UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE SELF VERSUS OTHERS:

SPATIAL LOCALIZATION AND TIMING OF TRAIT JUDGMENTS IN THE MPFC AND

PCC/PRECUNEUS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

MELODY REESE
Norman, Oklahoma
2020

THE SELF VERSUS OTHERS:
SPATIAL LOCALIZATION AND TIMING OF TRAIT JUDGMENTS IN THE MPFC AND
PCC/PRECUNEUS

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF PSYCHOLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Lauren Ethridge, Chair

Dr. Michael Wenger

Dr. Scott Gronlund

Dr. Robert Terry

Dr. Lei Ding

Acknowledgments

I would like to acknowledge my committee for their guidance in writing this dissertation and for their unwavering commitment to helping me succeed. Thank you: Dr. Lauren Ethridge, Dr. Michael Wenger, Dr. Robert Terry, Dr. Scott Gronlund, Dr. Lei Ding, and Dr. David Liu. I give a special thank you to Dr. Lauren Ethridge for her honest and invaluable support throughout my time in her lab, of which I am sincerely grateful for. I have appreciated your willingness to guide me through the design and implementation of a project with only tangential relevance to your lab. You have helped me gain confidence as a researcher, and your own work ethic and successes are incredibly inspiring.

Thank you to Dr. Michael Wenger for suggesting an HSMM-MVPA approach that bypasses my own concerns about making assumptions during exploratory aspects of this project. I admire your curiosity for research of all kinds; your enthusiasm is quite contagious. You have always believed in me and have offered many wonderful, collaborative opportunities to enrich my research career. Thank you for your encouragement along the way. Thank you, Dr. Lei Ding, for providing feedback regarding the appropriate source analysis methods and procedures for my EEG setup. I have not taken your classes or worked with you much, but I have enjoyed my time sharing the lab space with you and your team. Thank you to Dr. Scott Gronlund and Robert Terry for suggestions and comments made during the project proposal as well as for their continual kindness. Both of you bring a big smile to my face when I see you around campus. You are major contributors to my passion for cognitive, forensic, and quantitative psychology. I wish I could say more, but this is getting quite long. Lastly, I would like to thank Keegan McMillin for his advice and reassurance in the initial design of this study; Emma Auger for helping me work through bugs in the pre-processing pipeline; Lisa DeStefano for training me in

data pre-processing and basic Matlab programming; and the Department of Psychology for all

that they do. Thank you, all, for helping me grow.

Table of Contents

List of Tables

List of Figures

Abstract

Simulation theory (ST) states that people understand others through simulation, which counters the probabilistic reasoning view of theory theory (TT). When thinking about traits of a known other, people use self-referential thought. It is unclear which theory—ST or TT—best describes the method by which self-referential thoughts occur. A combination of event-related potential (ERP), event-related spectral perturbation (ERSP), source localization, and hidden semi-Markov model multivariate pattern analysis (HSMM-MVPA) techniques are hypothesized to disentangle self-vs-other information processing and distinguish competing theory of mind theories during a trait judgment task. EEG was recorded for 45 participants (30 females) ages 18-24 (M = 19.4) on resting and task measures, in which participants determined whether character and appearance words matched characteristics of the self and a close and distant other. Data analysis included repeated measures MANOVAs of reaction times, amplitudes and latencies generated from the parietal (PCC/precuneus) P300 and latter components of the frontal (mPFC) and parietal LSW. Time-frequency analysis included evoked and induced power through 100 Hz. ERP data was localized with MNE to verify location and timing assumptions for P300 and LSW. Lastly, HSMM-MVPA provided an alternative look at differences in number and duration of processing stages. The P300/LSW and source localization showed no differences between self, mother, and Fallon, which did not reflect prior BOLD activations. ERP data did not have the specificity to detect changes amid highly variable trials. Differences in self and mother were predicted by induced gamma ERSP, suggesting involvement of gamma in information integration or categorization. HSMM-MVPA models fit TT predictions and showed significant self-other differences in duration of processing and magnitude of peaks. Future research should clarify the role of the mPFC in self-referential thought and its relation to ST and TT with simultaneous fMRI and EEG and populations with impaired self-recognition such as ASD and schizophrenia.

Chapter 1: Background

Theory of Mind

Theory of mind is a well-studied concept stemming from philosophy of the mind during Descartes' era and elaborated upon by research in developmental, personality, and social psychology (Clarke, 2005). A person is said to have theory of mind when they understand that others have differing thoughts, feelings, and motivations from their own. In the psychology literature, this understanding of others is classified into five developmental stages: diverse desires, diverse beliefs, knowledge access, false belief, and hidden emotions (Peterson, Wellman, & Liu, 2005).

When a child has diverse desires, it means they have reached an understanding that people can have different desires from one's own. Diverse beliefs are similar, but this stage designates the understanding that people can have differing beliefs about the same situation. Next is knowledge access, which is the understanding that people may be ignorant about the truth. Building upon knowledge access, false belief is the knowledge that someone may believe something different from what is true. In the final stage, hidden emotions, a child understands that someone may feel a certain way but look as though they feel a different way. An example of this is when a person smiles to look happy when they are actually sad. The developmental order of these stages differs according to individual differences and cultural norms, although North American, Australian (Peterson, Wellman, & Liu, 2005), and German (Kristen, Thoermer, Hofer, Aschersleben, & Sodian, 2006) children generally follow the track outlined above (Goswami, 2011).

The relevance lies within the ability to distinguish actions, beliefs, and desires of oneself from that of others. The ability to understand other people suggests that humans have a clear-cut

separation of thoughts regarding ourselves versus thoughts about others. Thus, theory of mind is evidence for a distinction between self and other information and explains why young children rapidly change from a lack of understanding about others' differing thoughts, feelings, and desires to having a fully developed ability to mindread. Since most people have a fully developed theory of mind by age five, they have developed a distinct sense of self by that stage as well.

The two major categories of theoretical accounts for why and how people have theory of mind are called simulation theories and theory theories. Both have the overlapping idea that theory of mind is central to our social understanding.

Simulation Theory

The first set of theories to come about were simulation theories, in which people are said to understand others by predicting the other's feelings or reactions through simulation of oneself as the other, by simulating the feeling or reaction as if they were the other in the situation. This is essentially putting oneself in another's shoes, where neural activation is theorized to occur with pre-existing cognitive mechanisms for self-information processing and decision-making (Spaulding, 2012). With simulation theory, it is important to note that humans are not the only ones with theory of mind; other primates and even dogs have a rudimentary form of it. What makes humans unique is our cognitive ability to recognize that even though others have different thoughts, feelings, desires, and emotions, their experiences are comparable to our own (Mitchell, Macrae, & Banaji, 2006; Tomasello, 1999). Thus, according to simulation theory, your imagined predictions of another's thoughts, feelings, and behaviors would be comparable to what you would experience in the same situation.

In 2009, Frischen, Loach, and Tipper found evidence for simulation theory by showing that people can simulate someone else's frame of reference (2009). They demonstrated that, in a

single-person goal-directed reaching task, inhibition of a distractor corresponded to the salience, or closeness, of the distractor to the individual. But in a two-person version of the same task, reaction times were longer, and the distractors that were more salient to the agent of action rather than the self were inhibited, suggesting a simulated frame of reference of the other. Some simulation theorists believe that mirror neurons are evidence of such a primitive simulation heuristic, since mirror neurons activate both when observing someone perform a goal-directed action and doing the action oneself (Gallese & Goldman, 1998; Vogeley et al., 2001). Within the study of mirror neuron function, one of the first proposed theories was the theory of action understanding, which asserts that mirror neurons are the neural correlate of intention perception that allow us to understand others—to have empathy, a theory of mind, and a sense of morality (Rizzolatti & Craighero, 2004).

However, there is little consistent evidence that monkey mirror neurons are truly used for action understanding (Thompson, Bird, & Catmur, 2019), and in humans, mirror neurons are not necessarily required to exhibit action understanding (Hickok, 2009). One contributing factor to the losing debate over the theory of action understanding is defining what action understanding and goal-directed movement entail. A review by Thompson, Bird, and Catmur classifies studies of action understanding into three subcategories: action, goal, and intention identification. Their consolidated evidence suggests that mirror neurons are involved in action identification but not likely in intention or goal identification or perception (2019).

Moreover, mirror neurons activate for more than just goal-directed actions; upwards of 73% of them also activate for pantomimed versions of the same actions with no goal or intention attached (Kraskov, Dancause, Quallo, Shepherd, & Lemon, 2009) and object-free movement like lip smacking and tongue and lip protrusion (Ferrari et al., 2003). Additionally, majority of mirror

neurons activate only for specific features of an action. Some respond maximally for movement by the self instead of other or vice versa (Caggiano et al., 2011), left- or right-hand movement, according to directionality, or vicinity of an object (Caggiano, Fogassi, Rizzolatti, Their, & Casile, 2009). As such, mirror neurons are useful for perceiving certain features of movement much as visual neurons perceive features like edges and shapes. However, few if any mirror neurons have the capacity to generalize all features of an action performed by oneself and another, making intention perception unlikely (Thompson, Bird, & Catmur, 2019). As a result, action identification does not have the inferential qualities necessary to be the foundation of simulation or desire and belief understanding.

The theory for action understanding is one interpretation of mirror neuron function, and under this theory, mirror neurons are physiological evidence for simulation theory. However, as new evidence stacks against the role of mirror neurons in action understanding, the suitability for mirror neurons as evidence of simulation theory is questionable. Instead of mirror neurons, evidence for simulation theory should come from activation characteristics of mentalizing areas such as the medial prefrontal cortex. Simulation theory predicts that one's brain resources for self-mentalizing are also used to represent mental states of others (Goldman, 1992). Therefore, if regions known to activate for other-related information also activate for self-related information to a similar degree of magnitude and timing, then simulation of oneself as the other is a possible explanation for the findings. Experiments in auditory (Lima et al., 2015), visual (Kosslyn et al., 1993; Le Bihan et al., 1993), and motor (Schwoebel, Boronat, & Coslett, 2002; Meister et al., 2004) response have shown that imagined or simulated performance activates in regions of auditory, visual, or motor cortex that overlaps with overt performance with similar findings for socially-oriented judgments of faces and emotions (Wicker, Keysers, Plailly, Royet, Gallese, &

Rizzolatti, 2003; Singer, Seymour, O'Doherty, Kaube, Dolan, & Frith, 2004; Hesslow, 2012).

Furthermore, individual differences in the vividness of imagined states can be diagnostic of

behavioral outcomes and overt neural specificity (Lima et al., 2015).

     The mentalizing-based method of investigation is explored in the present study with a

cognitive model of simulation theory (Figure 1). In the presented model, the first-person mental

state, concept of the other, and perception of contextual information drives a simulated third

person mental state—a shifting egocentric frame (Perner, 1996). Fundamentally, the self is used

as an analog for the understanding the other, in which the same causal structures are used for self

and other during simulation (Apperly, 2008). From the simulated mental state representing the

other, a decision is made, and a simulated state of having made the decision while inhibiting self

information (Apperly, 2008; Zeng, Zhao, Zhang, Zhao, Zhao, & Lu, 2020) leads to a prediction

of the other's response.
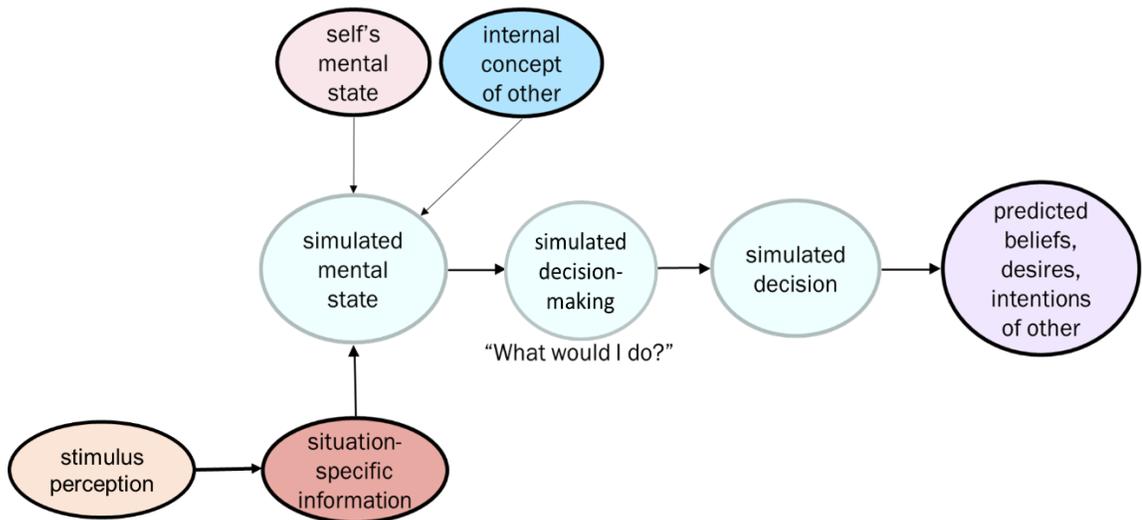
Figure 1. Proposed model of self-referencing in simulation theory

Theory Theory

The major opposing theory in theory of mind research is theory theory, which came about after simulation theory by way of Gopnik and Wellman in 1994. Gopnik and Wellman theorized that our understanding of others' mental states arises not from simulation but from experimental-like learning of the social environment (Gopnik & Wellman, 1994). This theory takes a more experiential approach than simulation theory, in which individuals use folk psychology to understand others' mental states (Gallese & Goldman, 1998). A common description of theory theory is that children are like scientists—curious about everything and constantly experimenting to uncover truths about the world. Understanding other people is no exception here. The major difference between simulation theory and theory theory is that simulation is a first-person process in which the simulator draws conclusions based on how he/she would think, feel, or act. Theory theory, on the other hand, is a third-person process, where the individual discovers empirical laws that govern human interaction and uses these to draw conclusions about how another person would respond.

Experiments such as one by Griffiths, Sobel, Tenenbaum, and Gopnik showed that children's understanding of others improved when provided new evidence (2011). The incorporation of information changed the child's conceptual framework just as theory theory would predict. Multiple studies have also shown that children use probabilistic reasoning; they combine prior knowledge and new evidence, likelihoods, and probabilities to make inferences (Sobel, Tenenbaum, & Gopnik, 2004; Kushnir & Gopnik, 2007; Gopnik & Wellman, 2012). Extending that line of reasoning to self and other inferences, children and adults incorporate prior knowledge and new evidence to make judgments or predictions about an other.

Neural activation during self- and other-related mentalizing would occur in either overlapping or distinct brain regions according to theory theory (Mahy, Moses, & Pfeifer, 2014), and both self and other would show timing and magnitude differences in the way that information is processed. The self would have faster processing than the other, since other-information would require a greater degree of probabilistic reasoning about the way in which the other would feel, behave, or desire compared to the self, and to an increasing degree with unfamiliarity, similar to that seen in Moran et al. (2011). When there is too little information about an other, the self may be used as a prototype for other-oriented predictions, potentially creating an inverted U-shaped curve for timing and reasoning requirements between self, close others, distant others, and complete unknowns (Kuiper 1981).

Sharing activation of precuneus/PCC and mPFC with theory of mind and self-judgment, autobiographical episodic memory is a potential mechanism for understanding relationships between experiences of the self and others in accordance with theory theory (Saxe, Moran, Scholz, & Gabrieli, 2006; Moreau, Viallet, & Champagne-Lavau, 2013; Rosa, Budson, Deason, & Gutchess, 2015). Interestingly, theory of mind, self judgment, and autobiographical memory-based tasks are interrelated to one another, emerge concurrently in development around age 3 and a half (Saxe, et al., 2006; Perner, Kloo, & Gornik, 2007; Spreng & Grady, 2009), and all involve the mPFC, PCC, and precuneus (Rabin & Rosenbaum, 2012). In a subset of patients with Alzheimer's Disease (AD) and frontal-temporal dementia, precuneus/PCC, mPFC, autobiographical episodic memory, and aspects of theory of mind and self-attribution are impaired, suggesting that these cognitive processes have shared functionality (Moreau, Viallet, & Champagne-Lavau, 2013). Additionally, involvement of autobiographical memory in theory of mind has been found to depend on familiarity, such that the reliance on shared experiences is

greater for close than distant others (Rabin & Rosenbaum, 2012). Personal closeness is the driving factor for mPFC differences between self and other rather than similarity or dissimilarity to the self.

A model of theory theory is proposed with the relationship between theory of mind, self, and autobiographical episodic memory in mind (Figure 2). Organization of the model was loosely adapted from Wickens and Carswell (2006). In the proposed model, the situation or stimulus is first perceived and irrelevant information is filtered out by selective attention. To initiate the social judgment, information from the internal concept of an other is retrieved and indirectly references the self through memories and exemplars. When combined with folk psychology, known information about other and what is partially referenced through self is appraised through a probabilistic reasoning process that weighs decision alternatives and ends in an understanding of the other's predicted response. Although developed separately, this model resembles the theory theory components of a computational model of theory of mind by Belkaid & Sabouret (2014).
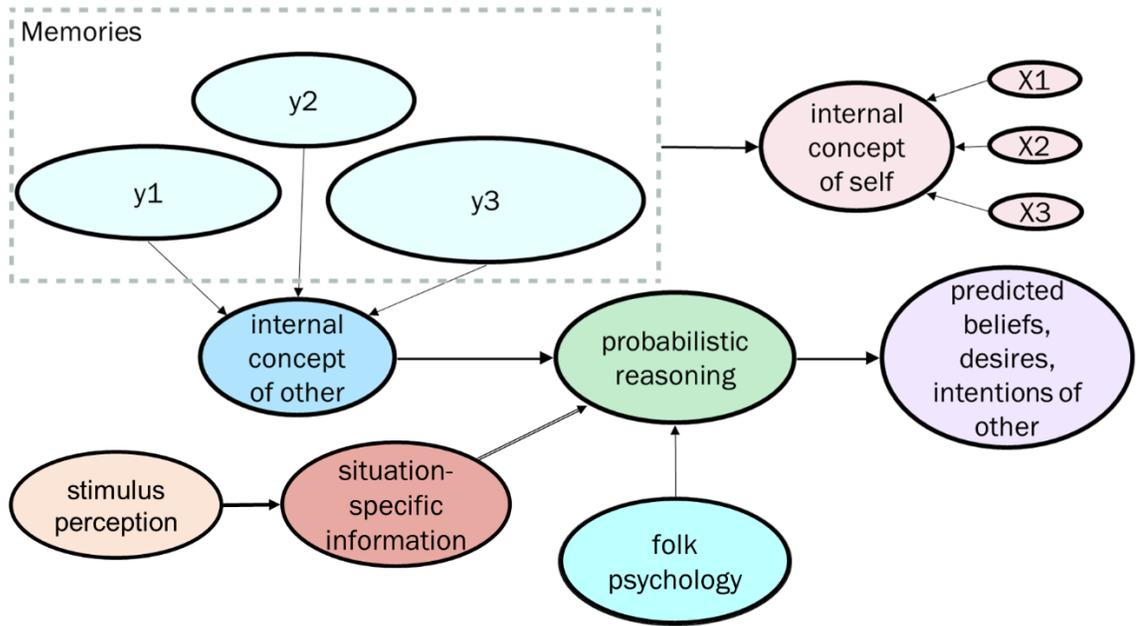
Figure 2. Proposed model of self-referencing in theory theory, where x and y represent stored information about the self and other, respectively.

The Self

Like theory of mind, the self is a long-studied concept dating back to 16th century philosophy. One of the major points of debate for this era was the mind-body problem, in which researchers sought to determine whether the mind and brain are separate. The dualistic explanation of the mind-body problem is that the mind is the brain. Thus, developing an understanding of others coincides with development of relevant brain regions. As Descartes described throughout his works from *The Rules* to *Meditations*, our thoughts and perceptions are merely patterns of brain activation (Clarke, 2005).

Descartes was also interested in the self-concept, as is partially evident by a famous quote attributed to his work: "cogito, ergo sum" or "I think, therefore I am" (Descartes, 1641). This claim came about through his own doubting—the fact that he could doubt meant he could think, and this meant he existed. Although very rudimentary, Descartes' statement is an

acknowledgement of self. As Gallup puts it, "it is our ability to conceive of ourselves in the first place that makes thinking and consciousness possible" (1998). That is, having a self is evidence for the existence of thought and consciousness—and perhaps thought and consciousness are evidence of a self-concept.

Since Descartes, Freud, Carl Rogers, and other pioneers took an interest in the topic of self-concept, the area developed rapidly since the 17th century, forming a modern self-concept theory positing that self-concepts are learned, organized, and dynamic (Purkey, 1988). Others have taken interest in the distinction between the self-concept and concepts of other people (Jenkins & Mitchell, 2011; Mitchell, Macrae, & Banaji, 2006; Moran, Lee, & Gabrieli, 2011). This self-other distinction through the lens of simulation and theory theory is the motivation for the current study.

