TEMPORAL CONNECTIVITY PATTERNS OF THE CORTICO-LIMBIC LEARNING AND REWARDS SYSTEM

by

Eliezer Yosef Kanal

B.S. in Psychology, University of Pittsburgh, 2004

Submitted to the Graduate Faculty of

the Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Eliezer Yosef Kanal

It was defended on

November 24, 2009

and approved by

Dr. Mingui Sun, Ph.D., Professor, Department of Bioengineering

Dr. Robert Boston, Ph.D., Professor, Department of Bioengineering

Dr. Charles Bradberry, Ph.D., Associate Professor, Department of Psychiatry

Dr. Julie Fiez, Ph.D., Professor, Department of Psychology

Dr. Robert Sclabassi, M.D., Ph.D., CEO, Computational Diagnostics, Inc.

Dissertation Director: Dr. Mingui Sun, Ph.D., Professor, Department of Bioengineering

TEMPORAL CONNECTIVITY PATTERNS OF THE CORTICO-LIMBIC LEARNING AND REWARDS SYSTEM

Eliezer Yosef Kanal, PhD

University of Pittsburgh, 2010

The human learning and rewards system is comprised of a number of cortical and subcortical neural regions, including the orbitofrontal cortex, striatum, and anterior cingulate. While modern neural imaging methods such as functional magnetic resonance imaging (fMRI) and functional positron emission tomography (PET) can successfully detect the activity of these regions, they cannot discern temporal activation patterns, due to the slow onset of the blood oxygen level dependent (BOLD) effect. Magnetoencephalographic imaging (MEG) is able to capture these temporal patterns but traditionally has been unable to detect activity originating from the deeper regions of the brain due to signal attenuation and high noise levels. The recently published exSSS method has shown significant promise extracting deep signals from MEG data. To elicit appropriate subcortical activity we utilized a previously published gambling task. This paradigm has been shown to differentially activate a number of subcortical regions within the rewards system, including the orbitofrontal cortex (OFC), striatum, and anterior cingulate cortex (ACC), based on reward-related feedback. MEG analysis using source localization methods in conjunction with source signal reconstruction techniques yielded neural activation time

courses for each of the regions of interest. Granger causality was used to identify the temporal relationships between each of these regions, and a possible functional connectivity map is presented. The behavioral paradigm was replicated using functional magnetic resonance imaging. fMRI activity patterns were similar to those previously reported in the literature using this paradigm. Additionally, the fMRI activation patterns were similar to those obtained via MEG source reconstruction of the exSSS-processed data. Our results support the literature finding that the rewards network is differentially activated based on feedback. Additionally, these results demonstrate the efficacy of the exSSS signal processing method for extracting deep activity, and suggest a possible use for MEG in the imaging of deep activity using other behavioral paradigms.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Learning and reward processing in the human brain occurs within a complex network of both higher- and lower-order systems. The process of detecting of the reward, associating it with a given stimulus, evaluating the relevance both of the reward and the determined association, and relating the newly learned association with previously formed ones requires a number of neural regions across the brain. This includes deeper regions such as the striatum, amygdala, hippocampus, and cingulate cortex, as well as more superficial regions, including the orbitofrontal cortex, the dorsolateral prefrontal cortex, and parietal cortex. That these regions are activated has been well known in the literature for decades; anatomical studies have demonstrated connectivity between the different regions of the network, and functional studies have shown that these regions are selectively activated in response to specific reward-related cues. For example, the striatum has been shown to play a crucial role in the initial acquisition of and subsequent maintenance of stimulus-response associations. The orbitofrontal cortex has been implicated in associating a given stimulus with a particular response, as well as updating information about an existing stimulus-response association if the nature of the association changes. Anterior cingulate cortex is often spoken of as a conflict monitoring region, with the functional responsibility of choosing between a number of possibly conflicting responses to a given stimulus.

However, the order of activation of the system, as well as the path taken by data flowing through the network, is as of yet unknown. Experiments in animal models using implanted electrodes or microdialysis techniques can help gain an understanding of components of the rewards system, but not the workings of the entire network as a whole. Additionally, human and animal reward systems can differ significantly. While modern neural noninvasive imaging methods such as functional magnetic resonance imaging (fMRI) and functional positron emission tomography (PET) can successfully detect activity in these regions, they cannot discern temporal activation patterns due to the slow onset of the blood oxygen level dependent (BOLD) effect. Electroencephalography (EEG) and magnetoencephalography (MEG) possess a temporal resolution on the order of milliseconds but are limited in space due to difficulties imaging the entire brain; deep brain sources are often difficult to extract from EEG and MEG data. Additionally, the electric fields detected by EEG are distorted by the tissue surrounding the brain, further complicating the difficulties in localizing detected signals.

Recent advances in MEG noise-reduction algorithms, however, have somewhat mitigated the difficulty in obtaining MEG signals from deep sources. MEG studies have historically been limited to only cortical imaging, due to both signal attenuation from deep neural sources as well as signal interference from superficial sources. Numerous techniques, on both the signal processing front and imaging front, have been proposed to solve this problem, but none have met with widespread acceptance, due to their difficulty in implementation, difficult to validate assumptions, or highly data-intensive nature.

Recently, a novel preprocessing technique called expanded signal space separation (exSSS) combines signal space separation (SSS) and beamspace methodology was proposed to enable the separation of recorded signals into deep and superficial

components. By modifying the leadfield matrices of the MEG sensor array, which dictates the relationship between the strength of a given neural magnetic source and its measured signal amplitude at the sensors, signals originating from deeper sources can be amplified while those coming from more superficial regions can be attenuated. Aside from allowing the imaging of deep structure activity, such a technique makes possible a wide variety of studies that were previously technically unfeasible. One specific application examines the temporal relationship between superficial and deep structures. The study of *in vivo* neural networks examines the activation patterns of various regions, but the relative time of onset of the different regions is often unclear, usually due to limits of the imaging modality. By utilizing the high temporal resolution of the MEG, we can construct a model describing the activation patterns of each identified neural source. The model can then be examined for dependencies between the different sources in an effort to examine the time of onset of each component, as well as feedback between the various components.

The advent of this technique allows for the unprecedented imaging of both deep and superficial neural sources at the high resolution characteristic of MEG. This, combined with existing methods to discern the actual waveform of neural source, allow us to recover and examine the waveforms of each neural source for almost the entire volume of the head.

The primary aim of this research is twofold. We would like to test the efficacy of the exSSS model in extracting deep activity from a whole-head neural recording. This will be accomplished by recording MEG data while the subject performs an experimental paradigm for which the activation pattern is known in the literature. The presence of the pattern will also be verified by replicating the paradigm in the original imaging modality. Additionally, the paradigm will be chosen such that it should activate the neural rewards system, enabling the examination of the activation

time courses for all relevant neural regions. The learning and rewards circuit is a model network for such an application. A significant body of research has examined the components of the reward network, and much is known about the mechanisms by which the components of the network become activated. However, very little is known about circuit's functional activation pattern in human subjects. An examination of the connectivity reveals a prominent striatal-OFC circuit.

As of yet, the exSSS method has not been tested on human data, but only on simulated dipoles. As such we will be examining the neural activity elicited during the rewards task using both MEG and fMRI. Through this multimodal approach we hope to gain a greater understanding two facets of the task and the neural activity it elicits. Firstly, we will examine whether we can replicate the literature reports of corticostriatal activation using this gambling task. Secondly, we will compare the fMRI and MEG activation patterns to see whether they are in agreement with each other, and examine possible sources of noise, systemic or otherwise, which may contribute to different findings.

We hypothesize that the gambling method will successfully elicit activity in the rewards pathway, and the detected patterns will correlate with the localization results of the exSSS-processed MEG data. Additionally, we hypothesize that the striatum will drive the activation of both the orbitofrontal cortex.

2.0 LEARNING AND REWARDS IN THE HUMAN BRAIN

The learning and rewards system is highly complex, both anatomically and functionally. The network contains a large amount of feedback and feedforward loops, error checking components, as well as storage systems. While much is yet unknown about the mechanisms of learning, the past few decades of research have resulted in a significant body of literature from which we can form strong hypotheses about the underlying mechanisms of learning.

2.1 ANATOMY AND PHYSIOLOGY

The human learning and rewards system is comprised of numerous cortical and deep structures. As can be seen in Figure 1, despite the omission of a number of relevant neural regions, this subset of the system is highly interconnected, and displays highly complex non-linear activation patterns. As such, our discussion will focus on three neural regions: the orbitofrontal cortex (OFC), the anterior cingulate cortex (ACC), and the striatum.

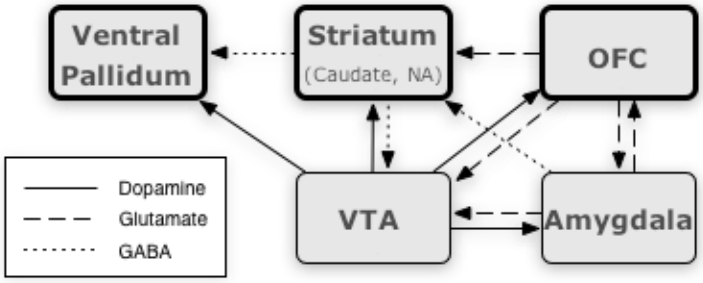


Figure 1: Some of the inter-regional connectivity within the learning and rewards pathway. Figure adapted from Figure 1 in Kalivas and Volkow [80].

2.1.1 Striatum

The striatum, located in the anterior of the basal ganglia, has been recognized as a crucial region in reward learning. The ventral region of the striatum, notably containing the nucleus accumbens (NA), has been shown to play a number of important roles in the cognitive and psychomotor aspects of learning. A number of animal experiments have demonstrated that the acquisition of stimulus-response (S-R) associations such as Pavlovian learning [74, 117] requires a functional NA. Additionally, the NA has been shown to drive the actual psychomotor response to a given reward [85, 118], as can demonstrated by the inappropriate learning experienced during selective NA shell activation [37].

The dorsal striatum, notably containing the head of the caudate nucleus, has been shown to maintain previously learned S-R associations [11, 35, 158, 160]. This maintenance entails continually re-establishing the previously-learned connection between an environmental stimulus and an expected outcome [76]. In the event that

the connection is no longer true, the caudate assists in extinguishing the association between stimulus and response [115]. This is particularly true within addiction research, where it has been established that the caudate maintains previously learned addictions [77, 158].

Additionally, each of these regions has been reported active performing a number of related functions. Error detection—the detecting of a discrepancy between the expected and actual outcomes—has been reported as a function of both of these regions, in that the ablation of either region in the animal model does not fully extinguish the animal's ability to detect discrepancies between expected and actual outcomes [139]. The determination of the relative magnitude, or "valence," of a given stimulus appears to involve a number of regions, including the caudate [41, 99]. Aside from valence, the striatum has also been implicated in determining whether a given reward is relatively "positive" or "negative", irrespective of magnitude [34].

It is worth noting that the basal ganglia contains a number of other regions crucial for rewards processing. The dorsal and ventral striatum receive substantial input from the substantia nigra (SN) and ventral tegmental area (VTA), respectively. These four regions collectively consist of one component of the rewards-related dopamine circuit, which has been the subject of extensive research [45, 60, 139, 140]. Outside the striatum, the amygdala has been implicated in the evaluation of the valence of a given stimulus [9], and hippocampal activity is often seen during learning and reward tasks [45, 57, 92].

2.1.2 Orbitofrontal cortex

In the cortex, the orbitofrontal cortex (OFC) has been shown to be highly involved in numerous aspects of rewards processing. Whereas the deeper regions mentioned

above seem to activate in response to any given rewarding stimuli, the OFC receives a input from a number of sensory regions, including gustatory, auditory, and visual, and has been shown to respond differentially to these signals based on their individual reward association [129, 152]. Reversal learning—in which animals are first taught to associate a given cue with one stimulus, and mid-experiment the association switches to another stimulus—has been shown to take place in the orbitofrontal cortex [131]. The OFC also has been seen to activate in response to the absence of an expected stimuli [92]. It is worth noting that while OFC activation during the viewing of rewarding stimuli is a common finding across modalities [71, 90, 164], for reasons likely due to paradigm differences this is not always the case [35, 99].

2.1.3 Anterior cingulate cortex

Anterior cingulate cortex (ACC) activity is often observed during decision tasks, and a number of hypotheses exist in the literature about the function of this activation. The most well known roles for the ACC are broadly-defined as conflict monitoring and outcome evaluation [13], consisting of the "overriding of prepotent responses, selecting a choice among equally permissible responses, [evaluating] errors" in a given choice situation, and "evaluating action outcomes" after feedback has been given [12]. Experimentally, ACC activation has been observed during a number tasks which can fall under these two monikers. Activation has been recorded during the evaluation of effort required and risks involved in reward seeking in both lesion studies [26] and behavioral experiments [42]. The ACC has also been implicated in determining the reward salience of a given response [86]. Some behavioral studies have elicited ACC involvement during the consideration of multiple possible outcomes to a given decision [14, 22]. Interestingly enough, a study examining humans with

naturally occurring ACC lesions found deficits in virtually all of the above, with a general "significant impairment of executive functions, including deficits in planning, monitoring of ongoing behavior, and strategy shifting" [121], all crucial functions of the learning and rewards network. Directly relevant to our research, numerous studies observed ACC activation during the feedback phase of gambling-type tasks [53, 73, 167].

2.1.4 Co-activated cortical regions

Other cortical regions are involved as well. The dorsolateral prefrontal cortex (DLPFC) has been implicated in arbitrary S-R association learning [123, 164]. This type of learning is characterized by the learned association of a specific response to a previously unassociated and unrelated cue. These associations are quite plastic, as subjects can learn, unlearn, and relearn arbitrary S-R associations very quickly [11]. Additionally, DLPFC activity has been observed during reward detection as well [162], and this activity is thought to be related to the formation of S-R associations, as mentioned above [162, 163]. The amygdala's role in emotion regulation is tightly linked to reward detection and recognition [9]. This role can be seen through a number of functions, including association formation [19], response formation [20], and response suppression [112]. Tracing literature has demonstrated that the hippocampus is a part of the learning and rewards network [3, 111], and numerous studies have seen hippocampal activation occurring concurrently with other rewards-related regions [52, 127].

2.2 REWARD SYSTEM CONNECTIVITY

One striking feature of the rewards network is the tremendous amount of apparent functional redundancy between the regions [25]. OFC and NA are both highly crucial for S-R learning, and OFC and striatum seem to be required for S-R maintenance. All three regions are necessary for error detection. While it is possible to extinguish S-R by lesioning one region and subjecting the animal to a very carefully designed experiment, the rewards network was inherently designed for redundancy, so that in the vast majority of practical situations learning can be executed by an animal or human lacking a subset of the system.

This redundancy becomes clearer when examining the anatomical connectivity between the regions (see Figure 1). The NA receives dense dopaminergic input from the ventral tegmental area (VTA) as identified by electrophysiologic recordings [44] and sends modulatory signals back to the VTA in a feedback mechanism [45]. Alexander et al. [3] used axonal tracing to identify neurons originating from the VTA, synapsing at the NA, continuing to the prefrontal regions, and then reversing the pattern, and labeled this the corticostriatal circuit. Gao et al. [50] used electrophysiological recordings in the rat model to demonstrate that stimulating the prefrontal cortex (PFC) elicits an inverse of the stimulation pattern in the VTA, suggesting a polysynaptic connection. Additionally, Shi [145] demonstrated a correlation between slow-oscillation bursting patterns in the rat VTA and similar firing patterns in the PFC, and also attributed the finding to a polysynaptic PFC-VTA connection. The PFC and NA are connected through the VTA; PFC stimulation increases extracellular dopamine in the NA core [84], but this release is only blocked by the injection of glutamate antagonists in the VTA [21], suggesting that the PFC-NA core connectivity is secondary to VTA activation via PFC glutamate which excites

VTA-NA dopamine neurons. The ACC receives significant input from the motor and premotor cortex [67], as well as secondary connectivity to the striatal regions [133]. As mentioned above, the OFC specifically receives significant input from various somatosensory regions. In the context of feedback networks, this implies a polysynaptic connection between these somatosensory regions and the striatal regions, as well.

Functionally, the corticolimbic network functions as a distributed reward processing network. Clinically, we can observe the severe impairment in decision making skills in individuals with damage to any part of the network. In one notable case, Eslinger and Damasio [43] described patient E.V.R., a patient with extensive OFC damage. He was described as possessing a normal IQ and being able to analyze ethical or social problems without difficulty, but showing profound impairment in any task requiring a decision [43, 137]. This behavior is typical of OFC damage [48]. Recently, there have been attempts to treat drug addiction by ablation of the NA [49, 68, 169]. While stereotactically lesioning the NA does help remove addiction, patients have been reported to undergo changes in personality, including difficulties in decision making skills [49]. Experimentally, this network has been verified by observing network behavior during learning and reward-related tasks. Liu et al. [99] had subjects perform a gambling task, and found that the entire network is activated, with each region responding to a particular aspect of the stimulus. He found that different subsets of the rewards network are activated throughout the task, and over the course of a single trial the entire network—including the OFC, caudate, and ACC—will be recruited to assist with rewards processing. Similarly, Boettiger and D'Esposito [11] used a complex pattern matching task in which the subject is required to learn sets of abstract spatial patterns to obtain a reward to demonstrate that learning and rewards processing activates the lateral PFC and striatal regions.

A number of such tasks have been used to demonstrate the network dependence of reward processing.

However, as suggested above, there is a degree of confusion regarding the roles of the particular nuclei. Depending on the activating paradigm the OFC seems to be activated during periods of reward expectation [71, 89] or not activated [35, 99]. The sensitivity of the NA to reward valence has been reported [110], but some studies have demonstrated NA activation independent of valence [34, 41]. On a different note, for the reasons mentioned above, virtually all human studies in the literature merely examine the presence or absence of activation patterns, without commenting on the temporal dynamics of these systems. Knowledge of the time-dependent relationships between these systems may significantly advance our understanding of the rewards system, and can help reduce the inconsistencies present in the literature.

Despite the problems with current imaging methods, we still form an educated guess regarding the temporal activation pattern. The VTA dopamine release upon relevant cue exposure is recognized as being the initial event in the learning pathway activation [36, 141]. As described above, the NA has the role of signal integrator, suggesting that it would be the final convergence point for the processed neural signals. The prefrontal cortex receives input from many sensory regions and outputs to the NA [3]. As such, it would follow that the prefrontal cortex plays a modulatory role in this network, as mentioned in the hypothesis.

2.3 GAMBLING PARADIGM

One reliable way to activate the rewards system has been to have the subject perform a gambling paradigm [2, 34, 53, 167, among others]. A variety of gambling

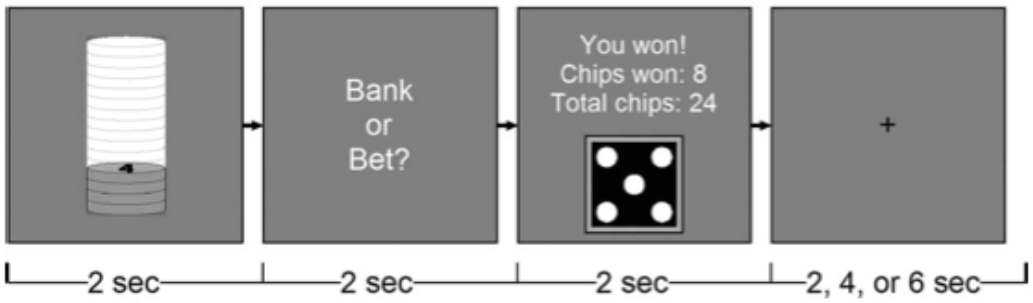


Figure 2: Schematic of a single MEG behavioral trial. The length of each segment is depicted on the schematic. Image taken from Figure 1B in Liu et al. [99].

tasks exist in the literature, but one task which activated the basal regions particularly strongly was the gambling task used by Liu et al. [99]. Liu and colleagues were attempting to "better understand the reward circuitry at different stages" of reward processing. They note how many of the existing papers in the literature offer seemingly conflicting reports of neural activity in the various regions of interest (ROI), likely due to differences between experimental paradigms. They specifically wanted to examine three aspects of reward processing; reward anticipation, outcome monitoring, and subsequent choice evaluation. To accomplish this goal they devised a gambling paradigm with four components to each trial; anticipation, response, feedback, and choice evaluation (Figure 2). Utilizing an event-related design, the researchers examined each component of the trial individually. The subject would be shown the money they would be gambling, then asked to choose either to bet or not bet that money. After the choice was indicated they would subsequently be shown the outcome. One important manipulation was that the outcome was always

shown, enabling subjects to feel regret for their choice. The possible outcomes are shown in Table 1.

Using these manipulations, they compared activity between three conditions. Each comparison corresponds to one of the characteristics of reward processing mentioned above. By comparing trials in which the subject chose to bet instead of choosing not to bet, we can examine activity during reward anticipation. Similarly, by examining differential activity between trials in which the subject won the gamble relative to losing trials we observe outcome monitoring. Finally, by comparing relatively—or "perceived"—correct trials to perceived incorrect ones, we can observe activity during choice evaluation.

The neural activation pattern resultant from this study (Figure 3) shows significant activation in a number of relevant deep regions of the brain; OFC, striatum, ACC, dorsomedial frontal cortex, and anterior insula, as well as less significant activation in the thalamus. This result is notable for our application in that all these regions are relatively deep. The lack of superficial activation can greatly enhance the signal emanating from deep regions. This result was one of the primary reasons why we chose this task for the present study.

Table 1: Visual depiction of possible perceived outcomes of gambling experiment.

	Possible Outcomes	
Perceived Correct	Bet & Win	Bank & Lose
Perceived Incorrect	Bet & Lose	Bank & Win

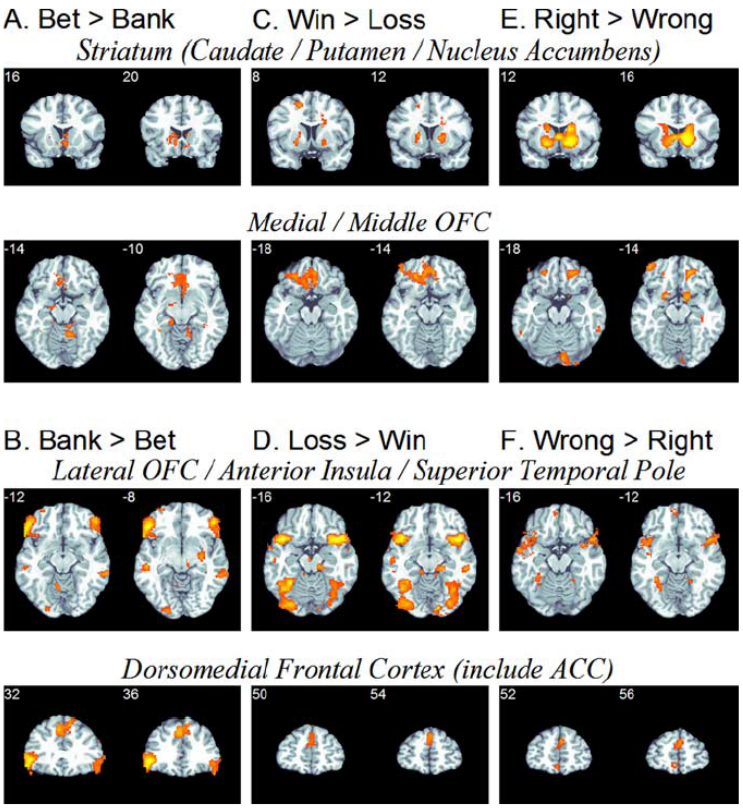


Figure 3: fMRI imaging results from Liu et al. [99]. A, C, and E show the striatum and medial OFC, and B, D, and F show the lateral OFC and insular cortex. The right side of the image is the right side of the brain. Figure and caption text taken from Figure 3 in Liu et al. [99].

Additionally, a significant signal differential was observed in the striatum and OFC during all three comparisons (Figure 4). A significant signal intensity differential between the two conditions of a given contrast (i.e., between bet and bank in the response contrast) is indicative of that region's responsibility in the given function.

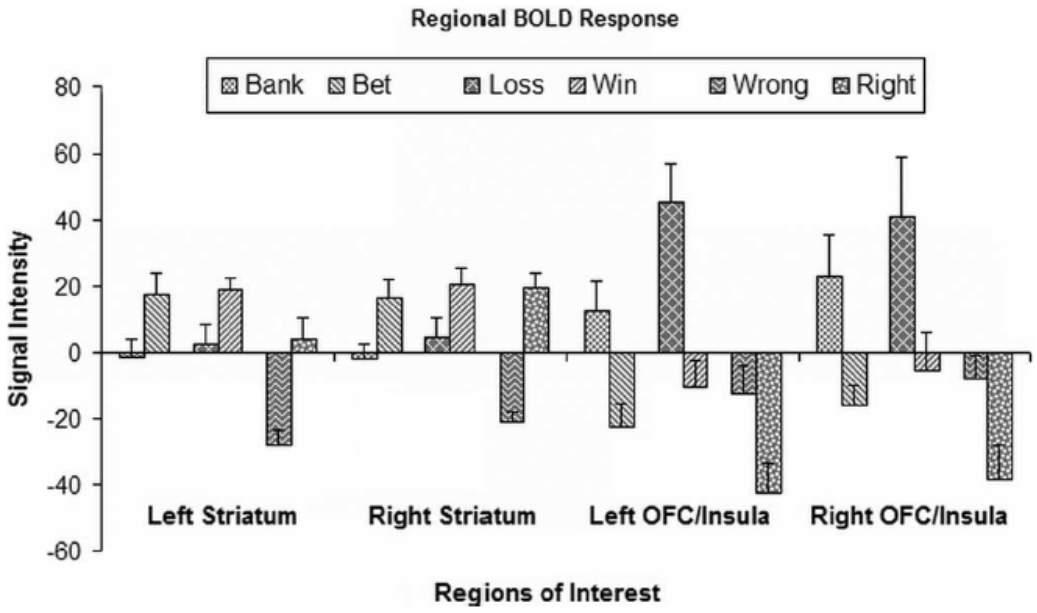


Figure 4: Averaged signal intensity across select regions of interest. Image taken from Figure 5 in [99].

