing participants to restrict saccades would engage inhibition related frontal activity when viewing wide field complex visual stimuli (Chikazoe et al., 2007). To avoid this potential confound in the data, two experiments were designed. In the first experiment, stimuli durations were shortened to 100ms. No instructions were given to participants related to eye movement but data revealed that such rapidly presented stimuli naturally dropped saccade probabilities to less than 10% without affecting accuracies in evaluating tool-object content. In a second experiment with the same participants, the stimuli durations were increased to 500ms, thereby allowing for the reemergence of saccades and fixations. This experimental design allowed contrasting the presence and absence of eye movements on the parietofrontal encoding of the interference effect of the manipulative grasp-posture in manner similar to the EEG study of Aim 1 (Natraj et al., 2013). The statistical analyses in Aim 1 was restricted to evaluating ERP voltage means via an ANOVA test within 200ms time-windows and over broad regions of interest (ROI) encompassing primarily bilateral parietofrontal and left motor electrodes. However, the experimental design here involved rapid stimuli that might elicit very fine spatiotemporal neural differences. Therefore, we considered a whole brain analysis to precisely capture the spatiotemporal dynamics of parietofrontal activations. To this end, data were analyzed in a mass-univariate fashion at each electrode-time pair using the framework of the General Linear Model (GLM). In essence, rather than perform just one ANOVA test, we carried out many tens of thousands of ANOVA tests at each and every electrode and time-point, with the appropriate corrections for multiple comparisons (Pernet et al., 2011). In principle, we used statistical methods similar to the analyses of large scale fMRI data (Smith, 2014). To ensure that participants were

accurate in the task given that the stimuli were rapidly presented, behavioral responses were collected. To ensure that participants were accurate in the task given that the stimuli were rapidly presented, behavioral responses were collected. The response hand used to record correct/incorrect tool-object relationships was counterbalanced in both experiments.

5.2 Hypotheses

We hypothesized that even with the rapid image presentations, participants would be able to perform the task and have accuracies significantly greater than chance. Similar to our prior behavioral data, we hypothesized that the manipulative grasp-posture and the spatial tool-object context would delay decisions on tool-object content (Borghi et al., 2012). With respect to eye movement, we hypothesized that the rapid stimuli within the 100ms experiment would negate participants from directing foveal attention towards scene features. Our major hypotheses concerned the spatiotemporal patterns of parietofrontal activity underlying the interference effect of the manipulative grasp-posture. In Aim 1, we had proposed the existence of a late appearing (400-600ms after image onset) right parietofrontal network to specifically encode the intent of the manipulative grasp-posture, along with a continual left parietofrontal network to encode the tool-object content itself. However, the results of Aim 2 had suggested that this right parietofrontal activity could be driven by enhanced attention over the manipulative tool-end. Given the inability of participants to direct foveal attention in the 100ms experiment within Aim 3, we hypothesized that the right parietofrontal activity (400-600ms post image onset) would be inhibited and instead, a dominant left parietofrontal network would underlie the interference effect of the manipulative grasp-posture. In the 500ms experiment, we hypothesized that the right parietofrontal activity (400-600ms post image onset) would reemerge as participants would have to time to execute saccades and direct attention.

5.3 Methods

5.3.1 Subjects

A total of 33 subjects between the ages of 20-30 years (average age: 23.7, S.D 1.8) were recruited for this study, who participated in both experiments (100ms and 500ms experiments). All subjects were right-handed based on the Edinburgh Handedness inventory (Oldfield, 1971). The experimental protocol was approved by the Institutional Review Board at Georgia Institute of Technology and each subject provided their written informed consent before the start of the experimental session. Subjects were healthy based on self-report, and had no history of neurological illness or injury. 5 subjects were excluded from the 100ms experiment either due to poor task performance (accuracy not above chance) or due to too much noise in the EEG data. 7 subjects were excluded from the 500ms experiment either due to poor task performance (accuracy not above chance), due to too much noise in the EEG data or as some could not complete the experiment due to prior commitments and time constraints.

5.3.2 Stimuli

Participants viewed the same static gray-scale stimuli as the prior studies. However, we dropped the static hand conditions from the experimental design similar to Aim 2, as this condition elicited largely similar neural and behavioral responses to the control no hand condition in our prior behavioral data and Aim 1 respectively (Borghi et al., 2012, Natraj et al., 2013). As with the previous Aims, a tool was placed across three contexts: correct (hammer-nail), incorrect (hammer-paper) and spatial (hammer-wood). These three tool-object contexts were orthogonally placed across three grasp-postures: no hand, functional tool-grasp (grasp hammer-handle) and a manipulative tool-grasp (grasp hammer-head). Figure 27 depicts a 3 $\times$ 3 matrix for a representative tool. There were similar 3 $\times$ 3 matrices for 20 other tools for a total of 180 unique stimuli across all conditions.

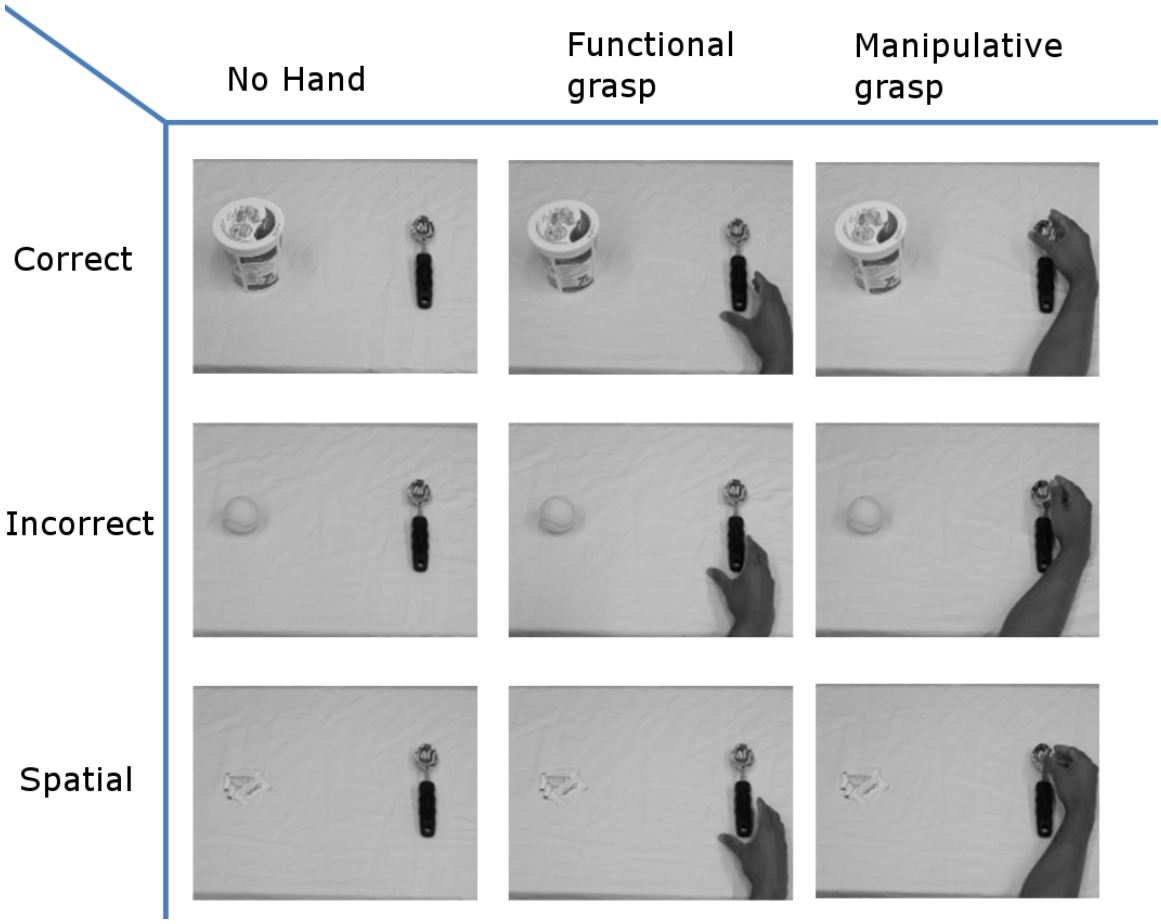

Figure 27: The stimuli were similar to Aim 1, Aim 2 and the prior behavioral data with the exception that the static hand conditions were not included in our hypotheses. Each column represents a particular grasp-posture and each row represents a particular tool-object context.

5.3.3 Experimental design

Each subject participated in two experiments within this study based on the latency of the stimuli. In the first experiment, the durations of the stimuli on screen were 100ms and in the second experiment, the durations were 500ms (Figure 28). Each stimulus was preceded first by a 2000ms fixation circle, followed by a 500ms warning cue (fixation cross) that the stimulus is upcoming. After image presentation, the fixation circle reappeared, continuing the cycle. Within each experiment, each subject viewed all the 180 images across all 9 conditions (20 tools placed across 9 contexts and grasps). To counter potential confounds of image familiarity, each stimulus was presented only once. Therefore, each subject viewed

20 stimuli per condition within each experiment. Within each experiment, the 180 images were randomized and presented across 3 blocks to provide adequate rest intervals for the subject. At the end of the first experiment, participants were given a longer break if needed. Participants were explicitly told to maintain fixation at the circle and cross whenever they appeared. Participants were told that they would see images of tool-object pairs with a tool defined as an entity in the right hemifield that could be used on the object in the left hemifield. Participants were instructed to evaluate whether tool-object pairs were functionally related or not and record their decision as accurately and quickly as possible by pressing a button on a response pad. To avoid the influence of the button press on parietofrontal and motor activity, the response hand was counterbalanced. Of the 28 subjects in experiment 1, 15 subjects used their right/left thumb to record an incorrect/correct tool-object relationship and the remaining 13 used their right/left thumb to record a correct/incorrect tool-object relationship. Of the 26 subjects in experiment 2, 13 subjects used their right/left thumb to record an incorrect/correct tool-object relationship and the other 13 used their right/left thumb to record a correct/incorrect tool-object relationship. Unlike Aim 1 and Aim 2 and similar to the prior behavioral work, this was an active experiment to ensure task compliance, especially given the rapid stimuli presentations. Prior to data acquisition, the images of tool-object pairs (without the grasp) were shown in a loop to the subjects to ensure that they were familiar with the tools and objects themselves. Any questions they might have had on the identity of a particular tool or object were addressed without giving away the relationship between the tool-object pair. The vast majority of the stimuli were familiar to the participants and the most common object that participants had difficulty recognizing was a tea strainer. However, once the identity of the tea strainer was revealed, participants were immediately able to recognize it. We followed this procedure to ensure that participants performed the task of evaluating tool-object content especially given that the stimuli durations were very short. A trial block of the 100ms experiment was also conducted to acclimatize the subjects with the speed of the stimuli to avoid any startle effect in the EEG data.

5.3.4 Data acquisition

To record neural activity, a 58 channel tin electrode EEG cap (Electrocap, Eaton, OH) in accordance with the international 10-20 system was used. The leads from each of the electrode on the EEG cap were connected to a data acquisition computer via Synamps 2 amplifiers (Neuroscan, Charlotte, NC). As well, two electrodes were placed above and below the left eye to record electrooculographic activity (EOG). Data acquisition was performed using a right ear reference at a sampling rate of 1000 Hz. The left ear was also recorded and was used offline in creating a linked ears reference. Participants behavioral responses were collected using a button box that was interfaced with the EEG data acquisition system using Stim 2 (Neuroscan, Charlotte, NC). The visual stimuli were presented on a 22" LCD screen at a distance of 3 feet from the participants center of field of view using custom scripts in Stim 2 (Neuroscan, Charlotte, NC). Given the geometry of the configuration, the viewing angles were therefore very similar to the experimental setup in Aim 2. The viewing angles were 29.5 degrees in the horizontal direction (left-right) 18.7 degrees in the vertical direction (up-down). The average horizontal separation between the tool-object centers were 18.6 degrees (1.3 degrees S.D). An infrared eye tracker (Gazepoint GP3, British Columbia, Canada) was used to measure eye movements across the 22" monitor at a sampling rate of 60Hz and was controlled using an additional laptop. The eye tracker had a tracking resolution ¡0.01 degrees, had a gaze position accuracy between 0.5 - 1 degree and generated an ascii file of the time course of gaze position over the 22" monitor. Unlike Aim 2, this was not a head mounted eye tracker and the infrared detectors was placed right below the monitor and was positioned to track the eyes. Using Gazepoint software, calibration was done by having the participants foveate at 9 predetermined locations corresponding to the center of the screen and the edges and midpoints of the sides of a rectangle along the boundary of the screen. Each experimental block started with the appearance of a blue box at the left edge of the screen that the participant was instructed to fixate on. The saccade to the box was recorded in both the EOG waveform trace and in the eye tracking data from the GP3 eye tracker and this was used to synchronize the data streams (EEG+behavior data and eye movement data) offline.

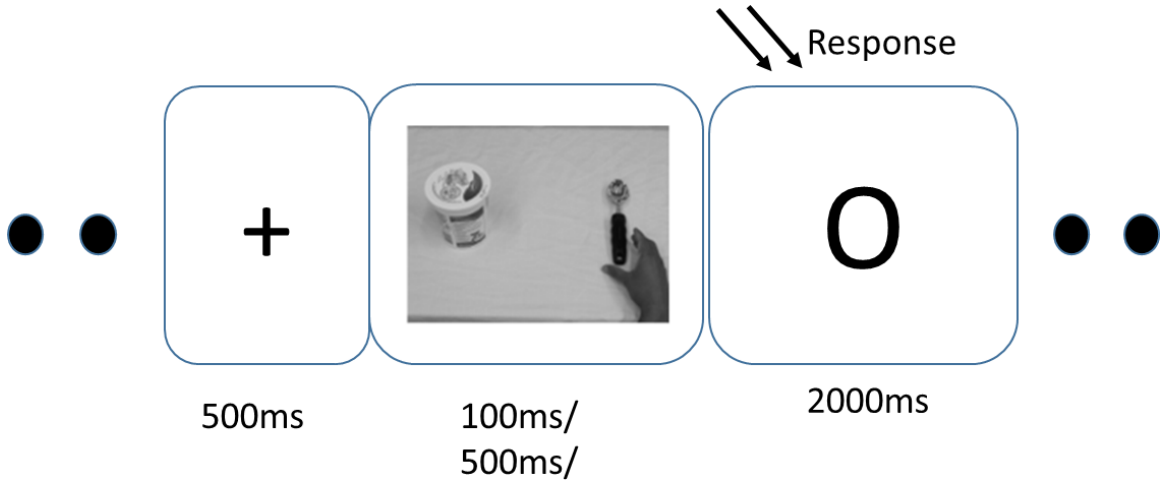

Figure 28: The experimental design and the timing of events in a trial

5.3.5 Data analyses

5.3.5.1 Behavioral response analyses

A response indicating that the correct tool-object context was functionally related was considered accurate. A response indicating that the incorrect and spatial tool-object contexts were not functionally related was considered accurate. Though the spatial context is ambiguous by design, the tool-object pairs are technically not functionally related. The accuracy for each condition within each subject was computed by diving the number of accurate responses to the total number of trials. To compute mean behavioral response times for each subject, trials with an inaccurate response within the correct and incorrect contexts were discarded. Given the ambiguous nature of the spatial context, no trials were discarded based on accuracy as this would cause a large loss in data. Within each condition, response latencies greater than 2 S.D from the mean were discarded as outliers. The mean response latency was computed from the remaining trials. Each subject therefore had 9 mean response latencies corresponding to each of the 9 conditions.

5.3.5.2 EEG data preprocessing

The raw EEG data was first low pass filtered from DC-30Hz. An epoch was then constructed by extracting data from 1000ms before to 1500ms after tool-object image onset. Each epoch

was therefore 2500 ms long (-1000 1500ms), with 0 corresponding to appearance of the tool-object image on screen and -500 corresponding to presentation of the cross indicating that an image was pending. Within the correct and incorrect contexts, trials that had inaccurate responses or response times greater than 2 S.D from the mean response time were discarded. Within the spatial context, trials were not discarded based on accuracy as this would cause a large loss in data given the ambiguous nature of the spatial context; however trials that had response times greater than 2 S.D from the mean response time were discarded. Using custom scripts in MATLAB (Mathworks, MA), the extracted and pruned epochs were then linear detrended and baseline corrected to the first 500ms (-1000 to -500ms), until onset of the warning cue. EEG epochs were then artifact corrected for eye movement and eye blink using the Recursive Least Squares algorithm (He et al., 2004, Gmez-Herrero, 2007, Natraj et al., 2013). Given the nature of the rapid stimuli, it is possible that the musculature around subjects' head might be tensed in anticipation (though subjects were instructed to be relaxed) as they focused on performing the task. Therefore, additional artifact correction was carried using the Independent Component Analysis (ICA) algorithm to remove residual noise and muscle-related artifacts (Delorme and Makeig, 2004, Hyvrinen et al., 2004). Typically noise factors out very well as it is independent of task related activity. For each subject, noise components were manually identified and removed (Figure 29).

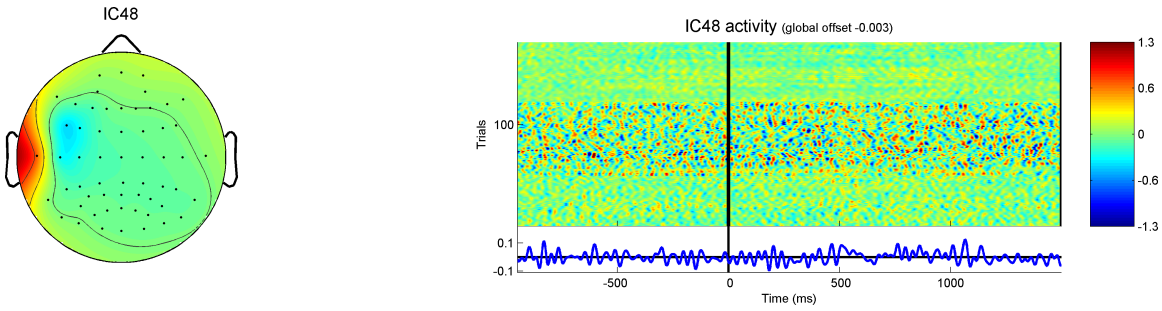

Figure 29: Correction of a muscle artifact component that is spatially located at the muscle around the left ear as shown in the head plot (left). It can be seen in the figure on the right that the activity of this component across trials was in general very noisy and consistent through time i.e. not dependent on the stimuli.

5.3.5.3 Eye movement data preprocessing

Fixation and saccades were determined using a twofold approach. First, given the requirement to precisely identify saccades, time-points immediately preceding deflections greater than 30uV in the EOG signal magnitude after image-onset were used to record saccade initiation (Kirchner and Thorpe, 2006). Next, these computed saccade initiation time-points were verified using the output of the GP3 eye tracker that generated ascii files of gaze position and automatically identified fixations. The sampling resolution of the GP3 eye tracker (60Hz) was much lower than that of the EOG signal (1Khz), so to overcome this difference, fixation onset as detected by the eye tracker were overlaid on the horizontal EOG waveform trace and saccade-onset as detected by the 30uV deflection was manually verified for each epoch. This approach is similar to the manual marking of movement onset in electromyography (EMG) signals (Mizelle et al., 2011). The corrected time-course of the eye tracking data was used to determine foveal gaze position by constructing Areas of Interest (AOI) in the stimuli by fitting a rectangle to the boundaries of the discrete scene features similar to Aim 2 (Natraj et al., 2015). The AOI were the object in the left hemifield, the manipulative tool-end and the functional tool-end. As we used the same stimuli as Aim 2 with roughly the same viewing angle, the AOI dimensions were roughly identical in both experiments. The average width of tools and objects was 4.5 degrees (2.04 degrees S.D) and the average height of tools and objects was 5.6 degrees (2.3 degrees S.D). With respect to the demarcation of the functional and manipulative tool-end, on average 34.2% (6.1% S.D) of the length of the tool from the top constituted the length of the manipulative tool-end and the bottom 65.8% (6.1% S.D) of the tool constituted the length of the functional tool-end. The width of the functional tool-end or manipulative tool-end increased in the presence of a functional or manipulative grasp-posture to accommodate the fingers and upper-wrist that were positioned next to the tool. It should be noted that the exact position of the grasp-posture was not controlled for across images and conformed more to the tool shape and size. Slack was given to the AOI to account for jitter in gaze position over the AOI, noise in the eye tracking system (Ambrosini et al., 2011) and to account for the fact that the observer may not fully enter an AOI given the rapid stimulus duration times.