Foundations of Self-Other Research

Using neuroimaging tools, researchers have investigated the self-other distinction that underlies theory of mind and the philosophies of self. The research has predominantly used functional magnetic resonance imaging (fMRI). Self-other judgment tasks based in fMRI often link self-referential thought to the dorsal and ventromedial prefrontal cortex (dorsal and ventral mPFC), the posterior cingulate cortex (PCC), and precuneus, which is a similar network of connectivity to both theory of mind and the default mode network (Northoff et al., 2006; Legrand & Ruby, 2009; Yerys et al., 2015).

The theme for many of these fMRI studies involves participants judging personality and appearance traits of themselves, someone emotionally connected to them, and someone familiar but not well known (Ochsner, et al., 2005; Heatherton et al., 2006; Mitchell et al., 2006; Moran et al., 2011). More specifically, the participant will often be sitting alone in an fMRI machine

with a computer screen in front of them. On the screen, the participant will be instructed to look at a fixation cross. An adjective will then pop up in a designated area below or above the cross, and the name of the person to be judged appears on the opposite side. When the adjective describes the person, the participant will press a button to indicate that yes, the adjective describes the person. When the adjective does not describe the person, the participant presses the appropriate "no" button to indicate that the adjective does not match the person.

The fMRI studies of this nature have indicated that the mPFC and regions such as the PCC and precuneus are activated when people make judgments about their own characteristics (Heatherton et al., 2006; Mitchell, Macrae, & Benaji, 2006; Moran et al., 2011; Northoff et al., 2006). The mPFC, PCC, and precuneus are also key regions in the default mode network (DMN) with compiling evidence of an association between resting state and self-referential processing through internally directed cognition (Molnar-Szakacs, & Uddin, 2013; Leech & Sharp, 2014).

Interestingly, the DMN and self-referential processing are abnormal in disorders like ASD and schizophrenia, which stem from deficits in shared functional connectivity of associated DMN and self-referential brain regions (Buckner, Andrews-Hanna, & Schacter, 2008). As an example, a study of schizophrenia found exaggerated DMN connectivity at rest and inadequate modulation of shared regions during self-referential processing (van Buuren, Vink, & Kahn, 2012). Studies in ASD have found atypical oscillatory activity in DMN, such as abnormally high theta, beta, gamma, and generally decreased alpha (Cornew, Roberts, Blaskey, & Edgar, 2012), suggesting a possible directionality of effect in individual differences of DMN activity in typically developing populations as well.

With the involvement of certain brain regions in dedicated self-referential processing, it is likely that self-referential information is processed in a fundamentally distinct way than

information about others. However, sections of the mPFC, like the dorsal segment, and the PCC also activate for other-related thoughts (Moran et al., 2011; Denny, Kober, Wager, & Ochsner, 2012), suggesting that self and other thought may be better differentiated through differences in timing and network connectivity rather than activity of individual brain regions. That being said, the ventral mPFC may be one single-region differentiator of self and other, making it a target for differentiating self and other networks and for development of biomarkers for self-related deficits in disorders like ASD and schizophrenia (Noel, Cascio, Wallace, & Park, 2017).

As previously mentioned, the uncertainty about the role of the mPFC, PCC, and precuneus in the self-concept comes from the fact that these regions also activate for other-related judgment. In a study by Moran et al., the PCC activated for appearance judgments of the self and close and distant others (2011). The close and distant others in the Moran study were the mother and President Bush, respectively. The dorsal mPFC activated for character judgments of the self, mother, and Bush and the ventral mPFC activated for appearance and character judgments for the self (2011). Surprisingly, they also found that the ventral mPFC activated for character judgments of the mother—the close other—despite past research suggesting the region was limited to self-related thought even for intimate others (Kelley et al., 2002; Heatherton, Wyland, Macrae, Demos, Denny & Kelley, 2006).

Based on behavioral survey data, Moran et al. (2011) concluded that this unanticipated activation of the ventral mPFC might occur because the close other's character traits help shape the participant's self-concept. From the neural data, however, it is unclear whether the ventral mPFC is activating for simulated self-thought during close other-character judgments or for indirectly referencing self by thinking about internal characteristics of someone close to us. One goal of the present study is to show that the ventral mPFC is a contributing part of a self-network

12

(Heatherton, Wyland, Macrae, Demos, Denny, & Kelley, 2006), where close-other information processing follows simulation theory or theory theory.

The P300

The P300, also called P3, is related to the mediation between perception and response, where the decision to categorize or classify a stimulus turns to an action (Verleger, Jaskowski, & Wascher, 2005). Since participants categorize character and appearance adjectives by person in the self-other judgment task before decision-making, there is an expected difference in P300 amplitude for self and other categories in relation to each other, resting state, and control judgments. Increases in P300 amplitude have also been associated with increases in selective attention, meaning that larger P300 amplitudes for one category or another may be associated with increased attentional resources toward that stimulus and the degree to which the information is processed (Sur & Sinha, 2009; Sowndhararajan, Kim, Deepa, Park, & Kim, 2018). Particular attention is paid to self-relevant stimuli over other kinds of information, and self-relevant information has previously been tied to increases in P300 amplitude compared to randomly generated control stimuli and sporadic red words (Gray, Ambady, Lowenthal, & Deldin, 2004)

P300 latency is believed to correspond to classification speed (Sur & Sinha, 2009), which may function as an indicator of categorization differences for character and appearance words related to the self and familiar and unfamiliar others. In other words, the P300 latency reflects stimulus-processing time, which may distinguish self- versus other-relevant stimulus processing in timing of cognitive mechanisms. If self-related and other-related information is processed differently, then we should see differences in the P300 latency for self-stimuli and mother's character, in addition to mother's appearance and Fallon-stimuli.

The P300 component has multiple frontal and parietal generators with the largest activity stemming from the centroparietal sources (Knight & Scabini, 1998; Linden, 2005). True generators of the P300 are often modality-specific and vary by subcomponent, such as the modulation of PFC for P3a but not always P3b (Swick, Kutas, & Neville, 1994; Linden, 2005). For instance, the P3a subcomponent occurs earlier in the time window with relation to novelty detection or stimulus probability, as is often seen in an oddball task. With limited EEG research in self-other judgment tasks, it is difficult to make specific localization predictions. The best guess given past fMRI findings during the self-other judgment task is that the PCC and precuneus are neural generators of the P300, with likely mPFC activity of the P3a (Esslen, Metzler, Pascual-Marqui, Jancke, 2008; Subramaniam et al., 2019), reflecting cognitive processing of the stimuli.

Late Slow Wave

Late slow waves (LSWs) are <10 Hz oscillations that occur approximately 300-1000 ms after a stimulus, with variation according to task structure and stimulus type (Sabbagh, 2013). The slow wave has been tied to working memory, where a stimulus is held in memory as the participant makes a judgment or decision. In general, the more negative LSW amplitudes are associated with increased information load. In the case of self-other, an other would be expected to require more information to make a judgment, meaning that LSW for other would be more negative than for self (Gray, Ambady, Lowenthal, & Deldin, 2004); more of an information load would have correspondingly longer latencies.

Generators of LSWs have not been thoroughly researched in the case of self-other and may vary by cognitive task, but theory of mind studies show LSWs in frontal regions more often stemming from medial frontal regions (Sabbagh, 2013). With evidence of LSW differences

between self and other judgments (Gray, Ambady, Lowenthal, & Deldin, 2004) and localization findings from theory of mind studies of understanding others (Sabbagh, 2013), differences in area amplitude and latency of self and other frontal LSWs are expected in the current study.

ERSP: Evoked and Induced Power

Although amplitudes and latencies of event-related potentials (ERPs) provide unique information about neural activity, there is one pitfall of ERP analysis. ERPs require data averaging, because of the otherwise indiscriminately small amplitudes and latencies of event-related activity amid ongoing background activity (Makeig, 1993). Averaging cancels out random activity that is not time- and phase-locked to the stimulus, leaving only event-related neural activity with a higher signal-to-noise ratio than before. However, averaging also masks trial-by-trial variability in oscillatory activity, which includes important information about changes in frequencies over time. Across many single trials, time-frequency measures such as total, evoked, and induced event-related spectral perturbation (ERSP) can use variable trial information to construct power values in each frequency band.

Event-related spectral perturbation (ERSP) is a measure that combines qualities of event-related synchronization and desynchronization to give baseline deviations in log spectral power across time and frequency in dB (Grandchamp & Delorme, 2011). Unlike ERP, ERSP does not include information about data polarity. Changes in spectral power, then, must be interpreted in the context that multiple underlying processes could be modulating the sources of neural activity in the positive or negative direction (Makeig, 2004). The benefit of ERSP measures over standard power measures like power spectral density (PSD) is the log-transformation that limits the influence of outliers in even a small number of trials (Izhikevich, Gao, Peterson, Voytek, 2018). Various self-other processing studies have demonstrated that self-relevant thought is most

saliently associated with evoked frontal and centroparietal alpha band power followed by frontal theta and gamma (Dastjerdi et al., 2011; Billeke, Zamorano, Cosmelli, & Aboitiz, 2013; Knyazev, 2013).

Induced power represents the power of event-related but not phase-locked activity. This non-phase-locked activity is theorized to reflect top-down processes like attention, memory, and decision-making, unlike evoked activity that reflects bottom-up perception and processing (Knyazev, 2013). More complex processing beyond stimulus perception such as attention, memory retrieval, and decision-making underlie self and other judgments, as proposed in the simulation theory and theory theory models. David, Kilner, and Friston (2006) refer to evoked and induced power as drivers and modulators, where drivers initiate stimulus response and affect neuronal assemblies while modulators engage mechanisms that alter responsiveness according to context. David et al. (2006) description of drivers and modulators maps onto the evoked bottom-up and induced top-down processing described by Knyazev (2013) and Mu and Han (2010).

Mu and Han incorporated induced power into a self-other judgment task like the task used in the current study. They found significant differences in non-phase-locked activity in theta, alpha, beta, and gamma oscillatory activity between the self, other, and control conditions. Their study was the first of its kind to connect non-phase-locked activity with self-referential processing (Mu & Han, 2010) and highlighted the importance of both evoked and induced power measures in distinguishing self and other. Such differences include phase-locked and non-phase locked magnitude and timing for processing associated with specific frequency bands for bottom-up and top-down processes. Averaging wide ranges of variability within-subjects in top-down processing may obscure differences by condition. Induced and evoked power highlight both non-phase-locked and phase-locked activity and avoid the potential obfuscation that can

occur when averaging trials for an ERP, although with a lower signal to noise ratio. All event-related measures together—the ERP and phase-locked and non-phase-locked power—offer rich information about the oscillatory dynamics of self-other EEG.

EEG Source Localization

EEG data are produced by clusters of neurons synchronous in geometry, orientation, and activity with local field activity summating to produce far-reaching signals that travel through bioelectric tissues through volume conduction and diffuse across the scalp (Makeig, 2004). By the time the signal reaches the scalp electrodes, there is no definitive way of knowing where in the brain the signals originated without simultaneous structural neuroimaging. This is the inverse problem. Methods have been developed to overcome the spatial hurdles of EEG, including algorithms that predict neural origins using the sensor data (Michel & Brunet, 2019). First, a forward solution is estimated from the distribution of scalp potentials brought about from underlying distributed source activity (Hallez et al., 2007). From the forward solution, an inverse solver can estimate sources of the scalp distributions (Grech et al., 2008; Song et al., 2015).

The objective of source analysis is to use sensor information from unknown brain sources and sometimes a priori knowledge to localize sensor information to the neural sources. However, methods requiring a priori assumptions about the number of dipoles may underestimate of the number of sources truly underlying the signal and bias source localization to the missing dipoles (Michel & Brunet, 2019). Additionally, assuming too many dipoles underlie the signal may result in false source readings. Non-parametric distributional source localization does not require a priori assumptions about the number of involved dipole sources, which avoids potential mislocalization (Pascual-Marqui, 2002). Thus, older algorithms like classical dipole source localization have been replaced by less problematic distributed source localization methods,

including minimum norm estimation (MNE), LAURA, and LORETA variants (Michel & Brunet, 2019). One of the most popular source localization methods is MNE, which is the method of focus (Gramfort et al., 2014; Wang et al., 1992)

MNE is an appropriate inverse method for EEG, where signals are generated near the surface from post-synaptic pyramidal cells in the cortex (Wang, Williamson, & Kaufman, 1992). MNE yields the same source solutions with and without a priori information, highlighting the robustness of the method even in the absence of theoretically motivated constraints or noisy data with few trials (Hauk, 2004). Although additional depth weights can be added to MNE, source output is generally restricted to one or several surfaces, where MNE is the most accurate. As comes naturally with the inverse problem, there are many current distributions that can equally explain EEG data. To choose a source solution, MNE finds the current distribution with the minimum ℓ_2-norm, which favors smaller and more distributed current estimates than the ℓ_1-norm (Lin, Belliveau, Dale, & Hamalainen, 2006). By constraining source estimates from the MNE inverse solution to the cerebral cortex, source activity reasonably reflects the resulting primary current of the EEG signal.

HSMM-MVPA

The simulation theory and theory theory model set up is ideal for a hidden semi-Markov model-based multivariate pattern analysis (HSMM-MVPA) approach designed by Anderson, Zhang, Borst, and Walsh (2016) and validated with real and simulated datasets. Prior applications of the HSMM-MVPA to EEG and fMRI could effectively recover processing stages with strong similarity to existing an ACT-R model of associated recognition (Anderson et al., 2016; Zhang et al., 2018). The underlying assumption of the HSMM-MVPA shared by classical and synchronized oscillation theory suggests that cognitive events are marked by a peak across

discrete brain regions caused by phasic bursts, phase resetting, or frequency band synchronization (Makeig et al., 2002; Shah et al., 2004; Klimesch, Sauseng, & Hanslmayr, 2007). Both theories share that the EEG signal is comprised of sinusoidal oscillations with uncorrelated background noise (Zhang, van Vugt, Borst, & Anderson, 2018).

From the sinusoidal oscillations, the HSMM-MVPA method identifies "bumps" or peaks with finite duration, amplitude, and topographical distribution that represent task-specific processing stages (Anderson et al., 2016). The bumps begin to rise at the onset of a significant event, and each bump is followed by a "flat" with a mean amplitude of zero representing the end of a processing stage. By identifying these patterns of peaks and valleys, the HSMM-MVPA model can provide the number, duration, magnitude, and topography of processing stages.

To be more specific, the MVPA identifies relevant patterns across sensors instead of assuming every sensor is contributing relevant information independently (Anderson et al., 2016). The hidden semi-Markov model then simulates the system with distinct processing stages mapped onto every trial with initial bumps and transitions occurring for variable time intervals. Mathematically, the HSMM is represented by $\lambda = (a, b, \pi)$, where a contains probabilities of a state transition given the current state, b is a matrix of probabilities of having a particular gamma distribution given a particular state, and π is the initial state distribution matrix with probabilities of starting in a given state (Yu, 2010). The semi portion of HSMM means that state transition probabilities are dependent on the amount of time elapsed since entering a state, which are assumed to be constant in classical HMM decision processes.

Both HMM and HSMM address three variations of problems (Yu, 2010; Cartella, Lemeire, Dimiccoli, & Sahli, 2015; Kang & Zadorozhny, 2016). Given model λ, one can determine the likelihood that λ generated a particular sequence of observations $O = \{o_1, \ldots o_n\}$,

$P(O \mid \lambda)$, or which sequence of states is optimal for producing the observed sequence. The expectation maximization algorithm (EM) is the final HSMM type that determines which model λ maximizes $P(O \mid \lambda)$. This is the focus of Anderson et al.'s HSMM-MVPA (2016).

With a selection of trained models, say ranging from 1-8 bumps, HSMM can determine which of the models most likely generated the series of observations under an optimal state sequence according to likelihood output as well as an estimated distribution over which state is occurring at time *t* (Cartella, Lemeire, Dimiccoli, & Sahli, 2015). Combining the estimable distributions and observed data, we can maximize the probabilities of having an observation sequence given the state sequence. By maximizing these probabilities, the model can maximize transition probabilities using the distribution of states. From there, with new transition and emission probabilities, the model can recursively estimate state probabilities until maximization is achieved. Figure 3 is an example of estimated probability distributions from trial 50 of the HSMM-MVPA self model. Likelihood output can then used to determine goodness-of-fit between models estimating 1-8 bumps, and leave-one-out cross validation can be used to determine which model best predicts data from each subject.
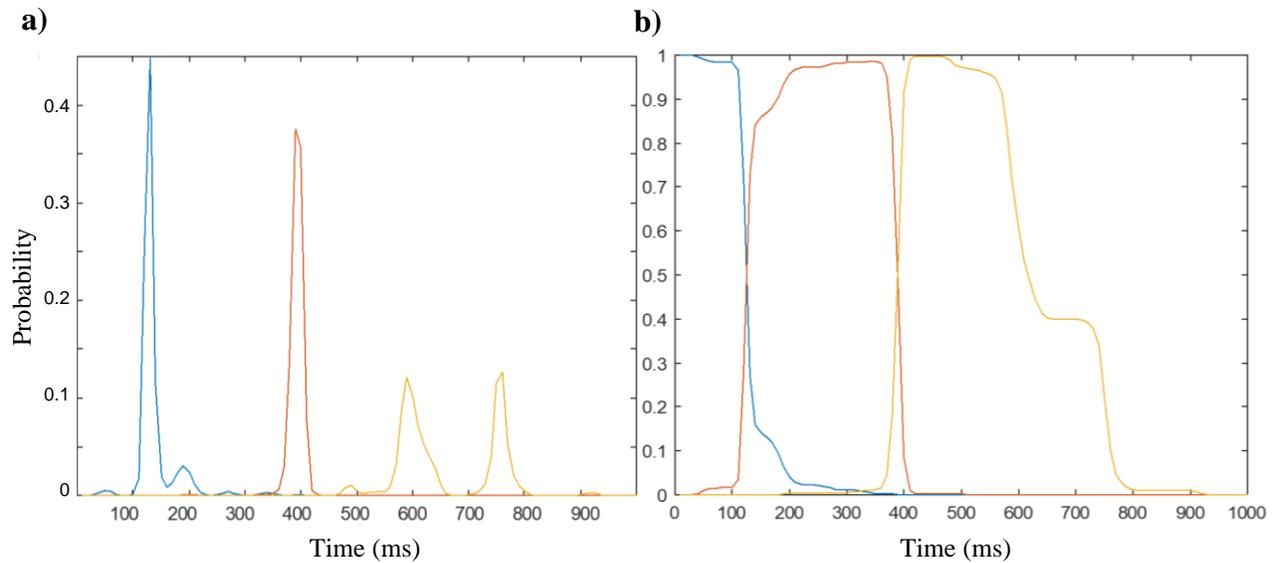
Figure 3. Example of estimated probability distributions of maximal bump timing from trial 50 of self. a) Probability that the peaks of processing stages are centered on a time point in a given trial. b) Probability that the time points fall within each of the processing stages for a given trial.

This HSMM-MVPA approach bypasses the need for theory that drives standard ERP and time frequency approaches in EEG analysis. Rather than pre-determining time windows to average across or defining peaks or power bands in an unstandardized field, HSMM-MVPA estimates parameters that maximize the probability of having the specified signal without losing between-trial variability in an average. The more objective approach is less susceptible to researcher bias as well as more likely to find differences that may be inaccurately captured in researcher-specified operationalizations.

The present experiment is an EEG adaptation of Moran et al. (2010), which found that character judgments of a close other (e.g., mother) are activated in a self-specific region of the brain. Two proposed explanations are given for the activation of mother's character judgments in a self-specific region of the brain based on the theory theories and simulation theories: 1) we indirectly reference the self through memories or exemplars when making judgments about others or 2) we simulate ourselves as the other. The proposed cognitive models based on each

21

theory differ in expected latencies and number of processing stages. Entering self, mother, and

Fallon trials separately into the HSMM-MVPA allows for comparison between the number,

duration, and magnitude of processing stages for each person condition. The HSMM-MVPA

results will either map onto the simulation theory or theory theory model or will suggest an

alternative interpretation.

Chapter 2: Overview of the Experiment

Purpose

Identifying the neural, structural, or biochemical mechanisms by which information about the self is treated differently from information about others may not be achievable with this experiment, although we progress toward that goal by evaluating results of self and other information obtained from two types of functional neuroimaging modalities—fMRI and EEG. The purpose of the experiment is multifold: to compare EEG and fMRI methodologies during a high-order cognitive task and provide an alternative interpretation of results that may disentangle the ventral mPFC's role in self-referential versus in-group thought through ERP and time-frequency analyses of the mid-frontal LSW.

Research Problems

The first major problem is that in Moran et al.'s (2011) results, the ventral mPFC activates when participants make judgments about their mother's character, when the region otherwise seems to be limited to self-related informational activation. Moran et al. (2011) attempted to explain the odd finding of ventral mPFC activation for the mother's character by claiming that the ventral mPFC is a region for processing of self-relevant information, and the mother's character is simply a part of the self-concept. However, the specific nature of self-relevant activation for a close other is left unclear.