Such differentials can be seen in the striatum during choice evaluation, as well as in the OFC/insular cortex during reward anticipation.

For our purposes, these findings suggested a paradigm that can activate deep reward-related regions. The strong signal differential between conditions, paired with the lack of superficial signal, made this paradigm an optimal paradigm choice for our experiment.

3.0 SIGNAL ANALYSIS TECHNIQUES

3.1 EXISTING NON-INVASIVE IMAGING TECHNOLOGY

Modern non-invasive functional human imaging studies typically involve one of four modalities; functional magnetic resonance imaging, (fMRI), functional positron emission tomography (PET), electroencephalography (EEG), and magnetoencephalography (MEG). Each of these methods are highly useful, as evidenced by their widespread use. A search for any of these modalities will yield thousands of published peer-reviewed studies. These four modalities represent two classes of imaging; detecting activity indirectly through changes in the blood-oxygen level, referred to as the BOLD effect, or observing the electric and magnetic fields generated by the movement of electrically charged ions at neuronal synapses.

3.1.1 BOLD effect-based imaging modalities

fMRI allows us to indirectly observe the activity of neurons by measuring the oxygen level in the surrounding blood vessels [24]. Neuronal action potential initiation and subsequent return to baseline activity requires a significant amount of energy. The energy expenditure of the sodium-potassium pumps acting to restore ion levels back to baseline is itself a significant energy drain, and necessitates an increase in the cell's

glucose metabolism, thus raising oxygen consumption [159]. This increase results in a increase in oxygen transported to that region of the brain [93]. Since oxyhemoglobin (Hb) is diamagnetic, there is a noticeable decrease in the magnetic susceptibility of the surrounding tissue in the presence of oxyhemoglobin relative to deoxyhemoglobin (dHb) [24]. As the MRI signal is highly dependent on the magnetic susceptibility of the tissue being imaged, this relative change acts as an endogenous contrast agent informing us of neuronal activity.

Functional PET also utilizes the BOLD effect, but to a different end. PET imaging involves the injection of a mildly radioactive agent (such as ^{15}O or ^{18}F) into the bloodstream and measuring the positron emission patterns during the decay process [125]. As neuronal activity increases, blood flow to the active region increases, and a greater number of positrons will be emitted from the active region, due to a higher cerebral regional blood flow (rCBF) [126]. Since the decay process involves the emission of two high-energy gamma rays at 180° angle from each other, the precise location from which the photons originated can be calculated based on the location and time at which the photons strike the detector surrounding the patient's head [69].

However, the BOLD effect has a major shortcomings when considering temporal dynamics. The change in blood flow caused by neuronal activity is believed to originate from the excitatory effect of glutamate, itself released during excitatory synaptic activity [6]. Glutamate, acting as both a neurotransmitter and a vasodilator, increases local blood flow as well initiates the release of other vasodilators. The combined effect of this cascade is a delayed increase in rCBF [39]. The delay period lasts approximately two seconds. After the flow increase, a very slight dip in contrast is sometimes visible. This dip is thought to be resultant from the sudden increase in blood volume [106]; more blood means more dHb, resulting in a slightly distorted

magnetic field [106]. However, the strongest contrast is not visible until at least five seconds with a full-width half-maximum (FWHM) response of approximately four seconds [33]. As such, the minimum temporal resolution available through fMRI imaging is two seconds, with significant temporal smoothing inherent in the signal, as evidenced by the relatively long FWHM value. Such resolution is not sufficient for the imaging of the dynamics within a neuronal network, as the entire network can often be recruited and subsequently deactivated on a millisecond timescale [64].

3.1.2 Ion motion-based imaging modalities

EEG and MEG utilize a different characteristic of action potentials to measure neural activity. Neuronal activity occurs through the movement of charged ions across synapses [8, chap. 12]. At sub-threshold levels, this movement is minimal. However, at threshold levels, an action potential is triggered, and ionic motion is increased significantly [83]. When action potentials occur simultaneously in approximately 50,000 uniformly-oriented cortical neurons (commonly found within a region of similar functional responsibility), the electric fields generated are strong enough to be detected by scalp EEG electrodes [132]. By contrasting the EEG signal at a given electrode with either a reference electrode or some baseline measure of brain activity, we can detect the deviations from resting neural activity. By examining this activity from a number of different sensor locations, we can estimate the source of the activity.

MEG functions on a similar principle. Rather than monitoring the electric fields elicited by action potentials, MEG detects the orthogonally oriented magnetic field [64]. The MEG device detects these fields using highly sensitive magnetometers and gradiometers [64] located on a voltage-free sensor array. This provides a benefit over EEG in that the electric fields can be significantly influenced by the meningeal

tissue surrounding the brain, significantly increasing the difficulty of EEG source localization, whereas the magnetic signals are unaffected by intervening organic tissue [107].

However, the fields being detected are incredibly tiny—on the order of microvolts for EEG [8] and femtotesla (10^{-15}) for MEG [64]. This results in the detected fields being highly susceptible to noise from a variety of sources, including heartbeats, eyelid motion, ambient light, and even the Earth's natural magnetic field [46, 87]. In the presence of such noise localizing the field source can be difficult. Additionally, the fields weaken exponentially as they travel through the neural tissue [161]. Sources on the neural cortex close to the sensors experience small attenuation, but deep sources can be significantly diminished, to the point where there is almost no observable field [119].

In the following sections, we will discuss a variety of noise reduction and source localization methods devised to help increase the signal to noise ratio. One major focus of this thesis is the development of the exSSS method, discussed in detail in Section 3.2.3. As this method applies only to MEG signal, the entirety of the following discussion will focus on MEG signal processing methods.

3.2 NOISE REDUCTION AND SOURCE LOCALIZATION OF MEG SIGNAL

Raw MEG data contains magnetic signatures of a number of sources, including many that can be classified as "noise." The origin of these sources can very widely. External sources include nearby medical devices, moving metal objects on the subject's clothing, or even ambient lighting. Noise inherent to the system can include faulty

or improperly calibrated magnetometer or gradiometer sensors. Noise is often biological as well, originating from eyeblinks, heartbeats, or swallowing. Depending on the application, many neural signals may be classified as noise, such as an unwanted signal from the motor cortex in a behavioral task that requires motion or the strong saccadic signal from eye motion.

A number of signal processing techniques have been devised to selectively remove these unwanted components while leaving the signal of interest unchanged. Many methods begin by representing the input signal as

$$\mathbf{b}(t) = \mathbf{H}(\mathbf{r}')\mathbf{J}(\mathbf{r}', t) + \mathbf{n}, \tag{3.1}$$

where $\mathbf{b}(t) = [b_1(t), \ldots, b_M(t)]^{\mathrm{T}}$ is a vector representing M–dimensional recorded signal, $\mathbf{J}(\mathbf{r}', t) = [\mathbf{J}_1(\mathbf{r}', t), \ldots, \mathbf{J}_N(\mathbf{r}', t)]^{\mathrm{T}}$ is a 6-D (position and orientation) vector containing the N sources, and $\mathbf{H}$ is a $6M \times N$ matrix of coefficients governing the linear mixing of the sources which convolve to produce $\mathbf{b}(t)$. These definitions will be used throughout the paper. As shown here, many models also include the term $\mathbf{n}(t)$ representing additive gaussian noise.

In the case of MEG signal analysis, the $M \times 1$ vector $\mathbf{b}$ is the signal recorded from M MEG sensors, the $N \times 1$ vector $\mathbf{J}$ is the signal originating from N sources, and $\mathbf{H}$ is the lead field matrix governing the relative strength of each source $\mathbf{J}_n$ at each of the m sensors [64, 156]. (The lead field matrix will be discussed in greater detail below.) By examining each signal component after processing, we can determine whether that component is likely to represent a noise process or a true neural signal. By discarding signals determined to be noise, we thus increase the signal-to-noise ratio in our data [156].

A number of noise removal techniques have been described in the literature. Principal component analysis (PCA) and independent component analysis (ICA)

decompose a dataset into "components," which can be either signal or noise [27, 103]. Through intelligent selection of the components they can then discard the noise components, thus retaining a cleaner signal [75]. Signal space projection (SSP) [156] utilizes previously recorded data samples and signal space transforms to project noise out of the dataset.

One major problem with such techniques is that they require the identification of "components" of the signal. Oftentimes sources of noise are difficult to quantify, and very commonly the signal of interest is weak and may be miscategorized as noise. The signal space separation (SSS) method [150] was developed as a method of removing noise, without relying on user classification of signal components. However, before a discussion on SSS, a thorough understanding of the concept if leadfields is required.

3.2.1 Leadfields

As noted above, raw MEG signal is acquired in the form of M time courses, where M is the number of magnetic field sensors. We refer to data in this form as existing in "MEG space". The neural sources themselves exist in 6-D space (three degrees in Cartesian coordinates, and three degrees specifying orientation), or the "source space". To convert from one space to another requires knowledge of the relationship between the sources and the magnetic field sensors. This relationship is mediated by the 6-D location of the sensor relative to the source. Biot-Savart's law of magneto-static fields describes how magnetic field strength degrades with distance, as well as explains how the orientation of the sensor will affect the recorded field strength.

Mathematically, the mapping from a given signal $b_m(t)$ from some sensor m at time t in MEG space (6-D space) to a given time-varying source or set of sources $\mathbf{J}(\mathbf{r}', t)$ at $\mathbf{r}'$ in source space (time-varying magnetic signal amplitudes) is expressed

as

$$b_m(t) = \int_\Omega \mathbf{H}_m(\mathbf{r}')\mathbf{J}(\mathbf{r}', t)d\Omega, \tag{3.2}$$

where Ω represents the entire source space, defining both location and orientation. $\mathbf{H}_m$, which represents the mapping variable between the two coordinate systems, is referred to as the leadfield matrix for sensor m [64, 128, 143]. As Equation 3.2 contains both an MEG space representation of our signal ($\mathbf{b}_m(t)$) and a source space representation ($\mathbf{J}(\mathbf{r}', t)$), this equation allows us to map from one space to the other, using the leadfield matrix. This matrix, specific to both a given source location $\mathbf{r}'$ and sensor m, describes the sensitivity of the given sensor to a source at location $\mathbf{r}'$. One further relationship we can discern is the second order relationships between the leadfields themselves, described as the Gram matrix:

$$\mathbf{G} = \int_\Omega \mathbf{H}'(\mathbf{r}')\mathbf{J}(\mathbf{r}')\mathbf{H}(\mathbf{r}')^{\mathrm{T}}d\Omega \tag{3.3}$$

Without loss of generality with respect to realistic source configurations, we can make $\mathbf{G}$ independent of our source configuration by setting $\mathbf{J}(\mathbf{r}') = 1$ for all source locations, resulting in

$$\mathbf{G} = \int_\Omega \mathbf{H}'(\mathbf{r}')\mathbf{H}(\mathbf{r}')^{\mathrm{T}}d\Omega. \tag{3.3a}$$

This relationship describes how highly the view of the magnetic field correlates between sensors, the importance of which will be explained below.

Intuitively, since the source space includes all regions which may potentially contain a magnetic source, the leadfield matrix should remain constant for a given MEG sensor configuration. Since the relationship between MEG space and source space is dictated by the laws of physics, the mapping between the two should not change.

It has been demonstrated, though, that careful manipulation of the leadfield can decrease errors in source localization. In order to understand this phenomenon, we must closely look at the definition of "source space." Generally speaking, the source space is any location at which a magnetic source could potentially be located. However, by restricting this definition to only include biologically tractable regions, such as only the source space located within the sensor array, we can obtain a more accurate leadfield matrix (see Figure 5). Through applying restrictions on the allowable source locations we can minimize the error in the dimensional mapping, and consequently also in the subsequent source localization. Manipulations of this type were first described by Taulu and Kajola [150] in their derivation of the Signal Space Separation (SSS) method.

3.2.2 Signal Space Separation

It was recently demonstrated that through the use of a coordinate transformation, MEG signal can be separated into two components, those originating from inside the physical sensor array and those external to it [150]. This separation can be achieved by expressing the magnetic field as the gradient of a harmonic scalar potential—a mathematical abstraction defined below—and utilizing the observation that the sensor array represents a source-free sphere in spherical space.

The calculation of magnetic field strength based on a source charge relies on two of Maxwell's equations:

$$\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t \tag{3.4}$$

$$\nabla \times \mathbf{B} = \mu_0(\mathbf{J} + \epsilon_0 \partial \mathbf{E}/\partial t) \tag{3.5}$$

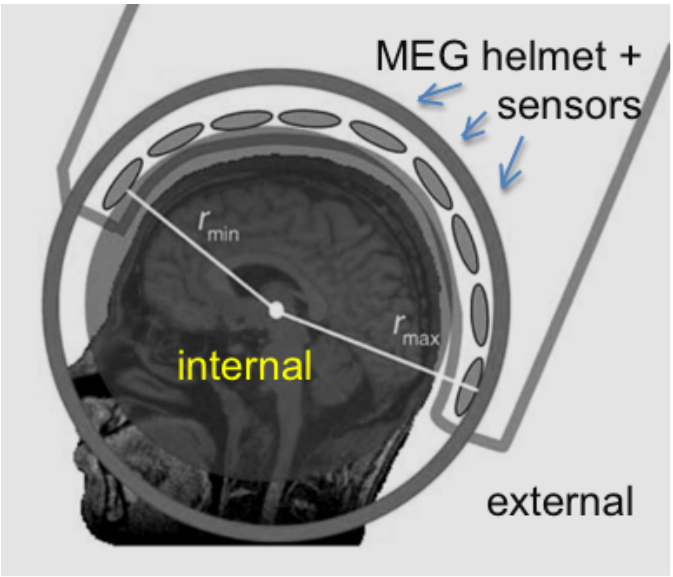


Figure 5: Schematic of internal and external regions as defined in the SSS method. Note that the internal region includes both the subject's head as well as the empty space between the head and the sensor array, and that the external region includes the neck and lower body of the subject. These potential noise sources cannot be accounted for in the SSS method and will need to be addressed using alternative noise removal mechanisms.

Time independence is demonstrated for equation (3.5) by considering the time-dependent second term, $\mu_0\epsilon_0\partial\mathbf{E}/\partial t$. Given the low frequencies typical of MEG signals, it can be shown that the time-dependent component of the signal is approximately three orders of magnitude smaller than the time-independent component, thus justifying their neglect [64]. The proof justifying the neglect of equation (3.4) is similar [64].We can thus represent Maxwell's equations using the quasi-static approximation [64],

$$\nabla \times \mathbf{H} = \mathbf{J} \tag{3.6}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \tag{3.7}$$

$$\nabla \cdot \mathbf{B} = 0. \tag{3.8}$$

We begin with the assumption that the sphere defining the sensor array is source-free, or $\mathbf{J} = 0$ at radius $r = R$, where R is the distance from the origin to the sensor array. As such, we can rewrite Equation 3.6 as $\nabla \times \mathbf{H} = 0$. Using the identity $\nabla \times \nabla\Psi = 0$, we define $\mathbf{H} = -\nabla\Psi$,where Ψ is termed the scalar potential. Substituting this into Equation (3.7), we arrive at

$$\mathbf{B} = -\mu_0 \nabla\Psi. \tag{3.9}$$

This states that the magnetic field at the sensors is a gradient of the scalar potential Ψ [146, 150]. Note that the concept of scalar potential is a mathematical abstraction constructed based on the assumption of a source-free sphere inherent to the MEG setup. Note also that by substituting the harmonic potential (Equation 3.9) into Maxwell's equation for the divergence of a magnetic field (Equation 3.8) we obtain the Laplacian $\nabla^2\Psi = 0$. We can express the scalar potential in the spherical harmonics domain, $\Psi(\varphi, \theta, r) = \Phi(\varphi)\Theta(\theta)R(r)$. The solution to Laplace's equation can then be obtained by separating the variables

$$\nabla^2\Psi = \frac{1}{r^2 \sin\theta}\left[\sin\theta\frac{\partial}{\partial r}\left(r^2\frac{\partial\Psi}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Psi}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial^2\Psi}{\partial\varphi^2}\right] \tag{3.10}$$

and solving the resultant harmonic equation. This has the known solution

$$\Psi(\mathbf{r}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\alpha_{lm}\frac{Y_{lm}(\theta,\varphi)}{r^{l+1}} + \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\beta_{lm}r^l Y_{lm}(\theta,\varphi), \tag{3.11}$$

where

$$Y_{lm}(\theta,\varphi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{l+m)!}}P_{lm}(\cos\theta)e^{im\varphi}$$

represents the normalized spherical harmonic function in which $P_{lm}(\cos\theta)$ is the associated Legendre function, α_{lm} and β_{lm} are the multipole moments of the internal and external current sources respectively, and r the radius [4, 78].

This result is very significant; by transforming the data to spherical coordinates and utilizing the scalar potential (Equation 3.9), we have separated the magnetic field into two components α_{lm} and β_{lm}—corresponding to internal sources and external sources, respectively—using nothing more than Maxwell's equations. Since by nature the signal of interest originates exclusively from within the brain, we can simply drop the second term from Equation 3.11, thus removing an entire source of noise from our dataset.

By substituting Equation (3.11) into our definition for scalar potential (Equation 3.9) and expressing the result using the modified vector spherical harmonic functions $\boldsymbol{\nu}_{lm}(\theta,\varphi)$ and $\boldsymbol{\omega}_{lm}(\theta,\varphi)$ [150] we obtain

$$\begin{aligned}\mathbf{B}(\mathbf{r}) &= -\mu_0\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\alpha_{lm}\frac{\boldsymbol{\nu}_{lm}(\theta,\varphi)}{r^{l+1}} - \mu_0\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\beta_{lm}r^l\boldsymbol{\omega}_{lm}(\theta,\varphi)\\ &\equiv \mathbf{B}_\alpha(\mathbf{r}) + \mathbf{B}_\beta(\mathbf{r}).\end{aligned} \tag{3.12}$$

In this way we have separate the recorded magnetic field into two components. Considering only the internal sources, we can now express the multipole moments α_{lm} using a lead field-like representation [4, 150]

$$\alpha_{lm} = \int_{v'}\boldsymbol{\lambda}^{\alpha}_{lm}(\mathbf{r}')\cdot\mathbf{J}_{\text{in}}(\mathbf{r}')dv', \tag{3.13}$$

where $\mathbf{J}_{\text{in}}$ represents the sources within the head (i.e., internal to the sensor array) and $\boldsymbol{\lambda}_{lm}^{\alpha}$ is a lead field-like vector directly related to vector spherical harmonic function $\mathbf{X}_{lm}(\theta, \varphi)$, defined as

$$\begin{aligned} \boldsymbol{\lambda}_{lm}^{\alpha} &= \frac{i}{2l+1}\sqrt{\frac{l}{l+1}} r^{l} \mathbf{X}_{lm}^{*}(\theta, \varphi) \\ \mathbf{X}_{lm}(\theta, \varphi) &= \frac{-1}{l\sqrt{l+1}} \left[\frac{m Y_{lm}(\theta, \varphi)}{\sin\theta} \mathbf{e}_{\theta} + i \frac{\partial Y_{lm}(\theta, \varphi)}{\partial \theta} \mathbf{e}_{\varphi} \right]. \end{aligned} \tag{3.14}$$

Similar equations can be constructed for $\boldsymbol{\lambda}_{lm}^{\beta}(\mathbf{r})$ [150]. It is important to recognize that the lead fields $\boldsymbol{\lambda}_{lm}^{\alpha}(\mathbf{r})$ and $\boldsymbol{\lambda}_{lm}^{\beta}(\mathbf{r})$ now represent separate datasets corresponding to the inner and outer signals, respectively. Each can be examined for activity that would be expected in that dataset. Our primary focus will be on the internal dataset representing neural activity.

On a technical note, we have been summating the associated Legendre function with l ranging from 0 to ∞, and m from $-l$ to l. In practice, this is computationally impossible, and the optimal value of l in practice has been the subject of research [150, 151]. Additionally, since the Legendre functions representing the inner and outer components are different, the optimal value of l will be different for each as well. We will refer to the optimal values for the inner and outer components of the signal as L_{in} and L_{out}.

Having demonstrated that any dataset containing a spherical source-free shell can be separated into inner and outer components, we can reformulate our problem specifically for application to MEG datasets. Given a set $\mathbf{x}$ of M sensors recording a set $\mathbf{s}$ of N signals, we can state

$$\begin{aligned} \mathbf{x} &= \sum_{l=1}^{L_{\text{in}}} \sum_{m=-l}^{l} \alpha_{lm} \mathbf{s}_{lm}^{\text{in}} + \sum_{l=1}^{L_{\text{out}}} \sum_{m=-l}^{l} \beta_{lm} \mathbf{s}_{lm}^{\text{out}} \\ &= [\mathbf{S}_{\text{in}} \mathbf{S}_{\text{out}}] \begin{bmatrix} \boldsymbol{\alpha}_{\text{in}} \\ \boldsymbol{\beta}_{\text{out}} \end{bmatrix}, \end{aligned}$$

where

$$S_{\text{in}} = [\mathbf{a}_{1,-1}, \ldots, \mathbf{a}_{L_{\text{in}},L_{\text{in}}}] \qquad \boldsymbol{\alpha}_{\text{in}} = [\alpha_{1,-1}, \ldots, \alpha_{L_{\text{in}},L_{\text{in}}}]^{\text{T}}$$
$$S_{\text{out}} = [\mathbf{b}_{1,-1}, \ldots, \mathbf{b}_{L_{\text{out}},L_{\text{out}}}] \qquad \boldsymbol{\beta}_{\text{out}} = [\beta_{1,-1}, \ldots, \beta L_{\text{out}}, L_{\text{out}}]^{\text{T}}.$$

That is, the recorded magnetic field can be decomposed into separate subspaces, each of which contains a distinct set of data from geometrically separated regions, thus achieving the goal of separation of external (presumably noise) signals from internal biological signals.

3.2.3 Expanded SSS (exSSS)

Ozkurt et al. [113] extended this work to enable not only separation of external and internal sources but also selected internal regions of interest from the rest of the background neural signal. This was accomplished by manipulating the leadfield-like $\boldsymbol{\alpha}$ coefficients obtained from the SSS algorithm (Equation 3.13) via a beamspace transform designed to amplify deep signals and attenuate superficial signal. In the following sections, I will introduce the concept of the beamspace transform and how it is used to manipulate leadfields in the exSSS algorithm to enhance deep neural signal within an MEG dataset.

3.2.3.1 Beamspace transformation Beamspace transforms are a class of spatial filtering algorithms, often used to localize or enhance spatial information within a given signal. Often, beamspace transforms function to reduce the dimensionality of a given dataset as well, thus simplifying further analysis as well as reducing data redundancy [128]. A beamspace transform, at its simplest, is expressed as the linear transform $\hat{\mathbf{b}} = \mathbf{T}^{\text{T}}\mathbf{b}$, where $\mathbf{T}$ is an $M \times L$ transformation matrix, M being

the number of sensors and $L < M$. The columns of $\mathbf{T}$ are assumed orthonormal without loss of generality. There are a number of ways to design $\mathbf{T}$ [18, 95, 170], but one method useful for MEG analysis maximizes the discarding of redundant data while minimizing data loss [128]. Mathematically, this is equivalent to maximizing $\mathbf{T}$ while minimizing the squared error of the transformation, $e^2_{\mathbf{T}}$. By varying the L, we can vary the aggressiveness our dimensionality reduction. The squared error can be stated as

$$e^2_{\mathbf{T}}(\mathbf{r}') = \|(\mathbf{I} - \mathbf{T}\mathbf{T}^{\mathrm{T}})\mathbf{b}(t)\|_2^2$$

where $\mathbf{b}(t) = [b_1(t), \ldots, b_M(t)]$ is the magnetic field recorded at all M sensors. This states that the error is the sum of the residual signal remaining after applying the transform to $\mathbf{b}$. Through Equation 3.2 we obtain $e^2_{\mathbf{T}}(\theta) = \|(\mathbf{I} - \mathbf{T}\mathbf{T}^{\mathrm{T}})\mathbf{H}(\mathbf{r}')\mathbf{j}(\mathbf{r}', t)\|^2$. By assuming that the dipole moments are constant through time, we may write $\mathbf{j}(\mathbf{r}, t) = \mathbf{j}(\mathbf{r})\boldsymbol{\gamma}(t)$, where $\boldsymbol{\gamma}$ represents the time-varying amplitude of the constant source. Considering the case of $\gamma(t) = 1$ for all t, we obtain

$$e^2_{\mathbf{T}}(\mathbf{r}') = \|(\mathbf{I} - \mathbf{T}\mathbf{T}^{\mathrm{T}})\mathbf{H}(\mathbf{r}')\mathbf{j}(\mathbf{r}')\|^2, \tag{3.15}$$

where $\mathbf{I}$ is the identity matrix. Given that we have no information about the source matrix $\mathbf{j}$, we assume $\mathbf{j}(\mathbf{r}') = 1$ for all $\mathbf{r}'$. If we subsequently minimize $\mathbf{T}$,

$$\min_{\mathrm{T}} \left\{ \int_{\Omega} e^2_{\mathbf{T}}(\mathbf{r}') d\mathbf{r}' \right\}, \tag{3.16}$$

we are designing $\mathbf{T}$ to be as small as possible while minimizing the data loss caused by the transformation. In matrix notation, this can be written

$$\max_{\mathrm{T}} \mathrm{tr}(\mathbf{T}^{\mathrm{T}}\mathbf{G}\mathbf{T}), \text{ subject to } \mathbf{T}^{\mathrm{T}}\mathbf{T} = \mathbf{I} \tag{3.17}$$

where $\mathbf{G}$ is the Gram matrix (Equation 3.3)

$$\mathbf{G} = \int_{\Omega} \mathbf{H}(\mathbf{r}')\mathbf{J}(\mathbf{r}')\mathbf{H}^{\mathrm{T}}(\mathbf{r}')d\mathbf{r}'. \tag{3.18}$$

As demonstrated earlier, if the source matrix is unknown, we assume $\mathbf{J} = \mathbf{I}$ (i.e., $\mathbf{J}(\mathbf{r}') = 1$ for all $\mathbf{r}'$), and this is equivalent to the Gram matrix (Equation 3.3).