5.3.5.4 Statistical analyses

Given the rapid stimuli presentations, we aimed to evaluate precise spatiotemporal patterns of neural activity in response to the images. Similar to Aim 1, we performed a within-subject parametric statistical test (repeated measures ANOVA) on the mean ERP signals. However, in Aim 1, we had averaged data across time-windows of interest which spanned 100-200ms and a priori only chose a subset of the 58 recording channels corresponding to bilateral fronto-parietal and left motor electrodes. While this approach was well suited to Aim 1 wherein the image was on screen for 2000ms, it is not designed to capture statistical differences with much finer spatial and/or temporal resolution. To this end, we employed a mass-univariate approach and performed the RM-ANOVA test (with post-hoc paired t-tests) at each and every electrode and time-point (channel-time pair (Pernet et al., 2011)) using the framework of the General Linear Model (GLM), in a manner similar to the analyses of large-scale functional neuroimaging data (Smith, 2014) using statistical functions in the LIMO and Fieldtrip toolboxes (Oostenveld et al., 2010, Pernet et al., 2011). Thus while we had performed only 3 ANOVA tests on mean ERP in Aim 1 (within three global time-windows and across 5 gross regions of interest), 136,358 ANOVA tests were performed here for every channel-time pair (58 channels, 2351 time points, -900ms to 1450ms with respect to image onset). Thus unlike Aim 1, we did not predetermine regions or time-points of interest and evaluated whole brain differences at each and every time-point. At the single subject level, the evoked responses for each condition were estimated as follows. First, trials that elicited an inaccurate response (within the correct and incorrect tool-object contexts) and trials wherein behavioral response latencies were greater than 2 S.D from the mean response were discarded. The evoked response at the remaining trails were then modeled using the following equation

$$Y = XB + E \tag{3}$$

In the above equation Y represented the matrix of all the trials at an electrode. There were as many rows in Y as there were trials and the number of columns corresponded to the

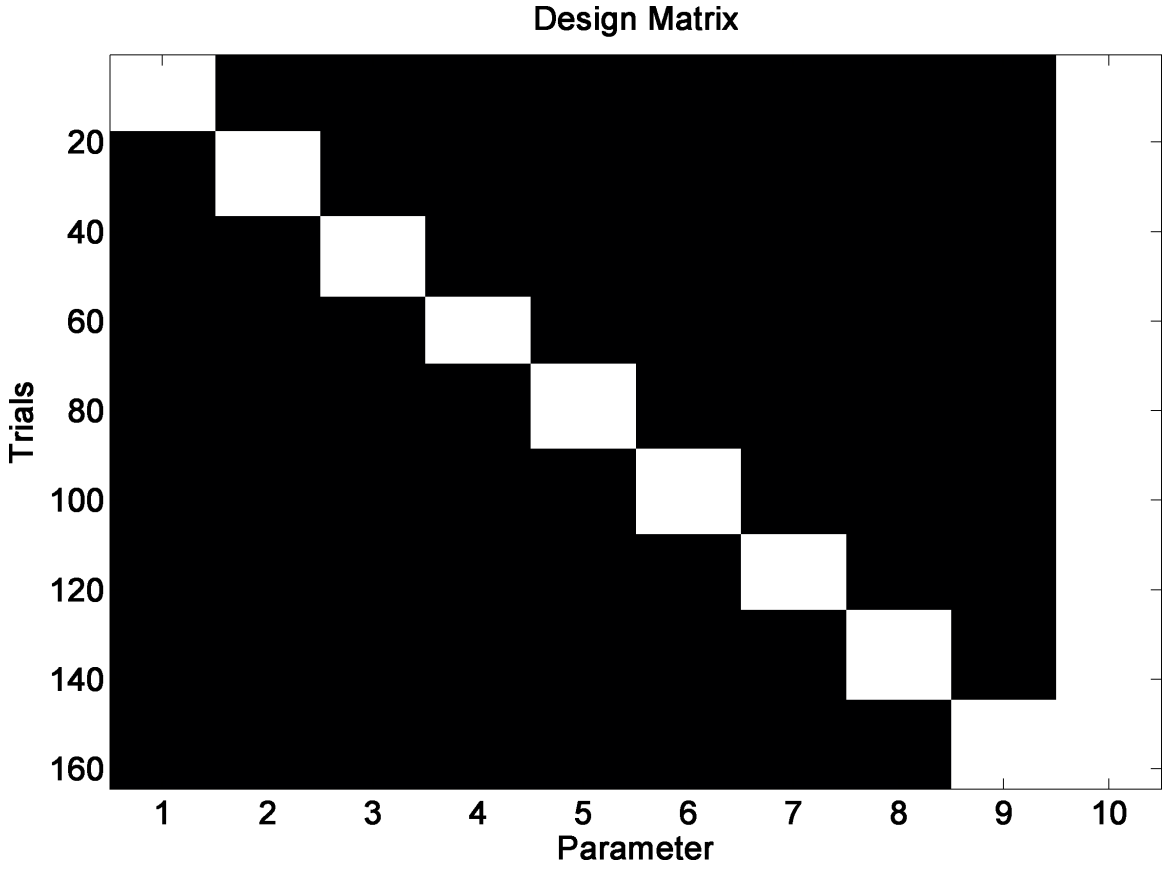

Figure 30: The design matrix X used in the GLM at a particular electrode. White signifies ones and black signifies zeros. There are as many rows in X as there are trials. The first 9 columns correspond to the condition-sorted trials of Y in Eq(3). The last column is all ones and is used to code the average activation across trials of all conditions.

time points in each trial. The rows of Y were sorted such that the first n_1 rows corresponded to the trials of condition 1, the next n_2 rows corresponded to the trials of condition 2 and so on. X was the design matrix and had as many rows as the number of trials. In the experimental design here, X has 10 columns, one for each condition and the last column to model the average activation across all 9 conditions. The first n_1 rows of column 1 were all 1 and represented the parametrized coding of the evoked response of condition 1 whereas the first n_1 rows of columns 2 through 9 were 0. Similarly, the next n_2 rows of column 2 were all 1 and represented the parametrized coding of the evoked response of condition 2 and so on. The last column of X was all 1 to represent the parameterized coding of the average effect across all trials of all conditions. B is the matrix of parameters (beta matrix in the fMRI literature) with as many rows as the number of parameters (10 in this study)

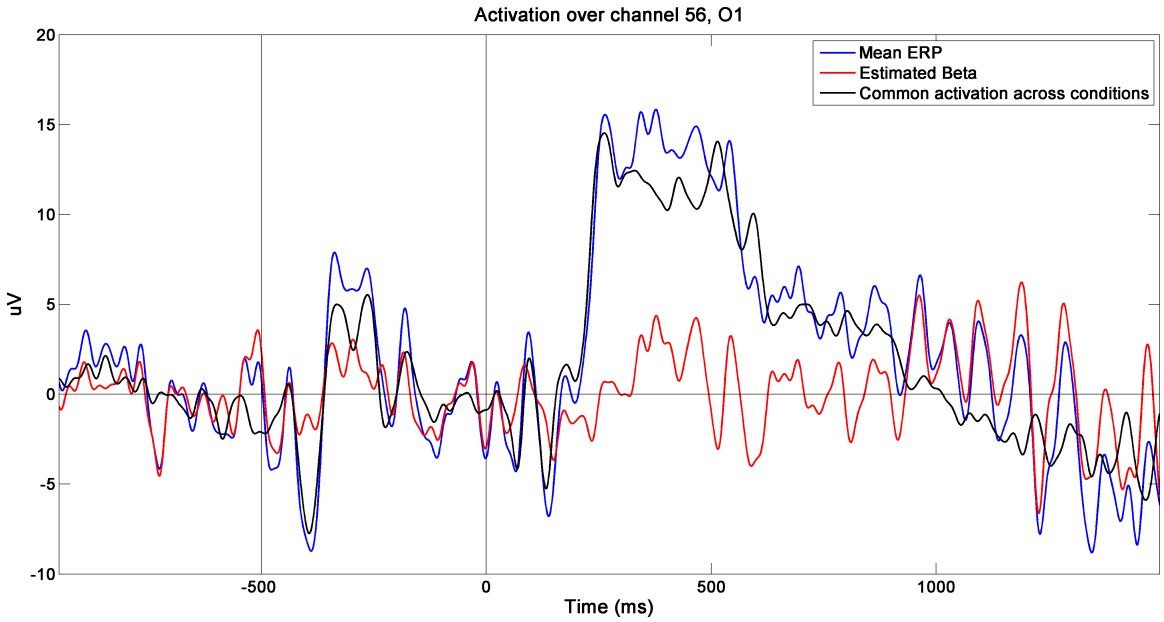

Figure 31: A plot comparing the parameter estimate (betas) with the trial average (mean ERP). -500ms corresponds to presentation of the fixation cross and 0 corresponds to the presentation of the stimulus. The parameter estimates aims to isolate activity that is condition specific whereas the mean ERP contains the common activation across all conditions in addition to the condition specific activation.

and with as many columns as the number of time points. The matrix E is of the same size as Y and denotes the error in the GLM. Each row of B therefore gives a parametric model of the evoked response unique to a specific condition and is obtained by the least squares solution given in the following equation.

$$B = (X^\top X)^{-1} X^\top Y \tag{4}$$

The advantage of estimating the parametric evoked response matrix B (the betas) over computing the trial averages is that the estimated parameters capture the within-subject variability from the individual trials and isolate the response unique to the particular condition. An example of the design matrix and parametric response are shown in Figures 30 and 31.

Each subject therefore had a parametric estimate of the evoked response for each of the 9 conditions at each of the 58 electrodes. Between-subject statistics were then performed at the group level at every electrode-time pair. A 2-way repeated measures within-subject

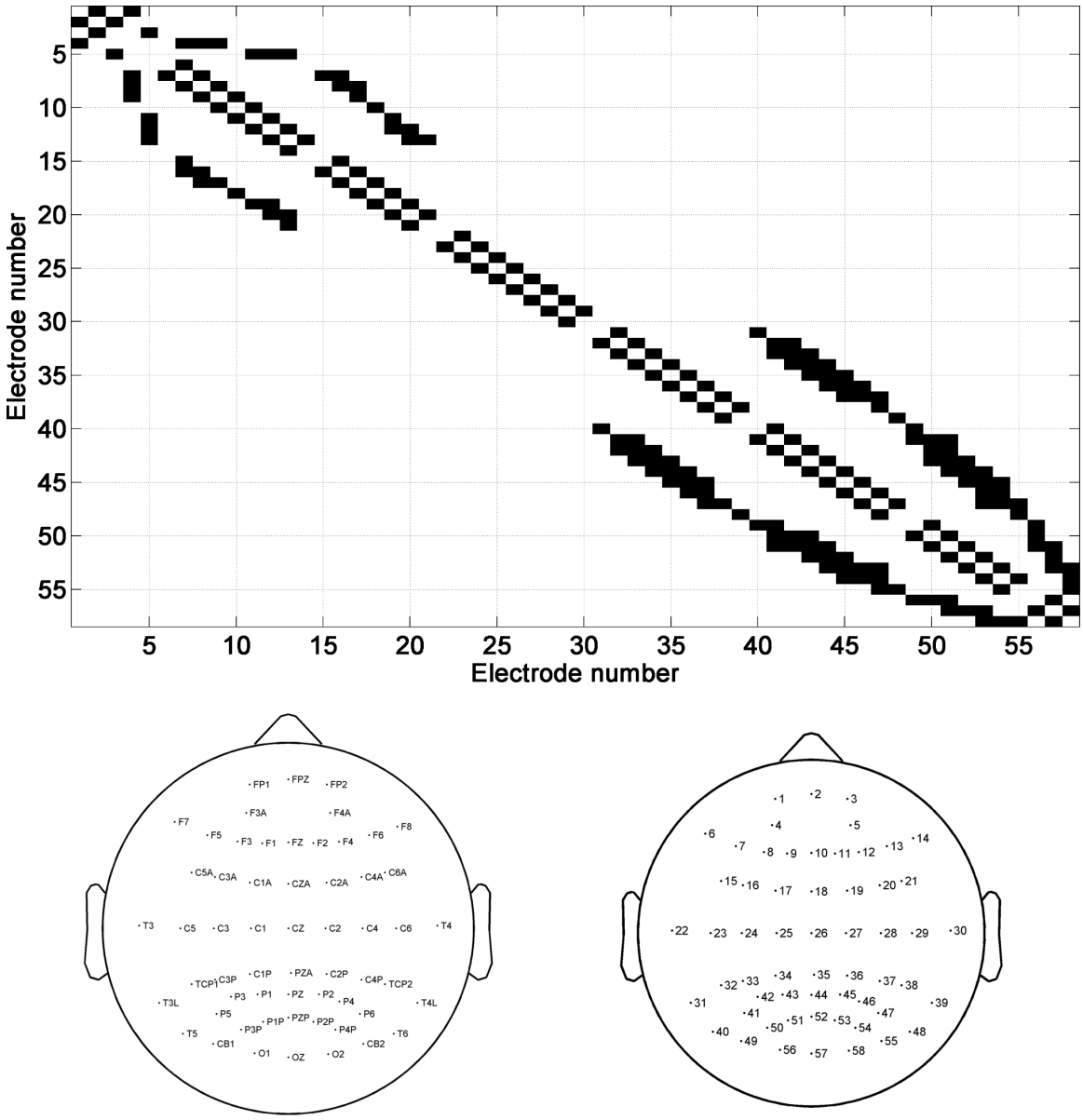

Figure 32: Matrix identifying electrode neighbors when performing multiple comparison corrections for the mass-univariate RM-ANOVA tests, using spatiotemporal cluster statistics. A black square at any row and column denotes that the electrode in that particular row/column is a neighbor of the electrode along the corresponding column/row. To reiterate, there are 58 recording channels and the correspondence between the channel number, name and position is also shown.

ANOVA was performed on the parameter estimates at each electrode and each time-point with factors Context (correct, incorrect, spatial) and Hand (no hand, functional, manipulative). Given the extremely large number of ANOVA tests, multiple comparisons is necessary to mitigate the probability of Type I errors. To this end, non-parametric spatiotemporal

cluster corrections were carried out (Maris and Oostenveld, 2007, Oostenveld et al., 2010, Pernet et al., 2011) which tend to be less conservative than the Bonferroni. Briefly, after computing an F- value at every channel-time pair, spatiotemporally contiguously significant F-values were aggregated and summed together to computer a cluster statistic. The significance of this cluster statistic was tested against the null distribution of cluster statistic values generated by 1000 random permutations of the dataset. Significance threshold for the cluster statistic was assessed at the 0.05 level. Electrode neighbors were identified based on the geometry of the 10-20 system (Homan et al., 1987) following the procedures recommended in (Pernet et al., 2011) and the matrix denoting neighboring electrodes is depicted in Figure 32. Spatiotemporal cluster correction therefore assumed that a significant statistical effect would be present both across spatially contiguous electrodes and across temporally contiguous time-points. Wherever there were significant effects of the ANOVA, post hoc paired t-tests were performed at all channel-time combinations that generated the significant effect. Similar to Aim 1, the no hand condition served as control for the manipulative grasp and functional grasp conditions individually. With respect to Context, the correct context served as control for the incorrect and spatial tool-object contexts individually. Significance for all the paired t-tests when inspecting a particular effect was assessed at the 0.05 level, with FDR correction for multiple comparison/Type I error (Genovese et al., 2002). All statistical tests were performed in MATLAB (The Mathworks, Boston, MA) and in SPSS (The IBM Corp).

5.4 Results

5.4.1 Behavioral response accuracy

With respect to the 100ms experiment, across all participants, the mean accuracy was 82.2 $\pm$ 1.69% (S.E), and the mean accuracy within each of the 9 conditions is depicted in Figure 33. Across conditions and participants, there were no differences in mean accuracy between the left/right button-press responses ($t(250) = 1.0396$, $p=0.3$); the data were thus sufficiently counterbalanced to the response hand. Data were then subsequently pooled and

entered into a 2-way repeated measures ANOVA with factors Context (correct, incorrect, spatial) and Hand (no hand, functional grasp, manipulative grasp). Results revealed a main effect of Context ($F(2,54) = 19.894$, $p < 1 \times 10^{-5}$) but no main effects of Hand ($F(2,54) = 0.521$, $p=0.597$) or any interaction effect ($F(4,108) = 0.424$, $p=0.791$). Post-hoc Bonferroni corrected paired t-tests revealed that participants were most accurate within the incorrect context when compared to the correct context (by $14.393 \pm 1.493\%$ (S.E), $t(83) = 9.641$, $p< 1 \times 10^{-5}$, $d_z = 1.05$) and spatial context (by $24.25 \pm 2.19\%$ (S.E)).$t(83) = 11.073$, $p< 1 \times 10^{-5}$, $d_z = 1.208$). Of the latter two, participants were more accurate in the correct context than the spatial context (by $9.86 \pm 3.02\%$ (S.E)).$t(83) = 3.258$, $p=0.0016$, $d_z = 0.355$).

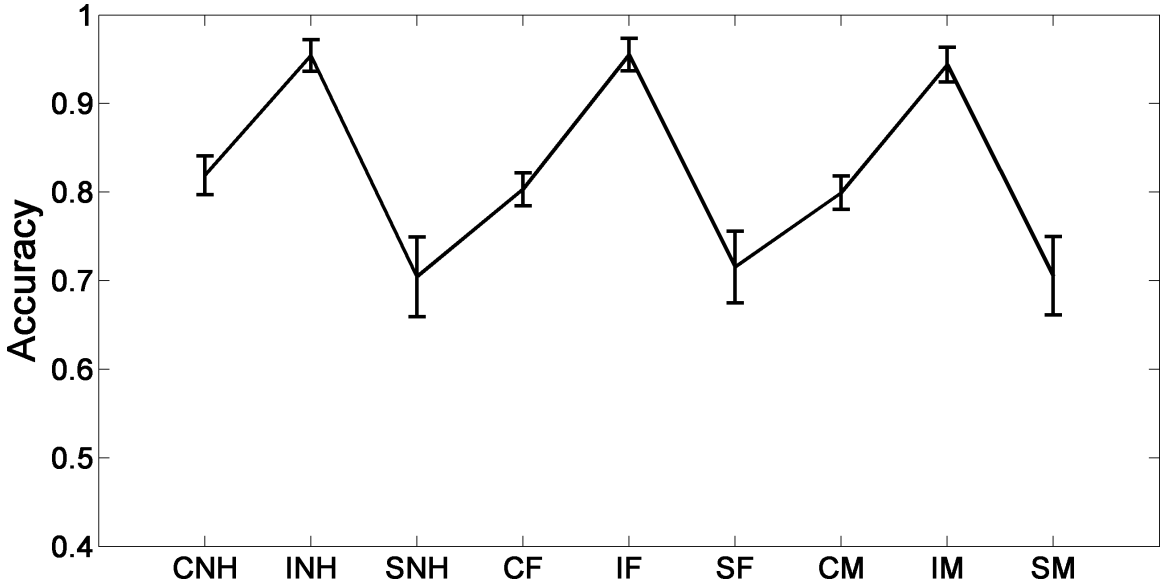

Figure 33: Mean accuracies and S.E bars for each of the 9 conditions within the 100ms experiment. C/I/SNH: correct/incorrect/spatial context-no hand. C/I/SF: correct/incorrect/spatial context-functional grasp. C/I/SM: correct/incorrect/spatial context-manipulative grasp

With respect to the 500ms experiment, across all participants, the mean accuracy was $84.16 \pm 1.53\%$ (S.E), and the mean accuracy within each of the 9 conditions is depicted in Figure 34. Across conditions and participants, there were no differences in mean accuracy between the left/right button-press responses ($t(232) = 0.2384$, $p=0.8118$); the data were thus sufficiently counterbalanced to the response hand. There were no differences in

accuracy between the two experiments, (t(484) = 1.15, p=0.2492). Data were then subsequently pooled and entered into a 2-way repeated measures ANOVA with factors Context (correct, incorrect, spatial) and Hand (no hand, functional grasp, manipulative grasp). Results revealed a main effect of Context (F(2,50) = 15.445, p$< 1 \times 10^{-5}$) and an interaction effect (F(4,100) =3.507, p=0.0101) but no main effects of Hand (F(2,50) = 0.265, p=0.768). Post-hoc Bonferroni corrected paired t-tests revealed that participants were most accurate within the incorrect context when compared to the correct context (by 13.369 ± 1.336% (S.E), t(77) = 10.009, p$< 1 \times 10^{-5}$, d_z = 1.133) and spatial context (by 25.47 ± 2.67% (S.E)).t(77) = 9.545, p$< 1 \times 10^{-5}$, d_z = 1.08). Of the latter two, participants were more accurate in the correct context than the spatial context (by 12.1 ± 3.6% (S.E)), t(77) = 3.364, p=0.0011, d_z = 0.381). With respect to the interaction effect, data revealed that the accuracy in response to the correct context-no hand condition was higher than the correct context-functional grasp condition (by 3.82 ± 1.573% (S.E), t(25) = 2.426, p=0.023, d_z = 0.476) and correct context-manipulative grasp condition (3.691 ± 1.54% (S.E), t(25) = 2.396, p=0.024, d_z = 0.47).

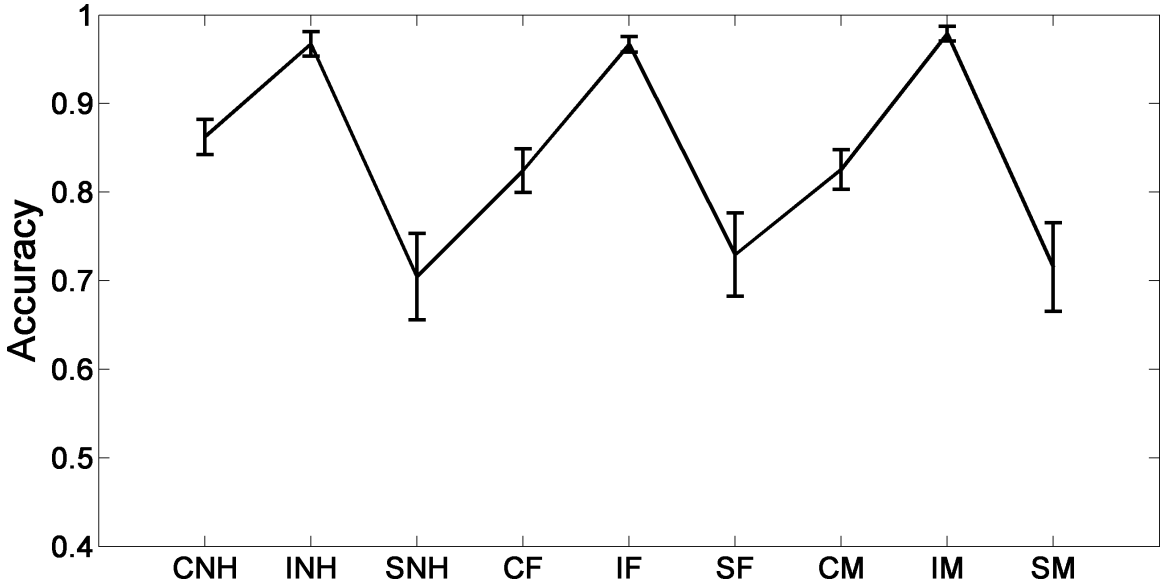

Figure 34: Mean accuracies and S.E bars for each of the 9 conditions within the 500ms experiment. C/I/SNH: correct/incorrect/spatial context-no hand. C/I/SF: correct/incorrect/spatial context-functional grasp. C/I/SM: correct/incorrect/spatial context-manipulative grasp

Overall, data showed that participants were highly accurate in both experiments though the stimuli durations were rapid. There were no statistical difference in accuracy between the two experiments, nor was there any influence of the response hand. The most salient result was the fact that in both experiments, participants were very highly accurate in identifying incorrect tool-object contexts ($> 90\%$). These results confirm our first hypothesis that participants would be accurate in the task even though stimuli durations are very short.

5.4.2 Behavioral response latency

With respect to the 100ms experiment, results showed that across all conditions and subjects, there were no statistical differences between left/right hands on mean response latencies ($t(250) = 1.7103$, $p = 0.089$); response latencies were therefore sufficiently counterbalanced. Subsequently, participants mean response latencies (Figure 35) were analyzed using the 2-way RM-ANOVA with factors Context (correct, incorrect, spatial) and Hand (no hand, functional grasp, manipulative grasp). Results revealed a main effect of Context ($F(2,54) = 14.649$, $p< 1 \times 10^{-5}$), a main effect of Hand ($F(2,54) = 22.231$, $p< 1 \times 10^{-5}$) but no interaction effect ($F(4,108) = 1.523$, $p=0.201$). Post-hoc Bonferroni corrected paired t-tests showed that the main effect of Context derived from the fact that spatial context elicited longer response times than the correct context (by 55.567 ± 9.98ms (S.E), $t(83) = 5.566$, $p< 1 \times 10^{-5}$, $d_z = 0.607$) and incorrect context (by 63.8 ± 7.96ms (S.E), $t(83) = 8.006$, $p< 1 \times 10^{-5}$, $d_z = 0.873$), while the latter two did not differ from each other ($t(83) = 0.865$, $p=0.390$). With respect to the main effect of Hand, results showed that the manipulative grasp conditions elicited longer response times than the no hand conditions (by 47.7 ± 7.08ms (S.E), $t(83) = 6.731$, $p< 1 \times 10^{-5}$, $d_z = 0.734$) and functional grasp conditions (by 30.54 ± 7.134ms (S.E), $t(83) = 4.28$, p=4.98e-5, $d_z = 0.467$). Among the latter two, the no hand conditions elicited slightly quicker response times than the functional grasp conditions (by 17.164 ± 6.587ms (S.E), $t(83) = 2.605$, $p=0.011$, $d_z = 0.283$).