Second, Moran et al. and other prior studies used fMRI to collect data, while this study utilizes EEG. There is uncertainty about EEG data having the spatial resolution necessary to distinguish the dorsal and ventral mPFC or precuneus and PCC. The sources are based on previous studies that used fMRI to gather data for self-other research (Heatherton, et al., 2006; Mitchell, Macrae, & Banaji, 2006; Moran et al., 2011). EEG is known to have lower spatial

resolution than fMRI, but fMRI has lower temporal resolution than EEG. Despite their differences, studies have shown that fMRI and EEG data are correlated to a relatively high degree, enough to reasonably expect similar data patterns. In particular, fMRI BOLD signals correlate strongly with EEG alpha and theta (inversely) and gamma power (positively) (Scheeringa, Petersson, Oostenveld, Norris, Hagoort, & Bastiaansen, 2009), which coincide with reasoning about social interactions (Blume, Lechinger, Guidice, Wislowska, Heib, & Schabus, 2015), DMN (Scheeringa, Petersson, Kleinschmidt, Jensen, & Bastiaansen, 2012), and self-referential processing (Knyazev, 2013). Low frequency oscillations like that of the LSW also strongly affect BOLD correlations and can influence local processing through modulation of gamma (Wang et al., 2012). Thus, certain lower and higher frequency bands of EEG may be methodologically parallel indicators of network activity seen in fMRI. In addition, a mouse study of oscillatory differences in dorsal and ventral mPFC showed that power levels of slow and fast frequencies could differentiate dorsal and ventral subregions, where dorsal regions had increased alpha, beta, and gamma power (Gretenkord, Rees, Whittington, Gartside, & LeBeau, 2017).

Given two major research problems—explaining unusual ventral mPFC activation for mother's character and source localizing to the ventral mPFC with EEG—general questions are as follows. Will adding timing data using EEG methodology help determine whether alternative or extended explanations are likely for the ventral mPFC's role in self-referential thought? Will EEG data lead us to the same conclusions as fMRI data using the same self-other judgment task?

Theoretical framework

Simulation theory can be summarized as a simulation of oneself as another, while theory theory is based more on exploration and discovery of self-generated theories that explain the real world. For simulation theory, drawing conclusions about what someone else may be thinking, for

example, would require the observer to simulate themselves as the observed and determine what one would do as an other. With theory theory, the process would be more probabilistic; one would take all known information about the observed and determine their most probable thoughts, feelings, and future actions.

Either of these theories may explain the odd findings by Moran et al. (2011). The research goal, in addition to replicating the original fMRI study with EEG is to use the replication and timing data to help determine which of these theories best applies. If ERPs are highly similar in shape, amplitude, and latency for the self and mother conditions, then simulation theory best addresses the results. If the mother waveform occurs later and with a significantly different magnitude than the self waveform, theory theory may provide the more reasonable explanation of underlying neural activity. More specifically, theory theory would predict that self information has an advantage in the brain, and that self information will be processed much faster and with more initial attentional resources and efficient processing.

Research Objectives

The goals of the study are to (1a) introduce novel ERP and (1b) time-frequency power measures to help further disentangle the physiological function of the ventral mPFC and address alternative explanations to Moran (2011). (2) Source analysis will determine whether conclusions made from P300 and LSW ERP components are spatially comparable to those made by fMRI data. (3) HSMM-MVPA will determine the number, timing, and duration of processing stages between conditions without a priori assumptions or theories.

The research questions are the following: (1a) Will ERP and (1b) time-frequency findings address uncertainties about Moran et al.'s (2011) explanation for why the ventral mPFC activated for the mother's character? (2) Will electroencephalography (EEG) data localize the

25

P300 and LSW with enough spatial resolution to see the differences between the dorsal and ventral mPFC, PCC, and precuneus seen in fMRI? (3) Without constraining the data according to theory, will HSMM-MVPA find differences in the number, magnitude, or duration of cognitive processing stages between conditions?

According to past research, (1a) The centroparietal P300 amplitudes will be larger for appearance conditions, as expected from larger PCC/precuneus BOLD responses in fMRI. LSW is predicted to have larger amplitudes for character over appearance judgments, as expected from larger dorsal mPFC responses for character judgments. The ventral mPFC has activated specifically for the self in many studies (Denny, Kober, Wager, & Ochsner, 2012) but also for close others in other studies (Mitchell, Banaji, & Macrae, 2005; Moran et al., 2011). Thus, the frontal P300/LSW is expected to be higher amplitude for self than other conditions, with close other being more similar to self than distant other. Since it is unclear whether the P300/LSW can distinguish between dorsal and ventral mPFC responses in magnitude, the component is predicted to represent a combination of dorsal and ventral mPFC activity with the largest amplitudes for the self-character condition, followed by self-appearance, mother-character, mother-appearance, Fallon-character, and Fallon-appearance. Self conditions are generally predicted to have a faster time course and largest amplitude in the P300/LSW time range for increased attentional resources and memory consolidation of self-referential information.

(1b) Evoked and induced power (Mu & Han, 2010) will differ between self and other, especially in alpha, theta, gamma (Knyazev, 2013) and beta (Park, Kim, Sohn, Choi, & Kim, 2018) frequencies according to past findings with self-referential processing in similar judgment tasks. Evoked alpha and theta power will be inversely related and gamma and beta will be positively correlated to past BOLD findings but patterns of neural oscillations will better

differentiate self and other across time, especially in later processing (Scheeringa, Petersson, Oostenveld, Norris, Hagoort, & Bastiaansen, 2009). For induced power, both low (theta, alpha) and high (beta, gamma) frequency neural oscillations of non-phase-locked activity will differentiate self and other, such that the magnitude of non-phase-locked neural activity will predict the degree of the self-reference over other-thought (Mu & Han, 2010).

(3) With the addition of timing information, ERP data will differentiate the self and other in source localization in the neural regions of interest with better timing discriminability than past fMRI data. Data from source analysis will show statistically significant differences in line with ERP expectations such that the centroparietal P300 and LSW will localize to the PCC/precuneus and frontal activity to the mPFC, where each component will have significant differences between self and other conditions.

(4) HSMM-MVPA predictions are as follows. In accordance with simulation theory, the self, mother, and Fallon models will have the same number of processing stages, will not vary in magnitude, and will not vary in duration of bumps. In accordance with theory theory, the self, mother, and Fallon models may have the same number of processing stages but will vary in duration and magnitude of bumps. The self will generally have faster processing stages than other and larger magnitude of bumps in processing through P300 latencies (300 ms) according to self-reference effects and potentially smaller magnitudes for self in probabilistic reasoning according to the salience of internal concepts of the self and other and degree of difficulty in reasoning.

Chapter 3: Method

Method

The goal of the experiment is the analysis of mid-frontal ERP amplitudes and latencies of the P300 and LSW, time-frequency analysis, spatial localization, and theory-independent HSMM-based multivariate pattern analysis to describe spatiotemporal characteristics of self-versus-other thought. Reaction time, ERP, and time-frequency data was processed using separate repeated measures MANOVAs, source statistics were calculated from the Monte Carlo method with clustering-based correction, and optimal HSMM-MVPA models were selected using sign tests. All results were qualitatively compared to fMRI data from the original Moran et al. (2011) study.

Participants

Data was collected from forty-seven college-aged adults from the University of Oklahoma participating for course credit. One participant (female, age = 19) had unusable data because of an acquisition error, and one dataset had an error during pre-processing. Forty-five datasets were included in analyses (15 males, age range 18-24, M_{age} = 19.3). Participants had to be at least 18 and have normal or normal-corrected vision to be included in the study. No other exclusion criteria were used.

Procedures

Within-subjects factors were trait (character, appearance) and person conditions (self, close other, distant other, uppercase), where the close other was the mother and distant other was the politically and morally neutral celebrity, Jimmy Fallon. Dependent variables for the ERP analysis included area amplitudes and latencies of frontal and centroparietal electrodes for the P300 and LSW. For the time-frequency analysis, dependent variables were evoked and induced

28

event-related spectral perturbation (ERSP) of delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and low (30-60 Hz) and high gamma (60-100 Hz) frequencies. Reaction times were separately analyzed in a univariate linear model. Afterward, source localization of the P300 and LSW ERPs and an HSMM-MVPA approach were performed.

First, each participant's baseline resting brain activity was measured with a 64-channel EGI brand (Electrical Geodesics, Eugene, OR) EEG net to account for individual differences in baseline activation. The baseline measurement consisted of 6 minutes of resting activity alternating eyes open or closed every minute. After resting data was collected, participants judged ('yes' or 'no') whether a series of descriptions matched characteristics of themselves, the close other, the distant other, and our control condition (uppercase) while whole-brain EEG data was gathered. Each participant was exposed to all four conditions for each of 90 character and 90 appearance words. Altogether, there were 720 total trials, 90 each of external and internal self, mother, Fallon, and uppercase conditions. Internal and external refer to the word being judged; internal is a character trait and external is an appearance trait. Adjectives were presented for 2000 ms on a black background below a fixation cross with the person to be judged (self, mom, Fallon) above the cross (Figure 4). The word display from top of the person word to bottom of the adjective word was 4.5 centimeters, center screen on a Dell E178FPb 17-inch LCD flat panel monitor with an approximate visual angle of 2.86° (distance from screen ≈ 90 cm).

In the uppercase blocks, participants judged whether character and appearance adjectives were upper or lowercase. The purpose of this was to ensure that participants were not randomly pressing buttons while also providing additional baseline data to account for brain activation during button pressing and non-social decision making. The person conditions also had randomized upper and lowercase adjectives to ensure that the uppercase condition could be used

as a button-pressing baseline without confounding case differences between people and case conditions.

To break down the 720 trials, there were 12 blocks of 60 trials with a jittered inter-stimulus interval of 1000-1500 ms and a self-paced break between blocks. The self, mother, and Fallon conditions were randomized prior to the study within each person block, as were the uppercase and lowercase adjectives for the uppercase blocks. The three uppercase blocks were randomly dispersed between person blocks, and impedance checks occurred at the beginning of the experiment and repeated after every four blocks for a total of three impedance checks across the duration of the experiment (see Appendix A for an overview of experimental procedures).
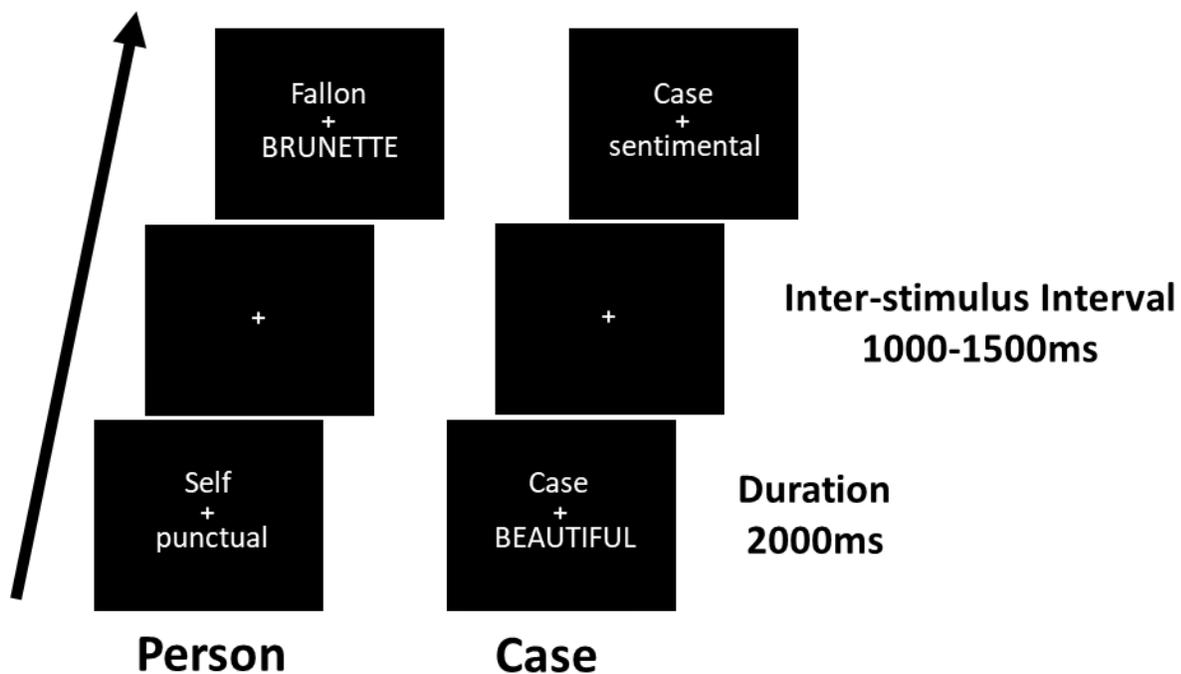
Figure 4. Example task stimuli. For the person and case conditions, targets (self, mom, Fallon, case) were above a fixation cross and traits (character or appearance words) were below. The fixation cross functions to minimize eye movement and, consequently, EEG artifact.

Data were analyzed using repeated measures general linear models (GLMs) of word (character, appearance) and person (self, mother, Fallon, case) conditions for reaction time and neuroimaging variables (P300 and LSW amplitude and latency and evoked and induced power). P300 and LSW amplitudes were then source localized with MNE. An HSMM-MVPA model for self, mother, and Fallon was developed for comparison to the proposed simulation theory (Figure 1) and theory theory (Figure 2) models. Comparison of the EEG findings to Moran et al.'s (2011) fMRI findings are discussed.

EEG Data

After the resting and task-related EEG data was collected, files were cut into four sections. One contained resting data, and the others included 20 minutes of data from each of three blocks separated by an impedance check. Pre-processing was performed according to a standardized pipeline utilizing EEGlab 14.1.1 (Delorme & Makeig, 2004), an open source toolbox for EEG analysis in Matlab (R2018b; The Mathworks, Natick, MA). Data had a 1000 Hz sampling rate for optimal temporal resolution of brain dynamics and were digitally filtered from 0.5 (12dB/octave slope; zero phase) to 100 Hz (24 dB/octave slope; zero-phase) with a 60 Hz notch filter and re-referenced to average reference. No more than 5% of total number of sensors were interpolated. Afterward, the data was processed via independent component analysis (ICA) using EEGlab 14.1.1 and high amplitude components associated with eye movement, muscle movement, heart rate, and other noise were removed from the data. After pre-processing, data from the three blocks were merged to create a single file per subject. To reduce dimensionality of the data and capture task-specific activity of the dorsal and ventral mPFC, PCC, and precuneus, sensors were limited to corresponding frontal [2, 3, 4, 6, 7, 8, 9, 11, 12, 13, 54, 59, 60] and centroparietal regions [26, 28, 31, 33, 34, 36, 38, 40, 42, 46], respectively (See Appendix B for

subsets on the electrode layout). Data was epoched from 500 ms prior to stimulus onset to 1500 ms post-stimulus onset for ERP, time-frequency, and source analysis and baseline corrected using the 500 ms pre-stimulus period. Descriptive statistics and within-subjects factors are reported in Appendix C, Table C1 andTable C2.

Reaction Time

A behavioral response such as reaction time reflects the cumulative duration of cognitive processing stages (Gray, Ambady, Lowenthal, & Deldin, 2004), while ERPs provide more detailed information regarding latencies of early and late components of cognition. The reaction time data was automatically collected at button press for each EEG trial using NetStation continuous recording software (Electrical Geodesics, Eugene, OR), and data was pulled from saved .txt files into Matlab for analysis. A repeated measures ANOVA was run with reaction time as the dependent variable and person and word as within-subjects factors. Post-hoc pairwise comparisons based on Sidak-corrected estimated marginal means showed differences in effect between each level of person and word. The GLM results were compared to Moran et al.'s (2011) reaction time findings (see Appendix D for table output).

ERP Analysis

ERP data was down-filtered to 20 Hz to reduce potential high-frequency interference, which was of particular concern around 30 Hz where reference noise was spotted during pre-processing. With the LSW occurring over the course of several hundred milliseconds, area under the curve measures were calculated with the open-source Matlab toolbox, ERPlab 7.0.0 (Lopez-Calderon & Luck, 2014) as an indicator of amplitude as opposed to peak measures. Since there was an unclear onset of the LSW and offset of the P300, the P300 and LSW were combined into one component, henceforth called P300/LSW. Using grand averages, individual ERPs, and

butterfly plots for all conditions as a guide, area amplitudes were calculated with a wide window from 200-1000 ms (Liesefeld, 2018).

The parietal and frontal P300/LSW were calculated separately with area amplitude using electrode clusters described above and shown in Appendix B. All conditions except fixation were measured with area amplitude, where positive and negative values were zeroed for frontal and centroparietal regions, respectively. Other ERPlab options for area calculation included rectified area and numerical integration. The ERPlab Wiki characterizes rectified area as the absolute value of area amplitudes, which transforms negatives into positives (see also: Luck, 2014). This is problematic if an ERP crosses the zero line before the cutoff—amplifying noise. The other option, numerical integration, subtracts negatives from positives to penalize ERPs for crossing the zero line sooner by shrinking the overall area value. By taking the subtraction approach, the difference in peak height is hidden, meaning it is unclear whether differences are due to latency or amplitude. Thus, zeroing was chosen as the most appropriate approach. Latencies for the P300/LSW were marked by the point under the curve where area is 50% on both sides using ERPlab (Lopez-Calderon & Luck, 2014).

All ERP and time-frequency variables were input as dependent variables in a multivariate GLM with person and word as within-subjects factors. The GLM, technically a repeated measures MANOVA, reduces the experiment-wise error rate compared to a series of univariate models, accounts for intercorrelations by creating a linear combination of the dependent variables (Huberty & Morris, 1989; Rencher, 2002) as well as incorporating covariance into the multivariate F (Garson, 2015) and having higher power (Rencher, 2002). The four multivariate tests (Pillai's Trace, Wilks' Lambda, Hotelling's Trace, and Roy's Largest Root) answer the broad question, "Does any of the neurophysiological data in this set of dependent variables vary

33

by person or word level after controlling for age and gender?" All four tests are reported, because they vary in purpose, robustness, and general methodology for testing multivariate effects. Pillai's Trace is the most robust to violations of MANCOVA assumptions, but all except Roy's Largest Root are generally robust (Olson, 1974; Rencher, 2002). Although Roy's Largest Root is at risk of inflated type I error, it provides unique information when compared to Hotelling's Trace (Olson, 1974). For instance, equal values indicate an effect driven primarily by a single dependent variable, a strong correlation between dependent variables, or that the effect is only minimally contributing to the model.

Once multivariate tests were run, univariate tests were reviewed only for the independent variables found to have significant effects in the multivariate analysis. This stepwise process increases power compared to running multiple univariate tests with no multivariate precursor (Rencher, 2002). Not only does the multivariate analysis have higher power than the univariate version, but by taking all dependent variables as a system or linear combination reduces family-wise error rate. As a result, first considering multivariate significance and using those p-values to inform univariate analysis is a method by which a researcher can include dependent variable inter-correlations while maintaining the ability to investigate the complex data on a more simplistic level with univariate and post-hoc analyses. The test of univariate within-subjects effects provides significance of each independent variable and interaction on each dependent variable.

The multivariate and within-subjects tests answer general questions about the relationship between independent variables, interactions, and covariates with the set of dependent variables and with each outcome individually. To further elucidate differences, post-hoc comparisons using pairwise comparisons from Sidak-corrected marginal means are used to compare the levels

of person and word. A deviation or custom contrast would be another option with the added ability to compare more than two levels of a factor at a time. However, pairwise comparisons are simple and satisfactory in this case. SPSS output including F, p-values, partial eta^2, and model R^2 are reported in the text with extended results in Appendix E. Parameter estimates are means of the dependent variables in a within-subjects design and are not included; pairwise comparisons of estimated marginal means show directionality.

Time-Frequency Analysis

Time-frequency variables were calculated using the same sensors as ERP analysis for evoked and induced power in the theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and low (30-60 Hz) and high (60-100 Hz) gamma frequency bands from -500 to 1500 ms. Data was notched at 30 Hz and 60 Hz to account for reference noise and electrical artifacts (Libenson, 2010). Power was calculated using the newtimef Morlet wavelet transform, with a linearly increasing cycle number from 1 cycle at 2 Hz to 30 cycles at 100 Hz. Total power measured via event-related spectral perturbation (ERSP) includes both evoked and induced activity. Separating evoked and induced power from the total allows a distinction to be drawn between phase-locked and non-phase-locked activity. The judgment paradigm may induce event-locked but not phase-locked cognitive processes such as probabilistic reasoning as the participants process properties of the self and other stimuli, and these processes may differ between conditions in magnitude and/or time course (Mu & Han, 2010; Papo, 2015). As a result, both induced and evoked power are included in the analyses to capture group differences in frequency band power of phase-locked and non-phase-locked activity. To calculate induced power (non-phase-locked activity), the power of the baseline-corrected averaged data (evoked or phase-locking activity) is subtracted from the un-corrected total ERSP (evoked and induced activity), which requires a

35

power calculation on single trial data. Dependent variables for multivariate GLMs include evoked and induced power for frontal and centroparietal electrodes in all frequency bands.