By defining $\mathbf{T}$ in this manner, we selectively keep the strongest signals from our source space. In this case, we chose to define $\mathbf{T}$ as a non-square matrix for this purpose. However, by manipulating $\mathbf{T}$ differently, we can amplify the signals emanating from specific regions of the source space, as defined by the Gram matrix.

3.2.3.2 exSSS Recall that the α_{lm} component of the SSS output can be represented using lead field-like manner, as depicted in Equation 3.13 (in which $\mathbf{J}_{\mathrm{in}}(\mathbf{r}')$ represented the sources and $\boldsymbol{\lambda}^{\alpha}_{lm}(\mathbf{r}')$ the leadfield). Let us denote the (l, m)-indexed SSS coefficient (l, m integers, $-1 \leq l \leq L_{\mathrm{in}}$, $-l \leq m \leq l$) as $a = 1, \ldots, p_{\mathrm{in}}$ and the (L, M)-indexed SSS coefficient (L, M integers, $-1 \leq L \leq L_{\mathrm{in}}$, $-L \leq M \leq L$) as $b = 1, \ldots, p_{\mathrm{in}}$. We can now construct the Gram matrix

$$\begin{aligned}
\mathbf{G} &= \int_{\Omega} \boldsymbol{\lambda}_{lm}(\mathbf{r}') \cdot \boldsymbol{\lambda}_{LM}(\mathbf{r}')d\Omega \\
&= \int_{\Omega} \boldsymbol{\lambda}_{a}(\mathbf{r}') \cdot \boldsymbol{\lambda}_{b}(\mathbf{r}')d\Omega \\
&= \int_{\Omega} \boldsymbol{\Lambda}(\mathbf{r}')\boldsymbol{\Lambda}(\mathbf{r}')^{\mathrm{H}}d\Omega,
\end{aligned}$$

where $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_{1,-1}, \ldots, \boldsymbol{\lambda}_{lm}]$. Our goal is to amplify signal from a given region of the signal space while attenuating the rest. This can formulated mathematically as

$$\max_{\mathrm{T}} \frac{\mathrm{tr}(\mathbf{T}^{\mathrm{T}}\mathbf{G}_d\mathbf{T})/v_d}{\mathrm{tr}(\mathbf{T}^{\mathrm{T}}\mathbf{G}_s\mathbf{T})/v_s}, \text{ subject to } \mathbf{T}^{\mathrm{T}}\mathbf{T} = \mathbf{I}, \tag{3.19}$$

where $\mathbf{G}_d$ and $\mathbf{G}_s$ represent the Gram matrices of the deep and superficial regions (regions where $\mathbf{r}' < \mathbf{r}$ and $\mathbf{r}' > \mathbf{r}$), and v_d and v_s correspond to normalizing constants for the deep and superficial regions, respectively. The solution to $\mathbf{T}$ can be shown to be the matrix whose columns are the eigenvectors of

$$\mathbf{G}_f = \frac{v_s}{v_d}\mathbf{G}_s^{-1/2}\mathbf{G}_d\mathbf{G}_s^{-1/2}. \tag{3.20}$$

corresponding to its largest eigenvalues [61]. The ath row and bth column of Gram matrix $\mathbf{G}_d$ (corresponding to the deep component) can be found as

$$(\mathbf{G}_d)_{ab} = \int_\theta \int_\varphi \int_{r=0}^{\hat{r}} \left(\frac{i}{2l+1}\sqrt{\frac{l}{l+1}} r^l \mathbf{x}_{lm}^*(\theta,\varphi) \right) \cdot \left(\frac{-i}{2L+1}\sqrt{\frac{L}{L+1}} r^L \mathbf{x}_{LM}^*(\theta,\varphi) \right) r^2 \sin\theta dr d\theta d\varphi \tag{3.21}$$

$$= \delta_{lL}\delta_{mM} \frac{1}{(2l+1)^2} \frac{l}{l+1} \frac{\hat{r}^{2l+3}}{2l+3}, \tag{3.22}$$

where $\hat{r}$ represents the radius of the enhanced deep region. (The origin of the enhanced region is always the geometric center of the MEG sensor array.) The progression from Equation 3.21 to 3.22 stems from the orthonormality of vector spherical harmonics. $\mathbf{G}_s$ is found in a similar manner:

$$(\mathbf{G}_s)_{ab} = \int_\theta \int_\varphi \int_{\hat{r}}^{R} \left(\frac{i}{2l+1}\sqrt{\frac{l}{l+1}} r^l \mathbf{x}_{lm} * (\theta,\varphi) \right) \tag{3.23}$$

$$\cdot \left(\frac{-i}{2L+1}\sqrt{\frac{L}{L+1}} r^l \mathbf{x}_{LM} * (\theta,\varphi) \right) r^2 \sin\theta dr d\theta d\varphi \tag{3.24}$$

$$= \delta_{lL}\delta_{mM} \frac{1}{(2l+1)^2} \frac{l}{l+1} \frac{R^{2l+3} - \hat{r}^{2l+3}}{2l+3}. \tag{3.25}$$

Note the difference between the integrands in Equations 3.21 and 3.22 . $\mathbf{G}_d$ is calculated by integrating from $r = 0$—i.e., the center of the sphere defined by the

MEG sensors—to $r = \hat{r}$, the radius of the deep region, whereas $\mathbf{G}_s$ is calculated by integrating from $r = \hat{r}$ to $r = R$, where R is the radius of the MEG sphere, $\|\mathbf{r}\|$. This is the differentiating factor between the two equations.

We can now solve Equation 3.20 through substitution, obtaining

$$(\mathbf{G}_f)_{ab} = \frac{v_s}{v_d} \frac{\hat{r}^{2l+3}}{R^{2l+3} - \hat{r}^{2l+3}} \delta_{lL} \delta_{mM}. \tag{3.26}$$

Since $\mathbf{G}_f$ is diagonal, all its eigenvectors constitute an identity matrix. This suggests a transformation matrix $\mathbf{T}$ that selects the $\boldsymbol{\alpha}$ coefficients corresponding to $r < \hat{r}$ and eliminating the other coefficients. In this manner, we can obtain a filtered dataset in which deep sources are significantly amplified while superficial activity is reduced [113, 114].

It is worth noting that a similar outcome can be produced in a variety of ways. For example, one could construct a beamspace transform which selectively amplifies deep source activity purely based on the 3-D location of each source space element within a boundary element model (BEM). In this way, the deeper elements of the leadfield matrix can be amplified without needing to obtain the $\boldsymbol{\alpha}$ coefficients. While this is possible, the main benefit of the exSSS method is that it approaches the problem of deep signal amplification from an analytical perspective, requiring only the knowledge of Maxwell's laws (see Section 3.2.2). Other methods, including the example given here, require knowledge of the spatial geometry of the source space, and approach the problem from a computational standpoint. In this way, the exSSS method makes as few assumptions as possible, resulting in a broadly applicable algorithm.

Using the exSSS method, we are able to amplify the deep neural activity to the extent that it is visible in the dataset. However, before attempting source localization, we still must define an accurate model of the head onto which we will attempt to localize sources. One common model is a simple spherical model. The simplicity

of this model simplifies the calculations required compute the inverse solution, and also obviates the need for a BEM model of the head. However, given that we were collecting MRI scans of our subjects, and given that modern computers are more than adequate for solving the inverse solution even over complex BEMs, we utilized a realistic head model for our calculations. The following sections describes how to calculate the forward and inverse problem using a realistic model of the head.

3.2.4 MEG Forward Solution using a Piecewise Non-homogenous Conductivity Model

3.2.4.1 Forward Problem Using the quasi-static representation of Maxwell's equations (see Section 3.2.2), the determination of the magnetic field $\mathbf{B}$ at an arbitrary point $\mathbf{r}$ is accomplished via the Biot-Savart equation:

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_o}{4\pi} \int \frac{\mathbf{J}(\mathbf{r}') \times \mathbf{R}'}{R^3} dv', \tag{3.27}$$

where $\mathbf{R} = \mathbf{r} - \mathbf{r}'$ and $R = |\mathbf{R}|$. In this context,

$$\mathbf{J}(\mathbf{r}') = \mathbf{J}^p(\mathbf{r}') + \sigma(\mathbf{r}')\nabla V(\mathbf{r}'), \tag{3.28}$$

where $\mathbf{J}^p(\mathbf{r}')$ represents the primary magnetic sources and $\sigma(\mathbf{r})V(\mathbf{r}')$ represents the volume currents in the surrounding conductive tissue induced by $\mathbf{J}$. We define $\nabla = \mathbf{e}_x\partial/\partial x + \mathbf{e}_y\partial/\partial y + \mathbf{e}_z\partial/\partial z$ and $\nabla' = \mathbf{e}_x\partial/\partial x' + \mathbf{e}_y\partial/\partial y' + \mathbf{e}_z\partial/\partial z'$. To model the forward solution, we would like a closed-form algorithm which can calculate $\mathbf{B}(\mathbf{r})$ given $\mathbf{J}^p(\mathbf{r}')$ for any $\mathbf{r}$ or $\mathbf{r}'$.

Using the identities $\mathbf{R}/R^3 = -\nabla\,(1/\mathbf{R}) = \nabla'\,(1/\mathbf{R})$ and $\mathbf{J} \times \nabla'(1/\mathbf{R}) = [(\nabla' \times \mathbf{J})/\mathbf{R}] - [\nabla' \times (\mathbf{J}/\mathbf{R})]$, Equation 3.27 can be rewritten as

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_o}{4\pi} \left[\int \frac{\nabla' \times \mathbf{J}}{\mathbf{R}} dv \; - \int \left(\nabla' \times \frac{\mathbf{J}}{\mathbf{R}} \right) dv' \right]. \tag{3.29}$$

By converting the second term Equation 3.29 to a surface integral and knowing that $\mathbf{J} \to 0$ as $\mathbf{R} \to \infty$, the term can be eliminated. Substituting Equation 3.28 for $\mathbf{J}$ and utilizing the identity $\nabla' \times \sigma \nabla V = \nabla' \sigma \times \nabla V = -\nabla' \times V \nabla \sigma$, we arrive at

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_o}{4\pi} \int \left[(\mathbf{J}^p + V \nabla \sigma) \times \frac{\mathbf{R}}{R^3} \right] dv'. \tag{3.30}$$

Note that in this representation we no longer need to calculate ∇V. This represents an analytical solution for $\mathbf{B}$ which relies only on $\mathbf{J}^p$ and V, and assumes knowledge of tissue conductivity values $\sigma(\mathbf{r})$. Making use of the quasi-static approximation of Maxwell's equation we can equate $\nabla \cdot \mathbf{J} = 0$ and relate $\mathbf{J}$ and V as follows:

$$\begin{aligned} \nabla \cdot (\nabla \times \mathbf{B}) = 0 = \nabla \cdot \mu \mathbf{J} = \nabla \cdot (\mathbf{J}^p - \sigma \nabla V) \\ \Rightarrow \nabla \cdot \mathbf{J}^p = \nabla \cdot (\sigma \nabla V) \end{aligned} \tag{3.31}$$

We cannot practically assume that $\sigma(\mathbf{r})$ is known for all $\mathbf{r}$. However, if we segment the tissue using a BEM and assume piecewise homogeneity of σ within each element, we are able to approximate the true conductivity distribution with fairly high accuracy. This requires the adaptation of Equation 3.30 to handle a range of discrete constant values for σ. By using the approximation $\sigma(\mathbf{r}) = \sigma_i$ for $i \in m, m = 1 \cdots M$ where M is the number of regions on the grid, Equation 3.30 becomes

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_o}{4\pi} \int_G \left(\mathbf{J}^p \times \frac{\mathbf{R}}{R^3} \right) dv' - \frac{\mu_o}{4\pi} \sum_{i=1}^{m} \sigma_i \int_{G_i} \left(\nabla' V \times \frac{\mathbf{R}}{R^3} \right) dv',$$

where G_i = different regions of conductivity and σ_i = conductivity of the G_i region (see Figure 6). Denoting the first term as $\mathbf{B}_o$ and utilizing the identities mentioned above, we arrive at

$$\begin{aligned} \mathbf{B}(\mathbf{r}) &= \mathbf{B}_0 - \frac{\mu_o}{4\pi} \sum_{i=1}^{m} \sigma_i \int_{G_i} \left(\nabla' V \times \frac{\mathbf{R}}{R^3} \right) dv' \\ &= \mathbf{B}_0 + \frac{\mu_o}{4\pi} \sum_{i=1}^{m} \sigma_i \int_{\partial \mathbf{G}_i} V \frac{\mathbf{R}}{R^3} \times d\mathbf{S}, \end{aligned}$$

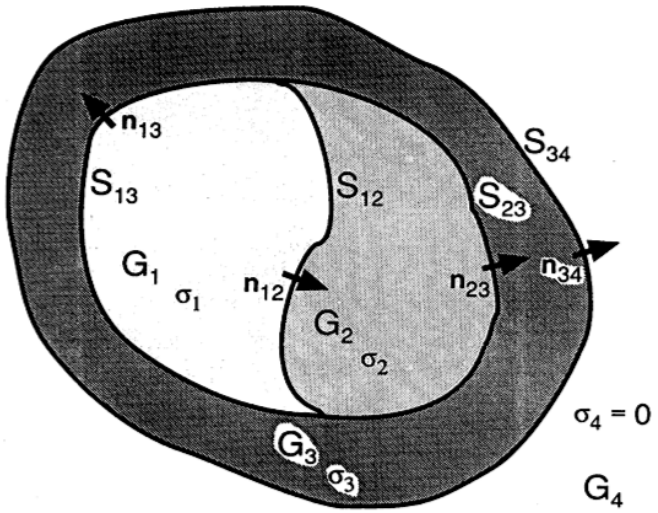


Figure 6: A schematic of the regions involved in a piecewise homogenous model of the brain. G_i = different regions of conductivity, ∂G_i = boundaries of G_i, σ_i = conductivity of the G_i region, S_{ij} = surface between regions G_i and G_j, $\mathbf{n}_{ij}(\mathbf{r}')$ = unit vector normal to surface S_{ij}. Image taken from Figure 18 in [64].

where ∂G_i = boundaries of G_i, S_{ij} = surface between regions G_i and G_j, $\mathbf{n}_{ij}(\mathbf{r}')$ = unit vector normal to surface S_{ij}. If instead of summing over boundaries of G_i we instead sum over the interfacing surfaces S_{ij}, we arrive at

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_0 + \frac{\mu_o}{4\pi}\sum_{ij}^{m}{}' (\sigma_i - \sigma_j) \int_{\mathbf{S}_{ij}} V \frac{\mathbf{R}}{R^3} \times d\mathbf{S}_{ij}, \tag{3.32}$$

which is a closed-form integral equation for $\mathbf{B}(\mathbf{r})$ relying only on $\mathbf{J}$ and V. At this point we could use Equation 3.31 to determine V from $\mathbf{J}$, but due to the difficulty in solving the harmonic equation we instead attempt to devise an integral equation for V. One method to simplify the problem [54] is to utilize Green's second identity,

$$\int_G (\Phi\nabla\cdot\nabla\Psi - \Psi\nabla\cdot\nabla\Phi)dv = \int_{\mathbf{S}} (\Phi\nabla\Psi - \Psi\nabla\Phi)\cdot d\mathbf{S}. \tag{3.33}$$

Defining $\Psi = V$ and $\Phi = 1/R$, we obtain

$$\sum_{i=1}^{m} \sigma_i \int_{G_i} \left(\frac{1}{R} \nabla \cdot \nabla V - V \nabla \cdot \nabla \frac{1}{R} \right) dv'$$
$$= \sum_{ij} \int_{S_{ij}} \left[\sigma_i \left(\frac{1}{R} \nabla_i V - V \nabla_i \frac{1}{R} \right) - \sigma_j \left(\frac{1}{R} \nabla_j V - V \nabla_j \frac{1}{R} \right) \right] \cdot d\mathbf{S}'_{ij}. \quad (3.34)$$

We can substitute the identity [165] $V\nabla^2(1/R) = -4\pi V$ in the left side of equation (3.34) and recall that $\sigma_i \nabla V = \mathbf{J}^p$, resulting in

$$\sum_{i=1}^{m} \int_{G_i} \left(\frac{1}{R} \nabla \cdot \mathbf{J}^p \right) dv + 4\pi\sigma_i V. \quad (3.35)$$

Through an algebraic manipulation of Gauss's theorem, we can manipulate the first term of Equation 3.34 as follows. Equating

$$\int \nabla \cdot (\mathbf{J}^p/R) dv = \int (\mathbf{J}^p/R) \cdot d\mathbf{S} = \int \left(\mathbf{J}^p \cdot \nabla \frac{1}{R} + \frac{1}{R} \nabla \cdot \mathbf{J} \right) dv,$$

we arrive at

$$\int \left(\frac{1}{R} \nabla \cdot \mathbf{J}^p \right) dv = \int (\mathbf{J}^p/R) \cdot d\mathbf{S} - \int \mathbf{J}^p \cdot \nabla \frac{1}{R} dv \quad (3.36)$$
$$= -\int \left(\mathbf{J}^p \cdot \nabla \frac{1}{R} \right) dv, \quad (3.37)$$

the second equality arising from the case where $\mathbf{J}^p$ vanishes on S, the boundary containing the sources. Equation 3.34 is now in the form of

$$4\pi\sigma V = \int \left(\mathbf{J}^p \cdot \nabla \frac{1}{R} \right) dv + \sum_{ij} (\sigma_i - \sigma_j) \int_{\mathbf{S}_{ij}} V \nabla' \frac{1}{R} \cdot d\mathbf{S}'_{ij}. \quad (3.38)$$

This final equation states that the voltage at any given point can be determined by summing the determining the strength of the voltage source at that radius and adding the summed effect of that voltage source on all elements of all surfaces within the head. It should be noted that this is an implicit equation for V which is best solved iteratively through the Gauss-Seidel method or other similar iterative algorithms [65].

3.2.5 Inverse Methods

The forward problem is crucial for our ability to calculate leadfields, as described above in Section 3.2.1. To localize a source based on measurements at the sensor array, though, we must attempt to determine the specific source geometry that would produce the measured field pattern, also known as solving the inverse problem [64, 136]. Since physical limitations only allow the sensors to see a subset of the possible sources, this problem is ill-posed and must be solved by applying approximations and assumptions. Still, there are many methods that attempt to solve this problem [28, 66, 103, 120, 128, 136, 143, 161].

Before any discussion of individual methods, it must be understood that the usefulness of any analysis technique is in part determined by its ability to resolve neural sources both accurately and precisely. The mean displacement error (average distance between the actual and reconstructed source; a measure of accuracy) for distributed source reconstruction has been reported as 2-3 mm for superficial sources and 7 mm for deep sources [94, 97]. Inasmuch as we are examining deep sources, we may assume that a given source identified by the above methods reflects an underlying source current within 7 mm of the reported location.

This margin of error is further compounded by the fact that distributed source localization reports sources as being present on a particular source manifold, which is only a subset of the actual 3-D space in which the sources exist [64]. Because of this, sources located outside of the manifold must be projected onto the manifold prior to localization, and thus prior to the application of the 7 mm error. As such, sources may be some distance from the reported region. This aspect of source localization will be addressed further in the discussion section.

3.2.5.1 Minimum Norm Estimation (MNE)

One of the early methods developed to solve the inverse problem was the minimum norm estimation method, MNE [136]. Let $\mathcal{F}$ be a function space containing all possible current source locations, and let Γ be a known set of points containing all square-integrable current distributions to which we will confine the sources $\mathbf{J}^p$, with the p superscript denoting primary current (see Section 3.2.4.1). We define the inner product $\langle J_1^p, J_2^p \rangle = \int_\Gamma \mathbf{J}_1^p(\mathbf{r})\mathbf{J}_2^p(\mathbf{r})d\Gamma$. Note that if $J_1^p = J_2^p$, $\|J^p\|^2 = \int_\Gamma |\mathbf{J}^p(\mathbf{r})|^2 d\Gamma$.

Recalling the definition of the leadfield (Equation 3.2), we can rewrite the leadfield equation using inner product notation,

$$b_i = \langle H_i, \mathbf{J}^p \rangle. \tag{3.39}$$

Note that H is defined for each individual sensor i; the i sensors only inform us about currents within the subspace of observable current source locations $\mathcal{F}' \in \mathcal{F}$. MNE attempts to find an estimate $\mathbf{J}^*$ of $\mathbf{J}^p$, confined to $\mathcal{F}'$ such that $\mathbf{J}^*$ is the current distribution with the overall smallest amplitude capable of explaining the measured signal $\mathbf{b}$. This is achieved by looking for a set of weights $\boldsymbol{\omega}$ such that

$$\mathbf{J}^* = \sum_i \omega_i H_i \tag{3.40}$$

$$b_i = \langle \boldsymbol{\omega}\mathbf{H}, \mathbf{J}^* \rangle = \langle \mathbf{H}, \mathbf{J}^p \rangle. \tag{3.41}$$

This minimization is achieved by inserting Equation 3.40 into Equation 3.39 [64, 136]:

$$\begin{aligned} b_i &= \langle H_i, \mathbf{J}_i^p \rangle \\ \tilde{b}_i &= \langle H_i, \mathbf{J}_i^* \rangle = \sum_k \langle H_i, \omega_k H_k \rangle = \sum_k \omega_k \langle H_i, H_k \rangle \\ &= \sum_k \omega_k G_{ik}, \end{aligned} \tag{3.42}$$

where $\mathbf{G} = \langle H_i, H_k \rangle$ is the second-order relationship between the leadfield vectors, or Gram matrix (depicted earlier in Equation 3.3).

$$\begin{aligned} \tilde{\mathbf{b}} &= \mathbf{G}\boldsymbol{\omega} \\ \boldsymbol{\omega} &= \mathbf{G}^{-1}\tilde{\mathbf{b}} \end{aligned} \tag{3.43}$$

Given that we are estimating $\mathbf{b} = \mathbf{HJ} + \mathbf{n}$ (Equation 3.1), the error of the MNE measure is stated as [30]

$$\begin{aligned} Err_{\mathbf{G}}^{-1} &= \|\mathbf{G}^{-1}\mathbf{b} - \mathbf{J}\|^2 \\ &= \|\mathbf{G}^{-1}(\mathbf{HJ} - \mathbf{n}) - \mathbf{J}\|^2 \\ &= \|(\mathbf{G}^{-1}\mathbf{H} - \mathbf{I})\mathbf{J} + \mathbf{G}^{-1}\mathbf{n}\|^2 \\ &= \|\mathbf{MJ}\|^2 + \|\mathbf{G}^{-1}\mathbf{n}\|^2, \text{ where } \mathbf{M} = \mathbf{G}^{-1}\mathbf{H} - \mathbf{I} \\ &= tr(\mathbf{MRM}^{\mathrm{T}}) + tr(\mathbf{G}^{-1}\mathbf{C}\mathbf{G}^{\mathrm{T},-1}), \qquad (3.44) \\ \mathbf{G}^{-1} &= \mathbf{RH}^{\mathrm{T}}(\mathbf{ARH}^{\mathrm{T}} + \mathbf{C})^{-1}. \qquad (3.45) \end{aligned}$$

where $\mathbf{R}$ and $\mathbf{C}$ are the covariance matrices of the signal vector $\mathbf{J}$ and noise vector $\mathbf{n}$, respectively, and $\mathbf{G}^{-1}$ the inverse operator. The transition from Equation 3.44 to 3.45 is accomplished by taking the derivative with respect to $\mathbf{G}^{-1}$, equating to zero, and solving for $\mathbf{G}^{-1}$.