With respect to the 500ms experiment, results showed that across all conditions and subjects, there were no statistical differences between left/right hands on mean response

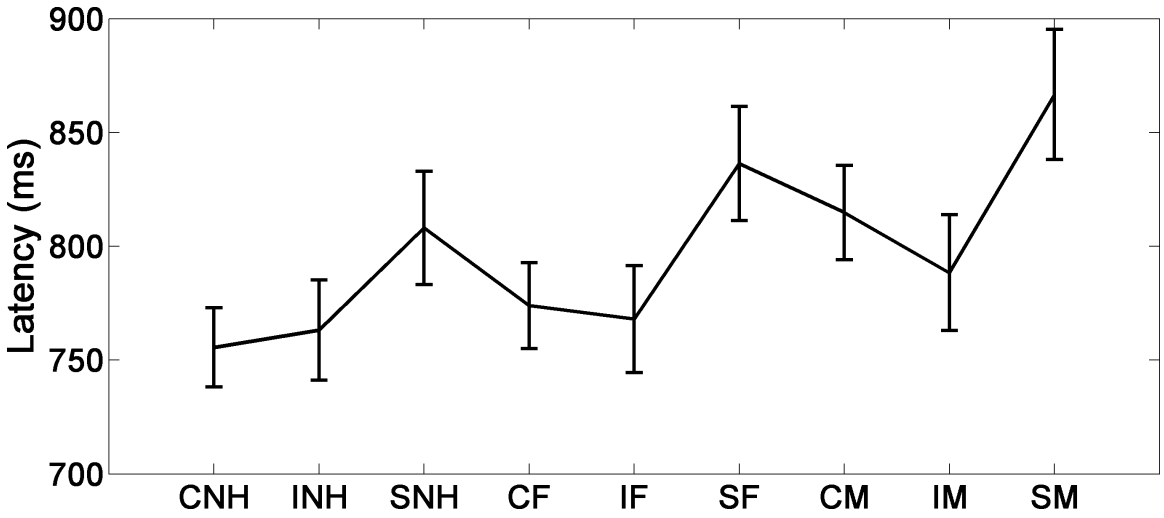

Figure 35: Mean latencies and S.E bars for each of the 9 conditions within the 100ms experiment. C/I/SNH: correct/incorrect/spatial context-no hand. C/I/SF: correct/incorrect/spatial context-functional grasp. C/I/SM: correct/incorrect/spatial context-manipulative grasp

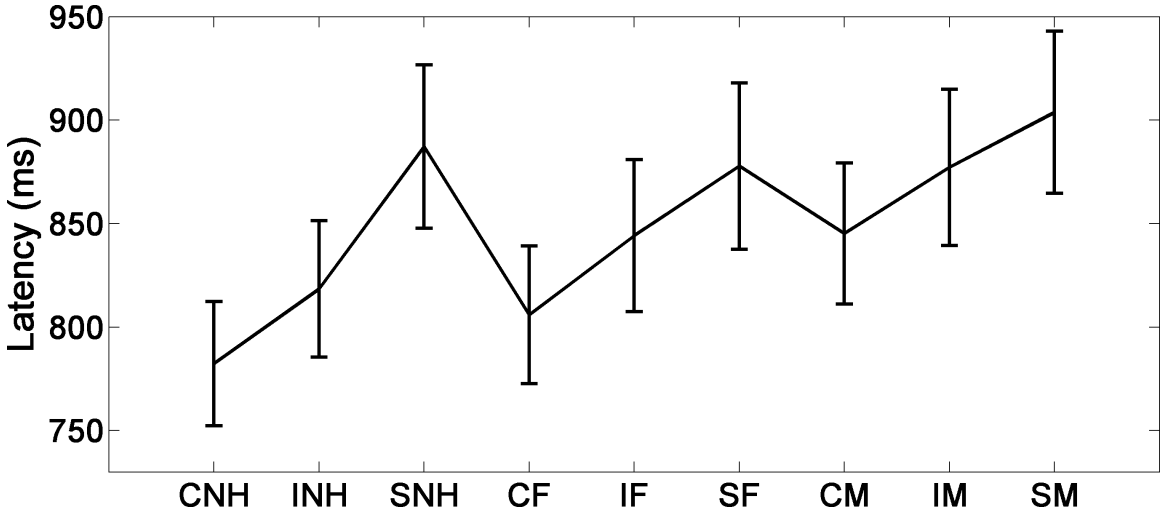

Figure 36: Mean latencies and S.E bars for each of the 9 conditions within the 500ms experiment. C/I/SNH: correct/incorrect/spatial context-no hand. C/I/SF: correct/incorrect/spatial context-functional grasp. C/I/SM: correct/incorrect/spatial context-manipulative grasp

latencies ($t(232) = 1.573$, $p = 0.1171$); response latencies were therefore sufficiently counterbalanced. In addition, across all participants and conditions, the average response time in the 500ms experiment (849.11 $\pm$ 185.13ms S.D) was significantly slower ($t(484) = 3.63$, $p=3.04\times10^{-4}$, $d_s = 0.33$) than the 100ms experiment (797.22 $\pm$ 125.69ms S.D). Subsequently, participants mean response latencies (Figure 36) were analyzed using the 2-way

RM-ANOVA with factors Context (correct, incorrect, spatial) and Hand (no hand, functional grasp, manipulative grasp). Results revealed a main effect of Context ($F(2,50) = 9.136$, $p < 1 \times 10^{-5}$), a main effect of Hand ($F(2,50) = 16.714$, $p < 1 \times 10^{-5}$) but no interaction effect ($F(4,100) = 1.898$, $p=0.117$). Post-hoc Bonferroni corrected paired t-tests showed that the main effect of Context derived from the fact that spatial context elicited longer response times than the correct context (by 78.48 ± 13.08ms (S.E), $t(77) = 5.996$, $p < 1 \times 10^{-5}$, $d_z = 0.679$) and incorrect context (by 43.038 ± 10.03ms (S.E), $t(77) = 4.29$, $p=5.13\times10^{-5}$, $d_z = 0.486$). Amongst the latter two conditions, the correct context elicited quicker response times than the incorrect context (by 35.44 ± 13.83ms (S.E), $t(77) = 2.563$, $p=0.012$, $d_z = 0.29$). With respect to the main effect of Hand, results showed that the manipulative grasp conditions elicited longer response times than the no hand conditions (by 46.076 ± 9.36ms (S.E), $t(77) = 4.92$, $p < 1 \times 10^{-5}$, $d_z = 0.734$) and functional grasp conditions (by 32.76 ± 7.294ms (S.E), $t(77) = 4.492$, $p=4.98\times10^{-5}$, $d_z = 0.467$), while the latter two did not differ from each other ($t(77) = 1.741$, $p=0.086$).

Overall, results here recreated our prior behavioral data (Borghi et al., 2012), thereby supporting our hypothesis on response latencies. Specifically, in both experiments the manipulative grasp condition, and the spatial tool-object context elicited the longest response times. There was no influence of the response hand on overall mean response times in both experiments. In addition, the 500ms experiment elicited longer response times than the 100ms experiment.

5.4.3 Saccades

Within the 100ms experiment, participants initiated saccades on average in 9.96 ± 1.6788 % S.E of trials. Within no trial did participants initiate a saccade within 100ms, similar to previously documented research on the fastest saccade initiation time (Kirchner and Thorpe, 2006). Within the 500ms experiment, participants initiated saccades on average in 74.88 ± 5.05 % S.E of trials. Therefore, participants were nearly 7.5 times more likely to initiate saccades in the 500ms experiment when compared to the 100ms experiment. This result validates our hypothesis that without any active instructions, participants would naturally

inhibit saccades in the 100ms experiment with the reemergence of saccades in the 500ms experiment.

5.4.4 EEG results

In both the 100ms and 500ms experiments, the mass-univariate RM ANOVA analyses revealed a main effect of Hand, a main effect of Context but no interaction effect. In the following sections, we focus on the main effects individually across both experiments.

5.4.4.1 Main effects of Hand

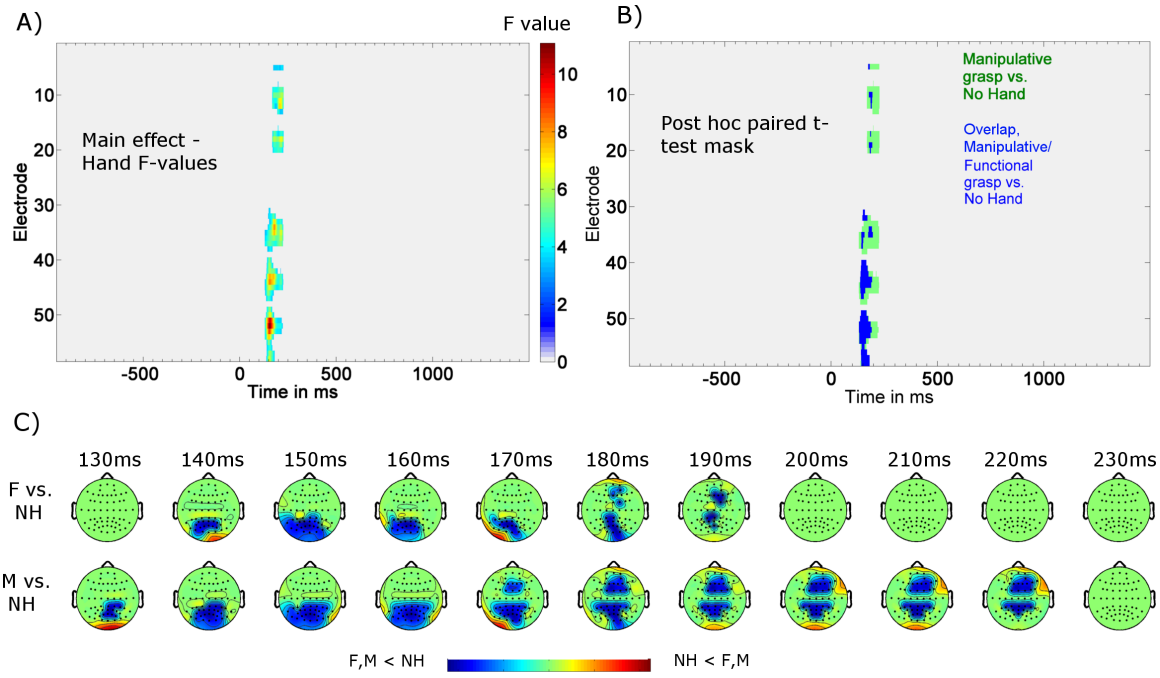

Figure 37: A) Statistical image depicting the F-values at channel-time pairs that exhibited a significant main effect of factor Hand (no hand, functional grasp, and manipulative grasp) from the RM-ANOVA test at the 0.05 level in the 100ms experiment. The F-values were corrected for multiple comparisons using spatiotemporal cluster statistics. B) Statistical image depicting the mask of the post-hoc paired t-tests that were performed at the significant channel-time pairs in A) contrasting the manipulative grasp and functional grasp condition individually to the no hand control (FDR corrected). Mask color represents channel-time pairs when both contrasts were significant (overlap) or when only one contrast was significant. C) Projection of normalized, significant t-values from the mask in B) onto the scalp electrodes at a few discrete time intervals, with a color bar representing the direction of the statistical difference. NH: No Hand, F: Functional grasp, M: Manipulative grasp. The actual t-values for the contrasts at every channel-time pair are shown in Videos 2 and 3.

With respect to the 100ms stimuli durations experiment, the results of the main effects

of the factor Hand at significant channel-time pair are shown in Figure 37. In Figure 37A, the F-values of significant channel-time pairs that exhibited the main effect surviving multiple comparisons (spatiotemporal cluster correction, 0.05 level) are depicted as a statistical image. At each significant channel-time pair, post-hoc paired t-tests on mean ERPs were performed, individually contrasting the functional grasp condition and manipulative grasp condition with the no hand condition as control (similar to Aim 1 (Natraj et al., 2013)). Given the many thousands of t-tests, the p-value was corrected using the FDR procedure at the 0.05 level as multiple comparison correction for Type I error (Genovese et al., 2002). A mask of the significant results of the t-tests that passed the FDR corrected threshold is plotted in Figure 37B and the projection of the mask on the scalp electrodes is shown in Figure 37C at a few discrete time-points. The actual t-values from the paired t-tests at every millisecond and across all electrodes for the significant channel-time pairs are shown in Video 2 for the functional grasp condition vs. the no hand condition contrast and in Video 3 for the manipulative grasp condition vs. the no hand condition contrast. Results showed the existence of a negative oriented cortical potential difference (the N100 potential) arising from centro-right parietal regions at 130ms post image onset and this difference subsequently propagated to centro right frontal regions till 228ms post image onset. It can be seen in Figure 37C and Videos 2 and 3 that only the manipulative grasp condition elicited a more prolonged and stronger centro-right ERP difference with respect to the no hand condition. The propagating potential is highlighted in the grand-average ERP plots (Figure 38) over a few select electrodes. The parietofrontal processing of the manipulative grasp-posture in the 100ms experiment refuted our hypothesis that the absence of saccades would negate right frontal activity. However, rather than appear around 400-600ms post image onset, similar to Aim 1, results revealed early ERP differences (130 to 228ms) characterized by a greater negative oriented cortical potential (the N100) originating at posterior electrodes and subsequently propagating to frontal electrodes. Importantly, a similar ERP difference was not observed for the functional grasp-posture.

With respect to the 500ms stimuli durations experiment, Figure 39A depicts the statistical image representing the main effects of factor Hand across all statistically significant

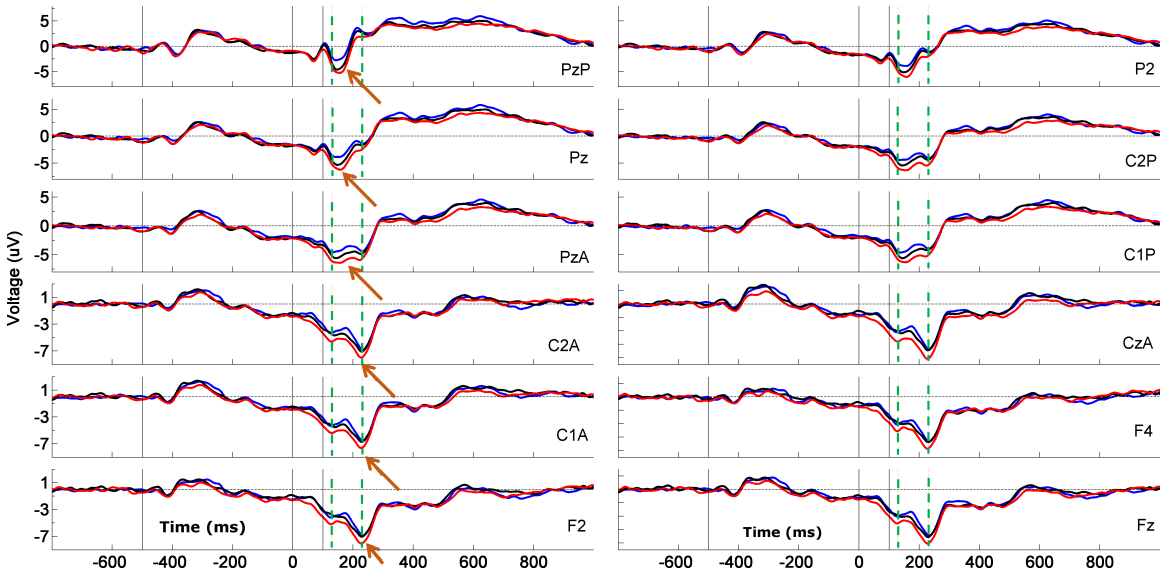

Figure 38: Plots depicting grand-average ERPs for the three Hand conditions: **manipulative grasp**, **functional grasp**, and the **no hand** control at a few centro-right parietofrontal electrodes in the 100ms experiment. For visualization of parieto-frontal propagation, the N100 is marked with an arrow in the first column and the vertical dotted green lines mark 130 to 228s post image onset (0 being image onset in the x-axis), the time period characterized by the main effect of Hand.

channel-time pairs that survived multiple comparisons correction (0.05 level). Figure 39B depicts the mask of the significant, FDR corrected (0.05 level), paired post-hoc t-tests at the significant channel-time pairs that exhibited the main effect. Similar to the 100ms experiment and Aim 1 (Natraj et al., 2013), the no hand condition served as control for the functional grasp-posture and the manipulative grasp-posture. Similar to Figure 38, the t-values in the post-hoc mask are projected onto the scalp-electrode space in Figure 39E. As the 500ms experiment was characterized by saccades, Figure 39C depicts the histogram or discrete probability densities associated with the initiation latencies of the first and second saccades across all conditions. Figure 39D depicts mean probable gaze position for the manipulative grasp condition and the no hand condition. The grand-average ERP themselves in the 500ms experiment for the three Hand conditions are shown in Figure 40 for a few significant electrodes.

It can be seen in Figure 39B that the statistical difference between the manipulative

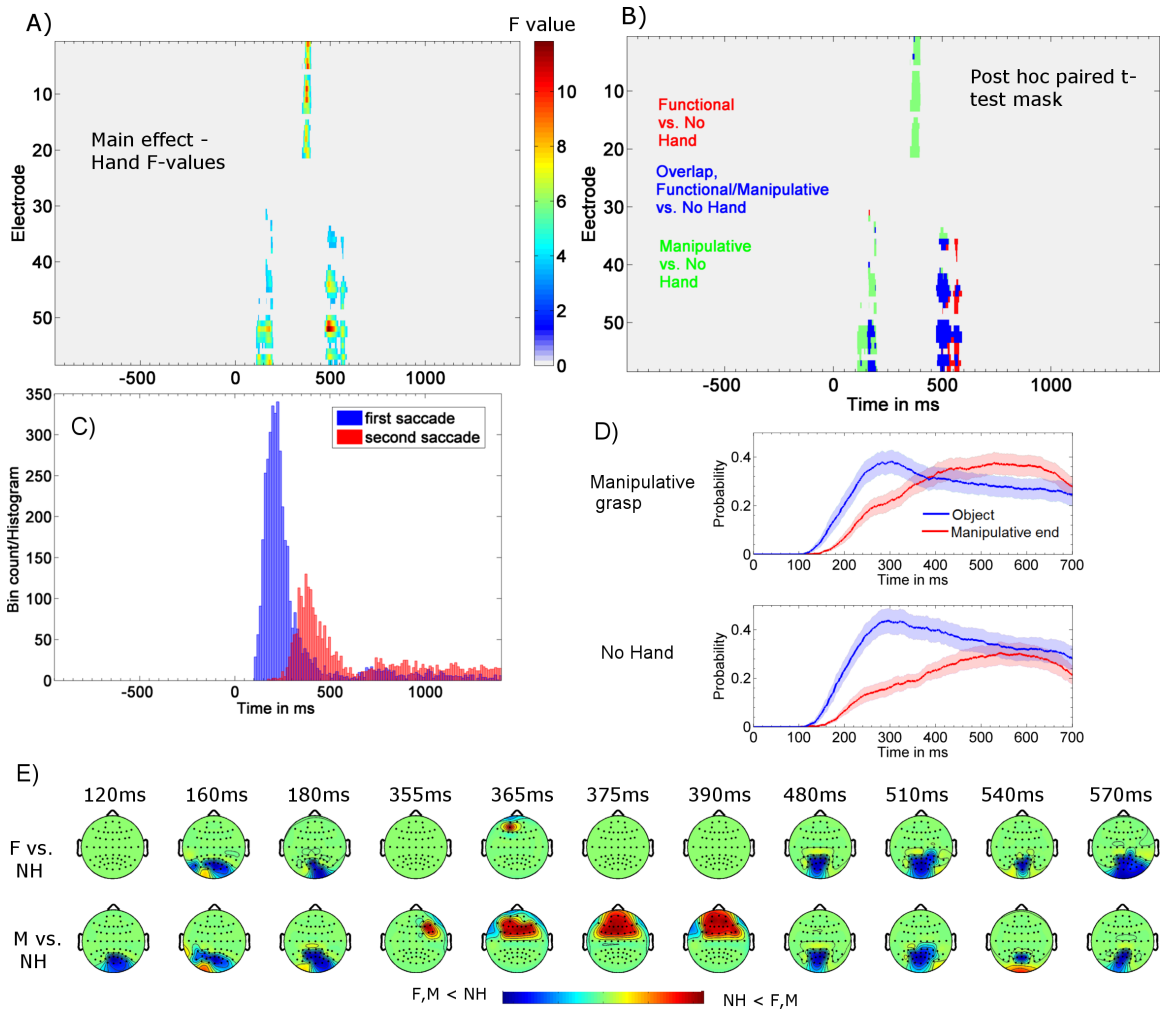

Figure 39: A) Statistical image depicting the F-values at channel-time pairs that exhibited a significant main effect of factor Hand (no hand, functional grasp, and manipulative grasp) from the RM-ANOVA test in the 500ms experiment. The F-values were corrected for multiple comparisons using spatiotemporal cluster statistics. B) Statistical image depicting the mask of the post-hoc paired t-tests that were performed at the significant channel-time pairs in A) contrasting the manipulative grasp and functional grasp condition individually to the no hand control (FDR corrected). Mask color represents channel-time pairs when both contrasts were significant (overlap) or when either contrast was significant. C) Histogram depicting discrete probability density function of first and second saccade initiation latencies. D) Mean probable gaze position (with S.E shading) for the manipulative grasp condition and no hand condition. E) Projection of normalized, significant t-values from the mask in B) onto the scalp electrodes at a few discrete time intervals, with a color bar representing the direction of the statistical difference. NH: No Hand, F: Functional grasp, M: Manipulative grasp. The actual t-values for the contrasts at every channel-time pair are shown in Videos 4 and 5.

grasp and no hand condition due to the greater N100 (Figure 40) emerged over centro-parietal regions similar to the 100ms experiment. It should be noted that parietal ERP