Evoked and induced power variables are entered into separate multivariate GLMs to reduce variable counts. Each multivariate GLM separately considered induced and evoked inter-correlations among frontal and centroparietal frequency bands. Output includes multivariate tests, within-subject effects, and pairwise comparisons (see Appendix F for full ERSP tables).

Source Analysis

Source localization was performed with MNE in Fieldtrip, an open source Matlab toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011; Donders Institute for Brain, Cognition and Behaviour, Radboud University, the Netherlands. See http://fieldtriptoolbox.org) for its simplicity, straightforwardness, common usage, and robustness to lack of a-priori information (Hauk, 2004). Given the low spatial resolution of EEG and limits of localization methods, differential timing of the ERPs in each condition will increase the likelihood of successfully separating the dorsal and ventral mPFC areas with source localization over separation at a single time point. Localization accuracy is imperative for forming appropriate conclusions, given the proximity of regions of interest—dorsal and ventral mPFC, PCC and precuneus. The targets were P300 and LSW area amplitude measures. Since the ending of P300 overlaps the beginning of LSW, P300 localization included activation at 300 ms, whereas LSW localization included 300-1000 ms activity to separate the components despite fuzzy boundaries.

Pre-processing. Data was first loaded into Fieldtrip after pre-processing, baseline correction, and epoching in EEGlab 14.1.1. High-amplitude components associated with eye movement, muscle movement, heart rate, and other noise were removed during pre-processing to avoid biasing localization (Kumar & Bhuvaneswari, 2012). Once loaded, the data was down-

filtered to 20 Hz to avoid convolution of artifact with neural signal. 0.5-20 Hz retained ERP

peaks of interest. Afterward, the data was averaged referenced and averaged from -500 to 1500

ms post-stimulus with a noise covariance matrix calculated from baseline.

     Forward solution. A 7 mm resolution source model was produced from a standard 3-

layer BEM head model provided by Fieldtrip after ensuring properly aligned electrodes (see

Appendix G, Figure G1 for electrode alignment). For EEG localization accuracy, the head model

included realistic boundaries for skull, scalp, and brain, which described a template geometry of

the head and estimated tissue conductivity at each layer (Michel & Brunet, 2019). EEG data is

highly sensitive to conductivity of the three layers, considering that electrical fields generated by

neurons must pass through these surfaces to reach electrodes on the scalp. Using a head model

that specifies levels of conductivity at each layer improves localization outcomes for EEG,

although MNE is robust to skull conductivity errors (Matt & Olaf, 2013).

     Electrode positions were determined from standard coordinates provided by EGI. The

template MRI from which the head model was derived came from Colin 27 data, which consists

of 27 averaged T1-weighted scans of one individual (Holmes et al., 1998). The source model

resolution was changed from the 1 mm default to 7 mm. The alteration reduced the number of

dipoles inside the brain from 1,996,960 at 1 mm to 5,810 at 7 mm to improve computational load

and remain within the recommended range of 3,000-6,000 solution points (Michel & Brunet,

2019).

     There is a trade-off to increasing or decreasing the number of dipoles fitted for source

estimation (Michel & Brunet, 2019). Having more dipoles significantly increases computational

load, leads to more issues with numerical precision, and increases the risk of having spurious

sources. Spatial resolution and accuracy are increased; however, there are diminishing returns on

spatial accuracy with limits set by the quantity and quality of data from the electrodes. Given that 64-channel scalp EEG has an estimated spatial resolution of 3 cm or worse (Gevins, Le, Martin, Brickett, Desmond, & Reutter, 1994; Burle, Spieser, Roger, Casini, Hasbroucq, & Vidal, 2015) and that 3,000-6,000 dipoles are recommended, a 7 mm resolution source model with 5,810 dipoles should be sufficient for localization.

From the source model, electrode positions, and head model, a leadfield was computed, which provides forward solutions for all dipoles in a channel by source matrix (Nolte, 2003; Oostenveld, Fries, Maris, & Schoffelen, 2011; Appendix G, Figure G2). The leadfield, head model, and ERP data was used as input into MNE source analysis, which returned a solution to the inverse problem in a 3D volume. All plots show activity from the cortical surface level in conjunction with MNE recommendations (Hauk, 2004).

Inverse solution and plotting. For MNE, lambda was set to 3 and data was scaled to source covariance and pre-whitened the leadfield using noise covariance. Source output was grand averaged for each condition before plotting and statistical analysis. P300 was plotted at 300 ms, and LSW was plotted in the 300-1000 ms range. Difference plots of source power were generated using surface plots interpolated with a sphere average and difference data computed using the (x1./x2)-1 operation on source grand averages. For source statistics, the non-parametric Monte Carlo method generated a t-map with a cluster-based correction to minimize family-wise error rate inflated by multiple comparisons (Maris & Oostenveld, 2007).

Cluster-based permutation tests. The Monte Carlo method first creates a random partition of the data for each of two conditions and then calculates the maximum of cluster-level summed t-values as the observed test statistic (Maris & Oostenveld, 2007; Popov, Oostenveld, & Schoffelen, 2018). The cluster-based test statistics are calculated by generating two-sided t-

values for every sample pair from the partitioned data, only keeping samples with whose t-value

is significant at the cluster-alpha threshold (.05), clustering according to spatiotemporal

adjacency, and setting the statistic to the maximum of cluster-level statistics taken from the sum

of t-values within clusters (Maris & Oostenveld, 2007). The steps are repeated many times to

construct a distribution of the statistic from which random draws are used to compute the

proportion of random partitions with a larger test statistic than the observed value. Steps were set

to 50,000. If p is smaller than critical alpha of 0.05, then the data from the conditions are said to

be significantly different. T-values from source statistics were plotted on the standard template

MRI derived from Colin 27 (Figure 6-7) following procedures outlined in Appendix G, Figure

G3. These surface t-maps were formed by interpolating the source data onto a surface mesh, in

which there were 10,920 voxels in functional data on a [20 26 21] grid and 346,499 vertices on

the cortical surface.

HSMM-MVPA

The hidden semi-Markov model-multivariate pattern analysis (HSMM-MVPA) approach

was adapted from but closely following Matlab code provided by Anderson, Zhang, Borst, and

Walsh associated with their 2016 publication. In their paper, they assert that processing stages of

the same type are assumed to be the same length despite condition or subject. This means that

variation in initial timing of the stage gives information about the response time of a subject to a

specific condition, and the duration of the stage gives an indication of whether the subjects are

undergoing the same types of processing in each condition. Considering the proposed models of

ST and TT, it is possible to determine which model best applies to self-other data processing.

Self, mother, and Fallon trials were placed into separate models to generate a self, mother, and

Fallon HSMM-MVPA that can be compared with respect to quantity, timing, and duration of stages.

Since any data that is not time-locked is hidden by the averaged waveform, single-trial data was required for the HSMM-MVPA approach to identify processing stages with variable-timing (Anderson et al., 2016). Epochs were 1 second in length, which, for most trials, cut off response-related processing (Appendix D, Table D1). Although 2 second epochs would ensure that all responses from all trials are included in the model, 2 seconds introduced a large amount of trial-to-trial variability in post-evoked neural data that was difficult for the HSMM-MVPA to filter through and gave many repetitive results. Many responses occurred in under 1.5 seconds, which leaves at least 0.5 seconds of unrelated neural activity to disrupt the model output. 1 second epochs had best fitting models with 3-4 bumps, while 2 second epochs had best fitting models of 8-9 bumps. It is unclear whether 8-9 is an overfit or inclusion of unrelated activity that muddies the interpretation of output. Models with 1 second and 2 second epochs both had duplication of topographies, however 1 second data generally had 1 repeat while 2 second data showed numerous repetitions. Thus, 1 second epochs were chosen despite occasionally including the tail-end of response processing in a handful of trials. The trials with longer time to response do not appear to be condition specific. Self, mother, and Fallon have 2 conditions each (character and appearance).

Hidden semi-Markov models are used to account for the timing variability of stages with the added benefit of being less computationally intensive as HMM-MVPA (Yu, 2010; Anderson et al., 2016). To reduce computational load, data was downsampled from 1000 Hz to 100 Hz, and a spatial PCA reduced the data to 10 components according to a scree plot, which accounted for 74.2, 74.6, and 74.4% of signal variance in the self, mother, and Fallon models, respectively.

Each component was z-scored for each trial to a mean of zero and standard deviation of 1. Instead of having 64 electrodes, the data then consisted of 10 normed PCA components with a constant between-trial variability of 1. To perform the same analysis as Anderson, Zhang, Borst, and Walsh (2016), bumps were specified to have 50 ms in width (5 10 ms samples), which makes a half-sine waveform. Bumps are always followed by flats, meaning that there will be n bumps and n + 1 flats in the signal. By specifying a narrow width, the bumps have more precision in signal identification, which is advantageous for detecting processing stages of variable lengths.

Since the bottom-up approach does not require a priori information about the number of processing stages, the model started with neutral parameters and re-estimated until convergence (Anderson, et al., 2016). Initial variables included 10 ms voltages across all 64 electrodes, start and end samples of the trials for specific subjects and conditions, conditions and subjects by trial and samples, electrode descriptions, and normed PCA dimensions set up independently of the code provided by Anderson and colleagues. These variables were used to run various aspects of the HSMM-MVPA script. Log-likelihoods for N-bump models, 10*N PCA values as magnitudes, and (N+1)*2 gamma parameter estimates for N+1 flats were output for models with 1 to 8 bumps.

As the basis for the HSMM-MVPA estimation, the gamma distribution predicts the waiting period before an event of interest in a skewed distribution with shape and scale parameters (Thom, 1958). In this application, shape is fixed to 2 and scale is freely estimated by the HSMM-MVPA to determine stage durations (Anderson et al., 2016). Stages are assumed to occur discretely such that no two stages occur at the same time. Additionally, the log transform for likelihood turns the product of densities into a sum, which reduces computational load and

avoids numerical issues with machine precision (Taboga, 2017). The model with the best fit is the one that maximizes log-likelihood. Log-likelihood in this case reflects the likelihood of the bumps being centered at each time point given the gamma parameters used to estimate bump locations (Zhang et al., 2018).

Because of an issue with duplicate topographies appearing in model output, several articles were evaluated for an explanation and solution to the inclusion of an extra, identical bump in output. After altering the number of PCA components to > 10 and < 10, changing bump durations from 50 ms to 100 ms, testing data on individual subjects, and detrending the data, it was clear that the duplications were not caused by overfitting, excessive noise, or temporally correlated noise. Adjusting code from 50 ms to 100 ms bump widths did improve output but did not eliminate the problem. Zhang et al. (2018) mentioned a similar phenomenon in their data, explained either by alpha ringing or an extra-long bump that was picked up twice or more in the MVPA algorithm. Their solution to constrain adjacent bumps to a maximum correlation of (T-50/150) resolved the issue, where T is the mean durations of "flats" or time periods between bumps. Nearby bumps (< 200 ms) could no longer be correlated, which avoided the problem of a long processing stage onset pulling into separate bumps. The likely explanation is that the P300/LSW-like peaks in the data were longer than the 50 ms expectation, which caused a "new" stage to appear around bumps 3 and 4 in a 5-bump model. With the correlation correction, the best bump models shifted from 5 to 3-4, which is further evidence that the repeated bumps were not independent stages.

Once the initial 1-8 bump models were estimated, leave-one-out cross-validation (LOOCV) was used to ensure that conclusions are generalizable (Anderson et al., 2016). LOOCV removes one subject from the dataset and uses the model to calculate the log-likelihood

of the missing subject's data. This process indicated how well the model predicted data outside of model training. LOOCV was repeated for all 45 subjects, giving log-likelihoods for all individuals and all 8-bump models. The log-likelihoods were then used to determine the optimal model for the greatest number of individuals. The balance between generalizability and parsimony is important to consider in conjunction with initial fit findings. Since LOOCV fits to data left out of model training, the effects of overfitting are mitigated (Anderson & Fincham, 2013; Anderson et al., 2016). In other words, log-likelihoods will not simply increase because there are more states or parameters in the data. A sign test was used to determine whether the initial N-bump model is better than the best fitting LOOCV model (Anderson et al., 2016). Sign tests are a non-parametric alternative to t-tests and assess the null hypothesis that the median difference between the 3 and 4-bump models is zero, where the 3-bump models have .5 probability of fitting a subject's data better than 4 bumps. Difference scores are calculated and used to assess the alternative hypothesis that the 4-bump model significantly improves log-likelihoods over 3 bumps. The final, chosen model predicted the largest proportion of subject data without sacrificing parsimony.

The HSMM-MVPA approach demonstrates how bumps in the signal can be localized to specific cognitive stages. However, identifying and classifying the number and duration of stages determined by the model is easiest when a cognitive architecture already exists. For Anderson et al., the HSMM-MVPA output was compared to an existing ACT-R model (2016). Instead of using a pre-existing cognitive model, the self, mother, and Fallon HSMM-MVPA were compared to the proposed ST and TT models to determine which theoretical framework best describes the data. Since the ST and TT models have not been validated, the interpretation of the HSMM-MVPA output will require future experiments to verify interpretation. Simulation theory

would assume that neural mechanisms for self would be a special case or subset of theory of mind functionality with similar magnitudes across self and other conditions. Otherwise, latencies and processing stages would not change across person. Theory theory would assume distinct mechanisms and latencies for processing of self and other, with potential similarity in number of stages but not magnitude or duration.

Chapter 4: Results

Behavioral results

Assumptions were tested prior to conducting the repeated measures (RM) ANOVA. Box plots showed no egregious outliers in reaction time by person or word. A Shapiro-Wilk test of normality held for all person conditions ($p > .05$), but not word conditions ($p < .05$). Sample sizes are equal within person conditions (N = 90) and word conditions (N = 180), meaning that results will be robust to normality violation (see Table C2 for sample size breakdown). According to Mauchly's test of sphericity, the sphericity assumption was violated for person ($p < .001$) but not for person*word ($p = .711$) with no data for word. This tests the null hypothesis that error covariance matrices of the orthonormal dependent variables is proportional to an identity matrix (IBM Corp., 2019). The Mauchly test cannot be violated for variables with only two levels, since it is testing for unequal covariances. For person and person*word, this test is highly sensitive to mild violations of normality and, combined with being low power, is not necessarily reliable (Garson, 2015). Regardless, to account for the violation, the conservative Huynh-Feldt correction was reported.

The 2 (character, appearance) by 4 (self, Fallon, mother, case) repeated measures ANOVA with Huynh-Feldt correction revealed main effects for person ($F(1.60, 70.7) = 613.83$, $p < .001$, partial $\eta^2 = .93$), word ($F(1, 44) = 59.16$, $p < .001$, partial $\eta^2 = .57$), and person*word ($F(3, 132) = 11.31$, $p < .001$, partial $\eta^2 = .20$; Appendix D, Table D2). Post-hoc pairwise comparisons using Sidak correction determined that all person conditions were significantly different from one another by at least $p = .013$ (Table D3). Case had the shortest reaction times (Table D1). Self had longer reaction times than mother but not Fallon, and Fallon had the longest latencies (Case < Mother < Self < Fallon). Appearance words had shorter latencies than character

words (character M = 1034 ± 17 ms; appearance M = 1000 ± 15 ms). Pairwise comparisons of person within each level of word and word within person were tested using a Sidak-corrected pairwise comparison of estimated marginal means to elaborate on person and word differences masked in the overall person*word interaction (Table D4). The Case < Mother < Self < Fallon pattern held for all but the comparison between self and Fallon for appearance words, which did not reach significance ($p = .137$). Last of all, case did not significantly differ by word condition ($p = .855$; Figure 5).
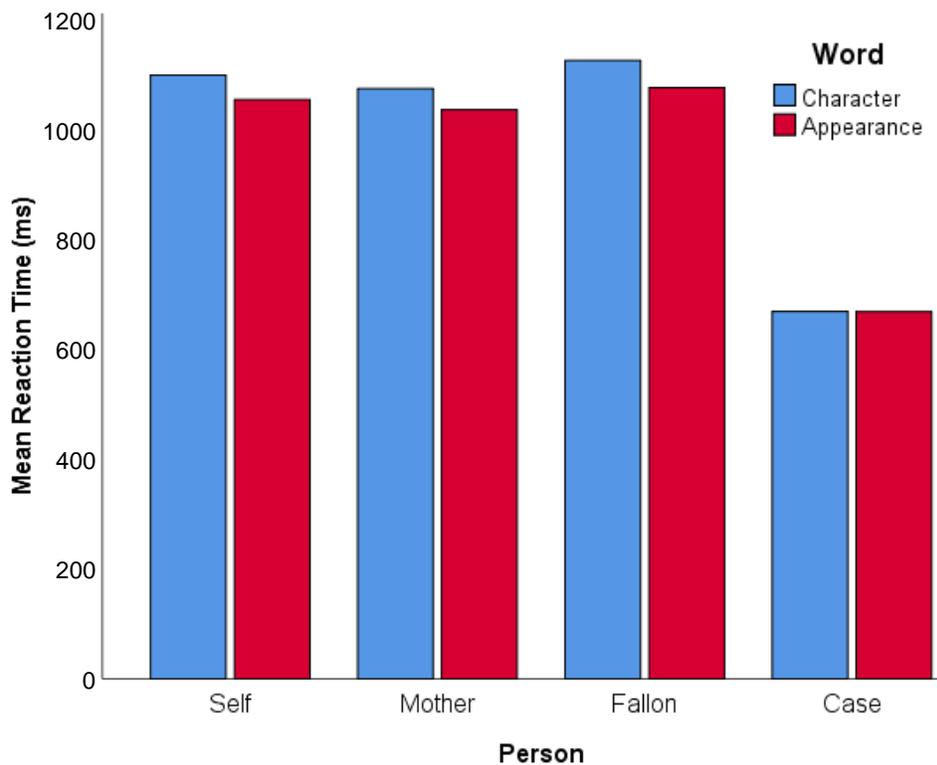
Figure 5. Mean reaction times by person and word condition.

ERP Amplitudes and Latencies

Spearman's rho correlations showed moderate correlations < 0.9 between dependent variables, which is an important balance for multivariate analyses that take intercorrelations into

account (Appendix E, Table E1**Error! Reference source not found.**). For the assumption of univariate normality, Shapiro-Wilk showed a violation for all word conditions ($p < .001$) on all ERP variables except centroparietal latency ($p > .330$). Person conditions violated normality for all but Fallon in frontal latency ($p = .066$) and self, mother, Fallon, and case in centroparietal latency ($p = .969$, $p = .323$, $p = .240$, $p = .986$). Since sample sizes are equal within person and within word, results are robust to normality violation. Mauchly's test of sphericity showed that the sphericity assumption was violated for person across all ERP variables ($p < .002$) and for person*word across frontal latency ($p = .026$). Therefore, at least one inter-correlation is non-zero and the conservative Huynh-Feldt corrected results were reported for the test of within-subjects effects. In the face of sphericity violations, a repeated measures (RM) MANOVA is recommended over multiple univariate RM-ANOVAs and warrants a multivariate approach for this data (Garson, 2015).

MANOVA output provides 4 multivariate tests to determine whether the independent variables have a significant effect on the system of dependent variables. Pillai's Trace takes the sum of effect sizes for the discriminant function and is highly robust when assumptions are not met (Olson, 1974; Rench, 2002; Grice & Iwasaki, 2007). Thus, Pillai's Trace is the primary test reported here. Multivariate output revealed that at least one ERP variable differed by person ($F(12, 393) = 5.59$, $p < .001$, Pillai's Trace = .438, partial $\eta^2 = .146$) and word ($F(4, 41) = 5.11$, $p = .002$, Pillai's Trace = .333, partial $\eta^2 = .333$) but not person*word ($F(12, 393) = 1.43$, $p = .151$, Pillai's Trace = .125, partial $\eta^2 = .042$; Table E2). According to raw discriminant function coefficients of the multivariate output, the best standardized weighted combination of ERP variables to discriminate person is frontal area (-.817) + frontal latency (1.12) + centroparietal area (.017) + centroparietal latency (-.631) and for person*word is frontal area (.268) + frontal

latency (-.302) + centroparietal area (-.907) + centroparietal latency (-.266). No multicollinearity

was found for area or latency that would bias discriminant function results (*VIF* ≤ 3.053 for all

variables; Poulsen & French, 2004). This means that the optimal difference between person

conditions would show low frontal and nearly zero centroparietal area, faster frontal latencies,

and slower centroparietal latencies with the greatest differentiation by frontal area and latency, as

would be expected from ventral mPFC BOLD findings. For person-word conditions, the optimal

difference would show high frontal area, low centroparietal area, and shorter latencies and is

likely drawing from differences in self, mother, and Fallon to case, which has shorter latencies,

more positive frontal, and less positive centroparietal area.