The lack of information about the signals $\mathbf{J}$ and $\mathbf{n}$ can cause the $\mathbf{R}$ and $\mathbf{C}$ matrices to be close to singular. To avoid problems with inversion, we add a regularization parameter to Equation 3.45 [30, 64, 120]:

$$\mathbf{G}^{-1} = \mathbf{RH}^{\mathrm{T}}(\mathbf{HRH}^{\mathrm{T}} + \lambda^2\mathbf{C})^{-1}. \tag{3.46}$$

This method is suitable for finding cortical sources, since the relative strengths of multiple cortical sources are typically similar at the sensors. However, the MNE

method contains a fundamental bias towards superficial sources. As the algorithm is constructed, an optimal source distribution is found by searching the infinitely large solution space for a source distribution which minimizes the error between a source pattern and the measured fields. While it may be true that a given magnetic field recording would be better explained by set of deep sources, the iterative method would sooner choose a significantly more complex superficial field pattern. As such, variants of the MNE method have been introduced to correct for this bias [29, 120]. One such method is the sLORETA method.

3.2.5.2 sLORETA The sLORETA method attempts to improve upon MNE via by separating the variances of the different components of the estimated signal more explicitly. We restate the assumption that the source covariance matrix S_J is completely uncorrelated, which allows us to restate Equation 3.46 as

$$\mathbf{W} = \mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{H}^{\mathrm{T}} + \lambda^2\mathbf{C})^{-1}. \tag{3.47}$$

We are minimizing the function to find an estimate $\hat{\mathbf{J}}$ of $\mathbf{J}$. Using Equation 3.47 to solve the MNE (Equation 3.41), we can calculate the variance of the scalp electrodes to be

$$\begin{aligned} \mathbf{J}_\Phi &= K\mathbf{J}_J K^{\mathrm{T}} + \mathbf{n} \\ &= KK^{\mathrm{T}} + \alpha H. \end{aligned} \tag{3.48}$$

Using this, we can calculate the error of the estimated source waveforms to be

$$\begin{aligned} S_{\hat{J}} &= TS_J T^{\mathrm{T}} \\ &= K^{\mathrm{T}}(KK^{\mathrm{T}} + \alpha H)^{-1}K. \end{aligned} \tag{3.49}$$

This accounts for more than a single source of variation; namely, the sources themselves and potential variation in the noise signal. Once this normalization has been accomplished, a voxel-by-voxel weighing of the signals occurs according to the following equation:

$$\gamma_\lambda = \hat{J}_l^{\mathrm{T}} \left([S_J]_{ll} \right)^{-1} \hat{J}_l, \tag{3.50}$$

which produces F-statistic-like output to modulate the estimated source.

3.2.5.3 dSPM Dynamic statistical parametric mapping, or dSPM, determines whether the activation observed via MNE or sLORETA is statistically significantly relative to baseline activity. This method examines the activity of each voxel at each time point and looks for statistical significance of the activation. Mathematically, we simply compute a z-score based on Equation 3.43 [29]:

$$\begin{aligned} \mathrm{var}(\hat{s}_j) &= \langle (w_i n(t))^2 \rangle = w_i C w_i^{\mathrm{T}} \\ z_i(t) &= \frac{\boldsymbol{\omega}}{\sqrt{w_i C w_i^{\mathrm{T}}}} \left(= \frac{\boldsymbol{\omega}}{\sqrt{\mathrm{var}(\boldsymbol{\omega})}} \right) \end{aligned} \tag{3.51}$$

We can then obtain a noise-normalized estimate of the current dipole at each voxel location i by summing the z score for each possible dipole orientation:

$$q_i(t) = \frac{\sum_{j \in G_i} \boldsymbol{\omega}^2}{\sum_{j \in G_i} w_j C w_j^{\mathrm{T}}}, \tag{3.52}$$

where G_i is the set of three dipole components for the given voxel location.

4.0 TIME SERIES ANALYSIS

One major benefit to the high temporal resolution of MEG is that we can view changes in neural activity in real-time. When combined with source localization and source waveform reconstruction, we obtain accurate, millisecond-scale time series for individual regions of the brain. These time series can subsequently be examined for a wide variety of temporal relationships. A number of techniques for such analysis have been developed over the past few decades. One technique that has gained recent popularity among neuroscientists [38] is Granger causality, an extension of the vector autoregressive technique, which examines directional temporal correlation between multiple neural signals.

4.1 VECTOR AUTOREGRESSIVE MODELS

Multivariate autoregression (commonly referred to as vector autoregression, or VAR) attempts to predict future values of a given set of time series using the past values as predictors [15]. For example, let $X_1(t)$ and $X_2(t)$ be two $n \times 1$ time series vectors

of separate events recorded during the same time span $T = 1, \ldots, t$. The univariate autoregression is formulated as follows:

$$X_\alpha(t) = \sum_{n=1}^{T} A^{\mathrm{T}} X_\alpha(t-n) + \varepsilon_\alpha, \tag{4.1}$$

where $\alpha = 1, 2$, A represents an $n \times 1$ vector of coefficients describing the contribution of each time point to the final prediction, and ε_α represents additive gaussian noise. The values of A can be obtained through any of a number of estimation method, most commonly the least squares method [10]. If we only consider a subset $p < T$ of the complete time span T in our regression, we refer to the autoregression as having a *lag* of p. The VAR is formulated in a similar manner, with the added factor of one time series influencing the other. For example, a two-dimensional time series $\mathbf{X} = [X_1, X_2]$ with lag $p = 5$ we would be formulated as follows:

$$\begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} = \sum_{n=1}^{5} \begin{bmatrix} a_{11n} & a_{12n} \\ a_{21n} & a_{22n} \end{bmatrix} \begin{bmatrix} X_1(t-n) \\ X_2(t-n) \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}, \tag{4.2}$$

In this formulation, the various A_{ijn} vectors represent the coefficients of the lagged variables (i.e., the contributions of each lagged measurement to the model). Note that each time series potentially relies on all other time series. This is more apparent if the equation is expanded into algebraic form:

$$\begin{aligned} X_1(t) = {} & [a_{111} X_1(t-1) + a_{121} X_2(t-1)] + \cdots \\ & + [a_{115} X_1(t-5) + a_{125} X_2(t-5)] + \varepsilon_1 \end{aligned}$$

Using this formulation, we can describe the relationship between variables for each component of a given system.

4.2 GRANGER CAUSALITY

Given the model described in equation (4.2), we can attempt to define temporal dependencies between the variables by examining the behavior of the model when various components are removed. Granger [59] describes a mechanism for determining temporal relationships between two variables coexisting in the same system as follows. Let U_t be all the information in the universe, $U_t - X_t$ denote all this information apart from the specified series X_t, and $\overline{U_t}$ be all past values of U_t. In this context, utilizing the notation used in [59] , we will define $P(X_t|Y_t)$ as the optimum least squares predictor of X_t using the information present in Y_t, with error $\epsilon(X_t|Y_t)$. The variance of the error is represented by $\sigma_\epsilon^2(X_t|Y_t)$. We then have the following definition [59]:

Definition (Causality)**.** If $\sigma_\epsilon^2(X_1|U_t) < \sigma_\epsilon^2(X_1|\overline{U_t - X_2})$, we say that X_2 is causing X_1, denoted by $X_{2t} \Rightarrow X_{1t}$. We say that X_{2t} is causing X_{1t} if we are better able to predict X_{1t} using all available information than if the information apart from X_{2t} had been used.

As such, if a including the time series X_2 in the VAR predicting X_1 decreases the variance of the VAR error, we say that X_2 Granger-causes X_1.

This type of relationship is tested in practice by examining the variance of the model described in equation (4.2) with and without the possible causal variable of interest. An F-test is performed to test the null hypothesis that $\mathbf{a}_{12} = [0, \ldots, 0]$, and if the F-statistic is greater than the critical value, we state that the time series X_2 Granger-causes X_1, or $X_2 \Rightarrow X_1$. If not, nothing can be inferred from the test. The test is then run for each X_k within $\mathbf{X}$; in our case, X_2.

Granger Causality analyses can uncover a number of interesting interactions between sets of time series. Firstly, since the prediction of future values of X_1 is conditioned only on past values of U_t (i.e., $\overline{U_t}$), we can utilize Granger causality to infer directionality between interactions. Given time series X_t and Y_t, determining that $X_t \Rightarrow Y_t$ does not imply that $Y_t \Rightarrow X_t$. Secondly, while the definition stated above examines the interaction between a given time series and the immediate history of itself and other series, there are a number of trivial modifications that can be made to the construction of the VAR which enable the examination of different temporal interactions, such as Instantaneous Causality, Causality Lag, Feedback, and other interactions [59]. Additionally, there have been a number of methods described which examine Granger causality in the frequency domain [7, 138].

4.3 TIME SERIES ANALYSIS IN FUNCTIONAL NEUROSCIENCE

The working hypothesis that functional responsibility within the brain is spatially segregated was formulate some decades ago [168], and significant research since then has lent itself towards supporting this concept. For many regions in the brain, functional neuroimaging has discovered that function defines structure, and structure bounds functional regions [147, 157]. Structures are not limited to a single function, just as a single function may be spread across multiple structures. However, in many applications, structure and function appear intimately related.

Working in this framework, we can attempt to map connectivity between structures to better understand how structure and function are related. A number of statistical tools have been utilized to help define these relationships, including coherence [107, 108], regression analyses [15], Granger causality [38, 59], directed [134]

and partial directed coherence [7, 138], among other methods [5, 52, 56, 105, 144]. These methods have been applied to fMRI, EEG, implanted EEG, and MEG data, as well as single neuron local field potential (LFP) data [1, 5, 56, 58, 70].

The Granger causality method has received significant attention, though, and many of the methods mentioned above are derivatives derived directly from Granger causality. Its strength lies in its enabling the examination of two critical components of any given time course; predictability and directionality. Within neuroscience, this translates into directional network maps [40]. Examining the activity pattern of a given region as a function of the neural signatures of related (or possibly related) regions, we can infer which regions are strongly connected, which weakly connected, which are connected through a poly-nucleic pathway, which are connected in a bidirectional feedback-type relationship, among others [38, 59]. By only examining the activity of one region as a function of the past values of a related region, we avoid impossible situations where future activation appears to cause an earlier signal.

One highly desirable outcome of Granger causality is the construction of possible functional connectivity maps between related regions. While the topic of mapping temporal relationships is itself an entire field of study comprising many possible relationship mechanisms [40], many neuroscience applications utilize only two types of relationships; unidirectional and bidirectional. The maps can be constructed from any neural imaging modality which can provide a set of time series, including EEG, MEG, and even BOLD-dependent modalities such as fMRI and PET [1, 56]. Despite their inherent over-simplicity, such maps can be of significant use in helping understand a complex neural network.

5.0 EXPERIMENT 1: EXAMINING REWARD-RELATED ACTIVITY VIA FMRI IMAGING OF GAMBLING TASK

Our primary overarching goal was to observe the activity of the human rewards pathways via MEG. We chose to elicit this activity via a rewards-related gambling paradigm (see Section 2.3). The reason for this choice was twofold. Firstly, as described earlier, the rewards system includes superficial components—which are more easily detected by the MEG scanner—as well as deeper nuclei. Superficially, the OFC mediates, in part, the formation and storage of modality-independent stimuli with their corresponding reward values [130]. Slightly deeper, the anterior cingulate cortex has been shown to be involved in decision making in the presence of conflicting or insufficient information [12], a function highly relevant to maximizing reward returns during a gambling task. Within the basal ganglia, striatal activity has been linked to reward detection as well as the formation and maintenance of S-R behavior [11, 117]. As such, through the use of a rewards-related paradigm we could effectively activate a number spatially distinct neural regions, superficially and deep.

Secondly, the rewards system has received a significant amount of attention from the research community, providing much insight into the possible workings of the network. Several researchers have noted a distinct similarity between the activity profiles of several reward-related regions both during experimental stimulation

[50, 51] and reward-related behavior [23, 122]. Gariano and Groves [51] noted a similarity between the firing of medial prefrontal activity, anterior cingulate cortex neurons, and midbrain dopamine neurons during neural activity induced by electrical stimulation. A more quantitative study later demonstrated that a significant portion of midbrain DA cells—approximately 30%—exhibit bursting following electrical stimulation of the PFC [153]. This directional PFC-striatal linkage has been observed during reward-seeking behavior as well. A study examining NMDA antagonists concluded that the NMDA-antagonist induced dopamine increases in the PFC were likely caused by a reciprocal PFC-VTA connection, beginning with a PFC-induced increase in VTA glutamatergic activity [149]. This study is particularly relevant, for it helps link function with anatomy; the existence of a dopaminergic mesocortical loop has been documented in both anatomical and functional literature [21, 148], but this study helps shed light on the activation sequence of this bidirectional network.

It should be noted, though, that evidence exists for a VTA-PFC activation sequence. In the same paper in which they observed evidence of a PFC-VTA-PFC paper Takahata and Moghaddam [149], the authors found evidence that PFC activation is secondary to the VTA, based on similarities between the reactions of rats receiving intra-PFC and intra-VTA infusions of AMPA antagonists. Additionally, Hollerman et al. [72] hypothesizes that from a functional standpoint it would be logical for both ACC and VTA to drive PFC function, inasmuch as striatal neurons provide reward information and behavioral contingencies, ACC provides error detection, and PFC incorporates the information of both [72]. However, the authors note that this is not a direct hierarchy, as information integration at the systems level is highly complex.

In our experimental setup, we are examining network activity during a complex reward task using non-invasive measures in the human. As such, it should be noted

that the preponderance of literature in the rat model may not be representative of the results we may find, due to a number of significant differences; differences in stimulation mechanism (direct current injection vs. regional recruitment due to behavioral performance), network recruitment (localized electrical stimulation vs. natural activation), task demands (no network demands during passive network monitoring vs. requirement to perform behavioral task), and imaging modality (highly sensitive, localized electrode recordings vs. diffuse whole-head activity). Alternatively, one may assume a bottom-up approach, that the underlying network dynamics are similar in both the human and rat models and that more complex functions build on similar underlying network dynamics.

In formulating our working hypothesis we tend to view the system from a functional perspective. The striatal regions, which act as both an integrator of information and a significant feeder to the frontal planning regions, would activate early in rewards processing. ACC, which acts as an expectancy monitor as well as initiating the execution of an action, would be activated numerous times over the course of the trial; providing expectancy information, comparing the stimulus to the expected result, and then as an initiator of activity. OFC, which acts as a monitoring and planning center, would be activated after receiving signal from the striatum and concurrently with ACC in its error-monitoring capacity. As such, the hypothetical network would look similar to that constructed by Hollerman et al. However, in formulating our network in this manner, we must acknowledge that the basic neuroscience research would tend to disagree with this formulation, as noted above.

To test this hypothesis, we utilized the experimental paradigm of Liu et al. [99]. While the efficacy of the exSSS method for extracting deep activity from MEG signal has been demonstrated both in theory and via simulations [113], it has not yet been validated in an experimental setting; indeed, this set of experiments marks

the first time this method has been tested in human subjects [81, 82]. Due to the complex nature of the MEG source localization process, we used functional magnetic resonance imaging (fMRI) to verify that the activity of interest was indeed being elicited, as well as to ascertain the location of the neural activity.

5.1 GAMBLING EXPERIMENT

Our experimental paradigm design was based on three primary criterion. Firstly, the paradigm needed to selectively activate the reward network. Secondly, the activity observed in the deep regions should be significant. As described above, there are a number of obstacle to observing deep activity in MEG; we wanted to ensure that our paradigm would elicit very strong activity so we would have the highest chance of detecting any such activity. Additionally, the paradigm should theoretically activate the regions of interest serially. With this design and with the millisecond-level temporal resolution of the MEG, we would be able to examine the activation of each region individually in the absence of activity from the rest of the network. Note that this aspect of the paradigm design exists mostly in theory, as no studies exist in the literature detailing the order of activation of the rewards network.

5.1.1 Methods

5.1.1.1 Participants We ran a total of five subjects in the MRI (two female), all right-handed (see Table 2). An informed consent form approved by the University of Pittsburgh was obtained from all subjects. Each subject was trained to proficiency in the task prior to entering the scanner. Participants were screened for mental health

Table 2: Description of subject population.

Subj. ID	Gender	Age	MEG	fMRI	Notes
4	Male	28	√	√	
5	Male	26	√		Subject entered fMRI scanner but cancelled scan due to claustrophobia. No data collected.
6	Male	27	√		
7	Female	33	√	√	MEG data discarded due to noise.
8	Male	26	√	√	
9	Female	29	√	√	
10	Male	29	√		Control subject.

and basic addiction disorders via questionnaire. All subjects were able to perform the task without difficulty. Analysis of behavioral data collected during MEG and fMRI runs showed an average of 171 correct trials and 145 incorrect trials per subject (st. dev. 6.34 and 5.85, respectively). One male subject experienced claustrophobia upon entering the scanner and cancelled the scan prior to data collection. The remaining four subjects completed the scan without event.

5.1.1.2 Design and Task We used the paradigm detailed by Liu and colleagues [99]. In this article they detailed a gambling paradigm with which they successfully elicited significant activity from the rewards network. The task proceeds as follows (see Figure 7 for a paradigm flowchart). Subject begin each block with a bank of ten gambling chips. At the first trial of each block, subjects are given the option to bet or bank a single chip. If they choose to bank, their wager is placed back in the bank, and the results of the subsequent die roll do not affect them. If they choose to bet,

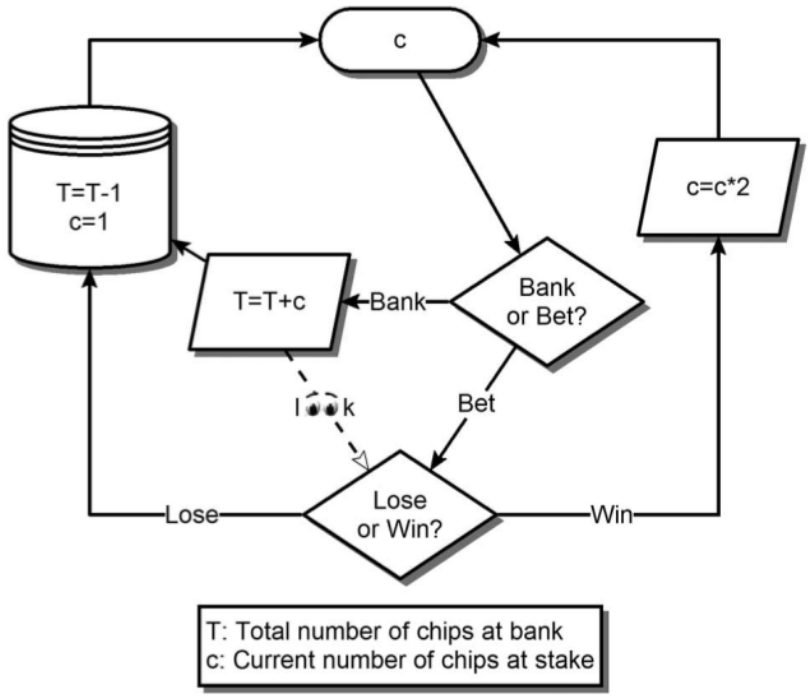


Figure 7: Subject begin each block with a bank of ten gambling chips. At the first trial of each block, subjects are given the option to bet or bank a single chip. If they choose to bank, their wager is placed back in the bank, and the results of the subsequent die roll do not affect them. If they choose to bet, if they win, their wager is doubled (i.e., they will now be betting on two chips). If they lose, their chip is confiscated. On the next round, they are presented with their current ante and given the "bet or bank" choice again. If they won the previous round, their ante consists of two chips (double the original ante), and if they win this round as well, their wager will again double to four chips. Subjects can win up to five times in a row at which point their bet will be automatically banked. If they lose at any point, their entire bet will be confiscated and they will need to use a chip from their bank to begin the next round. Image taken from Figure 1 in [99].

if they win, their wager is doubled (i.e., they will now be betting on two chips). If they lose, their chip is confiscated. On the next round, they are presented with their

current ante and given the "bet or bank" choice again. If they won the previous round, their ante consists of two chips (double the original ante), and if they win this round as well, their wager will again double to four chips. Subjects can win up to five times in a row at which point their bet will be automatically banked. If they lose at any point, their entire bet will be confiscated and they will need to use a chip from their bank to begin the next round. The overall goal is to maximize the number of banked chips.

This paradigm serves two purposes; the subject will have to make a rewards-related decision each round, and the subject will be anticipating receiving a reward or punishment each round. While this rich paradigm allows for a number of possible variables for investigation (chosen action (bank versus bet), outcome (win versus lose), response during "streaks" (trials in which subjects consecutively won or lost numerous times in a row)), Liu found significant differential activation of the striatal nucleus between trials in which the subject made the objectively "correct" choice—either through betting and winning or not betting and subsequently losing the die roll—versus when the subject made the objectively "incorrect" choice—either through betting and losing or not betting and subsequently winning.

The experiment was coded using E-Prime Studio 1.1 (Psychology Software Tools; Pittsburgh, PA). Behavioral analysis was completed in both the E-DataAid component of EPrime Studio and Microsoft Excel (Microsoft Co.; Redmond, WA). As the gambling nature of the study belies the defining of an objectively correct response prior to seeing the outcome, subject performance was assessed by both reaction time and the general trend towards earning more chips. All subjects practiced until they were judged proficient in the task by the study administrator before entering either the MEG or fMRI scanners.

5.1.2 Differences between the fMRI and MEG paradigms

The behavioral paradigm was identical to that used in the MEG (Section 5.1.1) with the exception of some aspects of paradigm timing and number of trials. In the MEG experiment the subject viewed a cue screen for two seconds (depicting their ante), a choice screen for either two seconds or until they responded (whichever came first), a feedback screen for two seconds, and a fixation screen for three seconds. In the fMRI experiment the cue, choice, and feedback screens were all displayed for a fixed two seconds, and the fixation screen was fixed at four seconds, for a total of ten seconds per trial. This change was effected to account for the significantly longer TI required in an fMRI study; whereas with the MEG we can scan at a resolution of 1000Hz, the typical scanning resolution for the fMRI is two seconds. As such, we timed our experiment so that the duration of each screen is equal to a single TR.

Additionally, the subject performed 320 trials in the MEG scanner, compared to 80 in the fMRI scanner. This is due to both the significantly higher SnR of the fMRI scanning modality as well as the high signal strength expected from the paradigm.

5.1.3 fMRI Data Acquisition and Preprocessing

fMRI data was recorded using the 3.0 Tesla GE Magnetic Resonance scanner in the UPMC Magnetic Resonance Research Center (MRRC). Image acquisition was carried out on a 3T Siemens Allegra MRI scanner. Thirty-four transaxial slices were acquired every 2s (FOV: 210, TE: 30, Flip angle: 70, Slice thickness: 3 mm), with a total of 150 EPI volumes collected per run. Three-dimensional anatomical MP-RAGE images and T2 structural in-plane images were collected for each subject. The purpose of this scan was twofold; obtain structural images of the subject for use in BEM construction and dipole visualization, and to validate the MEG findings

in fMRI. Since the MEG data analysis method being used is experimental, spatial information obtained from the fMRI activation pattern was used as a benchmark to both ensure that the intended activity is present as well as ensuring that our MEG analysis appears reasonably accurate as compared to the fMRI data.

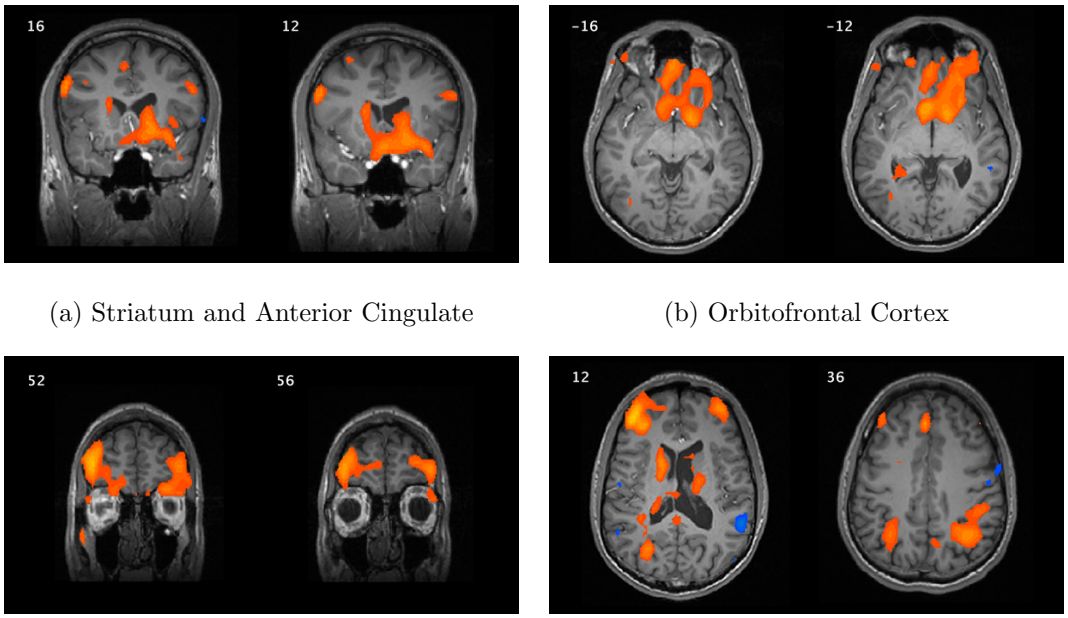
The behavioral paradigm was projected into the fMRI scanner room on a translucent plexiglass panel located approximately six inches from the end of the MRI bore. Subjects viewed the paradigm via a mirror mounted on the MRI head coil unit. Responses were made on using an identical hardware setup as that of the MEG.