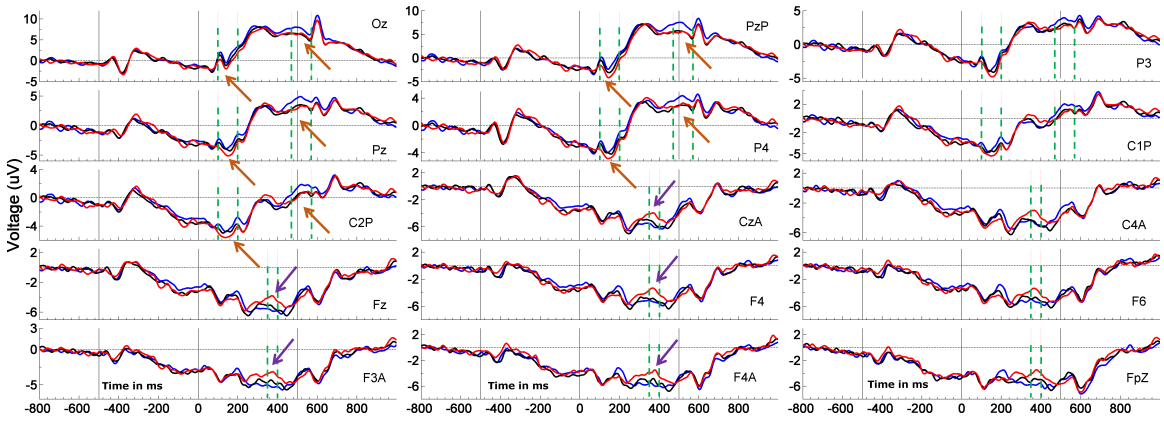

Figure 40: Plots depicting grand-average ERPs for the three Hand conditions: **manipulative grasp**, **functional grasp**, and the **no hand** control at a few occipital, parietal, premotor and frontal electrodes in the 500ms experiment. For visualization, ERP peaks are marked with an arrow in the first two columns. The vertical green dotted lines mark three time-windows of interest: 100 to 200s post image onset (over occipital and parietal electrodes), 350 to 450ms post image onset (over frontal electrodes) and 470 to 570ms post image onset (again over parietal and occipital electrodes). Saccades perturbed the propagation of the N100 differential between the **manipulative grasp** and **no hand** condition. When frontal differences did emerge at 350-400ms post image onset, the polarity of the difference was inverted from the earlier time window (100-200ms).

were temporally coincidental to saccades only at around 250ms or post image-onset (Figure 41), much after the early N100 ERP differences between conditions (100-200ms). The early parietal differences between conditions were more likely to be related processing the affordances of the grasp (similar to the 100ms) rather than the control of saccades. However, this statistical difference faded at about the time saccade initiation probability peaked (200-210ms) and unlike the 100ms experiment, this greater negative potential did not propagate to frontal electrodes. The functional grasp-posture also elicited a greater N100 when compared to the no hand, however this difference did not emerge as early as the manipulative grasp-posture. Frontal differences did emerge between the manipulative grasp and no hand conditions, at 350-400ms post image onset (Figures 39B and 40), at about the time the probability of initiation the first saccade was minimal and the probability of initiating the second saccade was peaking. It should be noted that across all conditions, participants were 1.6 times more likely to initiate only one saccade versus two. This later time-period (350-400ms) immediately preceded a differential gaze pattern (Figure 39D, 400-450ms) wherein

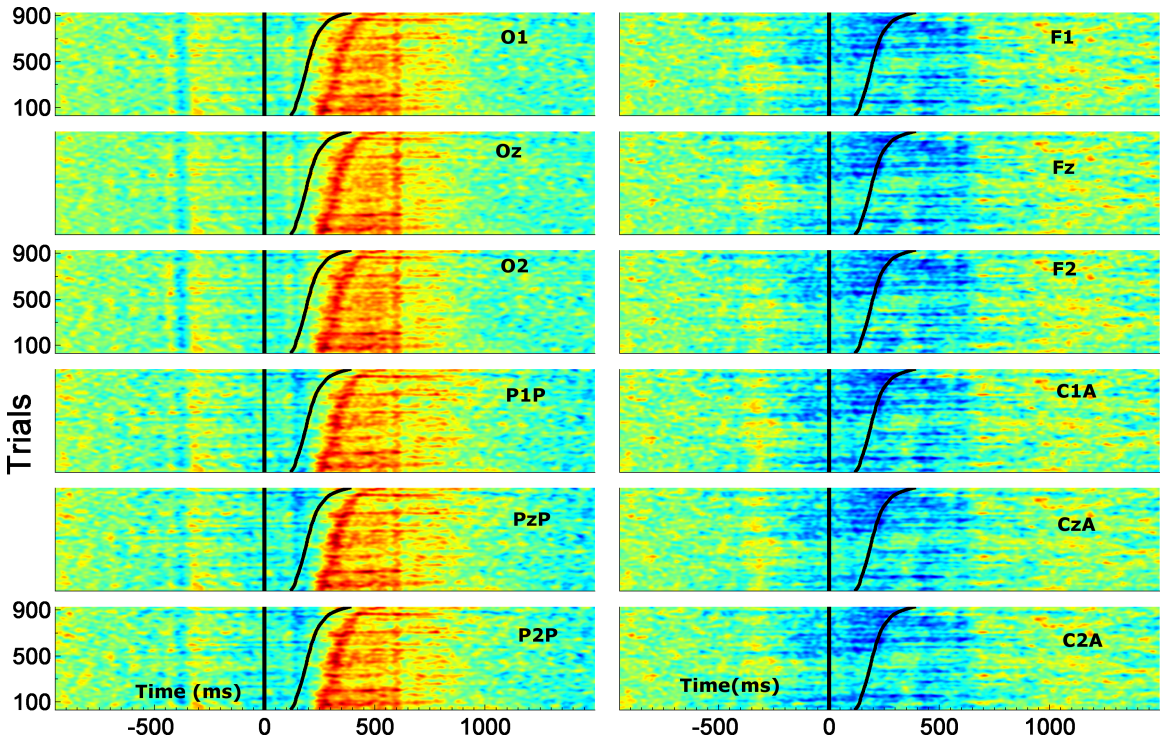

Figure 41: ERP image plots depicting the relation between first saccade initiation latency and ERP magnitude over occipito-parietal electrodes (left column) and frontal electrodes (right column). In each plot, an image of the ERP magnitude across trials is plotted along the rows, with red depicting positive microvolt values and blue depicting negative values. Trials have been sorted based on saccade initiation latency (thick black line) and then subsequently smoothed both across trials and time (rows and columns) with a Gaussian window (Delorme and Makeig, 2004). It can be seen that a temporal relationship exists over occipito-parietal electrodes in terms of the appearance of the positive ERP peak and saccade initiation latency. Importantly, this difference occurs well beyond the 100-200ms time window characterized by statistical differences in the N100 ERP between conditions. It can be seen that in the right column a similar relationship does not exist between ERP magnitude and saccade initiation latencies for frontal electrodes.

foveal gaze was more likely to be over the object in the no hand conditions than the manipulative grasp condition (by 7.66 $\pm$3.54% S.E, t(24) = 4.3108, $p = 2.396 \times 10^{-4}, d_z = 0.862$) and more likely to be at the manipulative tool-end in the manipulative grasp condition than the no hand conditions (by 9.48$\pm$3.056%, S.E t(24) = 4.3108, $p < 1 \times 10^{-5}, d_z = 1.243$). The 350-400ms frontal difference was characterized by an inversion in the polarity of the ERP difference, with the manipulative grasp-posture eliciting a greater positivity/lower negativity than the no hand condition (Figures 39E and 40). Importantly, a similar frontal ERP difference pattern was not observed for the functional grasp-posture vs no hand contrast.

A second occipito-parietal difference between the functional grasp and manipulative grasp condition relative to the no hand condition was observed at 470ms to 570ms post image onset (Figure 39B, Figure 39E and Figure 40). The polarity difference in this time later time window followed the early 100ms time-windows with the manipulative and functional grasp conditions exhibiting greater negativity/lower positivity than the no hand condition.

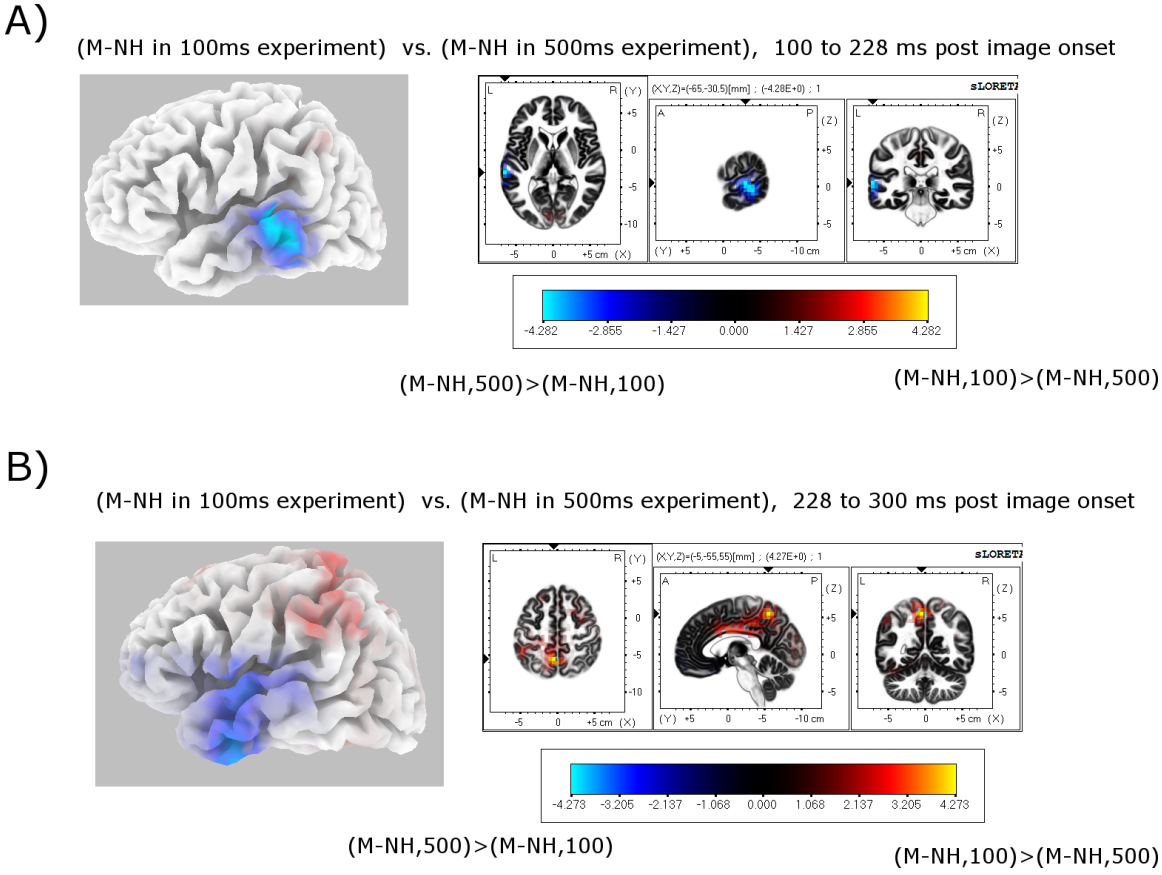

Figure 42: Statistical, source localized difference (at the 0.05 level using an independent samples t-test) in how the affordances of the manipulative grasp-posture is processed between the two experiments at A) 100-228ms post image onset and at B) 228-330ms post image onset. It should be noted that these statistical contrasts are a difference of differences (second level statistics (Smith, 2014)). The color bar in both plots represents t-values. M-NH: difference between the manipulative grasp-posture and no hand condition, computed by subtracting the mean ERP between the two conditions on a subject-by-subject basis. The MNI coordinates of the voxel/brain regions depicting peak statistical differences is also shown in the slices for both time-windows.

Finally we also sought to identify the neural generators underlying the difference in how the affordances of the manipulative grasp-posture is processed (relative to the no hand) between the two experiments. To this end, EEG data were source localized following the

procedures outlined in (Pascual-Marqui, 2002) and using the open source sLoreta software. Contrasts between the two experiments were performed using a difference of differences approach. At the first level, the difference in mean ERP between the manipulative grasp and no hand conditions was computed on a subject-by-subject basis within each of the two experiments individually. The output of the first step served to identify how the manipulative grasp-posture is processed with respect to the no hand within each of the two experiments, and the results were used as input to the second level. At the second level, the difference in how the affordances of the manipulative grasp-posture was processed between the two experiments was evaluated using two independent samples t-tests, one at 100-228ms after image onset and another at 228-300ms after image onset (Figure 42). These time-windows were based on when the statistical ERP differences between the two conditions emerged and also based on the peaks in the ERP themselves, similar to (Mizelle and Wheaton, 2010a, Mizelle et al., 2011). The output of the second level served to contrast how the affordances of the manipulative grasp-posture was processed within each experiment. Results (outlined in Table 1) revealed that encoding the affordances of the manipulative grasp-posture within the 500ms experiment greater engaged the left middle and superior temporal gyri early (100-228ms), whereas the 100ms experiment greater engaged the precuneus in the left parietal lobe later (228-300ms). The ability to continuously gather scene information in the 500ms experiment allowed greater engagement of left ventral regions early, whereas the requirement to continuously evaluate the manipulative grasp-postures affordances without any visual information elicited greater left dorsal activity at a later time window. Similar contrasts were carried out to evaluate how the functional grasp-posture was encoded. While we observed similar trends, none of the voxels reached the corrected 0.05 level threshold, in line with the ERP results which had shown greater neural engagement to process the manipulative grasp-posture.

5.4.4.2 Main effects of Context

Similar to the visualizations of the main effect of Hand, results depicting the main effects of Context are shown in Figures 43, 44, and Videos 6, 7 for the 100ms experiment and Figures

Table 1: Significant neuroanatomical differences in the encoding the affordances of the manipulative grasp-posture between the 100ms and 500ms experiment. Positive voxel values indicate the 100ms experiment greater engaged the corresponding brain region. Negative voxel values indicate the 500ms experiment greater engaged the corresponding brain region.

X, Y, Z (MNI)	Voxel Value	Broadmann, Lobe, Structure	Time-window
-65, -30, 5	-4.282	42, Left Temporal lobe, STG	100-228ms
-65, -35, 0	-4.211	22, Left Temporal lobe, MTG	100-228ms
-65, -40, -5	-4.204	21, Left Temporal lobe, MTG	100-228ms
-5, -55, 55	4.273	7, Left Parietal lobe, Precuneus	228-300ms

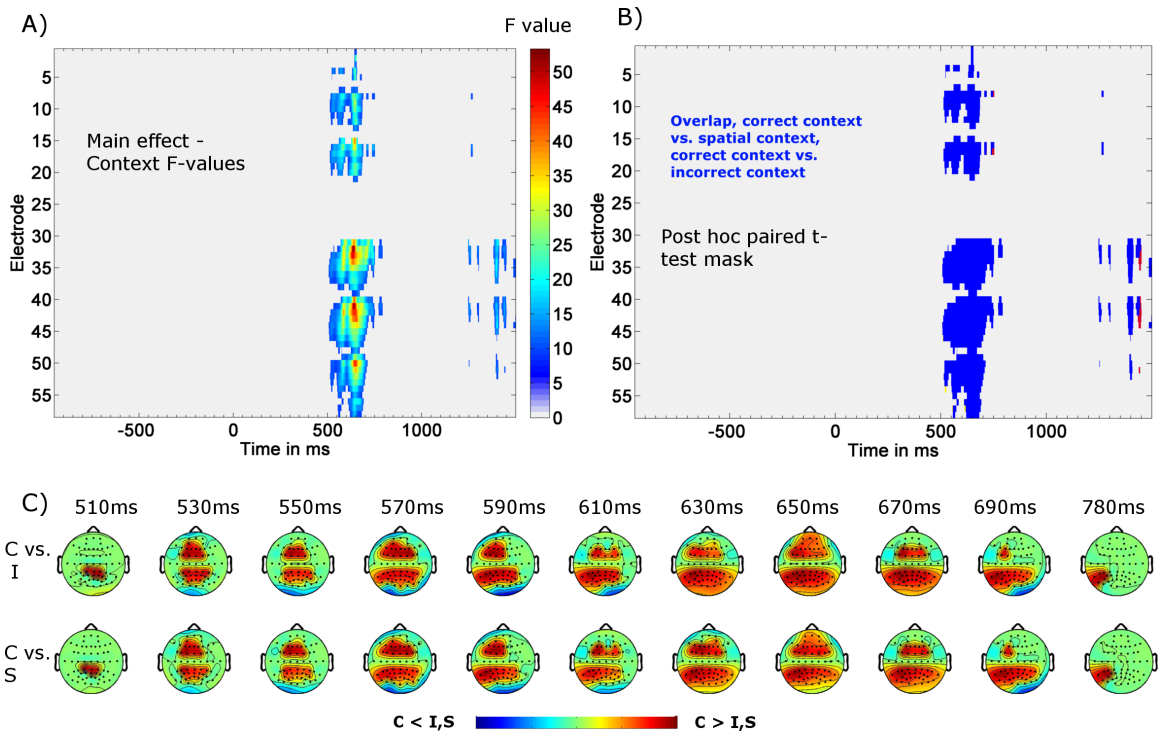

Figure 43: A) Statistical image depicting the F-values at channel-time pairs that exhibited a significant main effect of factor Context (correct, incorrect, spatial) from the RM-ANOVA test in the 100ms experiment. The F-values were corrected for multiple comparisons using spatiotemporal cluster statistics. B) Statistical image depicting the mask of the post-hoc paired t-tests that were performed at the significant channel-time pairs in A) contrasting the incorrect and spatial contexts individually to the correct tool-object context as control (FDR corrected). Mask color represents channel-time pairs when both contrasts were significant (overlap) or when either contrast was significant. C) Projection of normalized, significant t-values from the mask in B) onto the scalp electrodes at a few discrete time intervals, with a color bar representing the direction of the statistical difference. C Correct, I- Incorrect, S Spatial. The actual t-values for the contrasts at every channel-time pair are shown in Videos 6 and 7.

45, 46, and Videos 8, 9 for the 500ms experiment. Significant channel-time pairs were identified by performing spatiotemporal cluster correction on the F-values at the 0.005 level. We used a more stringent threshold here as Context elicited a much stronger spatiotemporal effect than factor Hand. While the main effects of the factor Hand were observed soon after image onset in both experiments, the main effects of the factor Context emerged much later on, at 508ms post image onset in the 100ms experiment and at 424ms post image onset in the 500ms experiment. Thus the neural encoding of Context (which was the experimental task) occurred much later on in time while the neural encoding of the grasp-posture occurred almost immediately after image onset, highlighting an automatic sensitivity of action encoding regions to grasp affordances. Post-hoc paired t-tests (FDR corrected, 0.05 level) were performed contrasting the incorrect and spatial contexts individually to the correct context as control. In both the 100ms and 500ms experiments, the correct tool-object context elicited a greater positive oriented ERP than the incorrect and spatial tool-object context. These greater ERP differentials were primarily focused over left parietal and frontal regions, though there tended to be bilateral parietal differences within the 500ms experiment. Data also revealed that within the 500ms experiment, the incorrect tool-object context elicited earlier ERP differences than the spatial tool-object context.

5.5 Discussion

There were many distinct results in Aim 3. First, results showed participants were able to perform the task with a high degree of accuracy ($> 80\%$) even with rapidly presented stimuli. Within the 100ms experiment, participants primarily utilized peripheral or extrafoveal means of capturing scene information, whereas fixations and saccades reemerged in the 500ms experiment. The major hypotheses in Aim 3 concerned with the role of saccades on the neural processing of the manipulative grasp-posture. Within the 100ms experiment wherein participants rarely executed saccades, the manipulative grasp condition elicited a larger, negative-oriented cortical potential when compared to the no hand condition, beginning at 130ms after image onset. This ERP difference originated at centro-right parietal

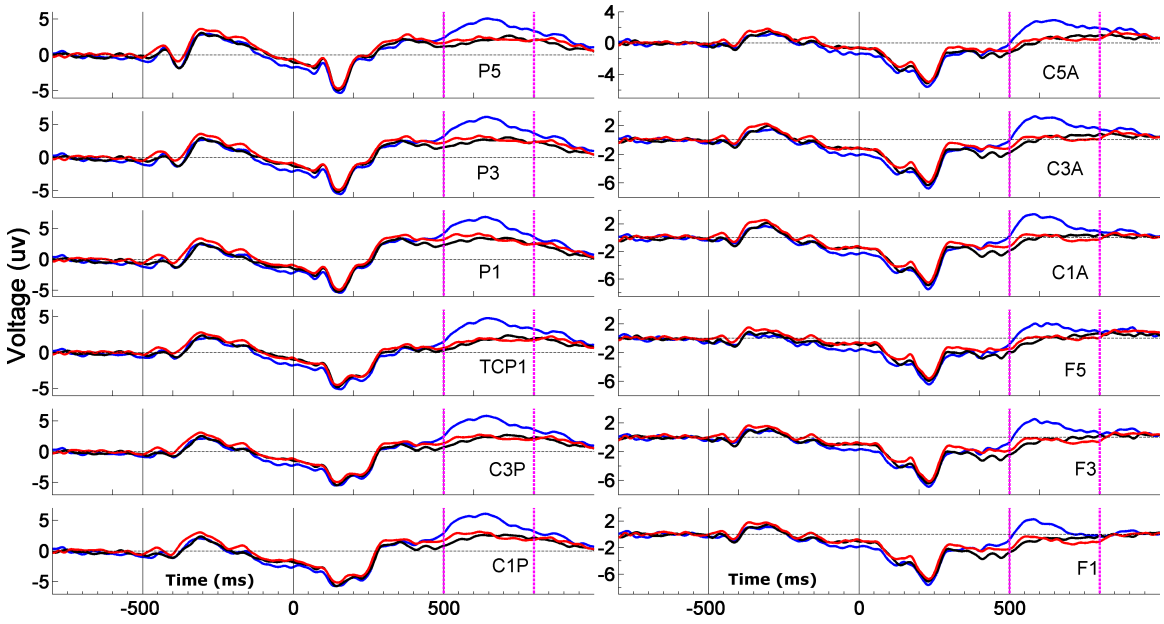

Figure 44: Plots depicting grand-average ERPs for the three Context conditions in the 100ms experiment: **spatial context**, **incorrect context**, and the **correct context** as control at left parietal and frontal electrodes in the 100ms experiment. The vertical magenta lines mark the time-window of interest: 500 to 800s post image onset. It can be seen that the correct context elicited greater positivity in the ERP than the incorrect and spatial contexts.

electrodes (N100 potential), and rapidly propagated to central and right frontal electrodes, lasting till 228ms after image onset. While the functional grasp-posture also elicited a differential N100, there was no propagation to frontal regions and in addition, the differences with the no hand condition were much less sustained and lasted only till about 200ms. We had hypothesized that there would be no significant right frontal differences between the manipulative grasp and no hand conditions in the 100ms experiment where saccades are rarely executed. This would have suggested that the right frontal response to the manipulative grasp-posture would have to be driven by a differential gaze pattern with respect to the no hand condition. Results in the 100ms experiment refuted our hypothesis given the centro-right parietofrontal ERP differential between the manipulative grasp-posture and no hand condition. However, rather than appear 400-600ms post-image onset (Natraj et al., 2013), the ERP difference was present only at a very early time window. Importantly, results within the 500ms experiment revealed that the frontal processing of the manipulative grasp-posture was not wholly independent of saccades, and the ability of the observer to