Although univariate output is automatically provided by SPSS for multivariate GLMs,

the univariate aspect of analyzing multivariate output is not required (Grice & Iwasaki, 2007) but

will be briefly reviewed. The univariate within-subjects test with Huynh-Feldt correction showed

significant differences by person for frontal area (F(2.43, 107.08) = 7.89, $p < .001$, partial η^2 =

.152), frontal latency (F(1.53, 67.11) = 5.55, $p = .011$, partial η^2 = .112) and centroparietal area

(F(2.23, 98.24) = 6.16, $p = .002$, partial η^2 = .123). Significant differences by word included

frontal and centroparietal latency (F(1, 44) = 12.91, $p = .001$, partial η^2 = .227; (F(1, 44) = 13.36,

$p = .001$, partial η^2 = .233; Table E3).

Person*word had a significant difference for centroparietal area in the univariate but not

multivariate model (univariate F(2.79, 122.89) = 4.02, $p = .011$, partial η^2 = .084; Table E3). The

most likely explanation is that frontal and centroparietal areas are highly and significantly

correlated, which is incorporated into the multivariate but not univariate analysis (Spearman's

rho = .801, p < .001; Appendix E, Table E2). In summary, person*word no longer significantly

contributed in the context of other dependent variables, namely frontal area, but had univariate

influence on centroparietal area. A correlation of .801 is potentially too high for the MANOVA, with the variables essentially being treated as a singularity during linear combination. The recommendation for dependent variables is a "moderate" correlation with vague and varying guidelines about what that entails. Researchers Grice and Iwasaki suggest that correlations up to .90 are acceptable for multivariate analyses (2007). Because of the potential for a biased multivariate outcome for person*word, the effect on centroparietal area was explored with this paragraph as a disclaimer.

Piecing apart the specific dependent variables that differ by person, word, and person*word, post-hoc pairwise comparisons showed that mother ($M_{Mother-Case}$ = .239 µV; p = .014), and Fallon ($M_{Fallon-Case}$ = .281 µV; p = .002) conditions had significantly larger frontal ERP areas than case, and self trended in the same direction ($M_{Self-Case}$ = .189 µV; p = .078) (Table E4). For frontal latency, self again trended ($M_{Self-Case}$ = -43.60 ms; p = .055) in the direction of mother, where latencies were shorter for mother than case likely due to the smaller amplitude and wider width of the case ERP ($M_{mother-Case}$ = -45.55 ms; p = .049). Centroparietal area showed significant differences by case for mother ($M_{Mother-Case}$ = .240 µV; p = .025) and Fallon ($M_{Fallon-Case}$ = .251 µV; p = .009) with near-trend-level differences for self ($M_{Self-Case}$ = .217 µV; p = .101). Areas were generally larger for person (self, mother, Fallon) than case. For word conditions, character significantly differed from appearance in frontal and centroparietal latency such that character words had shorter 50% area latencies (frontal $M_{Character-Appearance}$ = -18.81 ms, p = .001; centroparietal $M_{Character-Appearance}$ = -20.57 ms; p = .001; Table E5).

Looking exclusively at centroparietal area, pairwise comparisons of person*word showed that differences from case were greater for appearance judgments of mother ($M_{Mother-Case}$ = .360 µV; p = .004), and Fallon ($M_{Fallon-Case}$ = .309 µV; p = .008) than character judgments ($M_{Mother-Case}$

= .120 µV; p = .088; Fallon ($M_{Fallon-Case}$ = .194 µV; p = .089) with a trend in the opposite

direction for self (p = .081; Table E6). Again, this is only a univariate effect. According to

pairwise comparisons, no differences between self, mother, and Fallon were significant among

ERP variables.

Induced and Evoked Power

During assumptions testing, spearman's rho correlations between dependent variables

showed correlations < .90 for all induced and evoked variables. Induced low and high

centroparietal gamma had the largest correlation followed by induced centroparietal delta and

theta (Spearmans rho = .891; Spearman's rho = .887; Appendix F, Table F1). Both centroparietal

gamma variables are retained, however it is possible that a split between low and high gamma

frequencies is unnecessary. Sphericity is less of an issue for multivariate analyses, but the

assumption was tested out of transparency. Mauchly's test revealed a violation of the sphericity

assumption for all induced power variables by person (p < .004) except for frontal theta, alpha,

beta, and centroparietal alpha (p > .097). For evoked power variables, sphericity was only

violated for frontal delta (p = .032) and centroparietal delta (p < .001) and beta (p = .014).

Sphericity held all of person*word (induced: p > 110; evoked: p > .056). With the presence of

sphericity violation, all univariate results are reported with the Huynh-Feldt correction.

Induced Power. All four multivariate tests for induced power were significant by p < .001

for person (F(36, 369) = 2.48, p < .001, Pillai's Trace = .584, partial η^2 = .195) and person*word

(F(36, 369) = 1.47, p = .045, Pillai's Trace = .375, partial η^2 = .125) but not word (F(12, 33) =

.818, p = .631, Pillai's Trace = .229, partial η^2 = .229; Table F2). This suggests that at least one

dependent variable varies across levels of the independent variable person and the interaction

between person and word conditions. Expanding upon the multivariate within-person test results,

univariate tests show significant effects of person on frontal and centroparietal delta ($p_f = .033$, partial $\eta^2 = .074$; $p_{cp} < .001$, partial $\eta^2 = .159$), low gamma ($p_f = .010$, partial $\eta^2 = .097$; $p_{cp} < .001$, partial $\eta^2 = .185$) and high gamma ($p_f = .013$, partial $\eta^2 = .096$; $p_{cp} < .001$, partial $\eta^2 = .233$). The interaction showed significant differential effects of different levels of person*word on frontal high gamma ($p = .005$, partial $\eta^2 = .092$; Table F3).

One pattern for person-level effects included self, mother, and Fallon having greater low and high centroparietal gamma than case with an additional significance in delta frequencies for self and Fallon but not mother. Additionally, mother had less centroparietal and frontal delta and frontal low and high gamma than Fallon but only differed from self in low gamma. Therefore, low frontal gamma implied a pattern of Mother < Fallon = Self (Table F4).

Post-hoc person*word comparisons are summarized in the table below (Table 1 and more extensively displayed in Appendix F, Table F5). General patterns showed that significant differences were limited to delta and both ranges of gamma (30–60 Hz and 60–100 Hz). Mother's character only differed significantly by self character such that self responses had higher gamma power than mother responses. Self and mother had less high frontal gamma and centroparietal delta than Fallon for appearance, and self alone had less frontal delta than Fallon. Mother had consistently lower power than Fallon in frontal delta and low gamma with marginal significance for character and full significance for appearance. All conditions varied from case such that case had greater centroparietal delta for character and less centroparietal gamma for character and appearance, primarily in high gamma.

The discriminant function coefficients suggested an optimal linear combination of induced variables for person*word, with standardized coefficients fDelta(-.280) + fTheta(.180) + fAlpha(.157) + fBeta (-.173) + fGammaL(.056) + fGammaH(-.912) + cDelta (-.017) + cTheta (-

51

.061) + cAlpha(-.009) + cBeta(-.413) + cGammaL(-.343) + cGammaH(.458). Data showed

potential multicollinearity for all induced power variables except frontal and centroparietal delta,

which means there may not be a unique discriminant solution with inclusion of all variables and

variables may be poorly estimated ($VIF_{delta} \leq 8.338$ and $11.803 \leq VIF_{all} \leq 18.155$; Poulsen &

French, 2001). With that caveat in mind, weights may suggest that low frontal theta and

centroparietal delta and gamma best separates self, mother, Fallon, and case. Again, the

combination is most likely driven by differences in case from self, mother, and Fallon.

Interestingly but in line with pairwise results, the gammas seem to contribute more to the

person*word differences than other frequencies with opposite sign in frontal and centroparietal

regions.

Table 1. Induced power patterns of significance

Induced Power	Character		Appearance	
	Frontal	Centroparietal	Frontal	Centroparietal
Self = Mother < Fallon			*High γ	*δ
Self > Mother	Low γ			
	*High γ			
Self < Fallon			δ	
Mother < Fallon	δ		*δ	
	Low γ		*Low γ	
Fallon > Case			Low γ	*δ
			*High γ	β
Self = Mother > Case		*Low γ		
Self = Mother = Fallon < Case		*δ		
Self = Mother = Fallon > Case		*High γ		*Low γ
				*High γ

Sidak corrected, includes marginally significant (p < .10) and significant (p < .05) differences
with p < .05 denoted by *. δ = delta, β = beta, low γ = low gamma, high γ = high gamma.

Evoked Power. Multivariate output revealed significant differences for person (F(36, 369)

= 1.51, *p* = .034, Pillai's Trace = .385, partial η^2 = .128) but not word (F(12, 33) = .556, *p* = .861,

Pillai's Trace = .168, partial η^2 = .168) or person*word (F(36, 369) = 1.31, *p* = .113, Pillai's

Trace = .341, partial η^2 = .114; Table F2). Since at least one dependent variable varies at levels of person, univariate within-subjects tests were investigated.

Using a Huynh-Feldt correction, univariate analyses revealed a significant difference for centroparietal delta (p = .010, partial η^2 = .091). No other significant univariate differences were found for person, but there were marginally significant findings for frontal delta (p = .058, partial η^2 = .057), theta (p = .096, partial η^2 = .047), and low gamma (p = .081, partial η^2 = .050; Table F3).

Post-hoc pairwise comparisons showed more specifically that mother had more frontal delta (p = .016), centroparietal delta (p = .010), and marginally higher frontal low gamma (p = .053) than Fallon. No significant differences were found for self or case (Table F4).

Discriminant function coefficients suggested that evoked power best discriminates person by fDelta(.039) + fTheta(.073) + fAlpha(.430) + fBeta (.308) + fGammaL(.239) + fGammaH(.233) + cDelta (.651) + cTheta (-.915) + cAlpha(.426) + cBeta(-.790) + cGammaL(-.223) + cGammaH(-.035). Tests of multicollinearity indicated no violation ($VIF \leq 3.845$). Centroparietal theta, beta, and frontal and centroparietal alpha had the largest contribution to differences in person*word, which is informative considering that theta, alpha, and beta were found to differentiate self and other in EEG studies by Mu and Han (2010) and Park et al. (2018).

Source Localization

In source space, the cluster-based permutation test revealed significant (α = .05) two-tailed differences at 300 ms, the P300 latency, and 300-1000 ms, the LSW range, between the following conditions: Fallon and case character (P300 only), case character and appearance, self and case appearance, mother and case appearance, and Fallon and case appearance ($p < .025$ on each tail). Raising alpha to .10 for a significance level of .05 at the tails revealed an additional

difference between Fallon character and appearance LSW and new trend-level source activity in the aforementioned pairings. T-maps are plotted in Figure 6 and Figure 7 without masking for significance; significant regions were darkened using an overlay of images masked by $p < .025$ and marked with an asterisk. Image layering was done in GIMP 2.10.0 x64 for Windows. Since atlases cannot be used in conjunction with surface plots in Fieldtrip, conclusions about region-specific variation among conditions are cautiously considered. Sample statistics were chosen for clustering when less than critical alpha = .05. Every pairwise comparison had clusters, but few were significant at the cluster level.
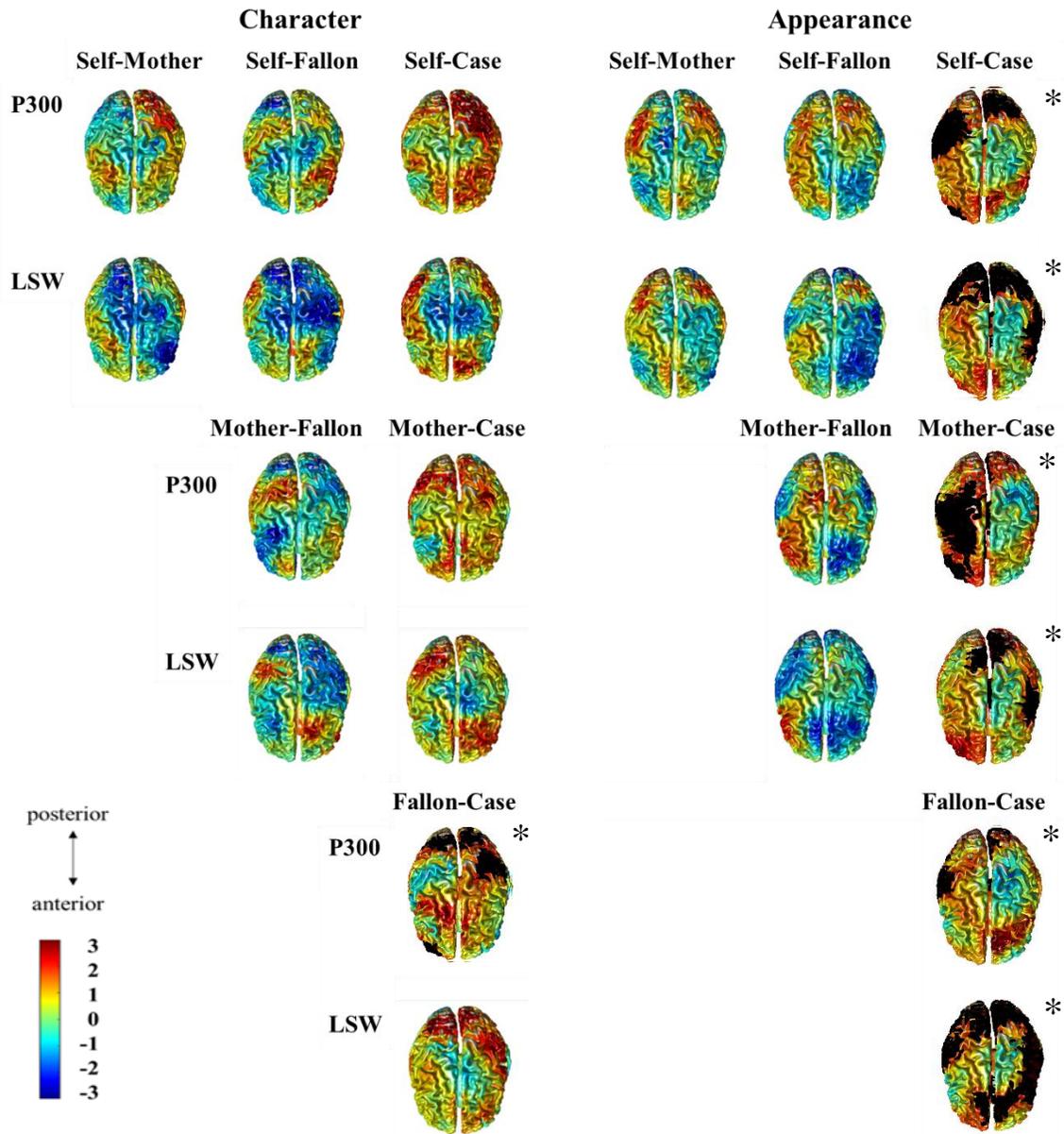
Figure 6. Plots of t-values in positive and negative clusters using Fieldtrip's ft_sourcestatistics with cluster-based permutation. The frontal lobe points downwards. Significant differences are denoted with an asterisk and a black overlay from images masked by $p < .025$, which were always in the positive two-tailed t direction.
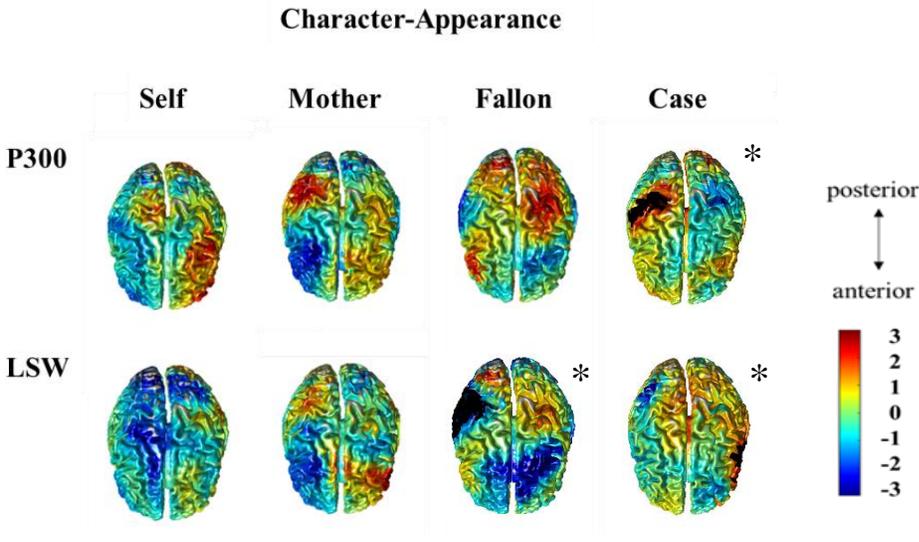
Figure 7. Unmasked t-maps for self, mother, Fallon, case, with character-appearance activity. Case P300 and LSW was significantly different between character and case in the positive *t* direction denoted with an asterisk and darkened overlay. Fallon had a marginally significant (*p* < .10) LSW difference for character-appearance in the right temporal lobe in the negative *t* direction.

HSMM-MVPA

Initial HSMM-MVPA log-likelihoods were highest for the 3-bump self and mother model and the 4 bump Fallon model. LOOCV showed optimal fit for the 4-bump Fallon model for 16 subjects and the 3-bump self and mother model for 12 and 23 subjects over all 1-8 bump alternatives (Figure 8). Log-likelihoods were similar for 3 and 4 bumps in each condition, therefore a sign test determined whether the 3 or 4 bump model had the best proportion of fit by subject for each person condition. For the MVPA application, the value change of the sign test determined whether a larger number of bumps was a logical choice over a more parsimonious model. The smallest positive or negative value was used for the test statistic to determine the probability of finding a value equally or more extreme. According to results for self and mother, 3 bumps had larger log-likelihoods than 4 bumps for 30 (*p* = .018) and 33 subjects (*p* = .001) out of 45, whereas 4 bumps had larger log-likelihoods for 28 subjects in the Fallon condition but did

not reach significance (p = .068). The sign test suggested that self, mother, and Fallon data did not have significant log-likelihood improvement with 4 bumps; the data is best explained by a 3-bump model.
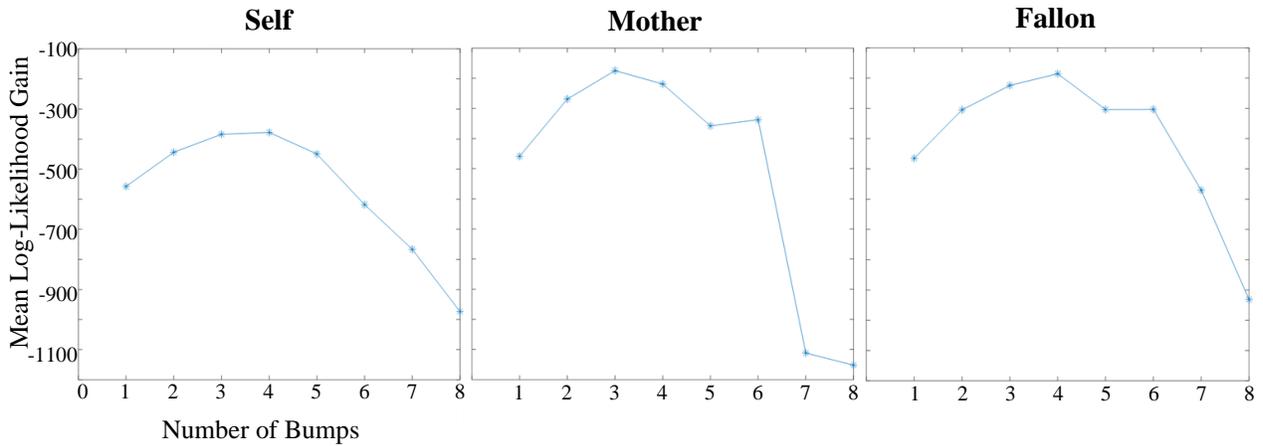
Figure 8. Mean log-likelihood gain of self, mother, and Fallon models using LOOCV