BrainVoyager QX 1.4 (Brain Innovation, Maastricht, The Netherlands) was used to analyze the recorded fMRI data [55]. Preprocessing routines included motion correction, slice scan time correction, spatial smoothing to 8 mm, and linear detrending. All data was Talairach aligned and transformed. Response activation patterns were modeled using the hemodymanic response function. Statistical analyses was completed within the framework of the general linear model (GLM). The conditions were modeled using a 2x2x2 design matrix (gambling risk (high vs low), subject response (bank vs bet), and outcome (win vs. lose)), with each trial fitting one of eight possible conditions. Separate covariates were included in the GLM for each possible condition combination to model each possible outcome. Due to the involvement of subject response, events could not be completely balanced, with some categories containing significantly more events than others. Using this GLM, a contrast was defined in which all "perceived correct" conditions were contrasted against all "perceived incorrect" conditions. We searched for areas in which activity was significantly different between the two conditions, with statistical significance was defined by meeting a threshold of $p < 0.001$ ($t > 3.85$). We examined activity in six primary ROIs, based on anatomical structures; bilateral OFC, cingulate, and striatum.

During the course of data processing using BrainVoyager, it became clear that the Talairach transformation was not functioning properly. Conversations with Brain Innovation technical support revealed that BrainVoyager was not properly reading some of the header information encoded in the raw fMRI files, and that was interfering with the processing of the files. This resulted in the GLM map overlays not coinciding properly with the actual subject anatomy, since the GLM map was constructed on non-Talairach transformed data, and then overlaid on a Talairach-transformed image. In order to correctly visually represent the data, the patterns were therefore transformed using a linear transform and re-overlaid on the structural images. Specifically, we constructed the GLM map, performed a linear transform on the data (increase vertical height to 120% of original, decrease width to 90% of original) using a standard image manipulation tool, and overlaid the resultant activity on the structural image. We felt justified in performing this manual transform since the Talairach transform is very nearly linear. The activity pattern depicted in Figure 8 was obtained and generated in this manner. This transform is not perfect, as evidenced by prefrontal activity appearing within the right orbit (Figure 8(b)). However, this does provide a close approximation of true activity.

5.2 RESULTS

We found BOLD activation in the "correct>incorrect" comparison in the left ACC, left striatum, and bilateral OFC, as well as bilateral activity in the parietal cortex and dorsolateral prefrontal cortex (Figure 8). These regions were all significant beyond the $p < 0.001$ threshold. Almost all regions were more strongly activated during the



(a) Striatum and Anterior Cingulate

(b) Orbitofrontal Cortex

(c) Lateral Frontal Cortex

(d) Striatum, Lateral Frontal Cortex, Anterior Cingulate, and Posterior Parietal Cortex.

Figure 8: fMRI signal resultant from "correct>incorrect" contrast using eight-parameter GLM as described in text. All activation depicted in image is significant at $p < 0.01$. The coordinate on the upper-left of each slice identifies the location of the slice in talairach space. The right side of the brain is the right side of the image. Note that due to technical difficulties, this is not the raw activity pattern generated by the GLM; see Section 5.1.3 for details.

"correct" condition than the "incorrect" condition (Table 3), with the exception of the lateral temporal regions.

We also examined whether these findings were also reflected in the activation time courses of these regions differed between the correct and incorrect comparisons. As hypothesized, we found significant activity in the striatum, lateral OFC, and

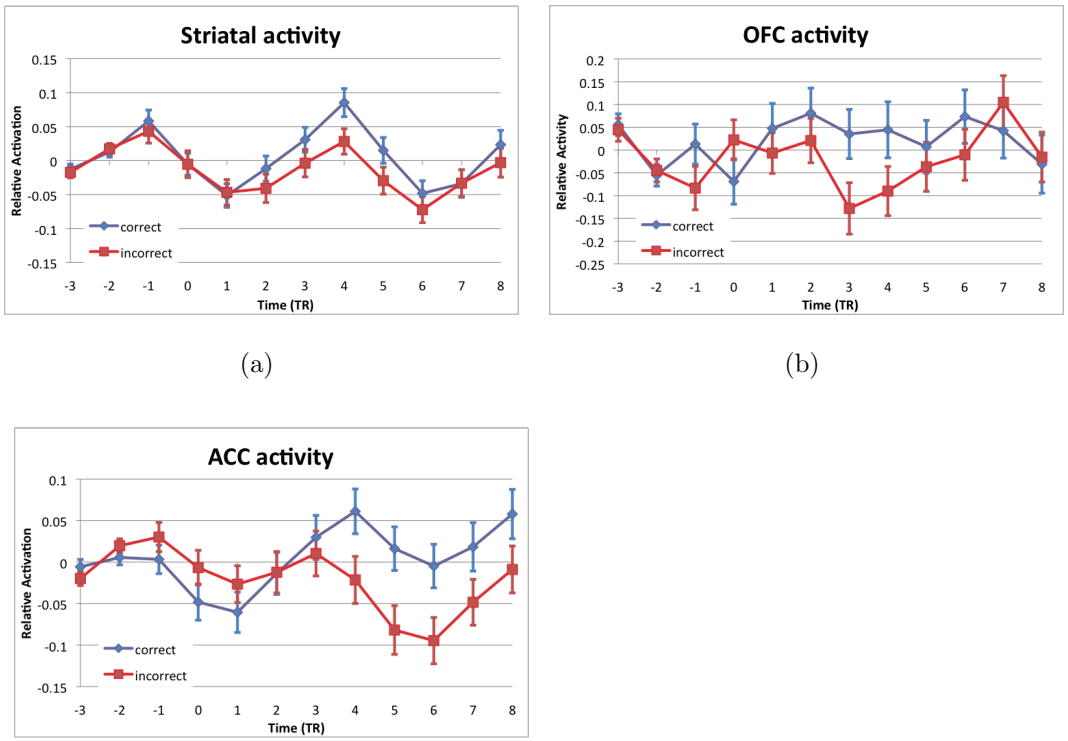
Table 3: List of peak voxels for fMRI activation clusters.

Label	BA	x	y	z	t
Correct > Incorrect					
Anterior Cingulate Cortex, left	*32*	-6	23	35	5.386
Dorsolateral Prefrontal Cortex					
left	*46*	-40	42	15	5.525
right	*46*	41	40	19	3.865
Motor Cortex, right	*6*	30	-8	39	3.871
Orbitofrontal Cortex, left	*11*	-1	37	-10	5.097
Parietal Cortex					
Inferior					
left	*39*	-30	-64	19	6.215
right	*39*	29	-64	28	6.129
Superior					
left	*7*	-30	-64	19	6.215
right	*7*	29	-64	28	6.129
Striatum, left	–	-13	4	7	5.396
Incorrect > Correct					
Posterior Temporal Cortex					
left	*10*	59	-37	16	5.782
right	*10*	-56	-39	15	3.214

ACC regions (Figure 8). Table 3 lists a number of other regions in which activity was observed.

In addition to examining the activation patterns we also examined significance in the action time courses of the ROIs (Figure 9). Notably, for the regions of interest, there exists a significant difference between the activation patterns of the different conditions, as expected.

One interesting observation is that a number of regions were *not* observed as active, that were observed in the original publication by Liu et al. [99]. Most no-



(a) (b)

(c)

Figure 9: Activity time courses obtained from ROI analysis of selected region depicted in Figure 8. The -1 point corresponds to the TR during which the subject viewed of the ante, the 0 point corresponds to the choosing between bank or bet, the 1 corresponds to the viewing of feedback, 2 and 3 correspond to the intertrial interval, and the 4 corresponds to the beginning of the next trial, with the pattern repeating from there on out. Note the increased differential activity between the conditions in all regions during the intertrial interval. While the differential shows up later than may be expected, this could be an related to the time taken by the subject to consider the outcome of the previous trial.

tably, activity was not observed in the medial OFC at all, and the ACC and lateral inferofrontal activity we observed appeared positive when viewed using the "correct $>$ incorrect", opposite that reported by Liu. As directionality of activation is often

used to infer function (i.e., more active when correct), this distinction is functionally relevant.

6.0 EXPERIMENT 2: EXAMINING REWARD-RELATED ACTIVITY VIA MEG IMAGING OF GAMBLING TASK

6.1 METHODS

The behavioral paradigm was identical to that used in the fMRI experiment (Section 5.1.1) with the exception of some aspects of paradigm timing and number of trials, described in detail in Section 5.1.2. Additionally, one MEG subject was run using a control version of the experiment. The function of the control was to ensure that the activity we were observing was indeed due to the subject viewing and processing the feedback, and not to some other confounding neural process. In the control experiment, the cue, "bank or bet" screen, and fixation were kept identical to that in the regular version. The feedback screen was modified so that any indicators as to whether the subject won or lost was removed. To that effect, the words "You Win!" (or "You Lose!") were replaced with the generic phrase "You XXX!". Additionally, the numerals indicating number of chips won and total chips were both replaced with a "XX". Finally, the die was replaced with a made-up die consisting of nine white dots. The overall effect was one of almost complete visual similarity to the regular task but completely uninformative when considering whether the subject actually

won or lost the round. Subjects were informed of the nature of the experiment, and performed the task prior to scanning to ensure that the task was understood.

It is worth noting that during the control experiment the subject would be able to determine whether they won or lost based on the appearance of the ante screen on the next round. However, this is temporally distant from the time points we analyzed (as discussed below) and should not interfere with our signal of interest.

6.1.1 Participants

Subject demographic information can be seen in Table 2. Six subjects (four male) were run in the MEG normal task and one subject (one male) was run in the MEG control task. Informed consent was obtained from all subjects. An analysis of behavioral data can be found in Section 5.1.1.1.

6.1.2 Data Acquisition and Preprocessing

Data was acquired using the MEG and MRI machines available at the University of Pittsburgh. The structural MRI data used for the boundary element model construction was obtained at the same time as the fMRI experiment described earlier.

6.1.2.1 Magnetoencephalography All MEG data was recorded using the 306-sensor Elekta NeuroMag MEG system in the University of Pittsburgh Medical Center (UPMC) Center for Advanced Brain Magnetic Source Imaging (CABMSI) with integrated magnetometers and orthogonal gradiometers. Prior to entering the scanner subjects were fitted with HPI positioning coils (Polhemus PATRIOT Digitizer, Colchester, Vermont) to enable coregistration of MEG data with realistic head models. Scan data was recorded at a temporal resolution of 2000 Hz. The subject was

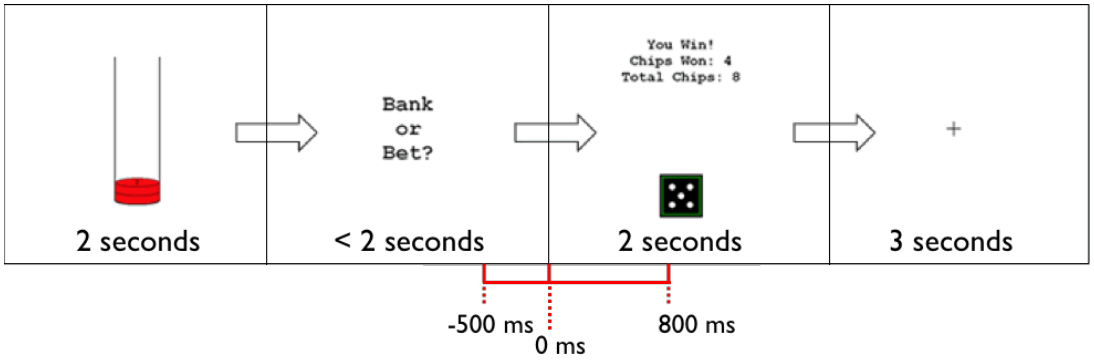


Figure 10: Schematic of a single MEG behavioral trial. The length of each segment is depicted on the schematic. The 1300 ms segment indicated underneath the schematic denotes the segment subjected to averaging, processing, and reconstruction. Note that the response phase lasts for $\leq$2 seconds; this is different from the length used in Liu et al. [99].

placed in a sitting position for the duration of the experiment approximately four feet from the display screen. All responses were indicated using the right-hand glove from the a 10-Button Fiber Optic Button Response System (Psychology Software Tools; Pittsburgh, PA). The paradigm was run for eight blocks of 40 trials, where each trial consisted of viewing the ante (2 seconds), making a bet/bank decision (2 seconds or until button press, whichever is shorter), viewing the results (2 seconds), and a fixation screen (3 seconds). A schematic of a single trial can be seen in Figure 10.

The MEG data itself was processed as follows. Data was first passed through a 1-40 Hz bandpass filter to remove low frequency drift and high frequency signals. While a number of researchers have examined the higher frequency (60-200 Hz) bands in MEG data [124], such signals attenuate with distance significantly quicker than low frequency signal, and it is highly unlikely that any signal would reach the sensor

array even when using the exSSS method. Data was subsequently averaged across perceived outcome as described in the methods section (5.1.1). The data was treated using the SSP noise reduction algorithm [156], using the SSP matrix generated by the NeuroMag system during the scan session, for further noise reduction. The exSSS and genexSSS methods were then applied to enhance deep activity [113]. Source localization was accomplished both by using sLORETA [120] and dSPM [29]. Regions of interest were defined on the BEM on a per-subject basis based on anatomical landmarks. Time series vectors were computed by averaging the activation value of each BEM mesh element for all elements within the region of interest (ROI). Time series analysis was completed in Matlab using the Causal Connectivity Analysis Toolbox [144] and Spatial Econometrics Toolbox [96]. All time series were normalized to 1 prior to analysis.

6.1.3 Boundary Element Model construction

The FreeSurfer suite (Martinos Center for Biomedical Imaging; Massachusetts General Hospital, Boston, MA) was used to create a boundary element model (BEM) of the scalp, white matter, and grey matter [31, 47]. This model was subsequently used to calculate the MEG forward and inverse problem. The BEM did not include skull and pial layers, as they are not necessary for MEG processing, due to the relative ease with which the magnetic signals penetrate biological tissue without interference [64]. The parameters used to create the BEM were those recommended by the MNE software User's Guide [63]. The BEM was used later in the analysis to visualize the inverse solution as generated by the MNE software.

6.2 RESULTS

The results of the MEG analysis are presented here at two different stages of processing: averaging and localizing. All data was processed as described in Section 6.1.2.1. For comparison purpose, we also examined an averaged dataset not subjected to the exSSS processing method, referred to here as the "unprocessed" dataset.

Many of the following figures depict magnetoencephalographic traces. Those unfamiliar with reading MEG data traces are referred to Figure 21 for a graphical orientation to the layout of the traces, as well as an approximate map of which neural region is represented by each sensor.

For the normal task, the averaged, unprocessed data contained significant peaks, in order of appearance, in the occipital (∼100 ms post-button press), parietal (∼230 ms), and motor (∼250 ms) areas (see Figure 11). These peaks were present in both the "correct" conditions. Minor peaks were observed in some subjects over the prefrontal regions at a number of time points (-45, 75, 170 ms prefrontal;). A strong temporal signal was present in almost all subjects at 150 ms. While the times mentioned here are not constant across subjects, they are mentioned so the reader can note the relative temporal difference between regional activation. These temporal differences were relatively constant between subjects.

Data from our control subject exhibited a slightly different pattern (see Figure 12), beginning with activity in the bilateral motor regions and left dorsolateral prefrontal cortex 35 ms prior to button press. A burst of activity was seen in the left prefrontal and anterior temporal region at ∼70 ms after button press, followed by strong visual activity in the occipital region 100 ms post-button press. A long period of activity was observed starting at 200 ms post-button press and lasting for approximately 200 ms. The most notable feature of the control data was the relative quiet in

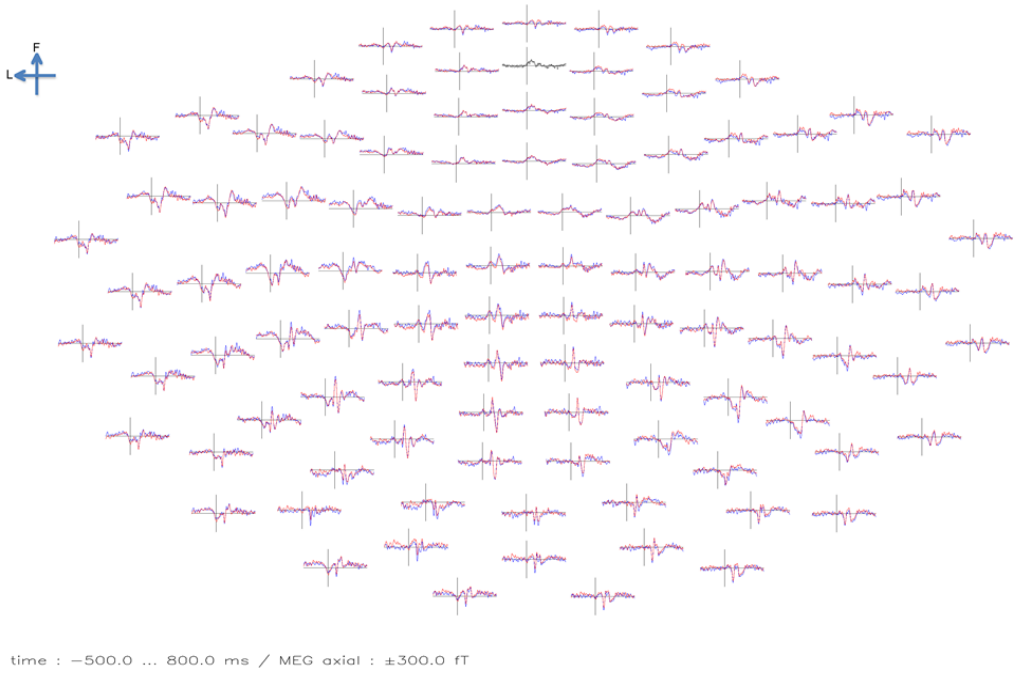


Figure 11: Magnetometers tracings from data after filtration via 1-40 Hz bandpass filter, SSP, SSS, and averaging across trials for a single subject. Red traces indicate objectively "correct" trials and blue indicate "incorrect," as discussed in the methods section. Note the presence of motor, visual, and parietal activity, all of which are necessary for the task (motor for button press, visual for eye motion to see cue, parietal for planning). Much of this disappears after differencing (see Figure 14).

the right temporal area from 150-250 ms; all normal subjects exhibited at least one large peak in this region during this time period (see Figure 13). Interestingly, there is significant activity in the left temporal region during this experiment, suggesting lateralized activation over the left temporal region.

To examine activity that was differentially strong in one condition relative to the other, we subtracted one condition from the other and examined the resultant waveforms. Notably, there were no significant peaks in the differenced data for the

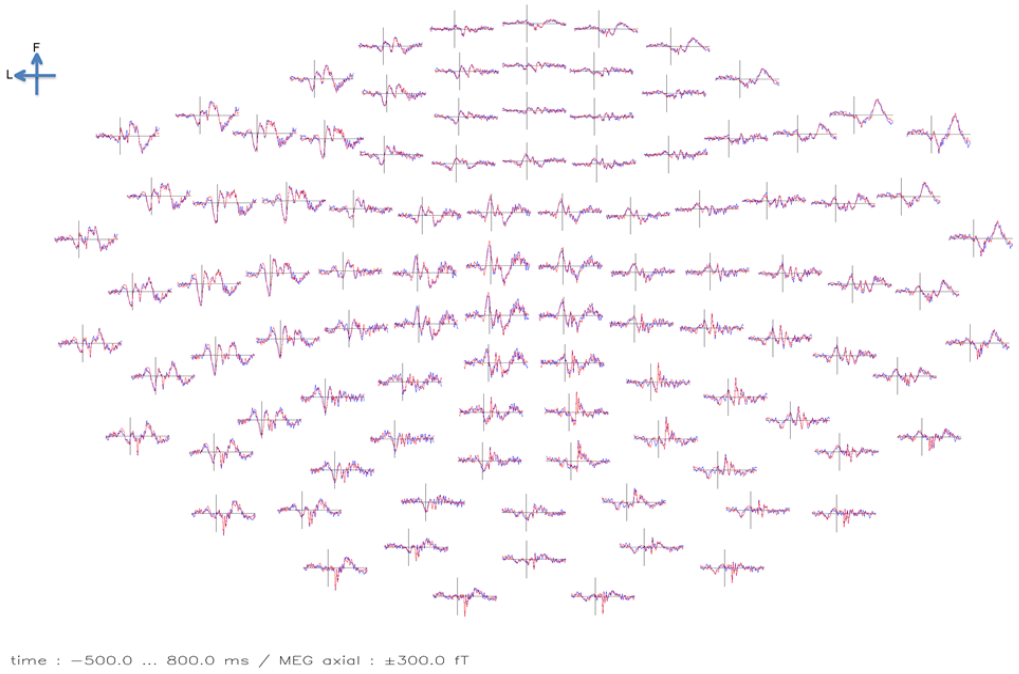


Figure 12: Magnetometer tracings from control data after bandpass filtering, averaging, SSP, and SSS processing. While extrapolation of the results is difficult due to this being single subject data, note that the general pattern of the non-differenced control activity is similar to that of normal subjects (Figure 11).

"unprocessed" dataset (Figure 14). The processed dataset—the one subjected to the exSSS spatial filtering mechanism—contained two broad waveforms present across the entire head of the subject at 100 ms and 250 ms (Figure 15). These waveforms were present almost exclusively in the magnetometers, with very little of this signal being present in the gradiometers. These signals were also highly variable across subjects, but the components mentioned above were present to a moderate degree in all datasets. Some subjects also exhibited a strong waveform at 450 ms in the sensors over the lateral inferior temporal regions.

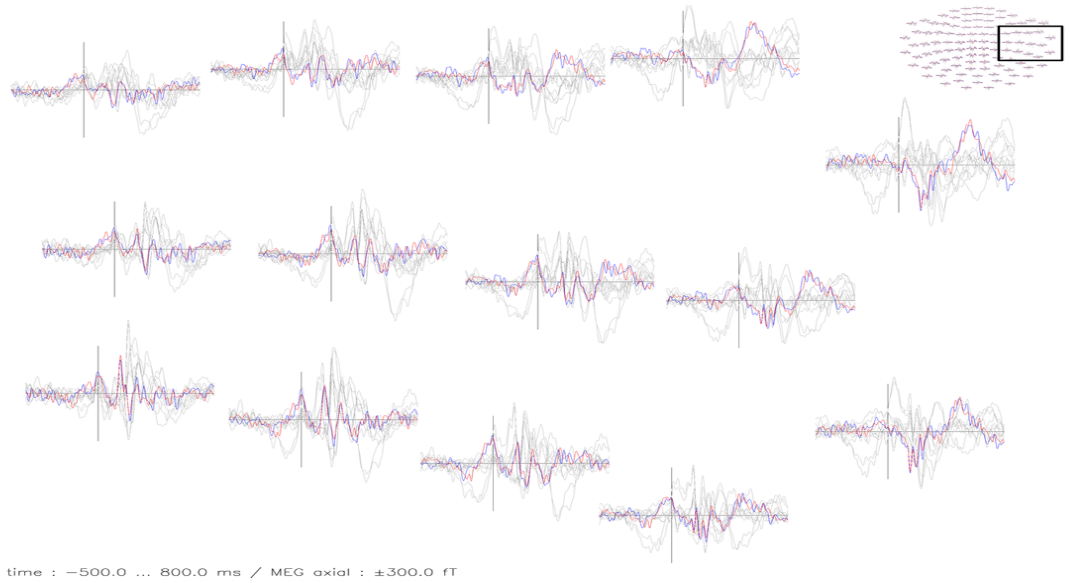


Figure 13: Right temporal magnetometers from the control subject. Red traces indicate objectively "correct" trials and blue indicate "incorrect" for the control subject, and the grey traces are the "correct" and 'incorrect" traces from all other subjects. As demonstrated in Figure 11, the difference between conditions is minimal, so all traces are similarly colored to simplify appearance. Note the relative lack of activity in these channels during the control task compared to the activity present during the normal task performance. This distinction will be made clearer in the localization results (Figure 19).

Utilizing the differenced dataset, source reconstruction was performed using dipole localization and the minimum norm method. In both case the sources were identified on the BEM mesh constructed from the individual subject's MRI images. MEG to MRI coregistration was accomplished using the HPI digitized points saved to the MEG datafile and locating those points in the corresponding MRI file. For subjects without MRI data, the reconstruction was performed on a comparable head, and the MEG data was transformed to the appropriate source space, thus minimizing localization errors.