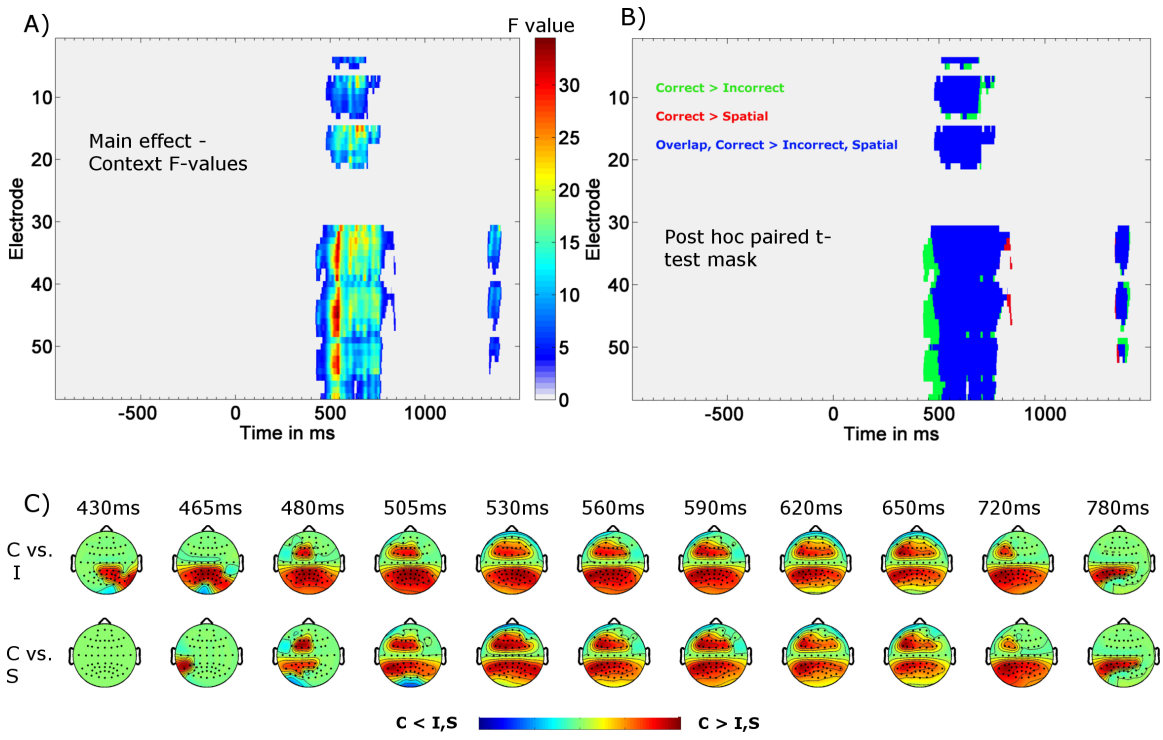

Figure 45: A) Statistical image depicting the F-values at channel-time pairs that exhibited a significant main effect of factor Context (correct, incorrect, spatial) from the RM-ANOVA test in the 500ms experiment. The F-values were corrected for multiple comparisons using spatiotemporal cluster statistics. B) Statistical image depicting the mask of the post-hoc paired t-tests that were performed at the significant channel-time pairs in A) contrasting the incorrect and spatial contexts individually to the correct tool-object context as control (FDR corrected). Mask color represents channel-time pairs when both contrasts were significant (overlap) or when either contrast was significant. C) Projection of normalized, significant t-values from the mask in B) onto the scalp electrodes at a few discrete time intervals, with a color bar representing the direction of the statistical difference. C Correct, I- Incorrect, S Spatial. The actual t-values for the contrasts at every channel-time pair are shown in Videos 8 and 9.

direct foveal attention towards scene features. Similar to the 100ms experiment, the manipulative grasp condition elicited a greater occipito-parietal N100 potential than the no hand condition, starting earlier than the 100ms experiment by 30ms. It should be noted that N100 in the 500ms experiment was related more to affordance understanding rather than gaze control as any temporal relationship between saccades and occipito-parietal ERP was observed at around 250ms, well beyond the time-window of statistical N100 differences between conditions. The early ERP differences between the manipulative grasp and no hand conditions faded when saccade initiation probability was maximum (approx. 210ms), and

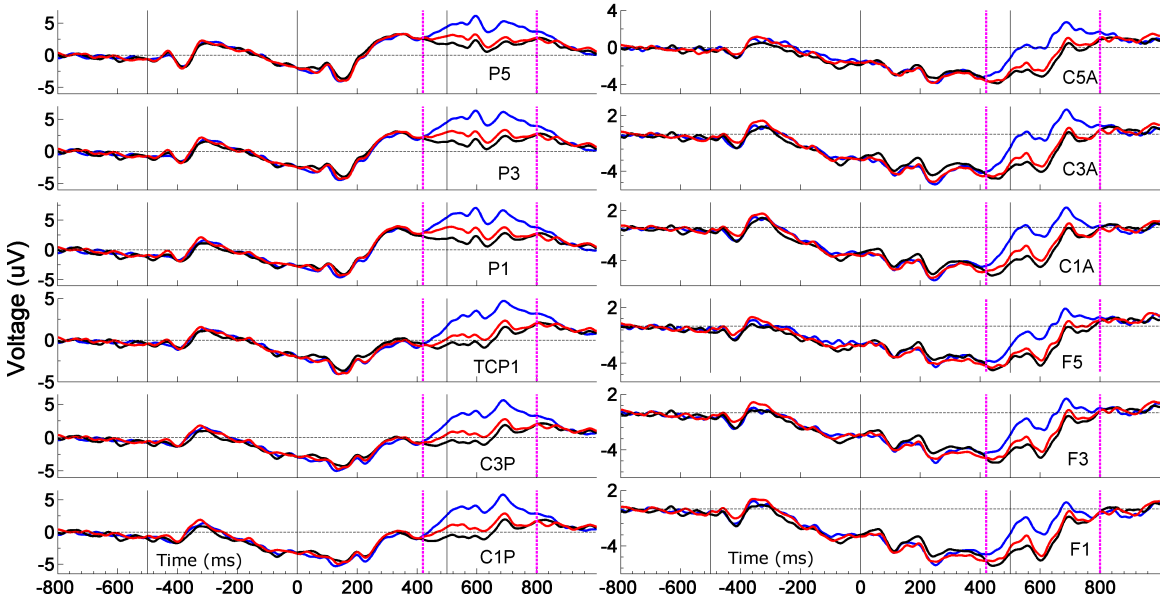

Figure 46: Plots depicting grand-average ERPs for the three Context conditions in the 500ms experiment: **spatial context**, **incorrect context**, and the **correct context** as control at left parietal and frontal electrodes in the 500ms experiment. The vertical magenta lines mark the time-window of interest: 420 to 800s post image onset. It can be seen that the correct context elicited greater positivity in the ERP than the incorrect and spatial contexts.

delayed the frontal propagation of the ERP differential to 350-400ms after image onset. At this time-window, first saccade probabilities were minimal and the probability of executing the seccond sacade was peaking. When the frontal differences did emerge, the polarity difference in the ERP was switched, with the manipulative grasp-posture eliciting a greater positivity/lesser negativity than the no hand condition. This frontal difference immediately preceded a differential gaze pattern (400-450ms) wherein attention was focused more on the object in the no hand conditions and on the manipulative tool-end in the manipulative grasp condition. Summarizing the key result, spatiotemporal activation patterns and ERP polarity characteristics with respect to processing the affordances of the manipulative grasp over frontal electrodes were influenced by the relative presence/absence of saccades. This key result does in part support our conjecture that frontal regions might serve both in attention mechanisms/eye movement control and affordance understanding. These spatiotemporal ERP difference patterns were unique to the manipulative grasp-posture and were not observed for the functional grasp-posture, in line with our overall hypothesis that

the manipulative grasp-posture elicits unique neural encoding mechanisms. Finally, when directly contrasting how the affordances of the manipulative grasp-posture was encoded in both experiments, source localization analyses revealed greater engagement of the left middle temporal gyrus in the 500ms experiment, 100-228ms post image onset, and greater engagement of the left precuneus in the parietal lobe in the 100ms experiment, 228-300ms post image onset. It should be noted that the encoding of the tool-object context itself in both experiments happened relatively much later in time, at roughly 500-800ms after image onset. In both the 100ms and 500ms experiments, the correct tool-object context elicited a greater positive oriented ERP than the incorrect and spatial tool-object context. These greater ERP differential was primarily focused over left parietal and frontal regions, though there tended to be bilateral parietal differences within the 500ms experiment.

5.6 Conclusion

Aim 3 integrated the conceptual underpinnings of Aims 1 and 2 specifically evaluating the overlap between action understanding and gaze control over parietofrontal regions. Given the common themes in all three Aims, further discussions integrating the salient results across all three Aims (and especially Aim 3 given its integrative nature) is detailed in the following chapter.

CHAPTER VI

INTEGRATION

6.1 Summary

The overall goal of the dissertation was to understand the how the affording properties of contextual and grasp-specific tool-use scenes influenced spatiotemporal patterns of parietofrontal activity as healthy right-handed participants evaluated the tool-object content in the images. In particular, we also evaluated the visual encoding of complex tool-use scenes and the effect of saccades and foveal attention on parietofrontal activations, given the anatomical overlap between gaze control and action understanding (Anderson et al., 1994, Mort et al., 2003, Vingerhoets et al., 2010). To address the goal of the thesis, we extended our prior behavioral study that evaluated reaction times when participants evaluated whether tool-object relationships were correct/incorrect when a tool was placed across three contexts of use (correct, incorrect, spatial) and four grasp-postures (no hand, static hand, functional grasp, manipulative grasp). Results had revealed that though the task required evaluating tool-object content, manipulative grasp-postures uniquely elongated decision times when compared to the no hand condition. A similar effect was not observed for the functional grasp and static hand conditions (Borghi et al., 2012). Results in Aim 1 revealed unique activations over right parietofrontal regions in response to the manipulative grasp-posture when compared to the no hand condition (Natraj et al., 2013) especially within correct and spatial tool-object contexts. Similarly, eye tracking results in Aim 2 revealed that the manipulative grasp condition elicited unique gaze patterns when compared to the no hand condition within correct and spatial tool-object contexts that afford a certain degree of action. Given the dual role of parietofrontal regions in both affordance understanding and gaze control, the purpose of Aim 3 was to evaluate the effect of visuospatial attention mechanisms on the parietofrontal processing of the manipulative grasp across all three contexts. Results revealed the frontal response to the manipulative grasp

when compared to the no hand condition was temporally modulated by saccades. In addition, there was greater and earlier engagement of the left middle temporal gyrus along the ventral stream to encode the manipulative grasp-posture when the observer had the ability to execute saccades and greater and later engagement of the left precuneus along the dorsal stream to encode the manipulative grasp-posture when the observer had to parse the scene extrafoveally without eye movements and rely on memory. Given the integrative nature of Aim 3 as it combines common concepts from the prior behavioral data and Aims 1 and 2 (Borghi et al., 2012, Natraj et al., 2013, Natraj et al., 2015), this chapter focuses on the interpretation of the results in Aim 3, detailed in the following sections. The key points with respect to spatiotemporal patterns in ERP differences between the manipulative grasp and no hand conditions are conceptualized in Figure 47.

6.2 Spatiotemporal patterns of parietofrontal ERP when encoding grasp-specific tool-use in the absence of eye movements

Processing of the manipulative grasp when compared to the no hand condition in the 100ms experiment of Aim 3 was characterized by a greater ERP negativity originating over occipito-parietal electrodes. Subsequently, the ERP difference propagated to frontal electrodes within 130-228ms post image onset. While the functional grasp also elicited similar occipito-parietal differences when compared to the no hand condition, the ERP difference did not propagate to frontal electrodes. In addition, the occipito-parietal response to the manipulative grasp occurred earlier and involved more electrodes than the functional grasp vs. no hand contrast. It should be noted that within the 100ms experiment, participants did not execute saccades in the vast majority of trials. Generally, the polarity of the ERP depends on the orientation of the neural generators with respect to the scalp and the direction of ionic flow that causes the post-synaptic potentials (Nunez and Srinivasan, 2006). In any case, it is commonly thought that irrespective of the polarity, an increase in the ERP magnitude reflects higher recruitment of the underlying neurons and a lesser temporal asynchrony in the firing pattern of the neurons (Kutas and Federmeier, 2011). The higher negativity in the ERP response to the manipulative and functional grasp when compared to the no hand control can therefore be interpreted as greater engagement of the

neurons in the underlying cortical layers, with a more robust response to the manipulative grasp-posture. The functional significance of the ERP negativity over occipital, parietal and frontal electrodes are detailed further as follows.

With respect to activation of electrodes over the visual cortex, it is possible that occipital activity could simply arise from the higher scene complexity when compared to the no hand condition (Kandel et al., 2000), due to increased retinotopic activity as participants extra-foveally attended to the grasp (Somers et al., 1999). However, it should be noted that the higher visual complexity arises from the presence of a biological agent (the grasp-posture) and recent motor control research has shown that activity in the visual cortex may underlie processing the affordances of grasp-postures (Gutteling et al., 2015) via cortical feedback mechanisms. For instance, the presence of a grasp-posture may engage higher-order (parietal) cortical neurons that process grasp-affordances via extra-foveal attention and loop back to lower-order neurons in the visual cortex, increasing the response of visual cortex neurons to the grasp-posture (van Elk et al., 2010, Gutteling et al., 2015). This line of reasoning is supported by the fact that significant differences between both the grasp conditions relative to the no hand control first appeared over parietal electrodes prior to occipital electrodes in the 100ms experiment (Videos 2 and 3). As a result, engagement of neurons under occipital electrodes may be due to both the increased visual complexity due to the presence of the grasp and a top-down driven enhanced visual processing of the affordances of the grasp per se. However, it has been well established that activity over parietal (Culham and Kanwisher, 2001, Culham and Valyear, 2006, Mizelle and Wheaton, 2010b) and temporal electrodes (Valyear and Culham, 2010, Mizelle et al., 2013) have been shown to process tool-grasps and functional tool-use affordances. In particular, the early posterior response to the presence of a grasp (both functional and manipulative) when compared to the no hand condition could represent an automatic engagement of grasp-specific neurons in the parietal cortex to process the motoric/functional grasp-affordances (Culham and Kanwisher, 2001, Culham and Valyear, 2006, Vingerhoets, 2014, Roth and Zohary, 2015) potentially also include mirror neurons in the inferior parietal cortex that underlie a mental simulation of the grasp itself (Chong et al., 2008). Activity of neurons under

temporal electrodes might represent encoding the semantic or conceptual properties of the grasp. The greater engagement of neurons under parieto-temporal electrodes likely represents encoding the affordances of the grasp-posture itself rather than a visual complexity phenomenon. While both the functional grasp and manipulative differed from the no hand control over occipital and parieto-temporal electrodes, it should be noted that the manipulative grasp-posture elicited earlier and more diffuse activity bilaterally and for a longer period of time (up to 228ms), unlike the functional grasp-posture (less than 200ms).

The key result with respect to the neural response to a grasp-posture in the 100ms experiment was the propagation of the posterior ERP difference to frontal electrodes. More specifically, the manipulative grasp uniquely elicited a more sustained frontal difference with respect to the no hand control (up to 228ms post image onset), whereas there was only a very transient frontal difference for the functional grasp vs. no hand contrast. Similar to previous studies and Aim 1, it is possible that frontal ERP difference could correspond to the engagement of the mirror neuron system (MNS), a set of bilateral prefrontal neurons that encode higher level motor cognition such as the intent or goal of tool-use gestures (Iacoboni et al., 2005, Aziz-Zadeh et al., 2006, Iacoboni and Dapretto, 2006, Natraj et al., 2013). The greater frontal N100 in the 100ms experiment in Aim 3 (100-230ms post image onset) is suggestive of a more pronounced engagement of the mirror neurons to encode the intent of the manipulative grasp-posture soon after image onset. As we had argued in Aim 2 (Natraj et al., 2015), it is unlikely that this effect is driven purely by the grasp per se, but more due to the affordances of the manipulative grasp-posture as it non-functionally interacts with the tool-end that has higher affordances with respect to the object (the operant tool-end). Indeed, participants were well primed to view the operant tool-end in the upper right hemifield and it is likely that a manipulative grasp-posture of the operant tool-end could have automatically primed grasp-specific mirror neurons over frontal regions to encode the action intent of the manipulative grasp-posture as it non-functionally interacts with the operant tool-end. Alternatively, the affordances of the functional grasp-posture are readily understood, potentially resulting in lesser automatic recruitment of mirror neurons over frontal regions. Apart from representing the mirror neuron system, the frontal response

to the manipulative grasp-posture may also correspond to diffuse activity of neurons in bilateral prefrontal cortices. These neurons may be driven by extra foveal attention towards the manipulative grasp and may serve to resolve the conflict (Jonides et al., 2002, Egner and Hirsch, 2005) presented by the atypical manipulative grasp-posture as it interferes with that part of the tool (the operant tool-end) that is vital to understand tool-object relationships (Natraj et al., 2015). It is possible that the mirror neuron system and conflict resolution system over frontal areas may be operating in parallel to encode the affordances of the manipulative grasp-posture. The fact that the response to the manipulative grasp-posture was observed so early after image onset highlights the sensitivity of grasp-specific neurons when encoding affordances. Such sensitivity could potentially have an evolutionary benefit in learning to rapidly recognize the intent of an observed hand-object action or even differentiate between tool-use and communicative gestures (Stout and Chaminade, 2012).

There are two differences in the frontal response to the manipulative gasp-posture between Aim 1 and the 100ms experiment in Aim 3. First, the frontal ERP difference in Aim 1 emerged 400-600ms post image onset and was constrained to right frontal electrodes, whereas in the 100ms experiment of Aim 3, the frontal ERP difference emerged earlier (130-228ms) and the ERP difference was diffusely spread over frontal and premotor electrodes, with a bias towards electrodes over right frontal areas. Second, the polarity of the ERP difference in Aim 1 was characterized by lesser negativity of the manipulative grasp-posture. However, the polarity of the ERP difference in the 100ms experiment of Aim 3 was characterized by a greater negativity of the manipulative grasp-posture when compared to the no hand control. We expand on these polarity differences and their functional significance in light of the result from the 500ms experiment in Aim 3, detailed in the following paragraphs.

6.3 Saccades modulate spatiotemporal patterns of parietofrontal ERP when encoding grasp-specific tool-use scenes

In line with our hypothesis, the 500ms experiment of Aim 3 evoked significantly more saccades in the majority of trials than the 100ms experiment. The early ERP responses to the manipulative and functional grasp-postures with respect to the no hand condition

was largely similar to the 100ms experiment, with early occipital temporal and parietal differences characterized by an increased negativity in the ERP (N100). Similar to the 100ms experiment, the response to the manipulative grasp-posture occurred earlier and over a larger number of electrodes than the functional grasp when compared to the no hand control. It should also be noted that significant ERP differences over posterior electrodes occurred earlier in the 500ms experiment (around 120ms) than the 100ms experiment (around 130ms). This delay could possibly be due to the fact that the stimulus transitioned to the fixation circle in the 100ms experiment (at 100ms post image onset) unlike the 500ms experiment wherein the stimulus was on screen till 500ms. Similar to arguments made in the previous section, the occipital response could correspond to encoding both the visual complexity of the scene and the affordances of the grasp-posture, whereas the parieto-temporal response likely represents the activity of grasp-specific neurons that encode the conceptual and motoric affording properties of the functional/manipulative grasp-posture. Unique to the 500ms experiment, the presence of saccades caused the posterior ERP differences to fade at about 200ms, right about the time when saccade initiation probabilities were peaking (200-210ms), possibly due to the fact that information processing is limited during saccades ((Bahill and LaRitz, 1984)). In addition, unlike the 100ms experiment, the ERP difference did not immediately propagate to frontal electrodes and instead was delayed to 350-400ms. The emergence of frontal differences (350-400ms) at right frontal and subsequently over bilateral premotor and bilateral frontal occurred after the majority of first saccades were completed and the probability of initiating the second saccade was rising. The first saccade therefore perturbed and delayed the propagation of the ERP difference from posterior to frontal regions. In addition, the frontal ERP difference was largely restricted to the manipulative grasp vs. no hand contrast, while the neural response to the functional grasp did not differ from the response to the no hand condition. When compared to the frontal difference in the 100ms experiment, there was also an inversal in the polarity of the ERP difference here as the manipulative grasp elicited lesser (rather than greater) negativity than the no hand control. This polarity difference was similar to the ERP difference between the same two conditions in Aim 1, 400-600ms post image onset over right frontal electrodes (Natraj

et al., 2013). As outlined earlier, irrespective of the polarity of the ERP, the magnitude is thought to reflect the level of engagement of the underlying neurons. As a result, the polarity inversal makes it unlikely that the frontal difference could correspond to a conflict resolution mechanism as this would imply that the no hand condition presents a greater conflict than the manipulative grasp condition. However, it is possible that given the ability to scan the scene, the observer could be encoding the relationship between the tool-object pair in terms of how the tool may engage on the object i.e. an object-oriented action priming effect (Thill et al., 2013, Natraj et al., 2015). In this scenario, the manipulative grasp-posture has lower affordances of engaging the tool on the object (Borghi et al., 2012, Natraj et al., 2013), and may necessitate the recruitment of mirror neurons to encode grasp-intent relative to engaging the tool on the object, thereby disrupting the object-oriented action priming effect. The aforementioned arguments are strengthened by the fact that in both Aims 1 and 2 (that have a similar experimental design) and the 500ms experiment in Aim 3, the ERP difference between the manipulative grasp and no hand conditions immediately preceded a differential gaze pattern between the two conditions wherein foveal attention was placed more on the object in the no hand conditions (object-oriented action priming), and more on the manipulative tool-end in the manipulative grasp conditions (understanding action-intent). From an action encoding perspective, understanding the intent of the manipulative grasp-posture (possibly via mirror neurons) in terms of how it allows engaging the tool on the object could have elicited a lower (negative-oriented) ERP magnitude given that this grasp-posture does not afford engaging the tool on the object. Alternatively, the no hand and functional grasp conditions do not inhibit the object-oriented action priming effect and thereby elicited largely similar neural responses to each other. Apart from the inhibitory effect of mirror neurons on the object-oriented action priming effect, the polarity inversal could also correspond to the N400 potential, that has been previously implicated in the attention-driven recall of semantic knowledge over prefrontal areas (a review of the N400 can be found at (Kutas and Federmeier, 2011)). Decoding tool-object relationships also involve a recall of their conceptual or semantic properties (Mizelle and Wheaton, 2010a) and a top-down recall of semantic knowledge may be mediated by frontal regions (Kutas

and Federmeier, 2011). The lower affordances of the manipulative grasp-posture could have interfered with the aforementioned cognitive mechanism, thereby resulting in a lower magnitude in the N400 when compared to the no hand control. It is also likely that the MNS and N400 might be operating in parallel, together resulting in a lower object-oriented action priming effect. The aforementioned recruitment of other neural systems to understand grasp-intent could potentially alter the underlying dipole or neural generator of the ERP signal, resulting in an inversal in the ERP differential for the manipulative grasp alone. Crucially, the polarity inversal in the ERP difference between the manipulative grasp and no hand conditions depends on the ability of the observer to continuously sample scene information and exhibit a differential pattern in spatial gaze position between the two conditions, as a similar polarity inversal at 400ms was not observed in the 100ms experiment of Aim 3. The buildup of the late-appearing frontal ERP difference is facilitated by the ability of the observer to continuously sample the scene. The aforementioned encoding mechanism (in the 500 experiment) is different from the 100ms experiment wherein the observer does not have the ability to scan the scene; as a result, the interference effect of the manipulative grasp-posture (in the 100ms experiment) may be more derived due to its non-functional interaction with the operant tool-end. These results broadly suggest of differential neural mechanisms to encode the manipulative grasp in the 500ms and 100ms experiments. As outlined earlier, it is likely that the former involves understanding grasp-intent relative to engaging the tool on the object and the latter arises from the interference effect relative to non-functionally grasping the operant tool-end.