Duration of Stages. Using a repeated measures ANOVA with post-hoc comparisons, Fallon significantly differed from self and mother in mean duration of stages 1, 2, and 4 using times from character and appearance and all stages, subjects, and electrodes (Person*Stage $F_{(6, 264)}$ = 160.92, p < .001, partial η^{2} = .785; Appendix H, Table H1; Figure 10). Self and mother had longer character and appearance latencies than Fallon for stage 1 (p < .001) during pre-cortical processing and shorter latencies thereafter ($p's$ < .022). For stage 3, around the P300 latency, Fallon significantly differed from mother but not self such that character and appearance latencies were shorter for mother (p < .047). Additionally, self differed from mother in stages 2 and 3 for character but not at all for appearance.
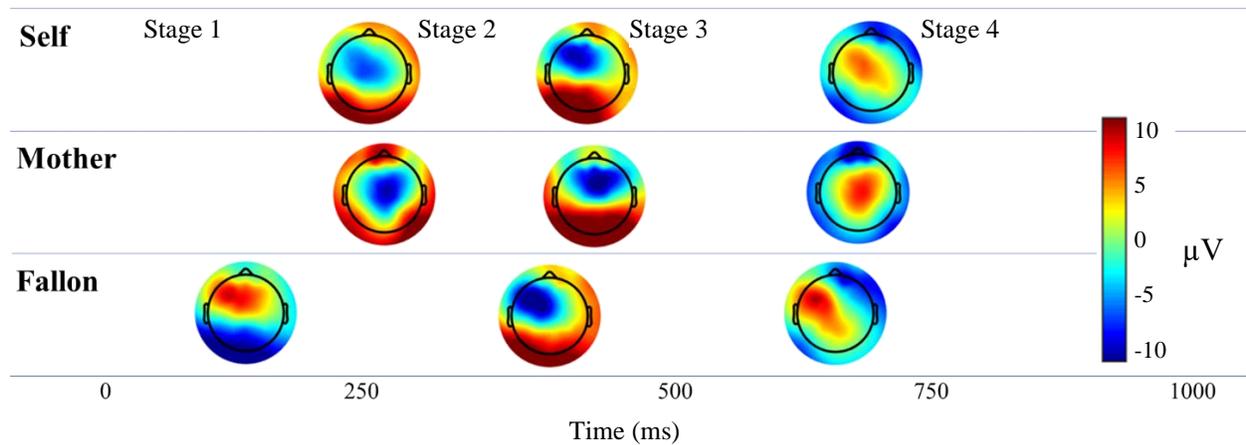
Figure 9. Bump topographies and stage durations for the 3-bump self, mother, and Fallon models collapsed across character and appearance.
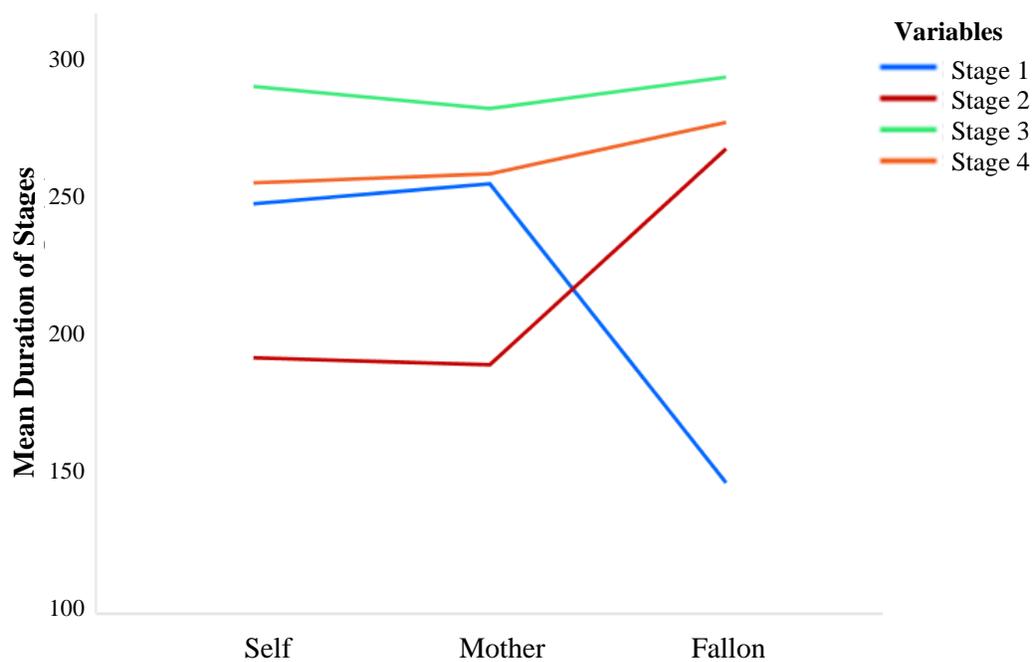
Figure 10. Mean duration of stages for 3-bump models of self, mother, and Fallon

Magnitude of Peaks. Using bump magnitudes averaged across the 10 PCA components, differences between person (self, mother, Fallon) and bump were assessed with a repeated measures Sidak-corrected ANOVA. Within-subjects effects of person, bump, and person*bump were significant at $p < .001$. The pairwise comparisons of person*bump showed that all

combinations of self, mother, and Fallon differed in bump magnitudes across all three bumps (p < .001; see Appendix H, Table H2). Estimated marginal means of self were significantly greater for mother and Fallon for bumps 1 and 2 and smaller for 3. Mother generally had smaller magnitudes than self but greater magnitudes than Fallon for bumps 1, 2, and 3. Fallon had the smallest bumps for 1 and 2 and the largest for 3. Bumps 1 and 2 suggested a Fallon < mother < self pattern of magnitudes, while bump 3 indicated self < mother < Fallon.

Electrode-Specific Activity. Using a single electrode from frontal and parietal areas (FZ and PZ, respectively) revealed that self and mother had more similar EEG activity than Fallon at FZ given the large positive frontal spike for Fallon around 150 ms (Figure 11). The first bump for parietal self around 250 ms broke a pattern from the others as well, but no test was used for statistical comparison.
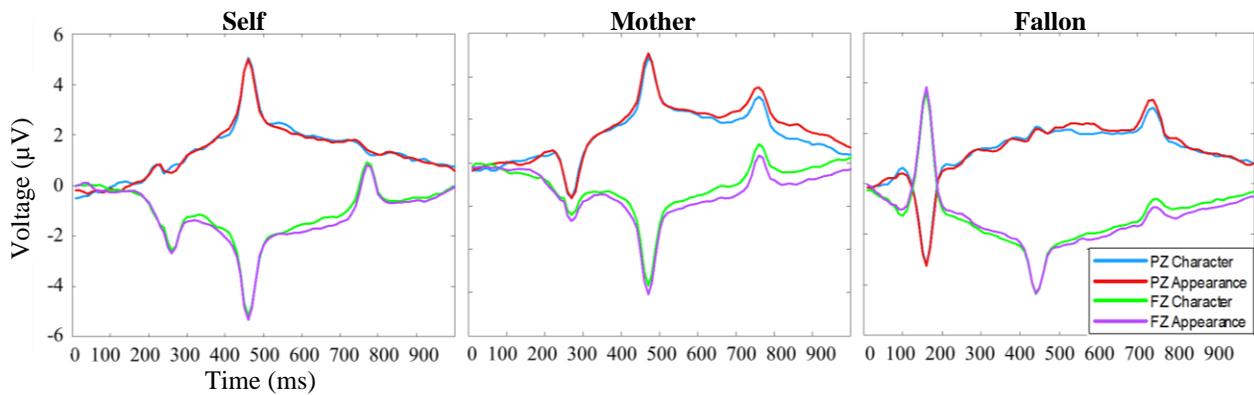
Figure 11. Character and appearance data for self, mother, and Fallon from frontal electrode FZ and parietal electrode PZ after adjusting to the average maximized likelihood bump locations across trials

Chapter 5: Discussion

<u>Mother judgments had faster RTs than self, Fallon judgments</u>

Compared to Moran et al. (2011), behavioral results show the same main effects for word and person and person*word interaction. Consistent with past findings, appearance words had shorter latencies than character words and the reaction time for distant other (Bush before, Fallon now) was significantly different from the close other (mother), where Fallon > mother. Moran et al. (2011) also found a difference between Bush and self (Bush > self), which held true for the current data (Fallon > self). Unlike Moran et al., there was an additional difference between self and mother such that mother had shorter latencies for character (M_{self} = 1146.03 ms; M_{mother} = 1120.25 ms) and appearance (M_{self} = 1099.50 ms; M_{mother} = 1080.06 ms) words than self, although Moran's et al.'s data trended in the same direction ($M_{self,char}$ = 1305.15 ms, $M_{self,appear}$ = 1267.24 ms; $M_{mother,char}$ = 1299.51 ms, $M_{mother,appear}$ = 1259.01 ms; 2011). Case was not present in the original study, but the significant difference between case and all other person conditions (self, Fallon, mother) was expected. Inclusion of the case condition thwarts concerns that person effects are simply due to differences in word length or concreteness between appearance and character words (Binder, Desai, Graves, & Conant, 2009; Moran et al., 2011).

ERP results in the next section address determine whether differences in reaction time are due to variation in processing in the 300-1000ms window. Interpretation concerning simulation and theory theory are discussed.

<u>ERPs Did Not Distinguish Person Conditions beyond Case</u>

Results showed that self had smaller frontal area than both mother and Fallon and an overall mean closest to the case condition. Additionally, mother and Fallon significantly differed from case with mother having a more similar mean to case than Fallon. Taken together, the

means of frontal area suggested a case < self < mother < Fallon pattern. The same pattern of means applied to centroparietal area, where case < self < mother < Fallon. However, self did not reach significance in the self-case comparisons, unlike mother-case and Fallon-case (Appendix E, Table E4).

For latencies, significance was only found in the frontal region for mother-case, with a trend for frontal self-case, suggesting that behavioral reaction time differences across person conditions were not due to timing of frontal or centroparietal activity. In summary, case ERPs had smaller area amplitudes than self, mother, and Fallon, which was expected (Figure 12). Both frontal and centroparietal regions had a trend of area means that suggested case < self < mother < Fallon, although significance was not reached for self-case, self-mother, self-Fallon, or mother-Fallon comparisons.

For word, character and appearance had an interesting dynamic, where character judgements had shorter latencies than appearance judgments in frontal and centroparietal regions (Table E5). It is likely that the narrower shape of the character ERP pulled the 50% area line earlier in time than the wider appearance ERP. This points to greater trial-by-trial variation in response to appearance judgments than character judgments. In other words, character judgments were more difficult overall—judging by their longer reaction times—but were more consistent in the degree of difficulty per judgment. Some appearance judgments were likely very straightforward (e.g., is Fallon brown-eyed?), while others were more challenging (e.g., is Fallon drab?). It is also possible but less likely that shifts in latency for character and appearance occurred outside of the specified regions or later in time post-LSW.

Although initial glances at the ERP results seem to suggest a simulation interpretation given the non-significance between self, mother, and Fallon in latency and area, it is important to

address other possibilities. The first issue is that area amplitude and area latency measures as defined may not have the specificity to pick up on group differences found in previous studies, which include differences in BOLD activity between self, mother, and Fallon. Variations that may exist are obscured by taking the overall area and disregarding other aspects of ERP shape. Additionally, it is possible that the largest group differences occur in the tail through 1500 ms (Figure 12). The pitfall of extending the range that wide is that it may capture unwanted activity beyond the LSW.
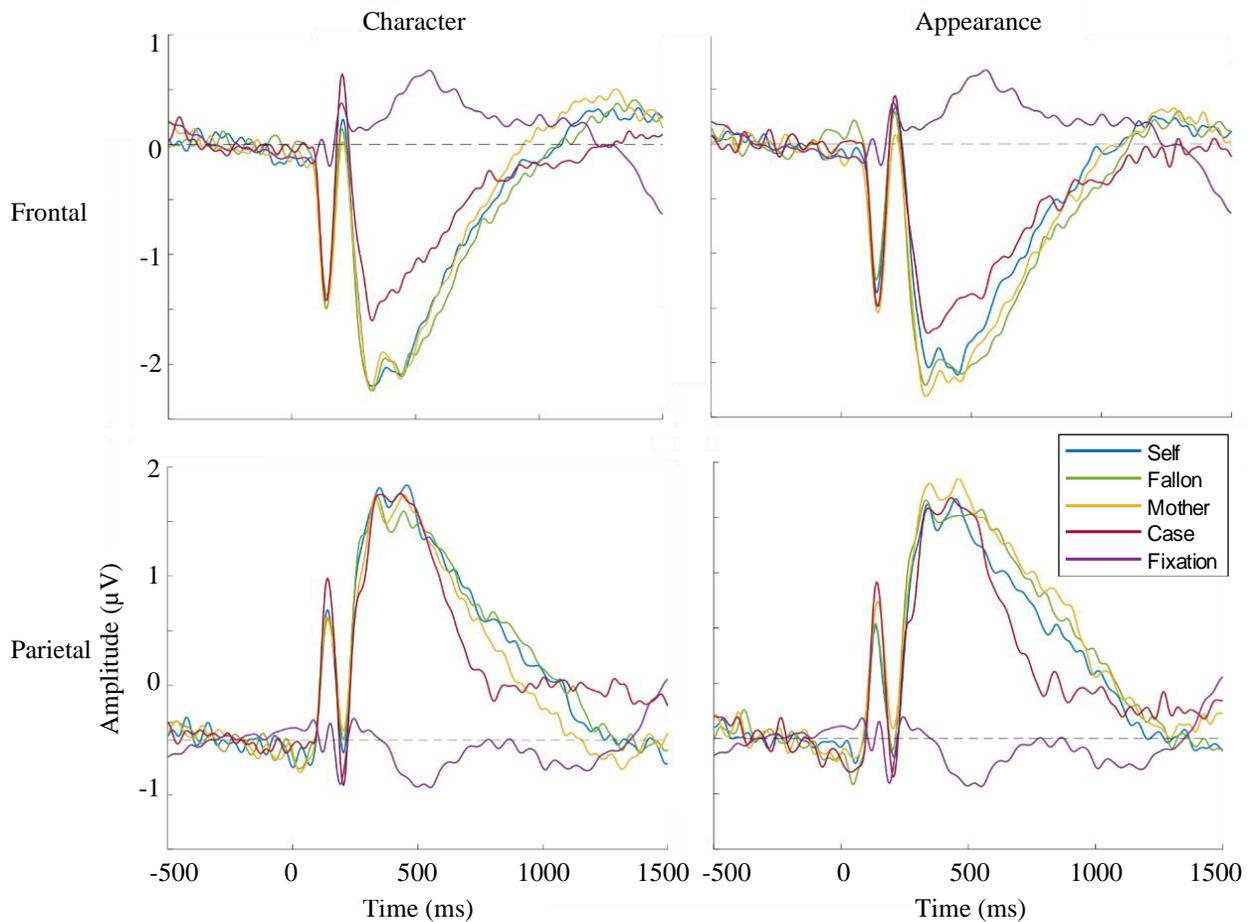
Figure 12. Frontal and centroparietal ERPs by person and word condition. Zero marks stimulus onset. Fixation is included for comparison of inter-stimulus activity with a visual cue to activity during the judgment task.

Beyond issues with ERP operationalization, an alternative explanation to the findings is that there were simply few, if any, real significant differences by person and word condition in the dataset. Past studies suggest otherwise (Heatherton, Wyland, Macrae, Demos, Denny & Kelley, 2006; Mitchell, Neil, & Mahzarin, 2006; Moran et al., 2011), although this particular sample of college-aged individuals may be an exception. Box plots did show outliers in the data, which are known to skew results of multivariate GLMs (Peña & Prieto, 2001). However, the outliers were not extreme enough to warrant removing and potentially ignoring real and important individual differences in cognitive processing and associated effects on neural activity.

Lastly, like power data, ERP variables were corrected for multiple comparisons using Sidak corrections, which are less conservative than Bonferonni corrections but still very conservative when tests are not independent (Abdi, 2007). The correction may be too penalizing for the EEG data, considering that tests are within-subjects and multivariate. A correction such as the Benjamini-Hochberg procedure may be a more reasonable approach to correct false discovery rate over the more traditional Bonferonni or Sidak-based corrections, where p-values less than the critical value (i/m)Q with i as rank, m as number of tests, and Q as false discovery rate (McDonald, 2014). As a final check to determine whether differences may exist in different electrode subsets, a multiplot figure was created to view grand average ERPs for each condition (person*word) across all 64 channels (Appendix E, Figure E1-Figure E2). After thorough investigation, differences in reaction time cannot be explained by timing or degree of cognitive processing between person conditions using ERP data.

Since ERPs were not significantly different by person aside from the control condition, ERP results appear to follow simulation theory over theory theory. This suggests that all person conditions are undergoing the processing steps through identical neural structures on similar

timescales. However, this also suggests that case, self, mother, and Fallon judgments had similar ERP area amplitudes and latencies in the frontal and centroparietal regions that showed group differences with fMRI. Overall BOLD concentrations differed by person condition across entire trials and in more precisely defined regions than EEG, while the electroencephalographic activity did not reliably differ across time. Information from frequency data, source space, and multivariate pattern analysis must be evaluated to determine whether simulation theory is plausible and consistent with all avenues of analysis in the study. In short, other avenues do not suggest simulation.

Instead, the lack of findings between self, mother, and Fallon and inconsistency with fMRI findings suggest a large degree of variability across trials and individuals. High variation is problematic for the ERP approach. By taking averages across epochs, differences between trials are masked in wideset ERPs that limit detection of change by condition. Essentially, there was too much trial-to-trial and person-to-person variability for ERPs to be effective at identifying significant differences across conditions to the same degree as fMRI BOLD concentrations. A technique that relies on single-trial data and preserves between trial variation such as HSMM-MVPA is more suitable for this trait judgment task and for other EEG tasks requiring higher-order cognitive processing with a variable timescale.

Induced and Evoked Power in Delta, Gamma Differentiated Person Conditions

Induced power. With regard to induced power, self, mother, and Fallon had significantly higher centroparietal delta power than case in character responses as well as less centroparietal gamma for character and appearance. Low frequency non-phase-locked oscillations such as delta arise from coherent activity between populations of neurons with an essential role in global processing and integration across neural regions (Mu & Han, 2010; Knyazev, 2012). Data

64

supports a related integration account, in which the self-relevance of information is evaluated by a subregion of the mPFC, which is functionally connected to and communicating with task-relevant regions like the PCC/precuneus (Cooper, Bassett, & Falk, 2017).

Higher frequencies such as gamma, on the other hand, correspond to optimization and global processing on a shorter timescale, allowing for a dynamic system of fast and slow oscillations constraining one another to coordinate activity across regions for required task processing. Delta modulates gamma in the hierarchy and has an essential function for slower, more complex computation such as required by character judgments or with increased memory load (Knyazev, 2012). Additionally, delta is associated with the P300 component, which suggests that delta activity may be related to attention, memory, and complex integration of information. The increase in centroparietal delta and corresponding decrease in centroparietal gamma comparing case to self, mother, and Fallon suggests that both frequency bands are part of a modulatory system for integrating social information that is absent or lessened in basic word processing. Frontal delta, high gamma, and centroparietal low and high gamma were chosen by the discriminant function. This suggests that these dependent variables maximally contribute to person distinctions.

Beyond case, self had enhanced induced power over mother in gamma bands (30-60 Hz and 60-100 Hz), in agreement with Mu and Han (2010). Self and mother both had less frontal high gamma and centroparietal delta than Fallon for appearance, which may suggest that Fallon appearance required a larger degree of information integration or complex processing than self and mother, considering that there is less information about Fallon on which to judge. In general, self and mother required less delta and gamma than Fallon for appearance and marginally for character judgment, but with self having higher gamma than mother for character judgments.

Researchers have proposed a developmental advantage of having an earlier recognition of mother over self which may explain the strong overlaps between mother and self concepts (Damon & Hart, 1982; Knyazev, 2013). Perhaps the concept of mother is a robust, relatively unchanging mental model, while the concept of self is ever-changing and, therefore, more complex to judge.

Evoked power. Mother was shown to have more frontal and centroparietal delta power than Fallon. Neither self or case differed from mother or Fallon in any evoked frequency bands. Considering that evoked power is more often related to stimulus processing (Knyazev, 2013), mother may have required more initial processing than Fallon. Although not used in the univariate or post-hoc comparisons, the discriminant function coefficients reflected past findings; frontal alpha and centroparietal delta, theta, alpha, and beta had the highest weights in the linear combination. Each of these frequency bands was expected to significantly distinguish self and other. This was not the case for significance, but the weights suggest an inherent importance of these frequencies, primarily in centroparietal theta. For instance, theta is tied to memory such that self has enhanced encoding and retrieval of episodic memory and an increase in memory load compared to self-irrelevant information (Doppelmayr, Klimesch, Schwaiger, Stadler, & Rohm, 1999). Although self did not differ from other in evoked power, the relative importance of theta and memory in differentiating person conditions in the centroparietal region is highlighted.

Summary. Overall, there were 12 induced and 12 evoked power variables (frontal and centroparietal; delta to high gamma) with several significant pairwise differences between person and word groups (see Appendix F, Figure F1 for induced and evoked power by person, word, and region; Table F4 and Table F5 for comparison tables). Ideally, the case condition would narrow down the most relevant relationships; if each person condition (self, mother, Fallon)

66

significantly differed from case, then differences between people would more likely be due to cognitive processing of person rather than spurious or person-irrelevant differences among groups. Consistent differences from case occurred for induced power of centroparietal delta for character judgments and centroparietal gamma for character and appearance judgments. No differences from case were found for evoked power. Timing of power provided a single new insight into the self character and mother character differentiation; self had higher induced gamma power than mother.

Source Differences Did Not Reliably Differ Between Self, Mother, and Fallon

Because of unknowns with localization accuracy and multiple comparisons, despite corrections, it is not appropriate to claim that a specific region significantly differs from group A to group B using cluster-based permutation (Hauk, 2004). The inverse problem suggests that there are infinite localization solutions, and although MNE is an inverse solver, the true underlying source activity may not be accurately captured. Therefore, results are described according to overall significant differences between groups and trends of difference in certain broad source regions.