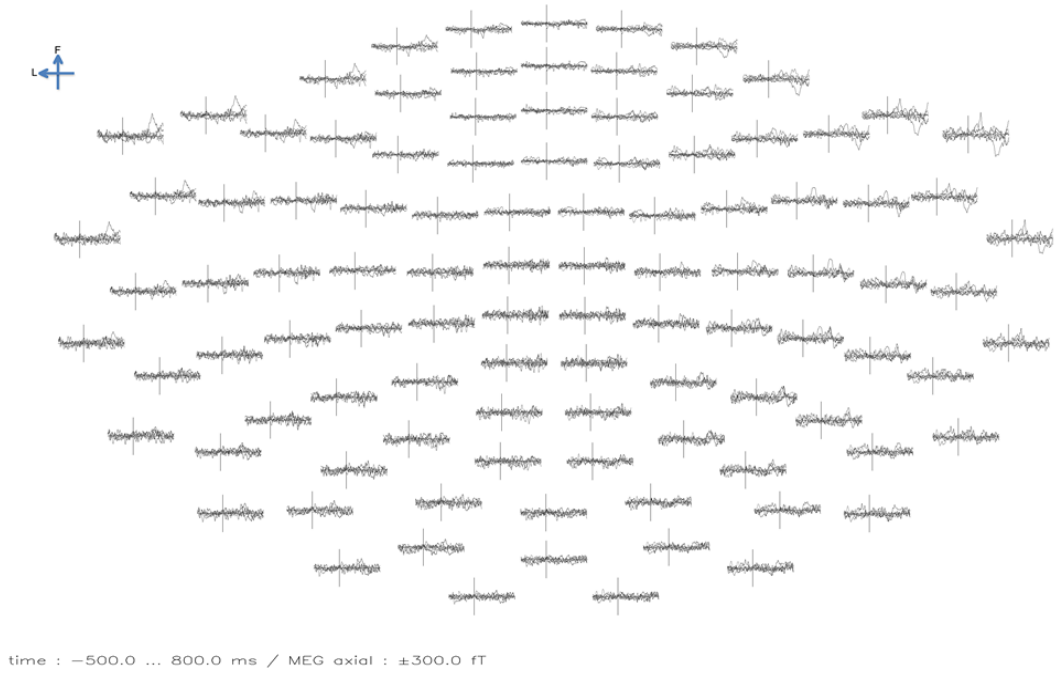


Figure 14: MEG magnetometer traces resultant from differencing the "correct" trials from "incorrect" trials depicted in Figure 11 for all subjects. Note the lack of significant peaks across almost all datasets. As can be seen in the frontal-temporal regions, one subject did exhibit strong, late activity. This trace appeared only in a single subject and bears further investigation.

Dipole localization revealed a number of dipoles located in the posterior cingulate regions and the white matter of the parietal lobe. Each of these dipoles presented with a $> 85\%$ goodness of fit. Both dipoles were active immediately after the button press. Parietal activity was observed again around 175-205 ms, followed the posterior cingulate dipole at 230-250 ms. These dipoles were highly variable between subjects, with three subjects showing the results reported here. The remaining subjects did not possess any notable dipole activity with a significant goodness of fit. As such,

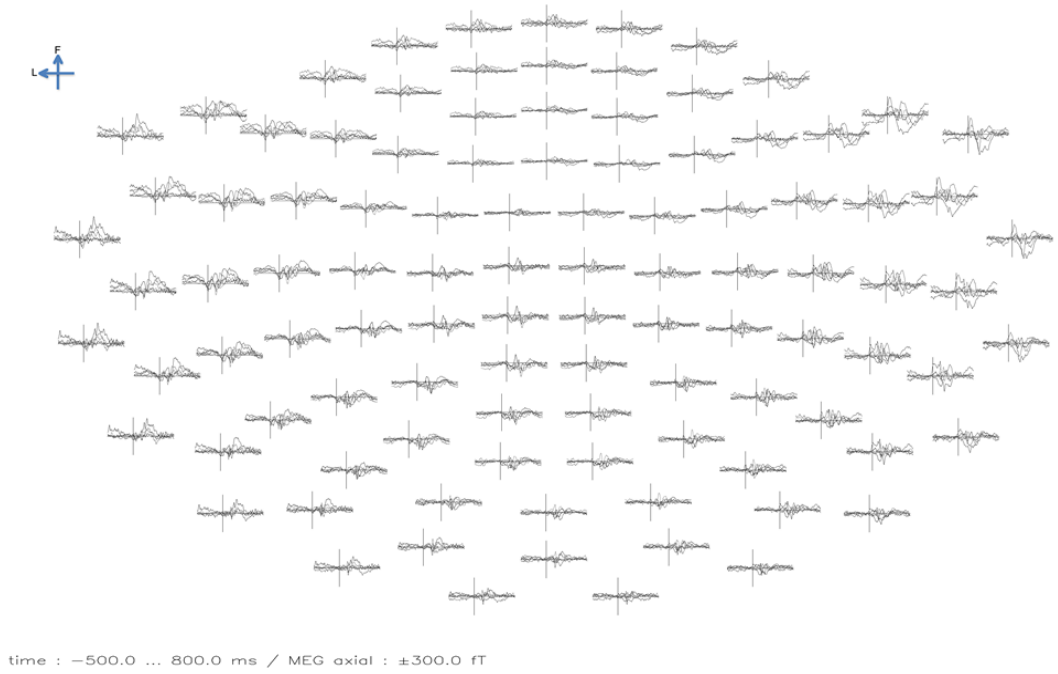
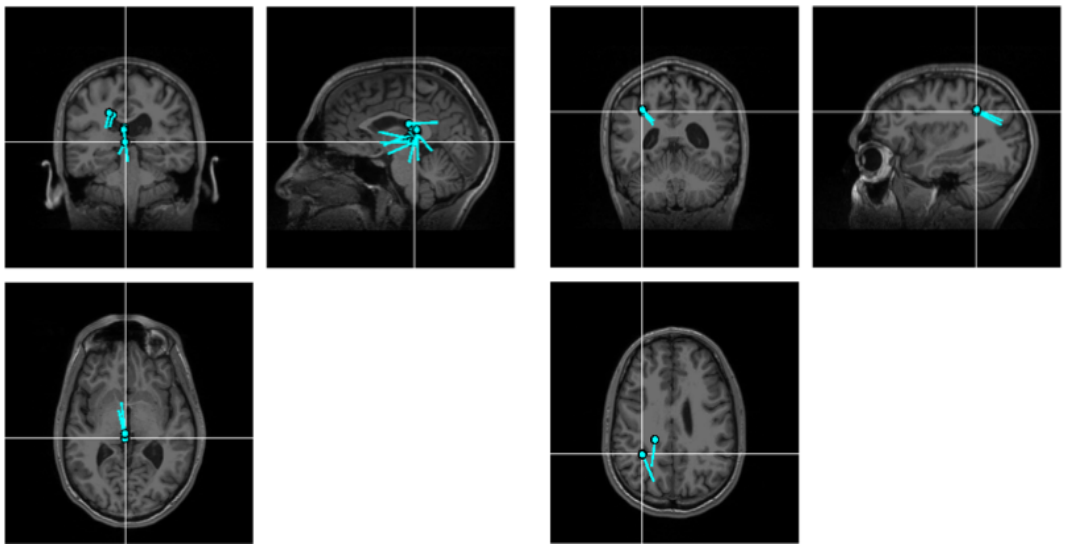


Figure 15: MEG magnetometer signal resultant from differencing the "correct" trials from "incorrect" trials depicted in Figure 11 and subsequently passing through the exSSS modified beamspace filter for all subjects. Note the presence of significant activity, particularly in the temporal regions, that was not present in the differenced data (Figure 14).

sources found via this technique were not used for source waveform reconstruction. (See Discussion, Section 7.3.3).

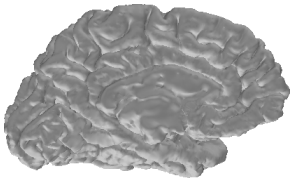
The minimum norm estimation method, when used on a realistic head model, provides an estimated current value for each voxel on the BEM at each time point. Three ROIs were defined based on their anatomical locations and proximity to subcortical structures (Figure 17). The estimated magnetic field strength of all voxels within each region was averaged together to obtain a single field strength value for



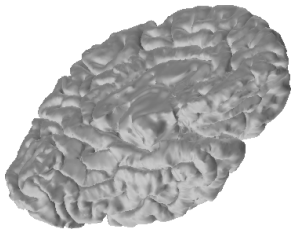
(a) (b)

Figure 16: Archetypal results of dipole localization in the (a) posterior cingulate and (b) parietal regions. Each subfigure contains three views of the same dipoles; coronal, sagittal, and transverse. On each view the plane of the other views are depicted as thin white lines, intersecting at a point visible on all three images. Note that the multiple dipoles present in each image are the result of fitting a single dipole to the magnetic field patterns at multiple timepoints.

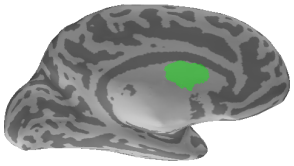
each ROI at each time point. The ROI time courses averaged across subjects can be seen in Figure 18. A subset of the averaged time courses were then used to construct a vector autoregressive (VAR) model. In order to maximize the relevant information while simultaneously minimizing the total information in the model, we defined our subset as 0 ms (button press) to 450 ms (post-button press). We uses the Akiaike Information Criterion to determine the appropriate lag for our VAR, which came out to 42 milliseconds. With this lag, we examined the relationships between the three



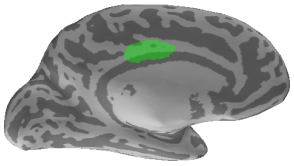
(a) BEM reconstruction of white matter, medial view



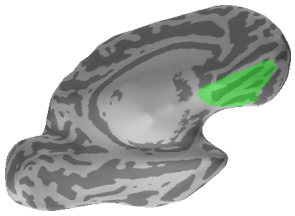
(b) BEM reconstruction of white matter, ventral view



(c) Inflated BEM representation of cortex. Green voxels were labeled by experimenter as striatum.



(d) Inflated BEM representation of cortex. Green voxels were labeled by experimenter as cingulate cortex.



(e) Inflated BEM representation of cortex. Green voxels were labeled by experimenter as OFC.

Figure 17: Image of white matter and inflated surface reconstruction from Subject 1. The ROIs were identified using this mesh and can be seen in subfigures (c), (d), and (e). Once can clearly observe a number of regions on the non-inflated meshes, including the head of the caudate and cingulate cortex. ROIs on other subjects were identified in a similar fashion. Due to limitations of the MNE software package, the ROI blocks can only be overlaid on the inflated cortex representation and not the white matter representation.

traces mentioned above. The Granger Causality analysis output is detailed in Table 4. These results suggest a significant ($p < 0.05$) Granger-causal relationship between the Striatum and OFC, and there is a non-significant trend towards a Granger-causal relationship between the cingulate cortex and OFC (Figure 20).

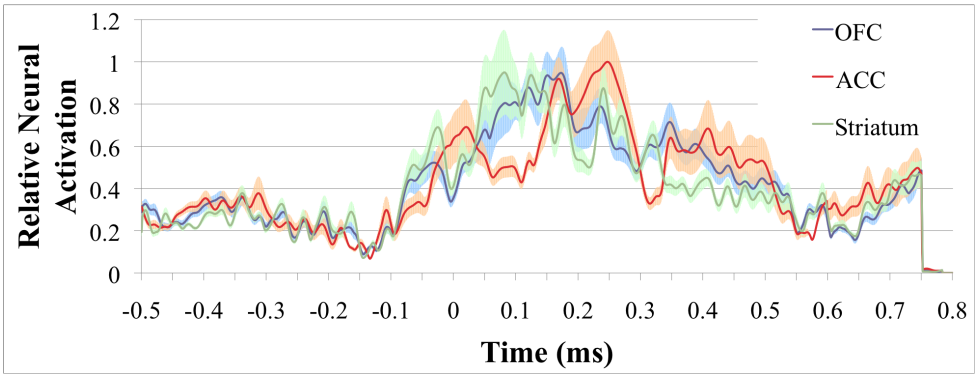


Figure 18: Activity traces for all three regions averaged across subjects. Note differential activity onset at 50 ms, with OFC and striatal activity increasing significantly relative to ACC. Striatal activity first peaks at 75 ms, then again at 120 ms, 170 ms, and later at 240 ms. The OFC reaches a single, broad peak around 170 ms, and continues a slow decline from there on out. The ACC peaks twice, once at 180 ms, and again later at 250 ms.

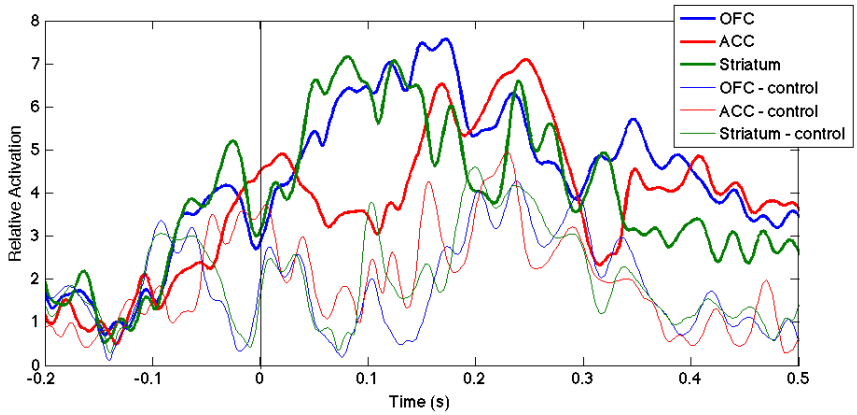


Figure 19: Activity traces for all three regions averaged across subjects, as well as the data from the control subject. With respect to the control subject, note significant differences between the datasets, particularly at 50-100 ms, and 300+ ms. This difference, coupled with the previously shown difference in raw data, is suggestive if a difference between neural activity in control and normal conditions.

Table 4: F-values from the Granger Causality test between the indicated regions. Given significance at F≤0.05, we conclude from these results that there is a significant Striatum→Cingulate and OFC→Cingulate Granger-causal relationship, as well as a Striatum→OFC Granger-causal relationship.

Region of interest	Possible Feedforward Region	Granger F-value
Cingulate Cortex	OFC	0.0014
	Striatum	0.0000
Striatum	Cingulate	0.9665
	OFC	0.9597
Orbitofrontal Cortex	Cingulate	0.4896
	Striatum	0.0047

For the control task, we performed minimum norm estimation for the same three regions described above, and obtained the activation patterns depicted in Figure 19. An F-test was used to determine whether there was a significant difference between the 0-450 ms section of the normal and control plots for each ROI. It was determined that all three ROIs exhibited significantly different traces in the normal condition relative to the control condition ($p > 0.001$). Dipole localization was not performed for the control task.

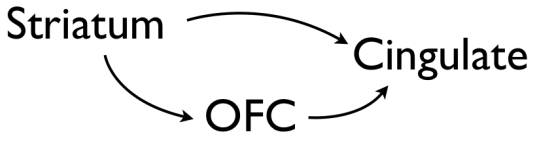


Figure 20: A pictorial view of the results depicted in Table 4. Only significant connections ($p < 0.05$) are listed on the map.

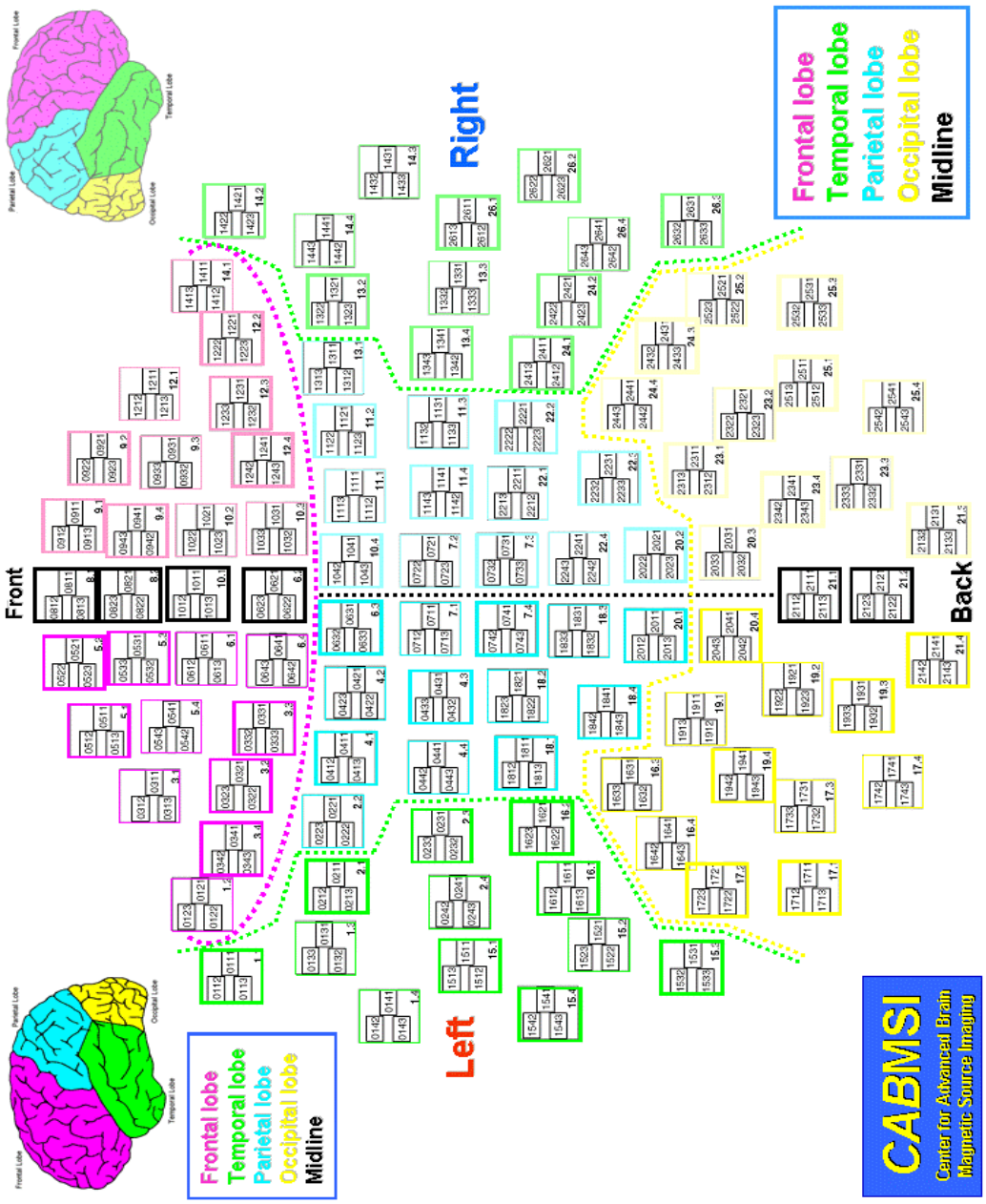


Figure 21: Schematic depicting sensor layout of the Elekta NeuroMag 306-sensor system on a typical human head. Each colored box represents three MEG sensors; two gradiometers and one magnetometer. The colored regions approximately represent the brain regions depicted in the figure, with demarcations as labeled. Figure obtained from CABMSI (UPMC, Pittsburgh, PA). Used with permission.

7.0 DISCUSSION

7.1 CHOICE OF PARADIGM

We chose to use this particular paradigm due to its serving two purposes; the subject will have to make a rewards-related decision each round, and the subject will be anticipating receiving a reward or punishment each round. In the original paper, Liu et al. [99] found significant differential activation of the striatal nucleus between trials in which the subject made the objectively "correct" choice—either through betting and winning or not betting and subsequently losing the die roll—versus when the subject made the objectively "incorrect" choice—either through betting and losing or not betting and subsequently winning.

It is worth noting that within the control paradigm the subject would be able to determine whether they had won or lost, despite an uninformative feedback screen, based on the ante screen shown at the beginning of the next round. However, as this is temporally distant from the time points we examined, this should not interfere with our signal of interest.

7.2 FUNCTION MAGNETIC RESONANCE IMAGING RESULTS

The main advantage of functional magnetic resonance imaging over MEG lies in the high spatial accuracy of the fMRI scanner throughout the entire head. As such, we used fMRI to verify that our paradigm did indeed elicit the activity of interest. In his paper, Liu et al. [99] found significant activation of the Striatum and OFC (medial and lateral), as well as the ACC, anterior insula, and temporal pole. Our replication of his experiment found similar regions of activation in the striatum, medial OFC, and ACC. This agreement between the two experiments suggests that this experiment represents a viable method for robustly activating the regions of interest, and in future experiments we can assume that given the identical task with identical instructions, identical neural activity would be elicited.

7.2.1 Activation differences between similar gambling paradigms in the literature

Surprisingly, we also found significant activity in some regions unreported by Liu, including highly robust activation in the PPC and dorsolateral PFC. Additionally, Liu reported finding robust activation of the lateral OFC and regions of the temporal cortex, areas in which we found no activity. As reported by Liu in the introduction to his paper, there are highly mixed findings regarding ROI activation during reward processing. For example, our results are in agreement with O'Doherty et al. [109] and Ullsperger and von Cramon [155], who cites lateral OFC activation during reward paradigms. However, a number of papers [104, 110], including the one on which this experiment is based [99], do not cite lateral frontal activation.

One oft-cited explanation is that discrepancies between activation patterns are due to paradigm differences. However, given that our study was in all respects identical to Liu et al. [99], we may be forced towards a different conclusion. One possibility is that there may exist differences in subject response to reward, and that some components of the learning and reward pathways may not be consistent between all subjects. This conclusions does have a precedent in the literature, both from "nature" and "nurture" standpoints. Quality of nutrition throughout life has been shown to have long-term effects on neural pathways [32]. Genetic factors have also been shown to have a significant effect on susceptibility to drug addiction [88], and some factors have been implicated in the neurogenesis of neural pathways [154]. Inasmuch as the learning and rewards system is closely related to the networks involved in addiction [80], postulating that the rewards system also maintains a genetic component is not a far leap. It is worth noting that the primary author of the gambling paper used in this experiment Liu et al. [99] did not consider the differences in activation to be of significant consequence, largely because there exists a precedence in the literature for activity in all mentioned regions during gambling tasks [98].

Minor inconsistencies notwithstanding, similar regions were identified as active in both the present and previous fMRI studies. As such, it would appear that the findings in the present study for the most part are in agreement with literature findings. Given that the paradigm used in the MEG is identical to that of the fMRI, the MEG study can be analyzed under the assumption that similar regions were activated in both the fMRI and MEG studies.

7.2.2 Differences between BOLD activation patterns and MEG activity

When examining results from fMRI and attempting to draw parallels to MEG activation, the source generators themselves need to be considered. fMRI activity stems from local BOLD responses, which itself arises from the increased blood flow to a region of high metabolic activity [24, 100]. MEG activity arises from the simultaneous activation of approximately 50,000 similarly-aligned neurons [132]. As such, the true source locations for each modality should be different. Ideally, the spatial differences between the two optimal solutions should be small, representing only the distance between the neuron generating the magnetic signal and the blood vessel providing oxygen to the neuron. However, as the two modalities are in fact observing different activity, it is prudent to note the difference in activity generators.

7.3 MAGNETOENCEPHALOGRAPHY RESULTS

7.3.1 Evidence of recruitment of relevant regions

One of the primary aims of this experiment was merely to see whether the activation patterns observed by Liu and colleagues in their fMRI experiment could be replicated in an MEG environment. Particularly, we wanted to examine whether we could detect both the superficial regions (visual activity, premotor and motor activity, parietal) as well as deeper regions (orbitofrontal cortex, striatal regions, cingulate). We were able to elicit strong visual and motor signals, most noticeable in the undifferenced data, which demonstrated both that the MEG is able to obtain signals of interest from this study, as well as suggest that the subjects were attending to the experiment (Figure 11). This is most strongly supported by the timing of the

waveforms; motor activity is present before 0 ms mark, suggesting premotor and motor cortex activation. Visual activity is observed 100 ms after the onset of the feedback screen, followed by parietal activity. This pattern suggests the observing of the visual scene followed by processing of that information. That we were able to elicit these signals is not in and of itself a significant finding, as many studies have demonstrated the MEG's ability to detect strong superficial signals [79, 101, 135]. Rather, this demonstrates that the subject is attending to the subject and that the information processing stream is active.

7.3.2 Efficacy of exSSS signal processing technique

The differenced dataset contained no notable activity prior to processing by the exSSS algorithm, whereas after processing distinct activity patterns can be seen(see Figures 14 and 15). We can draw two conclusions from this observation. Firstly, it seems that very little differential cortical activity exists in this experiment. This conclusion is supported by the Liu paper, which found virtually no differential superficial activity. Secondly, the exSSS method seems to successfully extract neural signals which were not present in the original dataset. We suggest that these signals are deep in origin, for reasons described in detail in the following section.

Prior to discussing the evidence for presence of a deep signal in our waveforms, we must address a more fundamental question, namely, is the proposed signal generator visible at all to the MEG sensors?As stated earlier, MEG sensors are able to detect approximately 50,000 identically aligned neurons firing in synchrony. If the proposed neural source does not contain such neural morphology, then no amount of signal processing will extract the neural signal from the dataset; it simply is not present, on account of its not being detectable by the sensors. In our case, we are primarily

attempting to examine three regions; OFC, ACC, and striatal cortex. In each of these regions, there exists a neural substrate for MEG-visible source generators. Both the OFC and ACC possesses large numbers of radially-oriented pyramidal neurons, which have been shown to be MEG source generators [64, 116]. The striatum has a number of highly ordered pathways both to, from, and within the striatum, many in many instances function as dipole generators [17]. As such, on a fundamental level, activity within these regions should be visible to the MEG.

A number of factors influenced our decision to label this signal as originating from deeper regions. Firstly, note the contrast between the general uniformity of the non-processed signal and the notable activity of the processed data. Statistical significance aside, based on a purely visual inspection, the exSSS method visually appears to have successfully extracted neural activity not otherwise present in the MEG traces. The statistical measure mentioned above provides support for this. As discussed in the background (Section 3.2.3), the exSSS method filters the data such that deep signals are amplified and superficial signals attenuated. The simple fact that signals that were not visible in the data were made visible after processing by a method which amplifies deeper signals provides support that the latent signal is deep in origin.