After the frontal differences faded at 400ms, parietal ERP differences reemerged at around 470ms when the majority of the second saccades were completed. At this later time-point, the underlying dipole or neural generator produced a positive-oriented ERP signal for all three hand conditions, characterized by a greater positivity for the no hand conditions when compared to the functional and manipulative grasp conditions (that were statistically more negative than the no hand control). As detailed earlier, a potential reason underlying this inversal in polarity could be the object-oriented action-priming effect that is highest in the no hand conditions (Borghi et al., 2012, Natraj et al., 2015). As a result, the

no hand conditions elicited greater positive ERP at around 500ms as they better facilitate the object-oriented action priming effect given that they also produced the fastest decision times. Conversely, parietal and occipital neurons might have been engaged to process grasp-affordances thereby resulting in a lower magnitude of the ERP (van Elk et al., 2010, Gutteling et al., 2015). It is important to note that it was only over frontal regions that the manipulative grasp uniquely differed from the no hand control, while both grasp conditions differed from the no hand over posterior electrodes. These results suggest the buildup of a positive-oriented ERP response to grasp-affordances that crucially depend on the ability of the observer to continuously parse the scene as a similar effect was not observed in the 100ms experiment. It should be noted that the buildup in the neural response over parietofrontal areas was unique to the processing of grasp-affordance as similar effects were not observed in the ERP responses to context-affordances. In the latter, the time-points characterized by significant ERP differences between the three contexts was independent of gaze control as they occurred roughly within same windows in both the 100ms and 500ms experiments (detailed further in subsequent paragraphs). Results here shed light on how saccades alter parietofrontal processing of grasp-specific tool-use affordances especially given the extensive anatomical overlap between gaze control and action perception over the aforementioned regions (Mort et al., 2003, Lewis, 2006, Vingerhoets et al., 2010).

6.4 The ability to foveate alters the engagement of the dorsal and ventral streams to process the affordances of the manipulative grasp

Finally, the direct contrast between the 100ms and 500ms experiment in Aim 3 revealed greater and earlier (100-228ms post image onset) activation of the left superior and middle temporal gyri in the ventral stream to process the affordances of the manipulative grasp-posture in the 500ms experiment and greater activation of the left precuneus in the dorsal stream at a later time window (228-300ms) to process the affordances of the manipulative grasp-posture in the 100ms experiment. Engagement of left ventral regions have been well implicated in semantic tool-use knowledge (Mahon et al., 2007, Valyear and Culham, 2010) and prior research in our lab has shown that the ventral stream may be well suited to encode the errors of hand-object actions (Mizelle et al., 2013). The ability to continuously

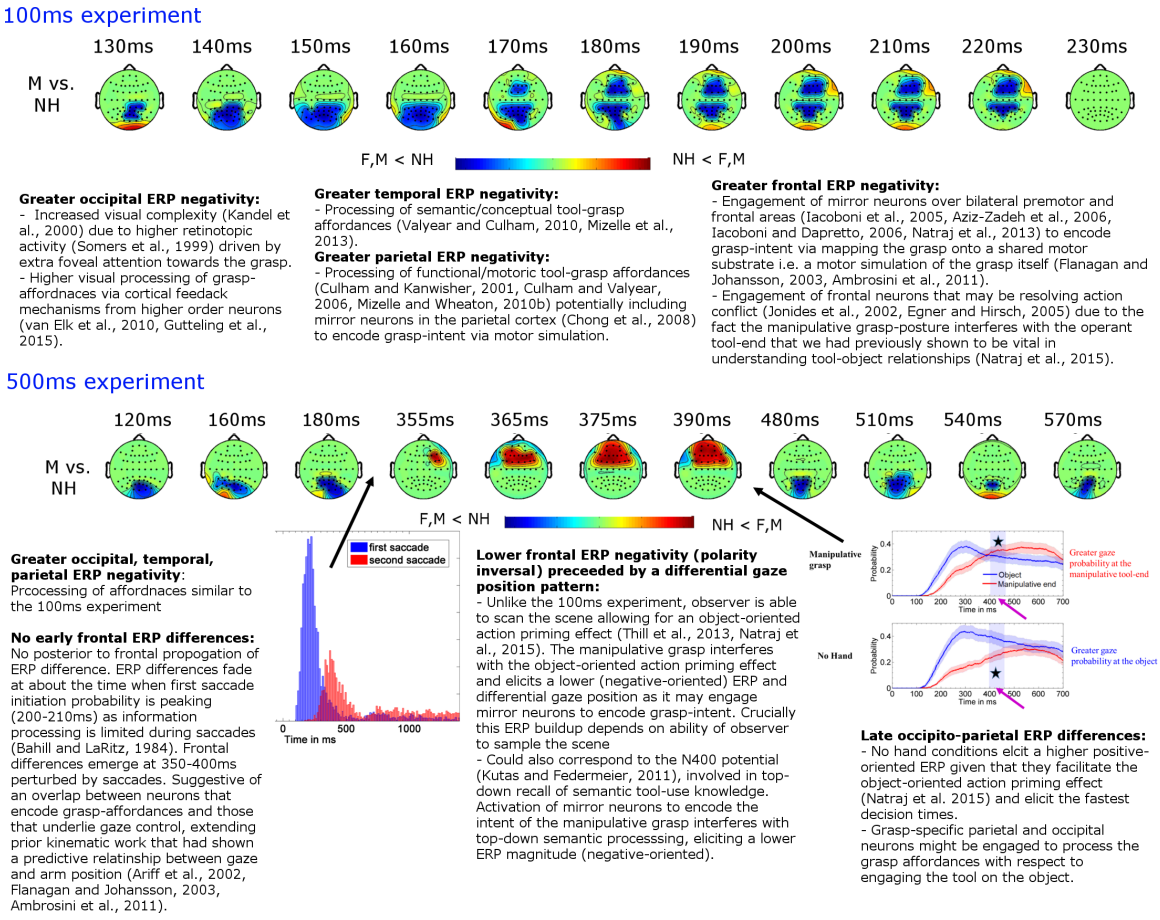

Figure 47: A brief bullet-point description underlying spatiotemporal ERP differences between the manipulative grasp and no hand conditions in the both the 100ms experiment (top row) and 500ms experiment (bottom row)

gather foveal information even before saccades were initiated (mean saccade initiation time was 227.9 $\pm$ 69.2ms) allowed for greater engagement of error encoding regions along the left ventral stream in the 500ms experiment. However, there was greater and later engagement of the left precuneus in the superior parietal cortex for the 100ms experiment wherein the observer had continuously parse the affordances of the manipulative grasp using short term memory. The left precuneus has been found to be active in many neuroimaging studies investigating the perception of hand-object movements whether static, dynamic or imagined (for review and examples, see (Lewis, 2006, Molnar-Szakacs et al., 2006, Vingerhoets et al., 2009)). It is possible that greater activity of the left precuneus in the 100ms experiment could represent a sustained mental effort in understanding the kinematic properties of the

manipulative grasp that occurs at a later time window than the ventral stream activation for the 100ms experiment. The results shed light on unique neural encoding mechanisms underlying the processing of the manipulative grasp-posture based on the ability of the observer to continuously parse scene information. The fact only left hemispheric anatomical generators were identified is line with prior research that have strongly implicated the left hemisphere to in action understanding (Johnson-Frey et al., 2005, Lewis, 2006).

6.5 Saccades do not modulate spatiotemporal ERP patterns underlying the perceptual judgement of tool-object context

It should be noted that the experimental task required participants to evaluate tool-object context and overall, results showed that the differential ERP responses to context occurred at roughly the same time in both experiments (about 500ms) and over roughly the same parietofrontal electrodes, and was largely unaffected by the difference between the two experiments. Parietofrontal neurons have been well implicated in understanding the graspability and manipulability of tools (Gentilucci, 2002, Johnson-Frey, 2004) and prior work has shown activation of these regions in understanding functionally appropriate tool-use (Mizelle and Wheaton, 2010b) with a distinct left hemispheric bias (Johnson-Frey, 2004, Johnson-Frey et al., 2005, Lewis, 2006). Indeed, even in the results here, the parietofrontal ERP differential between the three tool-object contexts was left-dominant. Inspection of the ERP signals suggested a positive-oriented polarity of the underlying neural generator with a greater positivity in the parietofrontal response to the correct tool-object context that has the highest affordances. The incorrect tool-object that affords least tool-object action elicited the lowest magnitude of positive-oriented parietofrontal ERP whereas the spatial or ambiguous context tended to be in between the other two contexts. The time scale of these ERP responses were well outside the first and second saccade windows in the 500ms experiment in Aim 3. The question then remains, as to why saccades modulated parietofrontal processing of the grasp-posture alone and not tool-object context. One potential reason could be that the observer may be engaging the MNS to decode grasp-intent by mapping the grasp onto his/her own shared parietofrontal substrate via motor resonance (Flanagan and Johansson, 2003, Iacoboni and Dapretto, 2006, Ambrosini et al., 2011).

Both the kinematic literature (Ariff et al., 2002) and action perception literature (Flanagan and Johansson, 2003, Ambrosini et al., 2011) have shown strong correlations between an observers eye movements and self-generated or observed hand movements. For example, in a biomechanics study done with the KINARM robot, participants performed reaching movements without visual feedback of their arm-position. The authors showed a lead-lag effect wherein current gaze position predicted future arm-position by about 200ms (Ariff et al., 2002). An example from the action perception literature is the eye tracking study done by Ambrosini et al. who showed that an observers gaze predicted the target of an actors hand-object action only when the actors hand was preshaped in an appropriate fashion (Ambrosini et al., 2011). Prior research is therefore suggestive of a tight neural coupling between eye and hand movements. If we assume that the observer may be engaging the MNS to understand grasp-intent by mapping the grasp onto a shared motor substrate, then it is possible that overlapping neurons underlying predictive eye movements might also be activated. As a result, when these neurons over frontal and premotor regions control gaze position, they perturb the parietofrontal ERP response to the grasp-posture alone while not influencing the ERP response to tool-object context.

6.6 Limitations and future directions

There were many outcomes and aspects of this dissertation that merit further investigation. Participants in all three Aims were always primed as to what constitutes a tool and object, and that the tool would always be placed in the right hemifield and that the object would always be placed in the left hemifield. It is worth investigating how neural and visual encoding mechanisms are influenced when participants are not primed in this manner since the definition of tool-object pairs in our stimuli can potentially be interchangeable (e.g. fork-pen). It is also worth investigating if neural and visual encoding mechanisms can be modulated by the handedness of the subject and perspective e.g. if the images were viewed from a left handed perspective or in an allocentric perspective (Kelly and Wheaton, 2013) and if the subjects themselves were left handers (Kelly et al., 2015). In addition, in lieu of a biological agent, the hand can be replaced by a prosthetic device (Cusack et al., 2011) and

it can be potentially investigated whether neural and visual encoding mechanisms would still be sensitive to grasp-affordances notwithstanding the mismatch between the agent and the observer. In particular, the conceptual theories underlying this dissertation can be utilized in a rehabilitation framework wherein it can be evaluated whether an amputee processes a prosthetic device in a non-functional posture similar to how participants in the all three Aims here processed the manipulative grasp-posture. Such an experiment would allow evaluating whether the amputee has completely internalized the prosthetic device as part of his/her own body schema (Cusack et al., 2015). While interesting in their own right, as such these aspects are outside the current scope and hypotheses of the dissertation that was specific to healthy right handers viewing right handed egocentric tool-object images.

It can be argued that the key result of the manipulative grasp-posture could be due to a visual interference effect wherein a greater portion of the arm is seen in the scene. It is unlikely the unique effects of the manipulative grasp could be driven due to such a visual saliency effect as results in Aim 2 and the 500ms experiment in Aim 3 showed that gaze was tightly constrained to the part of the arm near the operant tool-end and not the arm per se. Similarly, gaze was constrained to the part of the arm near the functional tool-end in the functional grasp conditions. All the images were also previously normalized for overall visual complexity (Borghi et al., 2012). Thus the effective AOI size for the functional and manipulative grasp condition were similar and it is more probable that the effects of the manipulative grasp-posture were driven by its affording qualities. The arm can also potentially be replaced by another long object (such as a block of wood) positioned in a manner similar to the various grasps to investigate whether the effects of the manipulative grasp is truly due to its affording qualities. The drawback to this approach is that as a separate entity, the piece of wood has certain affording properties that could potentially alter the affordances of the entire scene based on its relationship with the tool and object (e.g. screwdriver-nail with the wood positioned near the screwdriver). In addition, the human grasp-posture can be substituted with a robotic arm, prosthetic device, arm belonging to another animal (such as a chicken arm) or even a rubber arm, and it can be investigated whether artificial hand-like entities elicit similar interference effects as an actual human

arm. Such studies can shed light on how an observer encodes artificial agents into his/her own motor substrate. However this is an open-ended question and worth investigating in future studies.

The importance of the manipulative or operant tool-end over the functional tool-end is an aspect worth investigating as results in Aims 2 and 3 showed that participants primarily parsed the operant tool-end (spoon-bowl) and sparsely attended to the functional or graspable tool-end (spoon-handle). In addition, it was the manipulative grasp of the operant tool-end that was a gaze attractor unlike the functional grasp. A potential explanation could be the higher affordances of the operant tool-end with respect to the object, over the functional tool-end. For example, it is a spoons bowl rather than stem that actually engages with a cup of ice cream and a non-functional grasp of the bowl does not allow engaging the spoon on the ice-cream carton. Therefore, additional neural and visual resources might be allocated towards the operant or manipulative tool-end to decode the intent of a manipulative grasp. The importance of the operant tool-end over the functional tool-end follows recent work by van der Linden et al. who measured gaze landing positions as participants categorized images depicting stand-alone tools that were horizontally oriented (van der Linden et al., 2015). The authors found that while initial saccades landed near the tools center of gravity, refixations tended to be biased towards the operant tool-end than the graspable tool-end. Their experimental design was largely different from all three Aims here given the differing task requirements. More importantly, rather than a stand-alone tool, the tools in our stimuli were paired with other objects and viewed from an egocentric perspective. The pairing of the tool with other objects (from a right-handed egocentric perspective) resulted in a much greater ratio in the bias between the operant tool-end and functional tool-end (Natraj et al., 2015) when compared to van der Linden et al. (van der Linden et al., 2015). Results in this dissertation reinforces and strengthens the action-performing hypothesis that posits visuospatial attention to be biased towards the direction implied by the tool (Roberts and Humphreys, 2011, van der Linden et al., 2015), highlighting the importance of the operant tool-end over the functional tool-end. These unique differences between the functional tool-end and manipulative/operant tool-end on gaze control with

respect to evaluating tool-object relationships merits further investigating. This can be carried out, for example, by evaluating if visual preference for the manipulative tool-end over the functional tool-end persists when the functional tool-end is any other unrelated rigid structure (e.g. spoons bowl with a scissors graspable end) or if the functional tool-end clearly does not afford tool-object action (e.g. replacing the rigid stem of a spoon with a soft roll of paper). By manipulating the stimuli in this manner, the affordances of the tools operant end with respect to the object can be modulated and it remains to be seen the action-performing hypothesis is still satisfied.

In both Aims 2 and the 500ms experiment in Aim 3, eye tracking results revealed a stereotypical leftward bias in initial attention after image onset. Other visual perception studies have also documented this leftward bias in initial attention (for e.g., see (Dickinson and Intraub, 2009)). The authors, who were investigating the boundary extension phenomenon, found that when fixation was maintained at center prior to the appearance of an image with two random objects/object clusters, one in each hemifield, the first saccade exhibited a significant leftward bias. The experimental design here as well entailed participants to maintain fixation at center prior to viewing an image with a tool-object pair, one in each hemifield. The authors presented many plausible hypotheses for this leftward bias including right-hemispheric specialization for attention or the fact that English is read from left to right (Volberg and Hbner, 2004, Spalek and Hammad, 2005, Dickinson and Intraub, 2009). On the other hand, the leftward bias could also be due to inherent differences between the two experiments. Unlike the previous study (Dickinson and Intraub, 2009), the experiment here required right-handed participants to evaluate right-handed, egocentric tool-object pairs rather than random objects/object clusters. It is possible that the object-oriented action priming effect could have been a driving factor for the initial leftward bias as the object was always in the left hemifield. However, it should also be noted that some of the participants here were very strongly leftward biased. Thus rather than being reflective of any cognitive mechanism, the leftward bias may simply be a consequence of eye dominance (Bourassa, 1996, Ossandon et al., 2014). Recent research is starting to uncover the mechanisms underlying this bias and it appears to be driven in right-handers by

asymmetries in the allocation of initial attention once the two-object (one in each hemifield) stimulus appears (Ossandon et al., 2014). Given that this bias can potentially be perturbed (Foulsham et al., 2013), it is unclear what effect affordances might elicit on the leftward bias. The dataset in the 500ms experiment of Aim 3 offer a potential opportunity to further explore this leftward bias. For example, it can be evaluated if there is an interaction between stimulus type and initial saccade direction on pre-saccadic neural activity. While these aspects merit further investigation, as such they are outside the current scope and hypotheses of this dissertation.

There are a few final points to consider if the results of this dissertation were to be translated to future studies. All the studies in this dissertation fall in the domain of visual perception, but given the direct matching hypothesis (Flanagan and Johansson, 2003, Ambrosini et al., 2011), it is likely that similar brain regions might be recruited during actual tool-use. For instance, it can be hypothesized that an intentional manipulative tool-grasp might differ from a functional tool-grasp in the amount of activity elicited outside of canonical left parietofrontal tool-use neurons. However, it is unclear whether right parietofrontal activity might occur in the planning phase, in the movement execution phase or if they are synchronized to oscillations in left parietofrontal regions. Another point to consider is distinguishing the role of saccades vs. foveation on ERP responses. The direct link between saccades and ERP was observed in two scenarios. First, it was observed over occipito-parietal electrodes in the 500ms experiment of Aim 3 wherein first saccade initiation latencies predicted an ERP positivity at time-points outside of those involved in affordance processing. Second, saccades delayed the propagation of a posterior to frontal ERP difference between the manipulative grasp and no hand conditions in the 500 experiment. However, there was a clear posterior to frontal propagation of the ERP difference in the 100ms experiment of Aim 1 wherein saccades were sparsely executed. With respect to the effect of foveation on ERP responses, results showed that the ability of the observer to continuously gather visual input in the 500ms experiment (foveally and extrafoveally) altered the polarity of the ERP responses when compared to the 100ms experiment and allowed for the buildup of late parietal ERP differences that ostensibly depend on continuous

visual input. Further work is necessary to understand the interplay between affordances, saccades, fixations (foveal attention) and extra foveal attention on neural activity. This can be carried out in the same GLM framework outlined in the Methods section of Aim 3 by modeling saccades and fixations with separate regressors (Dandekar et al., 2012). It should also be noted that in Aim 3, we had counterbalanced the response hand used to record correct/incorrect decisions. It is unclear what effect using only a single response hand might have on ERP responses. It can be argued that since most of the ERP differences represent higher order motor cognition, they should be independent of the response hand. Alternatively, the counter argument can be made that as right-handed participants may be mapping the egocentric right-handed grasp onto their own motor substrates, neural circuits could be differentially primed when participants respond to the manipulative grasp with their right vs. left hand. The dataset of the 500ms experiment in Aim 3 offers an opportunity to explore the interaction between grasp-type and response hand on neural activity. However, as such this hypothesis is outside the scope of the dissertation especially given that behavioral data was counter-balanced. Finally, from an application perspective, results in this study can potentially be used in rehabilitative setting especially with the amputee population, as detailed earlier. However, concepts and results in this thesis are also applicable in settings that evaluate skill expertise. For example, the stimuli set can be broadened to include tools or gestures that might appear confounding to lay-subjects but might carry special significance to experts such as trained musicians or skilled workers. It can be then be evaluated whether the two groups (skilled vs. unskilled) show differences in predictive eye movements or in the neural encoding of context and grasp. In this regard, eye tracking shows promise given its relative ease of use and simplicity given advances in data acquisition techniques. Research is currently underway to evaluate whether eye movements when watching skillful hand-object actions transition from exploratory to predictive behavior as a novice gains expertise in the task.

APPENDIX A

BAYESIAN-MARKOV MODEL OF FOVEAL ATTENTION

To determine the foveal weighting of the AOI in an image, the AOI are considered as nodes in a graph. Consecutive gaze position within the boundaries of a node constitutes a fixation while transitions between the nodes constitute a saccade. The time course of foveal gaze position over the nodes of the graph can then be modeled as a first order discrete Markov process as prior research has shown this model to best describe similar eye movements over discrete elements of a static visual scene (Hacisalihzade et al., 1992). By the Markov property, the probable location of foveal gaze position one time step into the future is independent of the past history given the current position.

$$P\left(X_{t+1} = j | X_t = i, X_{t-1} = l ... X_0 = k\right) = P\left(X_{t+1} = j | X_t = i\right) \tag{5}$$

$$= \theta_{ij} \tag{6}$$

where X_t is foveal gaze position at time t

when $i = j$, θ_{ii} is the probability of a fixation at AOI i

when $i \neq j$, θ_{ij} is the probability of a transition from AOI i to AOI j

These transition probabilities over the graph can be represented by the transition matrix **P**. A representative transition matrix for a 3 node graph is:

$$\mathbf{P} = \begin{array}{c} \\ O_t \\ M_t \\ F_t \end{array} \begin{array}{c} \begin{array}{ccc} O_{t+1} & M_{t+1} & F_{t+1} \end{array} \\ \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{pmatrix} \end{array} \tag{7}$$

The matrix $\mathbf{P}$ is row stochastic; $\sum_{j=1}^{3} \theta_{ij} = 1$ for any row i. Every row sums to 1 and is a categorical probability distribution of foveal gaze transitions from one time step to the next. For example, the first row is the probability of foveal gaze position at time $t+1$ given that foveal gaze is currently over the object at time t. The second and third rows denote similar transition probabilities when foveal gaze position is at the manipulative end of the tool and the functional end of the tool respectively. The following a priori assumptions are made on $\mathbf{P}$ and hence the structure of graph.