Character-Appearance. Character activity had more nuanced differences than appearance across person conditions. In general, self, mother, and Fallon showed appearance differences from case, but only Fallon showed differences from case with character judgments. It is likely that the ERP data underlying the source activity does not have enough power to separate self and mother from case. That is, Fallon is differing from case more than self and mother but does not clarify whether self and mother have the same magnitude of activity. Additionally, case and Fallon are the only conditions to have differences across character and appearance. Case showed differences for P300 in the right centroparietal region, where character had increased source

67

activity opposite of the character-appearance findings for self, mother, and Bush in Moran et al. (2011). For LSW, case showed differences around the left temporal lobe, while the Fallon LSW trended toward less source activity in the right temporo-parietal region for character compared to appearance, as expected from prior BOLD findings.

Case. Generally, case had less source power than the other conditions judging by the source plots of t-values (Figure 6). Significant differences were primarily in appearance, where all person-responses (self, mother, and Fallon) differed from case, and only Fallon differed in character. Self character judgments appeared to have less activity in the centroparietal region for the LSW than case, although no overall differences were identified between self and case character at the cluster-level.

Examining source movies that show source power across time, positive frontal activity was very low in source power compared to other regions and often appeared in short bursts. Since source localization was computed on EEG data with known limitations in source resolution, it is possible that dorsal and ventral mPFC activity requires greater source power resolution to capture self and other differences. One reason that EEG in particular is limited in this sense is that the electrical signal underlying EEG is distorted as it passes through tissues to the electrode. Even with source localization of the EEG data, activity from the ventral and dorsal mPFC measured at the scalp surface could easily be overshadowed or misinterpreted after smearing. To take advantage of the superior timing of EEG with the required source specificity of fMRI, a simultaneous EEG-fMRI recording may be required to capture the intricacies of self-other differences in a judgment task.

68

HSMM-MVPA Distinguished Self and Other by Stage Durations and Bump Magnitudes

For the final phase of analysis, a sign test between 3- and 4-bump HSMM-MVPA models chose an optimal 3 bumps for self, mother, and Fallon conditions. Testing for timing and magnitude differences between the three models revealed significant differences between self, mother, and Fallon that were not found in other analyses of this data. Given differences between self, mother, and Fallon in timing and magnitude, the results are interpreted in terms of the proposed theory theory model (Figure 2).

Duration of stages. Self and mother differed from Fallon, such that self and mother had later character and appearance latencies than Fallon for stage 1 ($M_{self,character}$ = 248 ms; $M_{self,appearance}$ = 255 ms; $M_{mother,character}$ = 256 ms; $M_{mother,appearance}$ = 261 ms; $M_{Fallon,character}$ = 148 ms; $M_{Fallon,appearance}$ = 149 ms), which coincides with what was seen between mother and Fallon in evoked delta power. Considering that Fallon has a nearly 100 ms difference from self and mother in stage 1 and a very different topographical distribution, the Fallon model is likely selecting an entirely different processing stage than self and mother. Since stage 1 represents the time for the signal to reach the brain and undergo pre-cortical processing (Anderson et al., 2016), the Fallon model is likely picking up a visual N1-like sensory processing stage around 150 ms, while self and mother models are first picking up a P200 or P300-like stage beginning around 250 ms.

Although the stage 2 window for self and mother could correspond to either a P200 or P300-like peak, the specific nature of the stage is unknown. Both P200 and P300 have been tied to selective attention to stimulus features, information categorization, and sensitivity to self-referencing or self-relevance (Gray, Ambady, Lowenthal, & Deldin, 2004; Polich, 2007; Yu, Tu, Wang, & Qiu, 2010). The components are similar in theorized function, but the P200 is thought to reflect pre-perceptual processing, whereas the P300 reflects post-perceptual processing with

strong ties to both attention and working memory load (Gray, Ambady, Lowenthal, & Deldin, 2004; Wongupparaj, Sumich, Wickens, Kumari, & Morris, 2018).

With the function of P200 and P300 in mind, it is possible that as attention and self-relevance increase, the P200-like bump is augmented. In the case of self and mother information, attention and self-relevance would lead to a much larger peak for self and mother than for Fallon. It is possible that the modulation of the peak may be overshadowing the importance of an earlier stimulus processing bump for self and mother. For Fallon, the P200-like bump would be dampened by the degree of self-relevance lacking for a distant other, according to theory theory. In such a case, the Fallon model would have a small peak that may be overshadowed by stimulus processing peaks and explain the differences between the Fallon model and self, mother models. Therefore, stage 2 of the self and mother models may be described as an attention allocation or information categorization stage for which self had later latencies than mother. For Fallon, stage 2 is classified as early stimulus processing for the N1-like topography and latency.

When allowing a 4-bump model, self, mother, and Fallon topographies become more consistent (Figure 13). The first bump for self and mother models shifts to a topography and latency resembling the visual N1, introducing a standalone stimulus processing stage for self and mother. This would suggest that in a 3-bump model, the N1-like bump representing early stimulus processing is more salient than memory retrieval or probabilistic reasoning for the Fallon model, while the P200 or P300-like bump is more representative for self and mother. Additionally, all models gain a bump around 500 ms with 4 bumps, splitting stage 3 into stages 3 and 4. If the 4-bump model were accepted, the additional bump would suggest that the theory theory model is missing a stage. The probabilistic reasoning stage may be best described by separate reasoning and binary decision making processes (Johnson-Laird & Shafir, 1993).

3-Bump Model
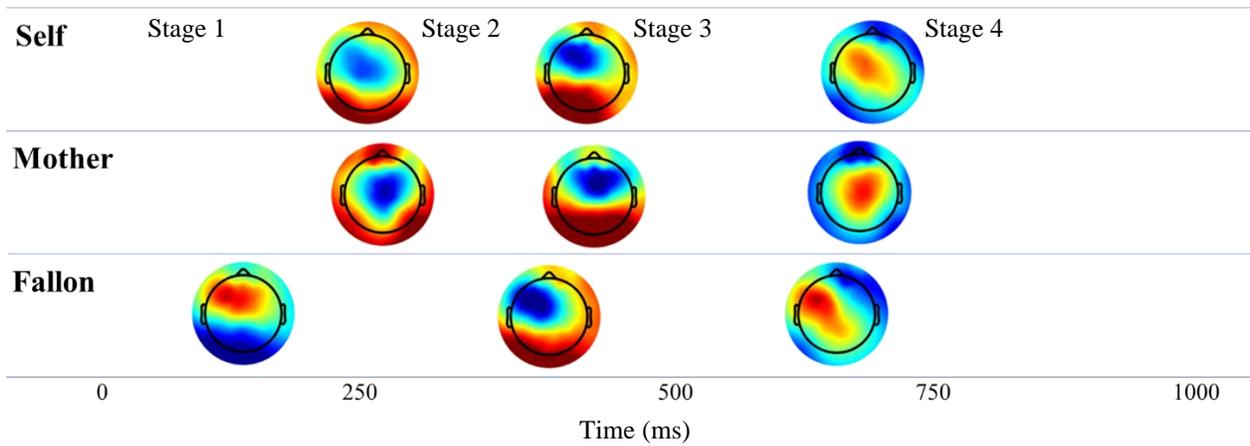
4-Bump Model
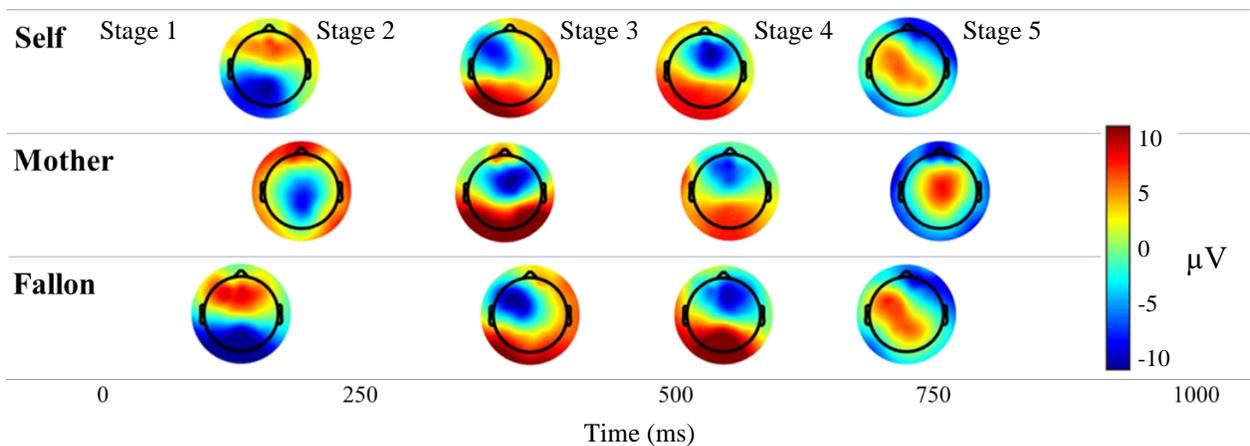
Figure 13. Topographies of self, mother, and Fallon for 3 and 4-bump HSMM-MVPA models. See Appendix H, Figure H2 for the 5-bump model topographies. The 3, 4, and 5-bump models have considerable similarity to the topoplots in Figure H1.

Beyond early sensory processing, self and mother had shorter latencies than Fallon for information retrieval (mother-Fallon only) and probabilistic reasoning stages by character and appearance, corresponding to time ranges 450-750 ms and 750-1000 ms, respectively. Information retrieval was best described by the pattern mother < self < Fallon, where mother differed from self and Fallon but self and Fallon did not differ from one another. The retrieval

71

stage may reflect LSW processing, which has been associated with episodic memory and extended working memory processes (Gray et al., 2004; Voss & Paller, 2017). Self also had later character latencies than mother for stages 2 and 3, which correspond to attention and categorization-based processing and memory retrieval stages around 250-450 ms for stage 2 and 450-750 ms for stage 3.

Differences in character response between self and mother but not Fallon could be explained by not having enough information about Fallon to make a reasonable probabilistic judgment. Consistent with theory theory, autobiographical memory may be engaged during judgment of mother but not Fallon or it may be engaged to a lesser degree for Fallon. This follows Rabin and Rosenbaum's study, in which they found evidence of self-referential activity in the mPFC during a mentalizing task of a personally close but not distant other (2012). In distant other situations, it is possible that self is used as a prototype or exemplar but processed in a probabilistic fashion (Kuiper, 1981; Karylowski, Konarzewski, & Motes, 2000). Essentially, in line with reaction time data, there is a U-shaped curve for using self as prototype in distant-other judgment. However, instead of faster reactions for self, the fastest reactions were for mother. By using self as a prototype or other schema in general, Fallon judgments are based less in autobiographical memory as self and personally close others and more on semantic and schematic memory that would explain differences between self and mother vs. Fallon (Oschner et al., 2005).

Character	Appearance
Stage 1 Fallon < Self = Mother	**Stage 1** Fallon < Self = Mother
Stage 2 Mother < Self < Fallon	**Stage 2** Mother = Self < Fallon
Stage 3 Mother < Self < Fallon	**Stage 3** Mother = Self < Fallon
Stage 4 Self = Mother < Fallon	**Stage 4** Mother = Self < Fallon

Magnitudes of Peaks. Bump magnitudes were averaged across all 10 PCA components and compared across person (self, mother, Fallon) and bump, collapsed across character and appearance. The within-subjects results showed that person, bump, and person*bump all had significant differences. All pairwise comparisons of self, mother, and Fallon at bumps 1, 2, and 3 were significant. Bumps 1 (information gathering and early processing) and 2 (memory retrieval) had a pattern of Fallon < mother < self, while bump 3 (probabilistic reasoning) indicated self < mother < Fallon as hypothesized.

> **Bump 1** Fallon < Mother < Self
> **Bump 2** Fallon < Mother < Self
> **Bump 3** Self < Mother < Fallon

In summary, differences in timing and magnitude for Fallon may be rooted in a greater reliance on semantic and schematic memory as opposed to autobiographical memory for personally close others and self (Rabin & Rosenbaum, 2012). For an unfamiliar other with little to no shared information with self, schema and common knowledge, or folk psychology are central to understanding. Having more shared experiences or episodic memories with a more familiar other will drive mentalizing about that other more so than semantic or schematic memory.

Given the number of stages and differences between self and other that were identified in the HSMM-MVPA approach, it is apparent that the theory theory model is more fitting for the data than the simulation model (Figure 1; Figure 2). For bump 3, approximately in the probabilistic reasoning stage of the theory theory model, self, mother, and Fallon significantly differed in peak magnitudes in the hypothesized pattern of ventral mPFC character activation: Self < Mother < Fallon. Durations of stage 4 associated with bump 3 also differed such that character showed a Self = Mother < Fallon pattern and appearance showed Mother = Self <

73

Fallon. The fluctuations of timing and magnitude across bumps and stages may explain why

mother had faster reaction times than self, as well as how influential time windows are on the

ability to distinguish self and mother latencies or activations.

Limitations and Assumptions

Source Localization with Low-Density Nets. Several limitations and assumptions were

encountered in this research and should be carefully considered. For example, the number of

electrodes and quality of data both influence source localization. Although the optimal 128- and

256-channel EEG nets give higher spatial resolution and more accurate localization solutions

than lower density nets, data from 64-channels have the advantage of demanding less memory

and computational power from an already computationally intensive process. Compared to

clinical nets that are often <32 channels, 64-channel nets also provide a reasonable compromise

between precision and time. However, the 64-channel net available for this study had an unequal

distribution of electrodes across the scalp surface, which partially violates best practice via

inadequate spatial sampling and thus may have distorted source solutions. Slight variations in

electrode positions from subject to subject could also contribute to the problem as well as

differences in scalp thickness and geometry, which could only be modeled with individual MRI

head models (Song et al., 2015).

Data quality is always important when artifact has potential to overshadow or bias results.

To improve the quality of solutions, data must be thoroughly pre-processed, including

appropriate filtering, referencing, resampling, ICA, and artifact removal. Data in this study was

cleaned with both automatic and manual artifact detection and ICA labeling using a standardized

Matlab pipeline, which ensures a large degree of consistency and thoroughness of pre-processing

between datasets (see EEG Analysis). However, 30 Hz noise was prevalent in the dataset, and although filtered out with a 30 and 60 Hz notch filter, may have left remnants of noise.

Lack of sensitivity. There are clear limitations in source analysis with using a template head model and MRI from Fieldtrip rather than individualized scans. Additionally, source analysis can be very computationally demanding, and the source model was restricted to a 7 mm source grid to reduce load. EEG data may have poor source accuracy, but the more data simplification, the more likely the source output will be less representative of reality (Cohen & Cuffin, 1991; Liu, Dale, & Belliveau, 2002).

Defining Time Windows. In traditional ERP and time-frequency approaches, there are limitations for knowing when a person is thinking about themselves or others during the character-appearance judgment task. Thus, the accuracy of traditional EEG approaches hinges on having appropriate time windows for capturing self- or other-related thoughts. Judgments may occur outside a window that is too narrow, but wide windows may bias results toward unrelated neural activity. However, individual differences in timing of late cognitive processes like P300 and LSW vary as cognitive events unfold, and subject-specific peaks are ill-defined. Therefore, ERP, time-frequency, and source analysis relied on grand averages to determine the optimal time window for minimizing extraneous data and maximizing the likelihood of capturing the P300 and LSW in every individual. A plot of superimposed individual ERPs verified that no subjects have ERPs occurring outside of the grand average time window.

With the HSMM-MVPA approach, subjective and theory-driven issues like defining time windows are surpassed. Instead of relying on theory, the HSMM-MVPA algorithm used pre-processed EEG data to find bumps that mark new stages of cognitive processing while minimizing potential researcher bias.

Information Processing. One critique of trait judgment tasks is that self, mother, and Fallon information could be processed differently simply because trait judgments of Fallon have not occurred before, requiring more or different reasoning and decision formulation than mother and self information. As a result, a future study should level the differences by requiring judgments about the self, a familiar, and an unfamiliar other in scenarios that have not been directly observed and cannot be easily predicted. In such a case, one cannot rely entirely on memories of oneself or another person to make judgments or predictions in a given situation. Moral dilemmas are an example of such scenarios, since it is impossible to know how someone, including oneself, would respond in situations they have never been in.

Conclusion

Although several measures were ineffective at distinguishing person or word conditions as expected, the present study addressed alternative explanations to Moran's finding that the ventral mPFC is activated for the mother's character in addition to self (2011). (1a) the P300 and LSW amplitudes did not match BOLD findings. The frontal and centroparietal P300/LSW did not show larger amplitudes for character over appearance and did not differ between self, mother, and Fallon in amplitude or latency. However, differences from case in frontal and centroparietal area amplitude suggested a case < self < mother < Fallon trend. The ERP data alone was not enough to suggest one theory of mind theory over another given methodological concerns, data trends, and mismatched EEG to fMRI findings.

(1b) Frequency band power has been highly correlated with BOLD responses, which implies that BOLD and EEG frequency band power should be similar and that differences between self and other may be explained by changes in frequency band power and timing of the EEG. In induced or non-phase-locked power, significant differences occurred primarily in delta

and low and high gamma, such that self had higher frontal gamma power than mother responses, self and mother had less frontal high gamma and centroparietal delta than Fallon in appearance, and self additionally had less frontal delta than Fallon. Although delta was not reported in past studies, gamma has been commonly found to be a differentiator of self and other (Mu & Han, 2010; Dastjerdi et al., 2011; Knyazev, 2013). These findings do not necessarily rule out simulation theory but do suggest that self, mother, and Fallon processing is differing, perhaps in integration or complex processing of information.

(2) P300 and LSW amplitudes did not reliably differentiate the self and other in source localization in the PCC, precuneus, and dorsal and ventral mPFC. The idea that timing discriminability would benefit source localization was impeded by the fact that P300 and LSW could not be easily differentiated in time. A time bin approach would solve this problem. T-maps of the source data did show significant differences between self, mother, and Fallon P3 and LSW and case in appearance. Further source specificity would require changes to the EEG set-up, study design, or use of a supercomputer. The self-other distinction is so nuanced in source space and probably sensor space that EEG generally may not be viable for the study of self-other thought in healthy populations. Neurocognitive disorders like autism spectrum disorder and schizophrenia have been associated with impaired self-recognition and may show more distinct differences between self and other than typically developing individuals in areas tied to impairment (Hobson & Meyer, 2005; Lyons & Fitzgerald, 2013). The ability to use EEG separate from fMRI and attain accurate but low-resolution source localization would be a cost-effective solution for labs with existing EEG setups or lower budgets to study intact and impaired self-referential thought. It is unlikely but possible that an upgrade to a 128 or 256-

channel net may be all that is needed to achieve self-other differences seen in prior research and in the HSMM-MVPA method described below.

(3) The HSMM-MVPA model independently selected 3-bump models for processing self, mother, and Fallon judgments during the self-other trait judgment task. Timing and magnitude findings of the self, mother, and Fallon HSMM-MVPA models fit with the theory theory model of mentalizing. Theory theory predicts that self information would be processed faster and with more initial attentional resources, in line with findings for the early processing and information gathering stages of processing. Theory theory would also predict that self would have less probabilistic reasoning load than other-based processing related to the degree of information available to reason about the self and other, which varies as a function of familiarity and indeed matches findings for the probabilistic reasoning stage of the model.

Topographies of the bumps were reasonably linked to topoplots of ERP data across time (Figure H1). Three distinct patterns appear in the topoplots that are associated with early stimulus processing, information gathering and memory retrieval, and probabilistic reasoning as chosen by the 3-bump self, mother, and Fallon HSMM-MVPA models with slight topographical variation between self, mother, and Fallon models. This final phase of EEG analysis was successful at identifying reliable self-other differences in both time and magnitude, with the added benefit of remaining objective. If mentalizing about others occurs through the probabilistic reasoning process following theory theory, then the odd activation of mother's character in Moran et al. (2011) can be explained by indirect self-referencing through episodic memory retrieval.

Although the decision to settle on simulation or probabilistic reasoning was not obvious across ERP, time-frequency, source analysis, and MVPA methods, the data does suggest that

simulation is not a likely explanation. Across all socio-cognitive processing, it is most probable that both simulation and theory theory are used in some capacity or combination for specific tasks. For example, the trait judgment task used in this study likely initiated direct judgments of self and other based on episodic and schematic knowledge as opposed to reflective judgments drawn from beliefs about how others may perceive us. These first and third person reflections are thought to stem from different underlying processes but with shared activation of the mPFC and PCC/precuneus (Oschner et al., 2005). Simulation may be more likely to occur in the reflective judgment, where judgments are based more on imaginative situations like, "If I were someone else, would I think I was kind?" Addressing that possibility is beyond the scope of this data.

Additionally, a proposed combination simulation-theory-theory perspective suggested that implicit simulation could be a default, fast model of understanding others, while folk psychology provides slower but more complex understanding when there is information about others to work from (Gallagher, 2007). If that were so, distant other judgments would have shorter reaction and processing times, which was not the case.