Secondly, the signature appears spread out across the coronal midline of the head. Cortical signals are often strongest in the sensors most directly superior to the region from where the signals originate, and while visible in neighboring signals, decrease exponentially in strength as a factor of distance. This signal appears strongest in the sensors along the rostral-caudal midline and quickly weakens as one examines sensors closer to the frontal and occipital poles. This magnetic field pattern suggests a source located close to the origin of the head (i.e., basal ganglia) with the poles oriented horizontally.

Thirdly, the timing of the activity across subjects (in the 150-300 ms range) matches the time one would expect for the reward circuitry to activate. Visual signal in our task was consistently observed across subjects between 100-115 ms. Activity which localized to the prefrontal regions (discussed below) was observed between 220-270 ms. As such, a 170 ms signal appearing broadly across the midline would fit well with the hypothesis mentioned earlier that information originates in the cortex appropriate to the modality (in our case, the occipital cortex for a visual stimulus), travels to the basal ganglia for processing and integration, and subsequently travels to the frontal regions.

Additionally, information traveling between the VTA and prefrontal regions have been reported in the literature as requiring between 80-120 ms, which was seen as proof that the functional pathway between the VTA and frontal cortex is polysynaptic [50, 145]. The delay we see is slightly shorter than this at 40 ms, which fits very well with this reported finding, given that the axonal pathway between the VTA and prefrontal regions is known to pass through the striatum. Indeed, this would appear to support the hypothesis that the functional rewards processing pathway includes the striatal regions. It is worth mentioning that due to its small size and large distance from the MEG sensors, we would not expect to observe VTA activity in the MEG.

Interestingly, the order of application of the various signal processing techniques is highly important. The traces in Figure 15 were only obtainable after processing in the order mentioned here—data cleaning (bandpass, SSP, averaging), SSS, differencing, exSSS. If the data is not differenced, or if the data is differenced after exSSS has been applied, the deep activity was not observed. This is likely due to that which was mentioned at the beginning of this section; the cleaner the input, the more accurate the output. By first differencing the conditions, we remove all extrinsic

signals, without affecting the latent deep signal. In this cleaned dataset the exSSS method is most effective in extracting the deep activity.

7.3.3 Dipole localization

Dipole analysis in this study was constrained by the fact that the activated region is spatially broad, and standard dipole localization techniques always will only yield a single focal point of activation. As such, researchers commonly use only a subset of the available EEG/MEG sensors when performing dipole localization. In our case, since the deep dipole is visible only across the entire head, we need to use the full set of MEG sensors when localizing, thus significantly increasing the likelihood that more than a one source will be present in the sensory array when conducting the dipole localization.

The presence of dipoles in the parietal region is not unexpected, given that parietal activity has often been observed in rewards-related experiments. The midline dipole near the cingulate cortex is somewhat suspect, particularly given that much of the activation observed in this task occurs bilaterally. As such, the likelihood that this dipole actually consists of two more lateral dipoles is quite significant. Additionally, alternate source localization analysis methods (discussed below) found a number of concurrently activated sources both anterior and posterior to the suggested cingulate dipole. Given these considerations, the deeper dipole is considered suspect and may likely not represent the actual locus of neural activation.

It should be noted that the activation patterns found in the fMRI component of this study can be used as a template of activation for dipole localization. (This is less true for the MNE solutions described below (Section 7.3.4), as the MNE results are rendered on a 2-D surface, making fMRI-MNE solution comparison more difficult.)

Generally, fMRI localization is more accurate than inverse localization methods. This is particularly true in our case given the use of the exSSS algorithm. In our analysis, all dipoles localized close to regions active in fMRI, but very few dipoles were overlapping with their fMRI counterparts. The parietal dipoles are slightly anterior to the location reported in fMRI, and the dipoles localizing to the caudate tail are slightly posterior to the fMRI striatal activation.

7.3.4 Minimum norm estimation

The localization methods described earlier provide an approximate current map across the surface of the cortex. Hence, unlike dipole localization, the current is not spoken to have "localized" to a particular region, but rather viewed as a probability map that the field originated from given section or sections of cortex. As such, ROI analysis consist of identifying the ROI on the BEM mesh and examining MEG activation—as identified by a given localization method—within that region. Since MEG, unlike fMRI, records only neural activity and no structural information, cross-subject averaging is a more difficult problem, and for this reason MEG activation data was examined on a single-subject basis.

Distinct inverse solutions can be found for each set of magnetic field measurements. Neural time course reconstruction consists of determining this solution for a given set of voxels for every time point. The time course for a given ROI can be determined by determining the activation time course for each voxel within the ROI and subsequently averaging the time courses across all voxels.

As described above, we examined the activity in three ROIs related to rewards processing—the striatum, the cingulate cortex, and the OFC. These regions were identified on the BEM mesh constructed from the structural MRI images and used

to determine neural activation. The ROIs themselves are depicted in Figure 17. Due to software limitations, the ROIs cannot be directly overlaid on the anatomically accurate white matter BEM mesh, and hence are depicted on an inflated representation of the brain.

Activity patterns for each of the three subjects were highly variable in time, with the peak activity for each individual regions varying by more than 50 ms between subjects. However, the temporal pattern of activity was generally constant between subjects, and the averaged activity traces (Figure 18) show a distinct, significant progression in peak activity from region to region.

One interesting observations was that for almost all subjects the reconstructed cingulate activity consistently appeared farther posterior than would be expected. This was determined to be an artifact of the exSSS method. The reasoning for this conclusion is that the effect of the exSSS method is achieved by manipulating the leadfield-like coefficients that make up the α coefficients (see Section 3.2.3). Recall that the leadfield matrix dictates the relationship between the source and it's field strength at the detector (see Section 3.2.1). Given this relationship, by modifying these leadfields we can "strengthen" the source signal at the detector. However, the source localization methods used for source reconstruction are, so to speak, "unaware" of this leadfield modification. As such, in order to correctly localize a signal which has been enhanced using exSSS, the localization method would need to correctly account for the re-weighing of the leadfield. As of yet no technique has been developed to compensate for this problem.

Another interesting finding across subjects was the lack of parietal signal present in the MNE solution of the exSSS processed datasets. This is to be expected, as the parietal signal is largely superficial, and should be minimized by the exSSS processing algorithm.

Each of the three regions was activated as early as 100 ms prior to button press, suggesting a possible recruitment of the reward circuitry during anticipation of the feedback. Anticipation signals are well-documented in the literature; subjects expecting to observe reward-related cues tend to display activity in the rewards networks [16, 42]. This observation lends more support to the efficacy of the exSSS method; the activity traces behave as would be predicted by the literature.

7.3.5 Normal vs. Control solutions

Minimum norm solutions to the inverse problem found significant differences between the activation patterns of the processed and unprocessed data (Figure 18). These findings support the conclusions stated above in Section 7.3.2. Note, however, that the processing of the ambiguous feedback activated the same ROIs as were activated during the normal task. This phenomenon has been previously reported in the literature. Using a gambling-style experiment, researchers presented subjects with two types of trials; a well-defined but risky decision, or an ambiguous decision. Using fMRI, the researchers found that while both trial types recruited similar regions, activity levels across the network were significantly more evenly distributed during the ambiguous condition as compared to well-defined condition [142].

7.3.6 Time series analysis: Granger Causality

The activation curves depicted in Figure 18 are suggestive of a feedforward network within the rewards network. The presence of this network, as well as its connectivity, has been demonstrated in the literature from a variety of angles. Shi [145] observed a polysynaptic connection between striatal and prefrontal neurons through correlated low-frequency oscillatory firing patterns in the two regions. Haber et al. [62] used

anterograde tracers injected into the rewards system to construct a three-dimensional map of anatomical connectivity. Numerous researchers have used functional imaging to observe various components of the network in a coactivated state [141, 166]. However, this study marks the first observation of system-wide activity with millisecond-level temporal resolution [81]. Using the source reconstruction techniques mentioned above we were able to reconstruct the activation traces of the ROIs. By constructing a VAR using the reconstructed activity traces, we can use Granger causality to determine whether a given region significantly affect the activation of other regions.

As depicted in Figure 20, we obtained a possible feedforward network connecting the Striatum, OFC, and cingulate cortex. This connectivity pattern supports our hypothesis that the striatum is one of the first recruited regions of the rewards network, followed by the OFC (see introductory text of Chapter 5). While it may appear that this is in disagreement with the preponderance of literature suggesting that prefrontal activity drives striatal activity (as discussed in the Chapter 5 intro), this is not the case. Recall that the data traces used in the Granger causality analysis were the result of, among other things, differencing between conditions. This particular processing step significantly affects the inferences which can be drawn from this dataset. There may be a number of neural processing steps occurring before the temporal region of interest (approximately 150-300 ms post-button press, as described in Chapters 5 and 6) which are invisible in the processed dataset. As such, we cannot conclude that the striatal activity broadly drives OFC activity; rather, we can only state that this experiment provides evidence that the neural processes underlying recognition of valence in reward follow a striatal-OFC pattern. Additionally, note further that this network pattern is likely highly task-dependent; in our task, very little long-term learning occurred, since each trial presented an independent, arbitrary choice to the subject ("bet or bank?").

In attempting to detect correlations between the different ROIs, we limited our analysis to milliseconds 150-320 within the broader -500 ms to 800 ms time series (refer to Figure 10). The reason for this limitation was twofold. Firstly, the signal of interest existed almost exclusively between milliseconds 150 and 320. Both raw MEG traces and reconstructed signals outside that range—particularly before -100 ms and after 400 ms—only showed noise-like random fluctuations in the traces. The second reason is that the construction of a VAR depends critically on the underlying time series. Including extraneous or unnecessary time points in the regression not only needlessly complicates the model, but also can adversely affect the quality of the regression. As such, we aimed to keep the series as small as possible while still including the entirety of the crucial components of the signal.

7.4 DATA CONSIDERATIONS

There are a number of issues that arose during data processing. These issues relate to the exSSS data processing method, the inverse localization methods, and the interaction between these two signal processing techniques. It is important that the reader be aware of these issues and how they affected the interpretation of data.

7.4.1 Localization Accuracy

One of the most crucial aspect of any neuroimaging technique is its ability to accurately localize neural activity to the exact region where that activity took place. Almost all MEG localization methods rely on the leadfield matrix to determine where the relationship between the recorded magnetic fields and the source being localized.

As described in the introduction (Section 3.2.2), exSSS utilizes a leadfield-like representation of the α_{lm} component of the recorded magnetic field (Equation 3.13) to enhance the signal. While the leadfield matrix itself is not manipulated, by changing the representation of the sources in this manner, we may affect the accuracy of the localization process, since the leadfield matrix may no longer represent a true solution to the forward problem [113].

This problem can be seen in our study when comparing the locus of activation in the cingulate region between the fMRI and MEG studies. In both the present fMRI study as well as in the reports by Liu et al. [99], the anterior cingulate region was activated during this task. However, as outlined in Figure 17, the cingulate cortex which appeared activated during the gambling task is significantly posterior to ROI from the fMRI studies. Given that the anterior cingulate is active during the task, given that ACC activation is often observed during behavioral tasks of this type (see Section 2.1.3), and given activation a few centimeters posterior to the expected region, we theorize that the observed activation is likely present in the anterior cingulate region, and any localization errors are due to one or more steps in our processing stream.

Recall that the data processing stream consisted of the following steps (see Section 6.1.2.1): 1-40Hz bandpass filter, averaging, noise reduction (SSS and SSP), and exSSS, followed by source localization. Of these, all but exSSS have been used extensively in the literature, and have been shown not to affect source localization accuracy [151]. As such, the novel exSSS method may be at fault.

From a theoretical standpoint, the underlying cause of the proposed location bias may be the beamspace transform performed during exSSS processing. Recall that the beamspace method enables the imposition of *a priori* information on our solution space [128]. This has the effect of modifying the magnetic field $\mathbf{b}$, as de-

scribed in Section 3.2.3. Normally, this is not a problem, since our goal in applying the beamspace transform is to alter the localization results. By modifying the **b** field, we hope to obtain a more compact, lower dimension dataset, which will produce more accurate localized sources. Additionally, the leadfields themselves are not being manipulated, only the recorded magnetic signal **b**. Thus, the mapping relationship between the sensors and the sources—wherever they may be located—has been preserved. In our algorithm, though, we are manipulating the location of the sources through a beamspace transform while simultaneously manipulating the leadfield-like matrix α_{lm}. It is possible that the combined effect of the beamspace modification and α_{lm} term modification can result in a change in the source-sensor mapping, thus resulting in incorrect source localization. However, this is conjecture, and significant further testing will be required to determine the cause of localization inaccuracies.

7.4.2 Deep source representation on BEM mesh

Minimum-norm source localization attempts to minimize the error between the field generated by a calculated set of possible source locations and the true actual recorded magnetic field. Due to the computational difficulty posed by attempting to calculate the inverse solution for a dense 3-D mesh, implementations of source localization algorithms fit the data to a two-dimensional surface representing the brain cortex [31, 47, 63]. This surface can be created through two methods. The first is to assume a spherical head model, and model the source as though the brain is a perfect sphere. This technique is sufficient for most cortical source locations [91]. Alternatively, one can construct a boundary element model from either a template or static MRI images of the subject's head and treat each mesh element as a separate discrete possible

source location. Importantly, using this second technique, separate surface meshes can be constructed for each hemisphere. However, in both of these methods, no subcortical structures are included in the model.

This very basic representation of the brain can lead to localization inaccuracies when attempting to localize subcortical sources. Inasmuch as the activity generated by the source was still recorded by the sensors, but the true region from which the signal originated does not exist in our solution space, the source will be fit to a different region or set of regions on the cortex, with increased localization error.

When using a spherical head model, the error introduced in this manner is significant, and great care should be taken to ensure that no deep sources are present when using this model for source localization. However, using a realistic BEM generated from MRI images—which, importantly, contain separate surface meshes for each hemisphere, including the medial side—can significantly reduce the magnitude of introduced error. This is due to the tendency for the model, in attempting to explain the recorded source distribution, to project the activation pattern to the region on the solution space closest to the true spatial location of the source. As such, given a realistic 3-D surface model of the lateral, medial, dorsal and ventral surfaces of each hemisphere, even deep sources are located physically close to a given patch of source and distant from others, allowing the error introduced through any projection to be minimized.

This problem is a foundational issue with the current source localization technology, and will only be solved when a minimum-norm method is developed that is compatible with realistic 3-D BEM meshes containing subcortical structures. However, until such a package is commercially available for research, care should be taken to ensure that any study which uses this type of analysis takes sufficient measures to ensure the absence of deep structure activation.

8.0 CONCLUSION

The work presented here describes the use of magnetoencephalography to obtain high temporal resolution waveforms describing the activation of a number of both deep and superficial brain regions during a gambling task. Using fMRI, we were able to verify the literature findings that this gambling paradigm serves to activate rewards-related regions in the brain, including the medial orbitofrontal cortex, striatum, and anterior cingulate. MEG recordings of neural activity acquired during the same task appeared to contain only noise, but after processing via the exSSS method we uncovered signals appearing to originate from subcortical brain regions. By applying inverse processing techniques, we obtained source waveforms for each of these regions. These waveforms were used to construct a vector autoregressive matrix, which in turn was used to describe Granger-causal relationships between each region.

We described a number of novel advances in the field of human rewards processing, as well as in the broader field of neural imaging. To our knowledge, the temporal functional relationship between the striatum, orbitofrontal cortex, and cingulate cortex has never been described in a human subject. This study is the first to describe the activation map that may exist between these regions.

Additionally, this study marks the first time MEG has been used to examine deep activity without requiring either *a priori* assumptions as to the expected wave-

forms (i.e., matched filter designs) or the construction of a multi-surface boundary element model with conductance information. Rather, we were able to extract deep activity by taking advantage of intrinsic properties of magnetic fields and the MEG environment using nothing more than the MEG leadfield matrix. This stems from the successful application of the SSS and exSSS algorithms to the applied research problem at hand. This marks a radical departure from traditional signal processing schemes, and may mark the development of the MEG as a platform technology for noninvasively obtaining high-temporal resolution waveforms of deep neural regions.

BIBLIOGRAPHY

[1] Abler, B., A. Roebroeck, R. Goebel, A. Hse, C. Schnfeldt-Lecuona, G. Hole, and H. Walter (2006). Investigating directed influences between activated brain areas in a motor-response task using fmri. Magnetic Resonance Imaging 24(2), 181 – 185.

[2] Akitsuki, Y., M. Sugiura, J. Watanabe, K. Yamashita, Y. Sassa, S. Awata, H. Matsuoka, Y. Maeda, Y. Matsue, H. Fukuda, and R. Kawashima (2003, Aug). Context-dependent cortical activation in response to financial reward and penalty: an event-related fmri study. Neuroimage 19(4), 1674–1685.

[3] Alexander, G., M. DeLong, and P. Strick (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 9, 357–81. 0147-006X (Print) Journal Article Review.

[4] Arfken, G. (1986). Mathematical Methods for Physicists (third ed.). Academic Press.

[5] Astolfi, L., F. Cincotti, D. Mattia, F. De Vico Fallani, A. Tocci, A. Colosimo, S. Salinari, M. G. Marciani, W. Hesse, H. Witte, M. Ursino, M. Zavaglia, and F. Babiloni (2008). Tracking the time-varying cortical connectivity patterns by adaptive multivariate estimators. IEEE Trans Biomed Eng 55(3), 902–913.

[6] Attwell, D. and C. Iadecola (2002). The neural basis of functional brain imaging signals. Trends Neurosci 25(12), 621–625.

[7] Baccala, Luiz, A. and K. Sameshima (2001). Partial directed coherence: a new concept in neural structure determination. Biological Cybernetics 84(6), 463–474.

[8] Barlow, J. S. (1993). The Electroencephalogram. MIT Press, Cambridge, MA.

[9] Baxter, M. G. and E. A. Murray (2002). The amygdala and reward. Nature Reviews Neuroscience 3(7), 563–573.

[10] Berndt, E. R. (1990). The Practice of Econometrics. Addison-Wesley Publishing Company, Reading, Massachusetts.

[11] Boettiger, C. and M. D'Esposito (2005, Mar). Frontal networks for learning and executing arbitrary stimulus-response associations. J Neurosci 25(10), 2723–32. 1529-2401 (Electronic) Journal Article.

[12] Botvinick, M. M. (2007, Dec). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. Cogn Affect Behav Neurosci 7(4), 356–366.

[13] Botvinick, M. M., T. S. Braver, D. M. Barch, C. S. Carter, and J. D. Cohen (2001, Jul). Conflict monitoring and cognitive control. Psychol Rev 108(3), 624–652.

[14] Botvinick, M. M., J. D. Cohen, and C. S. Carter (2004). Conflict monitoring and anterior cingulate cortex: an update. Trends in Cognitive Sciences 8(12), 539 – 546.

[15] Box, G. E. P. and G. Jenkins (1990). Time Series Analysis, Forecasting and Control. Holden-Day, Incorporated.

[16] Breiter, H., I. Aharon, D. Kahneman, A. Dale, and P. Shizgal (2001, May). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. Neuron 30(2), 619–39. 0896-6273 (Print) Journal Article.

[17] Brodal, P. (1992). The Central Nervous System: Structure and Function. Oxford University Press, New York.

[18] Burgess, K. and B. Van Veen (1996, Apr). Subspace-based adaptive generalized likelihood ratio detection. Signal Processing, IEEE Transactions on 44(4), 912–927.

[19] Cador, M., T. Robbins, and B. Everitt (1989). Involvement of the amygdala in stimulus-reward associations: interaction with the ventral striatum. Neuroscience 30(1), 77–86.

[20] Cardinal, R. N., J. A. Parkinson, J. Hall, and B. J. Everitt (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. Neurosci Biobehav Rev 26(3), 321–352.

[21] Carr, D. and S. Sesack (2000, May). Projections from the rat prefrontal cortex to the ventral tegmental area: target specificity in the synaptic associations with mesoaccumbens and mesocortical neurons. J Neurosci 20(10), 3864–73.

[22] Cavina-Pratesi, C., K. F. Valyear, J. C. Culham, S. Kohler, S. S. Obhi, C. A. Marzi, and M. A. Goodale (2006). Dissociating Arbitrary Stimulus-Response Mapping from Movement Planning during Preparatory Period: Evidence from Event-Related Functional Magnetic Resonance Imaging. J. Neurosci. 26(10), 2704–2713.

[23] Chang, J. Y., S. F. Sawyer, J. M. Paris, A. Kirillov, and D. J. Woodward (1997, May). Single neuronal responses in medial prefrontal cortex during cocaine self-administration in freely moving rats. Synapse 26(1), 22–35.

[24] Chen, W. and S. Ogawa (2000). Principles of BOLD Functional MRI, Chapter 10, pp. 103–113. Volume 1 of Moonen and Bandettini [102].

[25] Chudasama, Y., F. Passetti, S. E. V. Rhodes, D. Lopian, A. Desai, and T. W. Robbins (2003). Dissociable aspects of performance on the 5-choice serial reaction time task following lesions of the dorsal anterior cingulate, infralimbic and orbitofrontal cortex in the rat: differential effects on selectivity, impulsivity and compulsivity. Behavioural Brain Research 146(1-2), 105 – 119. The Rodent Prefrontal Cortex.

[26] Cohen, M., A. Heller, and C. Ranganath (2005). Functional connectivity with anterior cingulate and orbitofrontal cortices during decision-making. Cognitive Brain Research 23(1), 61 – 70. Multiple Perspectives on Decision Making.

[27] Comon, P. (1994). Independent component analysis, a new concept? Signal Processing 36, 287–314.

[28] Congedo, M. (2006). Subspace projection filters for real-time brain electromagnetic imaging. IEEE Trans Biomed Eng 53(8), 1624–1634.

[29] Dale, A., A. Liu, B. Fischl, R. Buckner, J. Belliveau, J. Lewine, and E. Halgren (2000, Apr). Dynamic statistical parametric mapping: combining fmri and meg for high-resolution imaging of cortical activity. Neuron 26(1), 55–67.

[30] Dale, A. and M. Sereno (1993). Improved localization of cortical activity by combining eeg and meg with mri cortical surface reconstruction: a linear approach. J Cogn Neurosci 5, 162–176.

[31] Dale, A. M., B. Fischl, and M. I. Sereno (1999, Feb). Cortical surface-based analysis. i. segmentation and surface reconstruction. Neuroimage 9(2), 179–194.

[32] Dauncey, M. J. (2009, Aug). New insights into nutrition and cognitive neuroscience. Proc Nutr Soc, 1–8.

[33] de Zwart, J. A., A. C. Silva, P. van Gelderen, P. Kellman, M. Fukunaga, R. Chu, A. P. Koretsky, J. A. Frank, and J. H. Duyn (2005). Temporal dynamics of the bold fmri impulse response. Neuroimage 24(3), 667–677.

[34] Delgado, M., H. Locke, V. Stenger, and J. Fiez (2003, Mar). Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. Cogn Affect Behav Neurosci 3(1), 27–38.

[35] Delgado, M., L. Nystrom, C. Fissell, D. Noll, and J. Fiez (2000, Dec). Tracking the hemodynamic responses to reward and punishment in the striatum. J Neurophysiol 84(6), 3072–7. 0022-3077 (Print) Clinical Trial Journal Article.

[36] Di Chiara, G. (1995, May). The role of dopamine in drug abuse viewed from the perspective of its role in motivation. Drug Alcohol Depend 38(2), 95–137. 0376-8716 (Print) Journal Article Review.

[37] Di Chiara, G. and V. Bassareo (2007). Reward system and addiction: what dopamine does and doesn't do. Curr Opin Pharmacol 7(1), 69–76.

[38] Ding, M., Y. Chen, and S. L. Bressler (2006, Aug). Granger causality: Basic theory and application to neuroscience.

[39] Dirnagl, U., U. Lindauer, and A. Villringer (1993). Role of nitric oxide in the coupling of cerebral blood flow to neuronal activation in rats. Neurosci Lett 149(1), 43–46.

[40] Eichler, M. (2005). A graphical approach for evaluating effective connectivity in neural systems. Philos Trans R Soc Lond B Biol Sci 360(1457), 953–967.

[41] Elliott, R., K. Friston, and R. Dolan (2000, Aug). Dissociable neural responses in human reward systems. J Neurosci 20(16), 6159–65. 0270-6474 (Print) Journal Article.

[42] Ernst, M., E. E. Nelson, E. B. McClure, C. S. Monk, S. Munson, N. Eshel, E. Zarahn, E. Leibenluft, A. Zametkin, K. Towbin, J. Blair, D. Charney, and D. S. Pine (2004). Choice selection and reward anticipation: an fmri study. Neuropsychologia 42(12), 1585–1597.

[43] Eslinger, P. and A. Damasio (1985, Dec). Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient evr. Neurology 35(12), 1731–41.

[44] Fallon, J. H. and R. Y. Moore (1978). Catecholamine innervation of the basal forebrain. iv. topography of the dopamine projection to the basal forebrain and neostriatum. J Comp Neurol 180(3), 545–580.

[45] Fields, H., G. Hjelmstad, E. Margolis, and S. Nicola (2007). Ventral tegmental area neurons in learned appetitive behavior and positive reinforcement. Annu Rev Neurosci 30, 289–316.

[46] Fisch, B. J. and R. Spehlmann (1999). EEG Primer: Basic Principles of Digital and Analog EEG (third ed.). Elsevier.

[47] Fischl, B., M. I. Sereno, and A. M. Dale (1999, Feb). Cortical surface-based analysis. ii: Inflation, flattening, and a surface-based coordinate system. Neuroimage 9(2), 195–207.