- A priori, the observer does not have a predisposed bias towards any AOI or subset of AOI, given that stimuli presentations are always presented in a randomized fashion. Therefore each row of $\mathbf{P}$ can be any valid distribution.
- The beliefs on the probability distribution of each row of $\mathbf{P}$ are updated at regular intervals as eye tracking data is collected.

These assumptions are formally implemented via Bayesian statistics (Minka, 2000, Bishop, 2006, Gupta, 2010). Each row of $\mathbf{P}$ is a categorical distribution and the conjugate prior for the categorical distribution is the Dirichlet distribution. The Dirichlet is a distribution over distributions; it provides a probability density function over the set of parameters that describe a categorical distribution. For the sake of notational simplicity, let $\boldsymbol{\theta} = \begin{bmatrix} \theta_1 & \theta_2 & \theta_3 \end{bmatrix}$ be the parameter that describes the transition probabilities when foveal gaze is at the object. The Dirichlet prior for this categorical distribution depends on $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix}$ and is given by

$$\mathrm{DIR}(\boldsymbol{\theta}|\boldsymbol{\alpha}) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \prod_{k=1}^{3} \theta_k^{\alpha_k - 1} \tag{8}$$

$$\alpha_0 = \sum_{k=1}^{3} \alpha_k \tag{9}$$

$$\Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u} du \tag{10}$$

The fraction with the factorial functions in (8) is a normalization constant to ensure that the Dirichlet is a valid distribution. Generally, any $\boldsymbol{\theta}$ of length L can be visualized as a probability vector that lies on a $L-1$ probability simplex, a surface characterized by the property that the sum of the coordinates of any point on it is 1. $\boldsymbol{\alpha}$ describes the shape of the probability density function on the simplex over all possible $\boldsymbol{\theta}$. For example, when $\alpha_k = 1 \ \forall \ k$, then $\text{DIR}(\boldsymbol{\theta}|\boldsymbol{\alpha})$ is a constant for for all possible $\boldsymbol{\theta}$. Thus, any distribution $\boldsymbol{\theta}$ is equally likely and can essentially be any point on the simplex. For other values of $\boldsymbol{\alpha}$, $\text{DIR}(\boldsymbol{\theta}|\boldsymbol{\alpha})$ is a probability density function over likely values of the parameter $\boldsymbol{\theta}$; the likely points on the simplex. Let $\mathbf{A}$ be the transition count matrix that records the number of one-step Markov transitions or occurrences from N sequential observations of eye tracking data.

$$\mathbf{A} = \begin{array}{c} \\ O_t \\ M_t \\ F_t \end{array} \begin{array}{c} \begin{array}{ccc} O_{t+1} & M_{t+1} & F_{t+1} \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \end{array} \tag{11}$$

Consider $\mathbf{a}$ to be the first row vector of $\mathbf{A}$ i.e., the one-step foveal gaze transitions from the object and for notational simplicity let $\mathbf{a} = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}$. Given N observations or all possible transition occurrences, the Bayesian estimate of $f\left(\boldsymbol{\theta}|\mathbf{a}, \boldsymbol{\alpha}\right)$ conditioned on the Dirichlet parameter $\boldsymbol{\alpha}$ is computed as follows:

The Dirichlet prior is chosen to model $\boldsymbol{\theta}$ the categorical transition probabilities

$$f\left(\boldsymbol{\theta}|\boldsymbol{\alpha}\right) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \prod_{k=1}^{3} \theta_k^{\alpha_k - 1} \tag{12}$$

The likelihood of the transitions is given by

$$f\left(\mathbf{a}|\boldsymbol{\theta}\right) = \prod_{k=1}^{3} \theta_k^{m_k} \tag{13}$$

$$m_k = \sum_{t=0}^{N-1} 1\left\{X_{t+1} = k | X_t = 1\right\} \tag{14}$$

$$\mathbf{m} = \begin{bmatrix} m_1 & m_2 & m_3 \end{bmatrix} \tag{15}$$

where m_k is the count of the transitions to AOI k, $\forall k \in (1,2,3)$ from the object ($k=1$) and by the Markov property, the transitions are independent. By Bayes' theorem, the posterior estimate of $\boldsymbol{\theta}$ is given by

$$f(\boldsymbol{\theta}|\mathbf{a},\boldsymbol{\alpha}) \propto f(\mathbf{a}|\boldsymbol{\theta})\, f(\boldsymbol{\theta}|\boldsymbol{\alpha}) \tag{16}$$

$$\propto \left(\prod_{k=1}^{3} \theta_k^{m_k}\right)\left(\prod_{k=1}^{3} \theta_k^{\alpha_k - 1}\right) \tag{17}$$

$$\propto \left(\prod_{k=1}^{3} \theta_k^{m_k + \alpha_k - 1}\right) \tag{18}$$

$$= \mathrm{DIR}(\boldsymbol{\theta}|\boldsymbol{\alpha} + \mathbf{m}) \tag{19}$$

Therefore the posterior estimate of $\boldsymbol{\theta}$ is also a Dirichlet density function over all likely $\boldsymbol{\theta}$ and its shape is defined by the number of transitions $\mathbf{m}$ and the parameter $\boldsymbol{\alpha}$. $\boldsymbol{\alpha}$ can be considered as the transition pseudo-count or prior belief that then determines the shape of the Dirichlet. While this density function provides a range of probable $\boldsymbol{\theta}$, a point estimate of the posterior is the central tendency or the mean of $\mathrm{DIR}(\boldsymbol{\theta}|\boldsymbol{\alpha} + \mathbf{m})$ given by

$$\mathrm{E}_{\mathrm{DIR}(\boldsymbol{\theta}|\boldsymbol{\alpha},\mathbf{a})}[\theta_k] = \frac{\alpha_k + m_k}{\sum_{k=1}^{L} \alpha_k + m_k} \tag{20}$$

The mean value of the parameter θ_k is thus computed by the ratio of the number of transitions to AOI k to the total number of transitions (including the pseudo-counts). When $\alpha_k + m_k > 1\ \forall k$, $\mathrm{DIR}(\boldsymbol{\theta}|\boldsymbol{\alpha} + \mathbf{m})$ is maximum at this mean value of $\boldsymbol{\theta}$ and when $\alpha_k + m_k = c$ $\forall k$, then the Dirichlet is symmetric about the uniform distribution. The shape of the Dirichlet is thus determined by the pseudo-counts and the observed transitions. As a result, for a stimulus with L AOI, entries in row i of the transition count matrix $\mathbf{A} \in \mathbb{R}^{L \times L}$ are the Dirichlet parameters that define the categorical distribution of Markov transitions from AOI i. Each row i of $\mathbf{P} \in \mathbb{R}^{L \times L}$ is the mean of the Dirichlet defined in row i of $\mathbf{A}$. In this Bayesian framework, $\mathbf{A}^{(t)}$ and $\mathbf{P}^{(t)}$ are updated after every observation $t = 0, 1, 2...N$, thereby estimating a homogeneous Markov chain at each time t given the prior data till $t-1$. This leads to the following framework governing the weighting of the AOI, generalizable to

stimuli with L discrete AOI.

A priori, at $t = 0$, each row distribution i of $\mathbf{P}^{(0)}$ is defined by a flat prior $f\left(\boldsymbol{\theta}_i^{(0)}|\boldsymbol{\alpha}_i^{(0)}\right) = \text{DIR}(\boldsymbol{\theta}_i^{(0)}|\boldsymbol{\alpha}_i^{(0)})$ where $\boldsymbol{\alpha}_i^{(0)} = \begin{bmatrix}1 & 1 & \dots 1\end{bmatrix}_{1\times L}$. The entry in the i^{th} row and j^{th} column of $\mathbf{A}^{(0)}$ ($a_{ij}^{(0)}$) and $\mathbf{P}^{(0)}$ ($\theta_{ij}^{(0)}$) are populated as follows

$$a_{ij}^{(0)} = \alpha_{ij}^{(0)} \tag{21}$$

$$= 1 \qquad \forall i, j \in \{1, 2, ...L\} \tag{22}$$

$$\theta_{ij}^{(0)} = \frac{a_{ij}^{(0)}}{\sum_{j=1}^{L} a_{ij}^{(0)}} \tag{23}$$

$$= \frac{1}{L} \qquad \forall i, j \in \{1, 2, ...L\} \tag{24}$$

At subsequent discrete steps $t = 1, 2, ...N$, $\mathbf{A}^{(t)}$ and $\mathbf{P}^{(t)}$ are updated in the following manner. If a transition is observed from AOI $i \rightarrow j$ at time t,

$$a_{ij}^{(t)} = a_{ij}^{(t-1)} + 1 \tag{25}$$

$$a_{kl}^{(t)} = a_{kl}^{(t-1)} + \delta \qquad \forall k \neq i\,,\, \forall l \in \{1, 2...L\} \tag{26}$$

$$\theta_{ij}^{(t)} = \frac{a_{ij}^{(t)}}{\sum_{j=1}^{L} a_{ij}^{(t)}} \qquad \forall i, j \in \{1, 2, ...L\} \tag{27}$$

This step increments a_{ij} in row i by 1 i.e., it updates the Dirichlet parameter that describes the categorical distribution of gaze transitions from AOI i. At the same time, every element of the other rows is incremented by a small random number δ. Theoretically δ can be defined in any manner (zero, constant, time-varying, or in proportion to $\boldsymbol{\theta}_i$ etc.). However for the purposes of Aim 2, δ is chosen to be ≈ 0.068. This incrementes the elements in the rows of $\mathbf{A}$ by 1 approximately every 300ms, accounting for extra-foveal attention. For example, if a fixation is observed over the first AOI at $t = 1$, then a_{11} is incremented by 1 and a_{kl} is incremented by δ (where $k \in \{2, 3..L\}\,,\, \forall l$). This has the effect of slightly tightening the Dirichlet around the uniform distribution in the absence of any observed transitions.

In the framework so far, two properties are implicit in the graphical structure of the Markov chain, reflected by the manner in which $\mathbf{P}^{(t)}$ is populated and updated. First, every AOI is assumed to be reachable to the observer since $\theta_{ij}^{(t)} > 0$ at any time step t in the

Bayesian update. Formally, the Markov chain is irreducible.

$$P(X_{t+1} = j | X_t = i) = \theta_{ij}^{(t)} \tag{28}$$

$$> 0 \qquad \forall t \in \{0, 1, 2...N\} \tag{29}$$

Second, the Markov chain is aperiodic as $\theta_{ii}^{(t)}$ in $\left(\mathbf{P}^{(t)}\right)^n$ is always > 0, $\forall n$ at any step t in the Bayesian update.

$$P(X_{t+n} = i | X_t = i) = \theta_{ii}^{(t)} \; in \; \left(\mathbf{P}^{(t)}\right)^n \tag{30}$$

$$> 0 \, \forall n \in \{1, 2, ...\} \tag{31}$$

In other words, gaze is not constrained to return to a current AOI only at multiples of some integer number of time steps. This allows for gaze jitter given motor noise in the planning of eye movements. Together, irreducibility and aperiodicity of the Markov chain guarantee the existence of a steady state distribution (DeGroot and Schervish, 2002) of foveal gaze over the L AOI at any time t of the Bayesian update, such that the following equations are satisfied (Ng et al., 2001).

$$\left(\mathbf{v}^{(t)}\right)^\top \mathbf{P}^{(t)} = \left(\mathbf{v}^{(t)}\right)^\top \tag{32}$$

$$\left[\left(\mathbf{P}^{(t)}\right)^\top - \mathbf{I}\right] \mathbf{v}^{(t)} = 0 \tag{33}$$

where $\mathbf{v}^{(t)} \in \mathbb{R}^{L \times 1}$ is the left eigenvector for the dominant eigenvalue of the transition matrix $\mathbf{P}^{(t)}$ (eigenvalue = 1). It is a column vector of steady state foveal gaze distribution over the L AOI.

$$\mathbf{v}^{(t)} = \begin{bmatrix} v_1^{(t)} \\ v_2^{(t)} \\ \vdots \\ v_L^{(t)} \end{bmatrix} \tag{34}$$

$$\sum_{k=1}^{L} v_k^{(t)} = 1 \tag{35}$$

$\mathbf{v}^{(t)}$ can be considered as a time-varying probability vector denoting the weighting or bias of foveal gaze over the AOI at time t given eye tracking data of fixation and saccades. Since

$\mathbf{P}^{(t)}$ is the point estimate (mean of each row) of $\mathbf{A}^{(t)}$, $\mathbf{v}^{(t)}$ is also the mean of the Dirichlet conditioned on the parameter $\boldsymbol{\gamma}$

$$f\left(\mathbf{v}^{(t)}|\boldsymbol{\gamma}^{(t)}\right) = \mathrm{DIR}\left(\mathbf{v}^{(t)}|\boldsymbol{\gamma}^{(t)} = \mathbf{v}^{(t)} \times (L+t)\right) \tag{36}$$

$$E_{\mathrm{DIR}(\mathbf{v}^{(t)}|\boldsymbol{\gamma}^{(t)})}[v_k^{(t)}] = \frac{\gamma_k^{(t)}}{\sum_{k=1}^{L}\gamma_k^{(t)}} \tag{37}$$

at any time t. For example at $t = 0$, the components of the dominant left eigenvector computed from $\mathbf{P}^{(0)}$ (that has every $\theta_{ij}^{(0)} = \frac{1}{L}\ \forall i, j$) is $v_k^{(0)} = \frac{1}{L}, \forall k$. The Dirichlet corresponding to this probability vector is given by (36,37)

$$f\left(\mathbf{v}^{(0)}|\boldsymbol{\gamma}^{(0)}\right) = \mathrm{DIR}\left(\mathbf{v}^{(0)}|\boldsymbol{\gamma}^{(0)} = \begin{bmatrix}1 & 1 & \dots & 1\end{bmatrix}_{1\times L}\right) \tag{38}$$

$$E_{\mathrm{DIR}(\mathbf{v}^{(0)}|\boldsymbol{\gamma}^{(0)})}[v_k^{(0)}] = \frac{1}{L} \tag{39}$$

Thus at $t = 0$, $\mathbf{v}^{(0)}$ is the uniform distribution that lies at the center of a $L-1$ probability simplex in $\mathbb{R}^L$ where each of the L axis ranges from 0 to 1 denoting the weighting of each AOI. $\mathbf{v}^{(0)}$ is also the mean of the Dirichlet $f\left(\mathbf{v}^{(0)}|\boldsymbol{\gamma}^{(0)}\right)$ that spans the entire simplex with equal probability and corresponds to the a priori assumption that the observer is unbiased towards any AOI. The trajectory of $\mathbf{v}^{(t)}$ through time on the simplex represents the time varying weighting of the AOI given fixations and saccades. If after N observations the AOI are weighted equally by the observer, then $\mathbf{v}^{(t)}$ should be within the 99% confidence interval of the Dirichlet $f\left(\mathbf{v}^{(N)}|\boldsymbol{\gamma}^{(N)}\right) = \mathrm{DIR}\left(\mathbf{v}^{(N)}|\boldsymbol{\gamma}^{(N)} = \begin{bmatrix}1+\frac{N}{L} & 1+\frac{N}{L} & \dots & 1+\frac{N}{L}\end{bmatrix}_{1\times L}\right)$ that is symmetric about the uniform distribution on the probability simplex. A pictorial representation of computing the AOI weight vector i.e. $\mathbf{v}^{(t)}$ is shown in Figure 47 and the mean trajectory from the data in Aim 2 is shown in Figure 48.

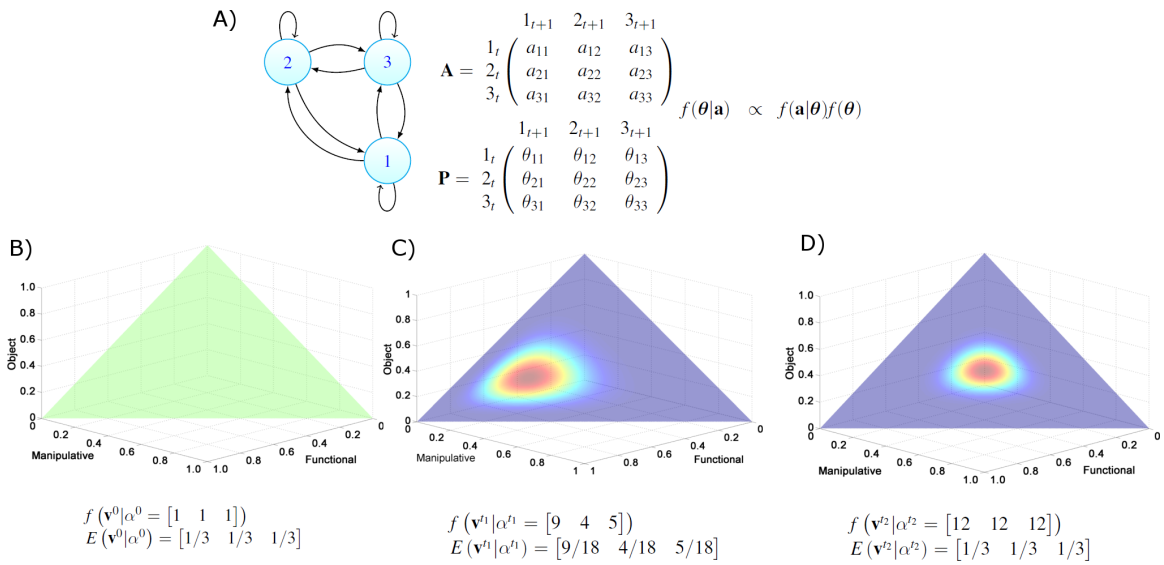

Figure 48: Computing the AOI weight vector or foveal bias. A) Markov-chain illustration of the 3 AOI in Aim 2. Each node corresponds to an AOI and the loops denoted fixation and saccade probabilities. Each row of $\mathbf{A}$ and $\mathbf{P}$ denotes the count and probability respectively of transitioning from one AOI to any other in one time-step. Fixation and saccade probabilities were estimated from the transition count matrix using Bayes' rule at each time step. The left eigenvector ($\mathbf{v}^\top$) of $\mathbf{P}$ at any time step is the AOI weight vector. B-D) Graphical representation of computing an AOI weight vector for a single trial. B) A flat Dirichlet prior given by the parameter α and distribution $f(\mathbf{v}|\alpha)$ was used for the AOI weight vector $\mathbf{v}$ before the start of the trial i.e., $\mathbf{v}$ could be any vector with uniform probability that lies on the simplex, a 2D surface. The co-ordinates of any point on the surface sum to 1. Each of the three axes represent the bias towards the 3 AOI (object, functional tool-end, manipulative tool-end). $\mathbf{v}$'s distribution (green) uniformly covers the entire surface, with a mean $\mathbf{v} = (E(\mathbf{v}|\alpha))$ given by the uniform distribution at the center. C) The posterior distribution and the mean of the AOI weight vector gets updated given samples of eye tracking data. D) Distribution of probable location of the AOI weight vector if the observer were to equally weight the AOI after 33 samples of eye tracking data given fixation counts and saccade frequencies. While the Bayesian update method gives a distribution for the AOI weight vector, the mean of the distribution was taken as the point-estimate of the trial's AOI weight vector.

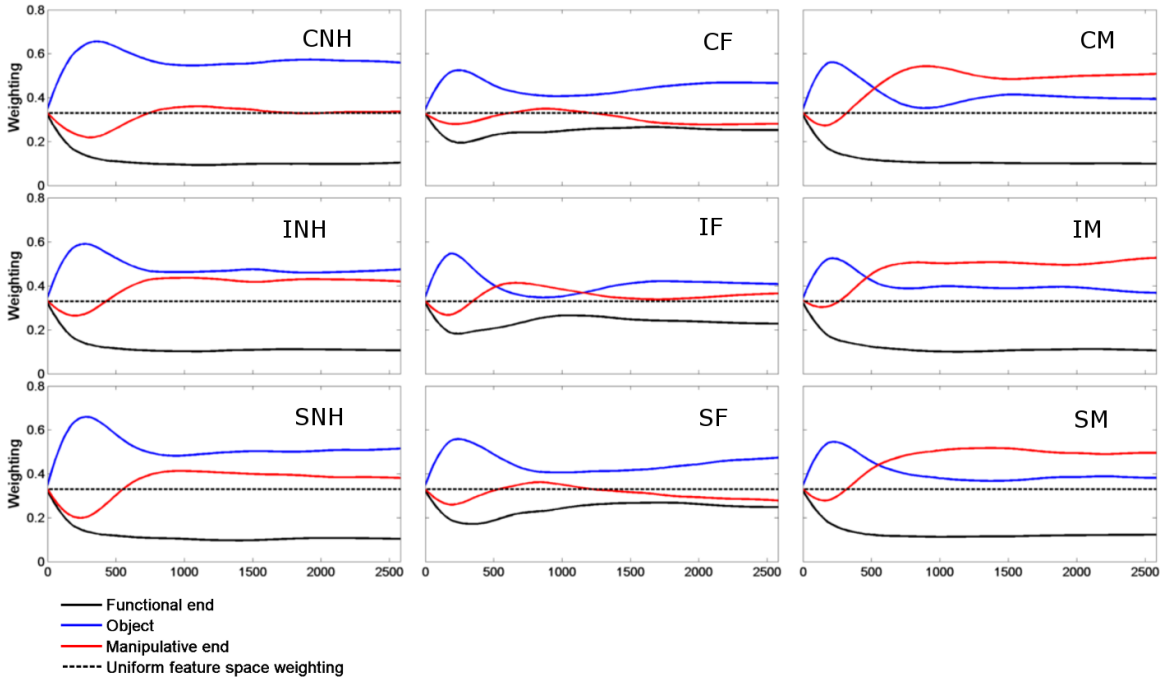

Figure 49: The mean time course of AOI bias estimated from the data in Aim 2 is shown, for each of the 9 conditions. For reference, the dotted line represents uniform bias towards the three AOI

REFERENCES

1. Almeida J, Mahon BZ, Caramazza A (2010) The role of the dorsal visual processing stream in tool identification. Psychological science 21:772-778.

2. Ambrosini E, Costantini M, Sinigaglia C (2011) Grasping with the eyes. Journal of neurophysiology 106:1437-1442.

3. Anderson T, Jenkins I, Brooks D, Hawken M, Frackowiak R, Kennard C (1994) Cortical control of saccades and fixation in man A PET study. Brain 117:1073-1084.