[48] Fuster, J. M. (1997). The Prefrontal Cortex: Anatomy, physiology, and neuropsychology of the frontal lobe. Lippincott-Raven, Philadelphia.

[49] Gao, G., X. Wang, S. He, W. Li, Q. Wang, Q. Liang, Y. Zhao, F. Hou, L. Chen, and A. Li (2003). Clinical study for alleviating opiate drug psychological dependence by a method of ablating the nucleus accumbens with stereotactic surgery. Stereotact Funct Neurosurg 81(1-4), 96–104.

[50] Gao, M., C. Liu, S. Yang, G. Jin, B. Bunney, and W. Shi (2007, May). Functional coupling between the prefrontal cortex and dopamine neurons in the ventral tegmental area. J Neurosci 27(20), 5414–21.

[51] Gariano, R. and P. Groves (1988, Oct). Burst firing induced in midbrain dopamine neurons by stimulation of the medial prefrontal and anterior cingulate cortices. Brain Res 462(1), 194–8.

[52] Gazzaley, A., J. Rissman, and M. Desposito (2004, Dec). Functional connectivity during working memory maintenance. Cogn Affect Behav Neurosci 4(4), 580–99. 1530-7026 (Print) Journal Article.

[53] Gehring, W. J. and A. R. Willoughby (2002, Mar). The medial frontal cortex and the rapid processing of monetary gains and losses. Science 295(5563), 2279–2282.

[54] Geselowitz, D. (1967). On bioelectric potentials in an inhomogenous volume conductor. Biophysical Journal 7, 1–11.

[55] Goebel, R., F. Esposito, and E. Formisano (2006, May). Analysis of functional image analysis contest (fiac) data with brainvoyager qx: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. Hum Brain Mapp 27(5), 392–401.

[56] Goebel, R., A. Roebroeck, D.-S. Kim, and E. Formisano (2003). Investigating directed cortical interactions in time-resolved fmri data using vector autoregressive modeling and granger causality mapping. Magnetic Resonance Imaging 21(10), 1251 – 1261.

[57] Goldstein, R., D. Tomasi, S. Rajaram, L. Cottone, L. Zhang, T. Maloney, F. Telang, N. Alia-Klein, and N. Volkow (2007, Feb). Role of the anterior cingulate and medial orbitofrontal cortex in processing drug cues in cocaine addiction. Neuroscience 144(4), 1153–9.

[58] Gow, D. W. J., J. A. Segawa, S. P. Ahlfors, and F.-H. Lin (2008, Nov). Lexical influences on speech perception: a granger causality analysis of meg and eeg source estimates. Neuroimage 43(3), 614–623.

[59] Granger, C. J. (1969, 438). Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37(3), 424.

[60] Graybiel, A. (1990, Jul). Neurotransmitters and neuromodulators in the basal ganglia. Trends Neurosci 13(7), 244–54. 0166-2236 (Print) Journal Article Review.

[61] Gross, J. and A. A. Ioannides (1999). Linear transformations of data space in meg. Physics in Medicine and Biology 44(8), 2081–2097.

[62] Haber, S., K. Kim, P. Mailly, and R. Calzavara (2006, Aug). Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. J Neurosci 26(32), 8368–76.

[63] Hamalainen, M. (2006, Dec). MNE software User's Guide, Version 2.5 (2.5 ed.). MGH/HMS/MIT Athinoula A Martinos Center for Biomedical Imaging.

[64] Hamalainen, M., R. Hari, R. Ilmoniemi, J. Knuutila, and O. Lounasmaa (1993). Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain. Reviews of Modern Physics 65(2), 413–497.

[65] Hamalainen, M. and J. Sarvas (1989). Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. Biomedical Engineering, IEEE Transactions on 36(2), 165–171.

[66] Hamalainen, M. S. (1992). Magnetoencephalography: a tool for functional brain imaging. Brain Topogr 5(2), 95–102.

[67] Hatanaka, N., H. Tokuno, I. Hamada, M. Inase, Y. Ito, M. Imanishi, N. Hasegawa, T. Akazawa, A. Nambu, and M. Takada (2003, Jul). Thalamocortical and intracortical connections of monkey cingulate motor areas. J Comp Neurol 462(1), 121–138.

[68] He, F., H. Guan, Z. Zhao, X. Miao, Q. Zhou, L. Li, D. Huang, A. Liu, and D. Miao (2008). Evaluation of short-term psychological functions in opiate addicts after ablating the nucleus accumbens via stereotactic surgery. Stereotact Funct Neurosurg 86(5), 320–329.

[69] Herscovitch, P., J. Markham, and M. E. Raichle (1983). Brain blood flow measured with intravenous h2(15)o. i. theory and error analysis. J Nucl Med 24(9), 782–789.

[70] Hesse, W., E. Moller, M. Arnold, and B. Schack (2003). The use of time-variant eeg granger causality for inspecting directed interdependencies of neural assemblies. J Neurosci Methods 124(1), 27–44.

[71] Hikosaka, O., M. Sakamoto, and S. Usui (1989, Apr). Functional properties of monkey caudate neurons. iii. activities related to expectation of target and reward. J Neurophysiol 61(4), 814–32. 0022-3077 (Print) Journal Article.

[72] Hollerman, J. R., L. Tremblay, and W. Schultz (2000). Involvement of basal ganglia and orbitofrontal cortex in goal-directed behavior. In H. Uylings, G. van Eden, J. de Bruin, M. Feenstra, and C. Pennartz (Eds.), Cognition, emotion and autonomic responses: The integrative role of the prefrontal cortex and limbic structures, Volume 126 of Progress in Brain Research, pp. 193 – 215. Elsevier.

[73] Holroyd, C. B. and M. G. Coles (2002, Oct). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev 109(4), 679–709.

[74] Hyman, S., R. Malenka, and E. Nestler (2006). Neural mechanisms of addiction: The role of reward-related learning and memory. Annual Review of Neuroscience 29(1), 565–598.

[75] Ikeda, S. and K. Toyama (2000). Independent component analysis for noisy data–meg data analysis. Neural Netw 13(10), 1063–1074.

[76] Ito, R., J. Dalley, S. Howes, T. Robbins, and B. Everitt (2000, Oct). Dissociation in conditioned dopamine release in the nucleus accumbens core and shell in response to cocaine cues and during cocaine-seeking behavior in rats. J Neurosci 20(19), 7489–95.

[77] Ito, R., J. Dalley, T. Robbins, and B. Everitt (2002, Jul). Dopamine release in the dorsal striatum during cocaine-seeking behavior under the control of a drug-associated cue. J Neurosci 22(14), 6247–53.

[78] Jackson, J. D. (1999). Classical Electrodynamics (Third ed.). John Wiley and Sons, Inc.

[79] Jones, A. P., D. G. Hughes, D. S. Brettle, L. Robinson, J. R. Sykes, Q. Aziz, S. Hamdy, D. G. Thompson, S. W. Derbyshire, A. C. Chen, and A. K. Jones (1998, Feb). Experiences with functional magnetic resonance imaging at 1 tesla. Br J Radiol 71(842), 160–166.

[80] Kalivas, P. and N. Volkow (2005, Aug). The neural basis of addiction: A pathology of motivation and choice. Am J Psychiatry 162(8), 1403–1413.

[81] Kanal, E., T. Ozkurt, R. Sclabassi, and M. Sun (2009, April). Detecting granger causality in the corticostriatal learning and rewards network using meg. In Bioengineering Conference, 2009 IEEE 35th Annual Northeast, pp. 1–2.

[82] Kanal, E., T. E. Ozkurt, M. Sun, and R. J. Sclabassi (2008). Meg imaging of prefrontal and striatal activity during a gambling task. In Neuroscience Meeting Planner, Number 520.12, Washington, DC. Society for Neuroscience.

[83] Kandel, E. R., J. H. Schwartz, and T. M. Jessell (2000). Principles of Neural Science (Fourth ed.). McGraw-Hill.

[84] Karreman, M. and B. Moghaddam (1996, Feb). The prefrontal cortex regulates the basal release of dopamine in the limbic striatum: an effect mediated by ventral tegmental area. J Neurochem 66(2), 589–98.

[85] Kelley, A. E., S. L. Smith-Roe, and M. R. Holahan (1997). Response-reinforcement learning is dependent on n-methyl-d-aspartate receptor activation in the nucleus accumbens core. Proc Natl Acad Sci U S A 94(22), 12174–12179.

[86] Kennerley, S. W., M. E. Walton, T. E. J. Behrens, M. J. Buckley, and M. F. S. Rushworth (2006, Jul). Optimal decision making and the anterior cingulate cortex. Nat Neurosci 9(7), 940–947.

[87] Kim, K. H., H. W. Yoon, and H. W. Park (2004). Improved ballistocardiac artifact removal from the electroencephalogram recorded in fmri. J Neurosci Methods 135(1-2), 193–203.

[88] Kirisci, L., R. Tarter, M. Reynolds, and M. Vanyukov (2006, Apr). Individual differences in childhood neurobehavior disinhibition predict decision to desist substance use during adolescence and substance use disorder in young adulthood: a prospective study. Addict Behav 31(4), 686–96. 0306-4603 (Print) Journal Article.

[89] Knutson, B., G. Fong, C. Adams, J. Varner, and D. Hommer (2001, Dec). Dissociation of reward anticipation and outcome with event-related fmri. Neuroreport 12(17), 3683–7.

[90] Knutson, B., A. Westdorp, E. Kaiser, and D. Hommer (2000, Jul). Fmri visualization of brain activity during a monetary incentive delay task. Neuroimage 12(1), 20–7. 1053-8119 (Print) Journal Article.

[91] Komssi, S., J. Huttunen, H. J. Aronen, and R. J. Ilmoniemi (2004). Eeg minimum-norm estimation compared with meg dipole fitting in the localization of somatosensory sources at s1. Clinical Neurophysiology 115(3), 534–542.

[92] Krawczyk, D. (2002, Oct). Contributions of the prefrontal cortex to the neural basis of human decision making. Neurosci Biobehav Rev 26(6), 631–64. 0149-7634 (Print) Journal Article Review.

[93] Kuschinsky, W. (2000). Regulation of Cerebral Blood Flow, Chapter 2, pp. 15–24. Volume 1 of Moonen and Bandettini [102].

[94] Leahy, R. M., J. C. Mosher, M. E. Spencer, M. X. Huang, and J. D. Lewine (1998). A study of dipole localization accuracy for meg and eeg using a human skull phantom. Electroencephalography and Clinical Neurophysiology 107(2), 159 – 173.

[95] Lee, T.-S. and M. Zoltowski (1991, Nov). Beamspace domain bearing estimation for fast target localization using an array of antennas. Signals, Systems and Computers, 1991. 1991 Conference Record of the Twenty-Fifth Asilomar Conference on 2, 913–917.

[96] LeSage, J. P. (1999, Feb). The Theory and Practice of Spatial Econometrics. http://www.spatial-econometrics.com/: University of Toledo.

[97] Lin, F.-H., T. Witzel, S. P. Ahlfors, S. M. Stufflebeam, J. W. Belliveau, and M. S. Hämäläinen (2006). Assessing and improving the spatial accuracy in meg source localization by depth-weighted minimum-norm estimates. NeuroImage 31(1), 160 – 171.

[98] Liu, X. (2009, Oct). Private correspondence.

[99] Liu, X., D. K. Powell, H. Wang, B. T. Gold, C. R. Corbly, and J. E. Joseph (2007, Apr). Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. J Neurosci 27(17), 4587–4597.

[100] Logothetis, N. K. (2003, May). The underpinnings of the bold functional magnetic resonance imaging signal. J Neurosci 23(10), 3963–3971.

[101] Miki, K., T. Kida, E. Tanaka, O. Nagata, and R. Kakigi (2009, Apr). The impact of visual movement on auditory cortical responses: a magnetoencephalographic study. Exp Brain Res 194(4), 597–604.

[102] Moonen, C. and P. A. Bandettini (Eds.) (2000). Functional MRI, Volume 1. Springer-Verlag, Berlin.

[103] Mosher, J. C., P. S. Lewis, and R. M. Leahy (1992). Multiple dipole modeling and localization from spatio-temporal meg data. IEEE Trans Biomed Eng 39(6), 541–557.

[104] Nieuwenhuis, S., D. J. Heslenfeld, N. J. von Geusau, R. B. Mars, C. B. Holroyd, and N. Yeung (2005, May). Activity in human reward-sensitive brain areas is strongly context dependent. Neuroimage 25(4), 1302–1309.

[105] Nolte, G., O. Bai, L. Wheaton, Z. Mari, S. Vorbach, and M. Hallett (2004). Identifying true brain interaction from eeg data using the imaginary part of coherency. Clin Neurophysiol 115(10), 2292–2307.

[106] Norris, D. G. (2006). Principles of magnetic resonance assessment of brain function. J Magn Reson Imaging 23(6), 794–807.

[107] Nunez, P. L., R. B. Silberstein, Z. Shi, M. R. Carpenter, R. Srinivasan, D. M. Tucker, S. M. Doran, P. J. Cadusch, and R. S. Wijesinghe (1999). Eeg coherency ii: experimental comparisons of multiple measures. Clinical Neurophysiology 110(3), 469–486.

[108] Nunez, P. L., R. Srinivasan, A. F. Westdorp, R. S. Wijesinghe, D. M. Tucker, R. B. Silberstein, and P. J. Cadusch (1997). Eeg coherency: I: statistics, reference electrode, volume conduction, laplacians, cortical imaging, and interpretation at multiple scales. Electroencephalography and Clinical Neurophysiology 103(5), 499–515.

[109] O'Doherty, J., H. Critchley, R. Deichmann, and R. Dolan (2003, Aug). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. J Neurosci 23(21), 7931–9.

[110] O'Doherty, J., M. Kringelbach, E. Rolls, J. Hornak, and C. Andrews (2001, Jan). Abstract reward and punishment representations in the human orbitofrontal cortex. Nat Neurosci 4(1), 95–102.

[111] O'Donnell, P. and A. Grace (1995, May). Synaptic interactions among excitatory afferents to nucleus accumbens neurons: hippocampal gating of prefrontal cortical input. J Neurosci 15(5 Pt 1), 3622–39. Using Smart Source Parsing May.

[112] Ohira, H., M. Nomura, N. Ichikawa, T. Isowa, T. Iidaka, A. Sato, S. Fukuyama, T. Nakajima, and J. Yamada (2006, Feb). Association of neural and physiological responses during voluntary emotion suppression. Neuroimage 29(3), 721–33.

[113] Ozkurt, T., M. Sun, and R. Sclabassi (2008a, June). Decomposition of magnetoencephalographic data into components corresponding to deep and superficial sources. Biomedical Engineering, IEEE Transactions on 55(6), 1716–1727.

[114] Ozkurt, T. E., M. Sun, and R. Sclabassi (2008b, April 4-6). Spatial filtering of meg signals for spherical regions in the source space. In Proc. 34th Northeast Biomedical Engineering Conference, Providence.

[115] Packard, M. and B. Knowlton (2002). Learning and memory functions of the basal ganglia. Annu Rev Neurosci 25, 563–93. 0147-006X (Print) Journal Article Review.

[116] Palomero-Gallagher, N., H. Mohlberg, K. Zilles, and B. Vogt (2008, Jun). Cytology and receptor architecture of human anterior cingulate cortex. J Comp Neurol 508(6), 906–926.

[117] Parkinson, J., P. Willoughby, T. Robbins, and B. Everitt (2000, Feb). Disconnection of the anterior cingulate cortex and nucleus accumbens core impairs pavlovian approach behavior: further evidence for limbic cortical-ventral striatopallidal systems. Behav Neurosci 114(1), 42–63.

[118] Parkinson, J. A., M. C. Olmstead, L. H. Burns, T. W. Robbins, and B. J. Everitt (1999). Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by d-amphetamine. J Neurosci 19(6), 2401–2411.

[119] Parkkonen, L. and J. Mäkelä (2002). Meg sees deep sources: measuring and modelling brainstem auditory evoked potentials. Proceedings of the 13th Int'l Conference on Biomagnetism, Jena.

[120] Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sloreta): technical details. Methods Find Exp Clin Pharmacol 24 Suppl D, 5–12.

[121] Peru, A., G. Pavesi, and M. Campello (2004). Impairment of executive functions in a patient with a focal lesion in the anterior cingulate cortex. evidence from neuropsychological assessment. Funct. Neurol. 19, 107–111.

[122] Peters, Y. M., P. O'Donnell, and R. M. Carelli (2005, May). Prefrontal cortical cell firing during maintenance, extinction, and reinstatement of goal-directed behavior for natural reward. Synapse 56(2), 74–83.

[123] Petrides, M. (1985). Deficits on conditional associative-learning tasks after frontal- and temporal-lobe lesions in man. Neuropsychologia 23(5), 601–14.

[124] Pulvermuller, F., N. Birbaumer, W. Lutzenberger, and B. Mohr (1997). High-frequency brain activity: its possible role in attention, perception and language processing. Prog Neurobiol 52(5), 427–445.

[125] Purves, D., A. G. J, D. Fitzpatrick, L. C. Katz, A.-S. LaMantia, J. O. McNamara, and S. M. Williams (Eds.) (2001). Neuroscience (2 ed.). Sinauer Associates.

[126] Raichle, M. E. (1994). Images of the mind: studies with modern imaging techniques. Annu Rev Psychol 45, 333–356.

[127] Roberts, A. C., D. L. Tomic, C. H. Parkinson, T. A. Roeling, D. J. Cutter, T. W. Robbins, and B. J. Everitt (2007). Forebrain connectivity of the prefrontal cortex in the marmoset monkey (callithrix jacchus): an anterograde and retrograde tract-tracing study. J Comp Neurol 502(1), 86–112.

[128] Rodriguez-Rivera, A., B. Baryshnikov, B. Van Veen, and R. Wakai (2006). Meg and eeg source localization in beamspace. Biomedical Engineering, IEEE Transactions on 53(3), 430–441.

[129] Rolls, E. (1996). The orbitofrontal cortex. Philos Trans R Soc Lond B Biol Sci (351), 1433–1444.

[130] Rolls, E. (2000, Mar). The orbitofrontal cortex and reward. Cereb Cortex 10(3), 284–94.

[131] Rolls, E., S. Judge, and M. K. Sanghera (1977). Activity of neurones in the inferotemporal cortex of the alert monkey. Brain Research 130(2), 229 – 238.

[132] Rowan, A. J. and E. Tolunsky (2003). Primer of EEG: With A Mini-Atlas, Volume 1. Butterworth-Heinemann.

[133] Rushworth, M. F. S., T. E. J. Behrens, P. H. Rudebeck, and M. E. Walton (2007, Apr). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. Trends Cogn Sci 11(4), 168–176.

[134] Saito, Y. and H. Harashima (1981). Recent advances in EEG and EMG Data Processing. Elsevier, Amsterdam.

[135] Salmelin, R., R. Hari, O. V. Lounasmaa, and M. Sams (1994, Mar). Dynamics of brain activation during picture naming. Nature 368(6470), 463–465.

[136] Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Phys Med Biol 32(1), 11–22.

[137] Saver, J. and A. Damasio (1991). Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. Neuropsychologia 29(12), 1241–9.

[138] Schelter, B., M. Winterhalder, M. Eichler, M. Peifer, B. Hellwig, B. Guschlbauer, C. H. Lucking, R. Dahlhaus, and J. Timmer (2006). Testing for directed influences among neural signals using partial directed coherence. J Neurosci Methods 152(1-2), 210–219.

[139] Schultz, W. (1998, Jul). Predictive reward signal of dopamine neurons. J Neurophysiol 80(1), 1–27. 0022-3077 (Print) Journal Article Review.

[140] Schultz, W. (2002, Oct). Getting formal with dopamine and reward. Neuron 36(2), 241–63.

[141] Schultz, W., P. Dayan, and P. Montague (1997, Mar). A neural substrate of prediction and reward. Science 275(5306), 1593–9.

[142] Schultz, W., K. Preuschoff, C. Camerer, M. Hsu, C. D. Fiorillo, P. N. Tobler, and P. Bossaerts (2008, Dec). Explicit neural signals reflecting reward uncertainty. Philos Trans R Soc Lond B Biol Sci 363(1511), 3801–3811.

[143] Sekihara, K., S. Nagarajan, D. Poeppel, A. Marantz, and Y. Miyashita (2001). Reconstructing spatio-temporal activities of neural sources using an meg vector beamformer technique. Biomedical Engineering, IEEE Transactions on 48(7), 760–771.

[144] Seth, A. K. (2005). Causal connectivity of evolved neural networks during behavior. Network 16(1), 35–54.

[145] Shi, W. (2005, Nov). Slow oscillatory firing: a major firing pattern of dopamine neurons in the ventral tegmental area. J Neurophysiol 94(5), 3516–22.

[146] Song, T., K. Gaa, L. Cui, L. Feffer, R. R. Lee, and M. Huang (2008). Evaluation of signal space separation via simulation. Med Biol Eng Comput 46(9), 923–932.

[147] Sporns, O., G. Tononi, and G. M. Edelman (2000, Feb). Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. Cereb Cortex 10(2), 127–141.

[148] Takahata, R. and B. Moghaddam (2000, Oct). Target-specific glutamatergic regulation of dopamine neurons in the ventral tegmental area. J Neurochem 75(4), 1775–8.

[149] Takahata, R. and B. Moghaddam (2003, Jun). Activation of glutamate neurotransmission in the prefrontal cortex sustains the motoric and dopaminergic effects of phencyclidine. Neuropsychopharmacology 28(6), 1117–1124.

[150] Taulu, S. and M. Kajola (2005). Presentation of electromagnetic multichannel data: The signal space separation method. Journal of Applied Physics 97(12), 124905.

[151] Taulu, S. E., J. Simola, and M. Kajola (2005). Applications of the signal space separation method. IEEE Trans Sig Proc 53(9), 3359–72.

[152] Thorpe, S., E. Rolls, and S. Maddison (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. Exp Brain Res 49(1), 93–115. 0014-4819 (Print) Journal Article.

[153] Tong, Z. Y., P. G. Overton, and D. Clark (1996). Stimulation of the prefrontal cortex in the rat induces patterns of activity in midbrain dopaminergic neurons which resemble natural burst events. Synapse 22(3), 195–208.

[154] Uhart, M. and G. S. Wand (2009, Jan). Stress, alcohol and drug interaction: an update of human research. Addict Biol 14(1), 43–64.

[155] Ullsperger, M. and D. von Cramon (2003, May). Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and

the cingulate motor area revealed by functional magnetic resonance imaging. J Neurosci 23(10), 4308–14. 1529-2401 (Electronic) Journal Article.

[156] Uusitalo, M. A. and R. J. Ilmoniemi (1997). Signal-space projection method for separating meg or eeg into components. Med Biol Eng Comput 35(2), 135–140.

[157] Van Essen, D. C. and H. A. Drury (1997, Sep). Structural and functional analyses of human cerebral cortex using a surface-based atlas. J Neurosci 17(18), 7079–7102.

[158] Vanderschuren, L., P. D. Ciano, and B. Everitt (2005, Sep). Involvement of the dorsal striatum in cue-controlled cocaine seeking. J Neurosci 25(38), 8665–70.

[159] Villringer, A. (2000). Physiological Changes During Brain Activation, Chapter 1, pp. 3–13. Volume 1 of Moonen and Bandettini [102].

[160] Volkow, N., G. Wang, F. Telang, J. Fowler, J. Logan, A. Childress, M. Jayne, Y. Ma, and C. Wong (2006, Jun). Cocaine cues and dopamine in dorsal striatum: mechanism of craving in cocaine addiction. J Neurosci 26(24), 6583–8.

[161] Vrba, J. and S. E. Robinson (2001). Signal processing in magnetoencephalography. Methods 25(2), 249–271.

[162] Watanabe, M. (1990). Prefrontal unit activity during associative learning in the monkey. Exp Brain Res 80(2), 296–309.

[163] Watanabe, M. (1992). Frontal units of the monkey coding the associative significance of visual and auditory stimuli. Exp Brain Res 89(2), 233–247.

[164] Watanabe, M. (1996, Aug). Reward expectancy in primate prefrontal neurons. Nature 382(6592), 629–32.

[165] Weisstein, E. W. (2008). Laplacian. MathWorld - A Wolfram Web Resource. http://mathworld.wolfram.com/Laplacian.html.

[166] Wise, R. and M. Bozarth (1984, Feb). Brain reward circuitry: four circuit elements "wired" in apparent series. Brain Res Bull 12(2), 203–8. 0361-9230 (Print) Journal Article.

[167] Yeung, N. and A. G. Sanfey (2004, Jul). Independent coding of reward magnitude and valence in the human brain. J Neurosci 24(28), 6258–6264.

[168] Zeki, S. and S. Shipp (1988, Sep). The functional logic of cortical connections. Nature 335(6188), 311–317.

[169] Zhang, B., J. Liao, Z.-R. Wang, Y. Chen, J.-w. Gu, L.-g. Chen, and J. Fang (2006). Study on short-term therapeutic effect and its impact factors of stereotactic surgery for treating opiate users with opiate dependence in sichuan. Wei Sheng Yan Jiu 35(5), 599–603.

[170] Zoltowski, M. and T.-S. Lee (1991, Mar). Maximum likelihood based sensor array signal processing in the beamspace domain for low angle radar tracking. Signal Processing, IEEE Transactions on 39(3), 656–671.