4. Ariff G, Donchin O, Nanayakkara T, Shadmehr R (2002) A real-time state predictor in motor control: study of saccadic eye movements during unseen reaching movements. The Journal of Neuroscience 22:7721-7729.

5. Aziz-Zadeh L, Koski L, Zaidel E, Mazziotta J, Iacoboni M (2006) Lateralization of the human mirror neuron system. The Journal of Neuroscience 26:2964.

6. Bahill AT, LaRitz T (1984) Why can't batters keep their eyes on the ball. American Scientist 72:249-253.

7. Bahill AT, Stark L (1979) The trajectories of saccadic eye movements. Scientific American 240:108-117.

8. Bishop CM (2006) Pattern recognition and machine learning: springer New York.

9. Borghi AM, Flumini A, Natraj N, Wheaton LA (2012) One hand, two objects: Emergence of affordance in contexts. Brain and cognition 80:64-73.

10. Bourassa D (1996) Handedness and eye-dominance: A meta-analysis of their relationship. Laterality: Asymmetries of Body, Brain and Cognition 1:5-34.

11. Buxbaum LJ (2001) Ideomotor apraxia: a call to action. Neurocase 7:445-458.

12. Caplan JB, Madsen JR, Schulze-Bonhage A, Aschenbrenner-Scheibe R, Newman EL, Kahana MJ (2003) Human oscillations related to sensorimotor integration and spatial learning. The Journal of Neuroscience 23:4726.

13. Castelhano MS, Mack ML, Henderson JM (2009) Viewing task influences eye movement control during active scene perception. Journal of Vision 9:6.

14. Chao LL, Martin A (2000) Representation of manipulable man-made objects in the dorsal stream. Neuroimage 12:478-484.

15. Chao LL, Weisberg J, Martin A (2002) Experience-dependent modulation of category-related cortical activity. Cerebral Cortex 12:545-551.

16. Chao ZC, Nagasaka Y, Fujii N (2010) Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkeys. Frontiers in neuroengineering 3.

17. Chikazoe J, Konishi S, Asari T, Jimura K, Miyashita Y (2007) Activation of right inferior frontal gyrus during response inhibition across response modalities. Journal of Cognitive Neuroscience 19:69-80.

18. Chong TT-J, Cunnington R, Williams MA, Kanwisher N, Mattingley JB (2008) fMRI adaptation reveals mirror neurons in human inferior parietal cortex. Current biology 18:1576-1580.

19. Corbetta M, Akbudak E, Conturo TE, Snyder AZ, Ollinger JM, Drury HA, Linenweber MR, Petersen SE, Raichle ME, Van Essen DC (1998) A common network of functional areas for attention and eye movements. Neuron 21:761-773.

20. Creem-Regehr SH, Lee JN (2005) Neural representations of graspable objects: are tools special? Cognitive Brain Research 22:457-469.

21. Cressie N, Wikle CK (2011) Statistics for spatio-temporal data: John Wiley & Sons.

22. Culham JC, Kanwisher NG (2001) Neuroimaging of cognitive functions in human parietal cortex. Current opinion in neurobiology 11:157-163.

23. Culham JC, Valyear KF (2006) Human parietal cortex in action. Current opinion in neurobiology 16:205-212.

24. Cusack WF, Cope M, Nathanson S, Pirouz N, Kistenberg R, Wheaton LA (2011) Neural activation differences in amputees during imitation of intact versus amputee movements. Frontiers in human neuroscience 6:182-182.

25. Cusack WF, Thach S, Patterson R, Acker D, Kistenberg RS, Wheaton LA (2015) Enhanced Neurobehavioral Outcomes of Action Observation Prosthesis Training. Neurorehabilitation and neural repair 1545968315606992.

26. Dandekar S, Privitera C, Carney T, Klein SA (2012) Neural saccadic response estimation during natural viewing. Journal of neurophysiology 107:1776-1790.

27. DeGroot MH, Schervish MJ (2002) Probability and statistics.

28. Del Giudice M, Manera V, Keysers C (2009) Programmed to learn? The ontogeny of mirror neurons. Developmental Science 12:350–363.

29. Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of neuroscience methods 134:9-21.

30. Di Russo F, Martínez A, Sereno MI, Pitzalis S, Hillyard SA (2002) Cortical sources of the early components of the visual evoked potential. Human brain mapping 15:95-111.

31. Dickinson CA, Intraub H (2009) Spatial asymmetries in viewing and remembering scenes: Consequences of an attentional bias? Attention, Perception, & Psychophysics 71:1251-1262.

32. Efron B, Tibshirani R (1993) An introduction to the bootstrap: CRC press.

33. Egner T, Hirsch J (2005) Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. Nature neuroscience 8:1784-1790.

34. Engel AK, Fries P (2010) Beta-band oscillations—signalling the status quo? Current opinion in neurobiology 20:156-165.

35. Falkenstein M, Hoormann J, Hohnsbein J (2001) Changes of error-related ERPs with age. Experimental Brain Research 138:258-262.

36. Flanagan JR, Johansson RS (2003) Action plans used in action observation. Nature 424:769-771.

37. Foulsham T, Gray A, Nasiopoulos E, Kingstone A (2013) Leftward biases in picture scanning and line bisection: A gaze-contingent window study. Vision research 78:14-25.

38. Fridman EA, Immisch I, Hanakawa T, Bohlhalter S, Waldvogel D, Kansaku K, Wheaton L, Wu T, Hallett M (2006) The role of the dorsal stream for gesture production. Neuroimage 29:417-428.

39. Gaymard B, Ploner C, Rivaud S, Vermersch A, Pierrot-Deseilligny C (1998) Cortical control of saccades. Experimental Brain Research 123:159-163.

40. Genovese CR, Lazar NA, Nichols T (2002) Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate* 1. Neuroimage 15:870-878.

41. Gentilucci M (2002) Object motor representation and reaching–grasping control. Neuropsychologia 40:1139-1153.

42. Geschwind N (1965) Disconnexion syndromes in animals and man. Brain 88:585-585.

43. Gibson J (1977) The concept of affordances. Perceiving, acting, and knowing 67-82.

44. Glover S, Miall RC, Rushworth MF (2005) Parietal rTMS disrupts the initiation but not the execution of on-line adjustments to a perturbation of object size. Journal of Cognitive Neuroscience 17:124-136.

45. Goh JO, Tan JC, Park DC (2009) Culture modulates eye-movements to visual novelty. PLoS One 4:e8238.

46. Goldenberg G (2003) Apraxia and beyond: life and work of Hugo Liepmann. Cortex 39:509-524.

47. Goldenberg G (2009) Apraxia and the parietal lobes. Neuropsychologia 47:1449-1459.

48. Gómez-Herrero G (2007) Automatic Artifact Removal (AAR) toolbox v1. 3 (Release 09.12. 2007) for MATLAB.

49. Gonzalez RC, Woods RE, Eddins SL (2009) Digital image processing using MATLAB.

50. Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. Trends in neurosciences 15:20-25.

51. Grèzes J, Tucker M, Armony J, Ellis R, Passingham RE (2003) Objects automatically potentiate action: an fMRI study of implicit processing. European Journal of Neuroscience 17:2735-2740.

52. Gupta M (2010) Introduction to the Dirichlet distribution and related processes. Technical Report UWEETR-2010-0006, Department of Electrical Engineering, University of Washington. URL http://www.ee.washington.edu/research/guptalab/publications/UWEETR-2010-0006.pdf.

53. Gutteling TP, Petridou N, Dumoulin SO, Harvey BM, Aarnoutse EJ, Kenemans JL, Neggers SF (2015) Action Preparation Shapes Processing in Early Visual Cortex. The Journal of Neuroscience 35:6472-6480.

54. Haaland KY, Harrington DL, Knight RT (2000) Neural representations of skilled movement. Brain 123:2306-2313.

55. Hacisalihzade SS, Stark LW, Allen JS (1992) Visual perception and sequences of eye movement fixations: A stochastic modeling approach. Systems, Man and Cybernetics, IEEE Transactions on 22:474-481.

56. Haith AM, Reppert TR, Shadmehr R (2012) Evidence for hyperbolic temporal discounting of reward in control of movements. The Journal of Neuroscience 32:11727-11736.

57. Halsband U, Lange RK (2006) Motor learning in man: a review of functional and clinical studies. Journal of Physiology-Paris 99:414-424.

58. Hayes MH (2009) Statistical digital signal processing and modeling: John Wiley & Sons.

59. He P, Wilson G, Russell C (2004) Removal of ocular artifacts from electro-encephalogram by adaptive filtering. Medical and biological engineering and computing 42:407-412.

60. Heilman KM, Maher LM, Greenwald ML, Rothi LJ (1997) Conceptual apraxia from lateralized lesions. Neurology 49:457-464.

61. Heilman KM, Rothi LJ, Valenstein E (1982) Two forms of ideomotor apraxia. Neurology 32:342-342.

62. Homan RW, Herman J, Purdy P (1987) Cerebral location of international 10–20 system electrode placement. Electroencephalography and clinical neurophysiology 66:376-382.

63. Humphreys GW, Yoon EY, Kumar S, Lestou V, Kitadono K, Roberts KL, Riddoch MJ (2010) The interaction of attention and action: From seeing action to acting on perception. British Journal of Psychology 101:185-206.

64. Hyvärinen A, Karhunen J, Oja E (2004) Independent component analysis: John Wiley & Sons.

65. Iacoboni M, Dapretto M (2006) The mirror neuron system and the consequences of its dysfunction. Nature Reviews Neuroscience 7:942-951.

66. Iacoboni M, Molnar-Szakacs I, Gallese V, Buccino G, Mazziotta JC, Rizzolatti G (2005) Grasping the intentions of others with one's own mirror neuron system. PLoS biology 3:e79.

67. Järveläinen J, Schuermann M, Hari R (2004) Activation of the human primary motor cortex during observation of tool use. Neuroimage 23:187-192.

68. Johnson-Frey SH (2004) The neural bases of complex tool use in humans. Trends in cognitive sciences 8:71-78.

69. Johnson-Frey SH, Maloof FR, Newman-Norlund R, Farrer C, Inati S, Grafton ST (2003) Actions or hand-object interactions? Human inferior frontal cortex and action observation. Neuron 39:1053-1058.

70. Johnson-Frey SH, Newman-Norlund R, Grafton ST (2005) A distributed left hemisphere network active during planning of everyday tool use skills. Cerebral cortex 15:681-695.

71. Jonides J, Badre D, Curtis C, Thompson-Schill SL, Smith EE (2002) Mechanisms of conflict resolution in prefrontal cortex. Principles of frontal lobe function 233-245.

72. Kalénine S, Buxbaum LJ, Coslett HB (2010) Critical brain regions for action recognition: lesion symptom mapping in left hemisphere stroke. Brain awq210.

73. Kandel ER, Schwartz JH, Jessell TM (2000) Principles of neural science: McGraw-Hill New York.

74. Kelly R, Mizelle J, Wheaton LA (2015) Distinctive laterality of neural networks supporting action understanding in left-and right-handed individuals: An EEG coherence study. Neuropsychologia.

75. Kelly RL, Wheaton LA (2013) Differential mechanisms of action understanding in left and right handed subjects: the role of perspective and handedness. Frontiers in psychology 4.

76. Kirchner H, Thorpe SJ (2006) Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. Vision research 46:1762-1776.

77. Kourtzi Z, Kanwisher N (2000) Activation in human MT/MST by static images with implied motion. Journal of cognitive neuroscience 12:48-55.

78. Króliczak G, Frey SH (2009) A common network in the left cerebral hemisphere represents planning of tool use pantomimes and familiar intransitive gestures at the hand-independent level. Cerebral Cortex bhn261.

79. Kutas M, Federmeier KD (2011) Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). Annual review of psychology 62:621.

80. Lakens D (2013) Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Frontiers in psychology 4.

81. Land M, Mennie N, Rusted J (1999) The roles of vision and eye movements in the control of activities of daily living. Perception-London 28:1311-1328.

82. Land MF, Hayhoe M (2001) In what ways do eye movements contribute to everyday activities? Vision research 41:3559-3565.

83. Levenstien MA, Yang Y, Ott J (2003) Statistical significance for hierarchical clustering in genetic association and microarray expression studies. BMC bioinformatics 4:62.

84. Levine TR, Hullett CR (2002) Eta squared, partial eta squared, and misreporting of effect size in communication research. Human Communication Research 28:612-625.

85. Lewis JW (2006) Cortical networks related to human use of tools. The Neuroscientist 12:211-231.

86. Liavas AP, Regalia PA (1999) On the numerical stability and accuracy of the conventional recursive least squares algorithm. Signal Processing, IEEE Transactions on 47:88-96.

87. Liepmann H (1900) Das Krankheitsbild Der Apraxie (" motorischen Asymbolie"): Karger.

88. Liepmann H (1980) Translations from Liepmann's Essays on Apraxia: London: Department of Psychology, University of Western Ontario.

89. Luck SJ (2014) An introduction to the event-related potential technique: MIT press.

90. Mahon BZ, Milleville SC, Negri GA, Rumiati RI, Caramazza A, Martin A (2007) Action-related properties shape object representations in the ventral stream. Neuron 55:507-520.

91. Makeig S (1993) Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. Electroencephalography and Clinical Neurophysiology 86:283-293.

92. Makeig S, Debener S, Onton J, Delorme A (2004) Mining event-related brain dynamics. Trends in Cognitive Sciences 8:204-210.

93. Maris E (2004) Randomization tests for ERP topographies and whole spatiotemporal data matrices. Psychophysiology 41:142-151.

94. Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG-and MEG-data. Journal of neuroscience methods 164:177-190.

95. Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annual review of neuroscience 24:167-202.

96. Minka T (2000) Estimating a Dirichlet distribution. Technical report, MIT.

97. Mishkin M, Ungerleider LG, Macko KA (1983) Object vision and spatial vision: two cortical pathways. Trends in neurosciences 6:414-417.

98. Mizelle J, Kelly RL, Wheaton LA (2013) Ventral encoding of functional affordances: a neural pathway for identifying errors in action. Brain and cognition 82:274-282.

99. Mizelle J, Wheaton LA (2010a) Neural activation for conceptual identification of correct versus incorrect tool–object pairs. Brain research 1354:100-112.

100. Mizelle J, Wheaton LA (2010b) Why is that hammer in my coffee? A multimodal imaging investigation of contextually based tool understanding. Frontiers in human neuroscience 4.

101. Mizelle JC, Tang T, Pirouz N, Wheaton LA (2011) Forming tool use representations: a neurophysiological investigation into tool exposure. Journal of cognitive neuroscience 23:2920-2934.

102. Molnar-Szakacs I, Kaplan J, Greenfield PM, Iacoboni M (2006) Observing complex action sequences: the role of the fronto-parietal mirror neuron system. Neuroimage 33:923-935.

103. Mort DJ, Perry RJ, Mannan SK, Hodgson TL, Anderson E, Quest R, McRobbie D, McBride A, Husain M, Kennard C (2003) Differential cortical activation during voluntary and reflexive saccades in man. Neuroimage 18:231-246.

104. Natraj N, Pella YM, Borghi AM, Wheaton LA (2015) The visual encoding of tool-object affordances. Neuroscience.

105. Natraj N, Poole V, Mizelle J, Flumini A, Borghi AM, Wheaton LA (2013) Context and hand posture modulate the neural dynamics of tool–object perception. Neuropsychologia 51:506-519.

106. Ng AY, Zheng AX, Jordan MI (2001) Link analysis, eigenvectors and stability. International Joint Conference on Artificial Intelligence 17:903-910.

107. Nieuwenhuis S, Forstmann BU, Wagenmakers EJ (2011) Erroneous analyses of interactions in neuroscience: a problem of significance. Nature neuroscience 14:1105-1107.

108. Nobre AC, Sebestyen G, Gitelman D, Mesulam M, Frackowiak R, Frith C (1997) Functional localization of the system for visuospatial attention using positron emission tomography. Brain 120:515-533.

109. Nunez PL, Srinivasan R (2006) Electric fields of the brain: the neurophysics of EEG: Oxford university press.

110. Ochipa C, Rothi LG, Heilman KM (1989) Ideational apraxia: A deficit in tool selection and use. Annals of neurology 25:190-193.

111. Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97-113.

112. Oostenveld R, Fries P, Maris E, Schoffelen J-M (2010) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Computational intelligence and neuroscience 2011.

113. Ossandón JP, Onat S, König P (2014) Spatial biases in viewing behavior. Journal of vision 14:20.

114. Pascual-Marqui RD (2002) Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. Methods Find Exp Clin Pharmacol 24:5-12.

115. Pernet CR, Chauveau N, Gaspar C, Rousselet GA (2011) LIMO EEG: a toolbox for hierarchical LInear MOdeling of ElectroEncephaloGraphic data. Computational intelligence and neuroscience 2011:3.

116. Pfurtscheller G, Da Silva FL (1999) Event-related EEG/MEG synchronization and desynchronization: basic principles. Clinical neurophysiology 110:1842-1857.

117. Picton TW (1992) The P300 wave of the human event-related potential. Journal of clinical neurophysiology 9:456-479.

118. Proverbio AM, Riva F, Zani A (2009) Observation of static pictures of dynamic actions enhances the activity of movement-related brain areas. PLoS One 4:e5389.

119. Ramayya AG, Glasser MF, Rilling JK (2010) A DTI investigation of neural substrates supporting tool use. Cerebral Cortex 20:507-516.

120. Raudenbush SW, Bryk AS (2002) Hierarchical linear models: Applications and data analysis methods: Sage.

121. Riddoch M, Humphreys G, Edwards S, Baker T, Willson K (2003) Actions glue objects but associations glue words: Neuropsychological evidence for multiple object selection. Nature Neuroscience 6:82-89.

122. Riddoch M, Humphreys G, Hickman M, Clift J, Daly A, Colin J (2006) I can see what you are doing: Action familiarity and affordance promote recovery from extinction. Cognitive neuropsychology 23:583-605.

123. Roberts KL, Humphreys GW (2011) Action-related objects influence the distribution of visuospatial attention. The Quarterly journal of experimental psychology 64:669-688.

124. Rolston JD, Wagenaar DA, Potter SM (2007) Precisely timed spatiotemporal patterns of neural activity in dissociated cortical cultures. Neuroscience 148:294-303.

125. Roth ZN, Zohary E (2015) Position and Identity Information Available in fMRI Patterns of Activity in Human Visual Cortex. The Journal of Neuroscience 35:11559-11571.

126. Rothi LJG, Heilman KM (1996) Liepmann (1900 and 1905): A definition of apraxia and a model of praxis. Classic cases in neuropsychology 1:111-122.

127. Ruchkin DS, Johnson R, Canoune H, Ritter W (1990) Short-term memory storage and retention: An event-related brain potential study. Electroencephalography and clinical Neurophysiology 76:419-439.

128. Simon JR, Berbaum K (1990) Effect of conflicting cues on information processing: the 'Stroop effect'vs. the 'Simon effect'. Acta psychologica 73:159-170.

129. Smith SM (2014) Overview of fMRI analysis. The British Journal of Radiology.

130. Somers DC, Dale AM, Seiffert AE, Tootell RB (1999) Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. Proceedings of the National Academy of Sciences 96:1663-1668.

131. Spalek TM, Hammad S (2005) The left-to-right bias in inhibition of return is due to the direction of reading. Psychological Science 16:15-18.

132. Steffensen SC, Ohran AJ, Shipp DN, Hales K, Stobbs SH, Fleming DE (2008) Gender-selective effects of the P300 and N400 components of the visual evoked potential. Vision research 48:917-925.

133. Stout D, Chaminade T (2012) Stone tools, language and the brain in human evolution. Philosophical Transactions of the Royal Society B: Biological Sciences 367:75-87.

134. Theodoridis S, Pikrakis A, Koutroumbas K, Cavouras D (2010) Introduction to Pattern Recognition: A Matlab Approach: A Matlab Approach: Access Online via Elsevier.

135. Thill S, Caligiore D, Borghi AM, Ziemke T, Baldassarre G (2013) Theories and computational models of affordance and mirror systems: an integrative review. Neuroscience & BioBehavioral Reviews 37:491-521.

136. Tranel D, Damasio H, Damasio AR (1997) A neural basis for the retrieval of conceptual knowledge. Neuropsychologia 35:1319-1327.

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

138. Tucker M, Ellis R (2001) The potentiation of grasp types during visual object categorization. Visual cognition 8:769-800.

139. Valyear KF, Culham JC (2010) Observing learned object-specific functional grasps preferentially activates the ventral stream. Journal of Cognitive Neuroscience 22:970-984.

140. van der Linden L, Mathôt S, Vitu F (2015) The role of object affordances and center of gravity in eye movements toward isolated daily-life objects. Journal of Vision 15:8-8.

141. van Elk M, van Schie HT, Neggers SF, Bekkering H (2010) Neural and temporal dynamics underlying visual selection for action. Journal of neurophysiology 104:972-983.

142. Van Gompel RP (2007) Eye movements: A window on mind and brain: Elsevier.

143. Vingerhoets G (2014) Contribution of the posterior parietal cortex in reaching, grasping, and using objects and tools. Frontiers in psychology 5.

144. Vingerhoets G, Acke F, Vandemaele P, Achten E (2009) Tool responsive regions in the posterior parietal cortex: effect of differences in motor goal and target object during imagined transitive movements. Neuroimage 47:1832-1843.

145. Vingerhoets G, Honoré P, Vandekerckhove E, Nys J, Vandemaele P, Achten E (2010) Multifocal intraparietal activation during discrimination of action intention in observed tool grasping. Neuroscience 169:1158-1167.

146. Volberg G, Hübner R (2004) On the role of response conflicts and stimulus position for hemispheric differences in global/local processing: an ERP study. Neuropsychologia 42:1805-1813.

147. Wheaton L, Fridman E, Bohlhalter S, Vorbach S, Hallett M (2009) Left parietal activation related to planning, executing and suppressing praxis hand movements. Clinical neurophysiology 120:980-986.

148. Wheaton LA, Hallett M (2007) Ideomotor apraxia: a review. Journal of the neurological sciences 260:1-10.

149. Wheaton LA, Nolte G, Bohlhalter S, Fridman E, Hallett M (2005a) Synchronization of parietal and premotor areas during preparation and execution of praxis hand movements. Clinical Neurophysiology 116:1382-1390.

150. Wheaton LA, Shibasaki H, Hallett M (2005b) Temporal activation pattern of parietal and premotor areas related to praxis movements. Clinical neurophysiology 116:1201-1212.

151. Yarbus AL (1967) Eye movements during perception of complex objects. In: Eye movements and vision, pp 171-211: Springer.

152. Yarbus AL, Haigh B, Rigss LA (1967) Eye movements and vision: Plenum press New York.

153. Yoon EY, Humphreys GW, Riddoch MJ (2010) The paired-object affordance effect. Journal of Experimental Psychology: Human Perception and Performance 36:812.