Future Research. Since participants were healthy and conditions were highly similar, varying only in person and word type, neural differences were small but biologically relevant. To ensure that there is enough power to assess these minute changes with respect to condition, measures must be precise in time, robust to noise, and otherwise properly defined. Using traditional EEG analyses, data could be re-analyzed with a more focused subset of electrodes around the medial prefrontal cortex and PCC/precuneus to address concerns over capturing nearby but irrelevant activity in an overly spread electrode subset. Even without data smearing that happens as pyramidal cell activity travels to the scalp, the selection of electrodes has a great influence on what is defined as frontal activity. The subset chosen here may be misrepresenting

the dorsal and ventral mPFC or PCC/precuneus. Although a single frontal subset was chosen, as opposed to separate dorsal and ventral selection, the P300 and LSW were hypothesized to distinguish ventral and dorsal according to their neural generators and differences in timing. Because the P300 and LSW were indistinguishable in the waveform, the ability to then distinguish areas of the mPFC using that approach was severely limited.

The most ideal solution for addressing both source and timing aspects of self and other differences would be simultaneous EEG and fMRI or a more suitable EEG setup for localization such as higher density nets and precise electrode positions. Having a 128-channel or higher density net would allow for a more selective and precise grouping of electrodes with EEG, and a greater benefit from fitting a larger number of dipoles (Michel & Brunet, 2019). Without collecting new data, the dataset could be investigated as a time series with 10 ms bins of activity across averaged and single trials. Since the original hypotheses involved a curiosity about timing differences between self and other, the time bins may be better suited to address timing changes across the ERP and better able to separate contributions of the P300 and LSW without specifying onset and offset in large time windows.

Differences may also be more apparent in self or other-relevant trials, where the subject answered "yes" that the trait adjective describes the person, as evidenced by (Mu and Han, 2010). Future research could consider the neurophysiological effects of relevance on self-other responses taking familiarity into account. Additionally, the task itself could be adapted into a more traditional theory of mind task, in which participants make judgments regarding an other's expected mental state or predicted response in a scenario. These adjustments would address any concerns over the word judgment task being an oversimplification or limited aspect of

mentalizing in theory of mind. For instance, considering only direct, first-person judgments of self and other and not reflective, third-person judgments.

Neurological Disorders. Having a more comprehensive understanding of the dorsal and ventral mPFC better in a healthy population is an important stepping-stone for understanding abnormalities in these regions associated with self or theory of mind deficits in ASD or schizophrenia such as trouble referencing self when understanding others (Sass and parnas, 2003; Veluw & Chance, 2014). As an example application, ASD involves deficits in autobiographical episodic but not semantic memory that could potentially determine whether unfamiliar others are being differentially judged with semantic as opposed to episodic memory and to what degree that differs for personally familiar others compared to self (Lind & Bowler, 2009). On the contrary, if the degree of semantic and episodic memory contribution in a typically developing population was known for self-other judgment across familiarity, then it would be possible to determine whether individuals with ASD use compensatory processes for impaired episodic memory that may explain difficulty but not complete absence of ability to distinguish self and other or undergo probabilistic theory of mind reasoning.

References

Abdi, H. (2007). The Bonferonni and Šidák Corrections for Multiple Comparisons. In Neil Salkind (Ed.), *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage.

Albiero, P., Matricardi, G., Speltri, D., & Toso, D. (2009). The assessment of empathy in adolescence: A contribution to the Italian validation of the "Basic Empathy Scale." *Journal of Adolescence*, *32*(2), 393–408.

Anderson, Norman H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, 9(3).

Apperly, I. A. (2008). Beyond simulation-theory and theory theory: Why social cognitive neuroscience should use its own concepts to study "theory of mind". *Cognition, 107*(1), 266-283.

Belkaid, M., & Sabouret, N. (2014). A logical model of Theory of Mind for virtual agents in the context of job interview simulation. ArXiv:1402.5043 [Cs]. http://arxiv.org/abs/1402.5043

Billeke, P., Zamorano, F., Cosmelli, D., & Aboitiz, F. (2013). Oscillatory brain activity correlates with risk perception and predicts social decisions. *Cerebral Cortex* doi:10.1093/cercor/bhs269

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19, 2767 –2796.

Blume, C., Lechinger, J., Giudice, R., Wislowska, M., Heib, D., Schabus, M. (2015). EEG oscillations reflect the complexity of social interactions in a non-verbal social cognition task using animated triangles. *Neuropsychologia*, 75, 330-340.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). *The Brain's Default Network*: *Anatomy, Function, and Relevance to Disease*. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38. https://doi.org/10.1196/annals.1440.011

Burle, B., Spieser, L., Roger, C., Casini, L., Hasbroucq, T., & Vidal, F. (2015). Spatial and temporal resolutions of EEG: Is it really black and white? A scalp current density view. *International Journal of Psychophysiology*, *97*(3), 210–220. https://doi.org/10.1016/j.ijpsycho.2015.05.004

Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, 217(4), 783–796. doi: 10.1007/s00429-012-0380-y

Caggiano, V., Fogassi, L., Rizzolatti, G., Pomper, J.K., Their, P., Giese, M.A., & Casile, A. (2011). View-based encoding of actions in mirror neurons of area f5 in macaque premotor cortex. *Current Biology, 21,* 144-148.

Caggiano, V., Fogassi, L., Rizzolatti, G., Their, P., & Casile, A. (2009). Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *Science, 324*, 403-406.

Cartella, F., Lemeire, J., Dimiccoli, L., & Sahli, H. (2015). Hidden Semi-Markov Models for Predictive Maintenance. *Mathematical Problems in Engineering,* 2015, 1–23. https://doi.org/10.1155/2015/278120

Chacon, C. (2009). Depression: The role of cultural factors and perception of treatment. Retrieved from http://scholarworks.sjsu.edu/etd_theses/4000/

Clarke, D. (2005). Sensation: Ideas as brain patterns. In *Descartes' Theory of Mind* (pp.45-77). Oxford, UK: Oxford University Press.

Cohen, D., & Cuffin, B. N. (1991). EEG versus MEG localization accuracy: Theory and experiment. Brain Topography, 4(2), 95–103. https://doi.org/10.1007/BF01132766

Cohen, S. & Hoberman, H. (1983). Positive events and social supports as buffers of life change stress. *Journal of Applied Social Psychology*, 13(2), 99-125.

Cooper, N., Bassett, D. S., & Falk, E. B. (2017). Coherent activity between brain regions that code for value is linked to the malleability of human behavior. *Scientific Reports*, 7(1), 43250. https://doi.org/10.1038/srep43250

Cornew, L., Roberts, T. P. L., Blaskey, L., & Edgar, J. C. (2012). Resting-State Oscillatory Activity in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *42*(9), 1884–1894. https://doi.org/10.1007/s10803-011-1431-6

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, 9(1), 28–30. doi: 10.1038/nn1611

Dastjerdi, M., Foster, B.L., Nasrullah, S., Rauschecker, A.M., Dougherty, R.F., Townsend, J.D., et al. (2011). Differential electrophysiological response during rest, self-referential, and non-self-referential tasks in human posteromedial cortex. *Proc.Natl. Acad.Sci.U.S.A*. 108,3023–3028. doi:10.1073/pnas.1017098108

David, O., Kilner, J. M., & Friston, K. J. (2006). Mechanisms of evoked and induced responses in MEG/EEG. *NeuroImage*, *31*(4), 1580–1591. https://doi.org/10.1016/j.neuroimage.2006.02.034

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85.

Davis, A. L., & Krishnamurti, T. (2013). The problems and solutions of predicting participation in energy efficiency programs. *Applied Energy*, *111*, 277–287.

Delorme, A. & Makeig, S. (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics (pdf, 0.7 MB) *Journal of Neuroscience Methods* 134:9-21

Denny, B.T., Kober, H., Wager, T.D., & Ochsner, K.N. (2012). A meta-analysis of functional neuroimaging studies of self and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *J Cogn Neurosci, 24*(8), 1742-1752.

Descartes, R. (1641). *The Philosophical Works of Descartes.* Meditations of First Philosophy. Meditation II: Of the nature of the human mind; and that it is more easily known than the body. Transl. by Haldane, E. in 1911.

Doppelmayr, M., Klimesch, W., Schwaiger, J., & Stadler, W. (2000). The time locked theta response reflects interindividual differences in human memory performance. *Neuroscience Letters*, 4.

Esslen, M., Metzler, S., Pascual-Marqui, R., & Jancke, L. (2008). Pre-reflective and reflective self-reference: A spatiotemporal EEG analysis. *NeuroImage, 42*(1), 437-449.

Euston, David R., Gruber, Aaron J., McNaughton, Bruce L. 2012. The role of medial prefrontal cortex in memory and decision making. *Neuron*, 76(6), 1057-1070.

Ferrari, P.F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience, 17,* 1703-1714.

Field, T. M., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitation of facial expressions neonates. *Science*, 218(4568), 179–181. https ://doi.org/10.1126/scien ce.7123230

Fincham, J. M., Lee, H. S., & Anderson, J. R. (2020). Spatiotemporal analysis of event-related fMRI to reveal cognitive states. *Human Brain Mapping*, *41*(3), 666–683. https://doi.org/10.1002/hbm.24831

Fox, N.A. & Bell, M.A. (1990) Electrophysiological indices of frontal lobe development. *NY Academy of Sciences, 608*, 677–704.

Fu, G., Xiao, W. S., Killen, M., & Lee, K. (2014). Moral judgment and its relation to second-order theory of mind. *Developmental Psychology*, 50(8), 2085–2092. doi: 10.1037/a0037077

Gallagher, S. (2007). Simulation trouble. *Social Neuroscience*, *2*(3–4), 353–365. https://doi.org/10.1080/17470910601183549

Gallese, V., Gernsbacher, M.A., Heyes, C., Hickok, G., & Iacoboni, M. (2011). Mirror neuron forum. *Sage*, 6(4) 369-407.

Gallese, V. & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Science*, 2, 493-501.

Gallup, G. (1998). Self-awareness and the evolution of social intelligence. *Behavioural Processes*, 42, 239-247.

Garson, G. D. (2015). Multivariate GLM, MANOVA, and MANCOVA. 32. Goldman, A.I. (1992). In defense of simulation theory. *Mind & Language, 7* (1-2), 104-119.

Gevins, A., Le, J., Martin, N. K., Brickett, P., Desmond, J., & Reutter, B. (1994). High resolution EEG: 124-channel recording, spatial deblurring and MRI integration methods. *Electroencephalography and Clinical Neurophysiology, 90*(5), 337–358. https://doi.org/10.1016/0013-4694(94)90050-7

Gopnik, A., & Wellman, H. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds), *Mapping the Mind: Domain Specificity in Cognition and Culture* (257-293). New York: Cambridge University Press.

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, *138*(6), 1085–1108. https://doi.org/10.1037/a0028044

Goswami, U. (2011). *The Wiley-Blackwell Handbook of Childhood Cognitive Development.* 2nd ed. UK: John Wiley & Sons.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Parkkonen, L., & Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *NeuroImage*, *86*, 446–460. https://doi.org/10.1016/j.neuroimage.2013.10.027

Grandchamp, R., & Delorme, A. (2011). Single-Trial Normalization for Event-Related Spectral Decomposition Reduces Sensitivity to Noisy Trials. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00236

Gray, H. M., Ambady, N., Lowenthal, W. T., & Deldin, P. (2004). P300 as an index of attention to self-relevant stimuli. *Journal of Experimental Social Psychology*, *40*(2), 216–224. https://doi.org/10.1016/S0022-1031(03)00092-1

Greene, Joshua D., Sommerville, Brian R., Nystrom, Leigh E., Darley, John M., Cohen, Jonathan D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293, 2105-2108.

Gretenkord, S., Rees, A., Whittington, M.A., Gartside, S.E., & LeBeau, F.E.N. (2017). Dorsal vs. ventral differences in fast up-state-associated oscillations in the medial prefrontal cortex of the urethane-anesthetized rat. *Journal of Neurophysiology, 117*(3), 1126-1142.

Griffiths, T. L., Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2011). Bayes and Blickets: Effects of Knowledge on Causal Induction in Children and Adults: Cognitive Science. *Cognitive Science*, *35*(8), 1407–1455. https://doi.org/10.1111/j.1551-6709.2011.01203.x

Halder, T., Talwar, S., Jaiswal, A. K., & Banerjee, A. (2019). Quantitative Evaluation in Estimating Sources Underlying Brain Oscillations Using Current Source Density Methods and Beamformer Approaches. *Eneuro*, *6*(4), ENEURO.0170-19.2019. https://doi.org/10.1523/ENEURO.0170-19.2019

Hauk, O. (2004). Keep it simple: A case for using classical minimum norm estimation in the analysis of EEG and MEG data. *NeuroImage*, 21, 1612-1621.

Heatherton, T. F., Wyland, C. L., Macrae, C. N., Demos, E. K., Denny, B. T., & Kelley, W. M. (2006). Medial prefrontal activity differentiates self from close others. *SCAN*, I, 18-25. doi: 10.1093/scan/ns1001

Herman, J. L., Stevens, M. J., Bird, A., Mendenhall, M., & Oddou, G. (2010). The tolerance for ambiguity scale: Towards a more refined measure for international management research. *International Journal of Intercultural Relations*, *34*(1), 58–65.

Hesslow, G. (2012). The current status of the simulation theory of cognition. *Brain Research*, *1428*, 71–79. https://doi.org/10.1016/j.brainres.2011.06.026

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21(7), 1229-1243.

Hobson, R.P. & Meyer, J.A. (2005). Foundations of the self and other: A study in autism. *Developmental Science, 8*, 481-491.

Holmes CJ, Hoge R, Collins L, Woods R, Toga AW, Evans AC. "Enhancement of MR images using registration for signal averaging." J Comput Assist Tomogr. 1998 Mar-Apr;22(2):324–33. http://dx.doi.org/10.1097/00004728-199803000-00032

Huberty, C. J., & Morris, J. D. (n.d.). *Multivariate Analysis Versus Multiple Univariate Analyses*. 7.

IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp.

Jenkins, A. C., & Mitchell, J. P. (2011). Medial prefrontal cortex subserves diverse forms of self-reflection. *Social Neuroscience*, *6*(3), 211–218. doi: 10.1080/17470919.2010.507948

Johnson-Laird, P. (1993). The interaction between reasoning and decision making: An introduction. *Cognition*, 49(1–2), 1–9. https://doi.org/10.1016/0010-0277(93)90033-R

Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3) a brief measure of dark personality traits. *Assessment*, *21*(1), 28–41.

Kang, Y. & Zadorozhny, V. (2016). Process Discovery Using Classification Tree Hidden Semi-Markov Model, IEEE 17th Int. Conf. on Information Reuse and Integration, pp. 361-368.

Karylowski, J. J., Konarzewski, K., & Motes, M. A. (2000). Recruitment of Exemplars as Reference Points in Social Judgments. *Journal of Experimental Social Psychology*, 36(3), 275–303. https://doi.org/10.1006/jesp.1999.1405

Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*(5), 785–794.

Klimesch, W., Sauseng, P., Hanslmayr, S., Gruber, W., & Freunberger, R. (2007). Event-related phase reorganization may explain evoked neural dynamics. *Neuroscience & Biobehavioral Reviews*, 31(7), 1003–1016. https://doi.org/10.1016/j.neubiorev.2007.03.005

Knight, R. T., & Scabini, D. (1998). Anatomic bases of event-related potentials and their relationship to novelty detection in humans, *Journal of Clinical Neurophysiology*, *15*(1), 3–13. https://doi.org/10.1097/00004691-199801000-00003

Knyazev, G. G. (2012). EEG delta oscillations as a correlate of basic homeostatic and motivational processes. *Neuroscience & Biobehavioral Reviews,* 36(1), 677–695. https://doi.org/10.1016/j.neubiorev.2011.10.002

Knyazev, G. G. (2013). EEG correlates of self-referential processing. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00264

Kosslyn, Stephen Michael, Nathaniel M. Alpert, William L.Thompson, Vera Maljkovic, Steven B. Weise, Christopher F. Chabris,Sania E. Hamilton, Scott L. Rauch, and Ferdinando S. Buonanno.1993. Visual mental imagery activates topographically organizedvisual cortex: PET investigations. *Journal of Cognitive Neuroscience*, 5(3): 263-287. doi:10.1162/jocn.1993.5.3.263

Kraskov, A., Dancause, N., Quallo, M.M., Shepherd, S., & Lemon, R.N. (2009). Corticospinal neurons in macaque ventral premotor cortex with mirror properties: A potential mechanism for action suppression? *Neuron, 64*, 922-930.

Kristen, S., Thoermer, C., Hofer, T., Aschersleben, G., & Sodian, B. (2006). Skalierun von "theory of mind" aufgaben (Scaling of theory of mind tasks). *Zeitschrift fur Entwicklungspsychologie und Padagogische Psychologie, 38*, 186-195.

Kuiper, N. A. (1981). Convergent evidence for the self as a prototype: The "inverted-u RT effect" for self and other judgments. *Personality and Social Psychology Bulletin*, *7*(3), 438-443.

Kumar, J.S., & Bhuvaneswari, P. (2012). Analysis of electroencephalography (EEG) signals and its categorization—A study. *Procedia Engineering, 38,* 2525-2536.

Kushnir, T., & Gopnik, A. (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology*, *43*(1), 186–196. https://doi.org/10.1037/0012-1649.43.1.186

Le Bihan, D., Turner, R., Zeffiro, T. A., Cuenod, C. A., Jezzard, P., & Bonnerot, V. (1993). Activation of human primary visual cortex during visual recall: A magnetic resonance imaging study. *Proceedings of the National Academy of Sciences*, *90*(24), 11802–11805. https://doi.org/10.1073/pnas.90.24.11802

Leech, R. & Sharp, D.J. (2014). The role of the posterior cingulate cortex in cognition and disease. *Brain, 137*, 12-32.

Legrand, D. & Ruby, P. (2009). What is self specific? Theoretical investigation and critical review of neuroimaging results. *Psychological Review, 116* (1), 252-282.

Lennox, R. D., & Wolfe, R. N. (1984). Revision of the self-monitoring scale. Retrieved from http://psycnet.apa.org/psycinfo/1984-27678-001

Libenson, M. H. (2010). Chapter 7—Filters in the electroencephalogram. In *Practical Approach to Electroencephalography*. Philadelphia, PA: Saunders.

Liesefeld, H. R. (2018). Estimating the Timing of Cognitive Operations With MEG/EEG Latency Measures: A Primer, a Brief Tutorial, and an Implementation of Various Methods. *Frontiers in Neuroscience*, *12*, 765. https://doi.org/10.3389/fnins.2018.00765

Lima, C. F., Lavan, N., Evans, S., Agnew, Z., Halpern, A. R., Shanmugalingam, P., Meekings, S., Boebinger, D., Ostarek, M., McGettigan, C., Warren, J. E., & Scott, S. K. (2015). Feel the Noise: Relating Individual Differences in Auditory Imagery to the Structure and Function of Sensorimotor Systems. *Cerebral Cortex*, *25*(11), 4638–4650. https://doi.org/10.1093/cercor/bhv134

Lin, F.-H., Belliveau, J. W., Dale, A. M., & Hämäläinen, M. S. (2006). Distributed current estimates using cortical orientation constraints. *Human Brain Mapping*, *27*(1), 1–13. https://doi.org/10.1002/hbm.20155

Lind, S. E., & Bowler, D. M. (2009). Recognition Memory, Self-Other Source Memory, and Theory-of-Mind in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 39(9), 1231–1239. https://doi.org/10.1007/s10803-009-0735-2

Liu, A. K., Dale, A. M., & Belliveau, J. W. (2002). Monte Carlo simulation studies of EEG and MEG localization accuracy. *Human Brain Mapping*, *16*(1), 47–62. https://doi.org/10.1002/hbm.10024

Liu, Y., Huang, H., McGinnis-Deweese, M., Keil, A., Ding, M. (2012). Neural substrate of the late positive potential in emotional processing. *The Journal of Neuroscience, 32*(42), 14563-14572.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8, 213.

Luck, S. J. (2014). "Quantifying ERP Amplitudes and Latencies." *An Introduction to the Event-Related Potential Technique* (2 ed.). Cambridge, MA: MIT Press.

Lyons, V. & Fitzgerald, M. (2013). Atypical sense of self in Autism Spectrum Disorders: A neuro-cognitive perspective. Recent Advances in Autism Spectrum Disorders – Volume I, Michael Fitzgerald, *IntechOpen*, DOI:10.5772/53680. Available from https://www.intechopen.com/books/recent-advances-in-autism-spectrum-disorders-volume-i/atypical-sense-of-self-in-autism-spectrum-disorders-a-neuro-cognitive-perspective

Mahy, C. E. V., Moses, L. J., & Pfeifer, J. H. (2014). How and where: Theory-of-mind in the brain. *Developmental Cognitive Neuroscience*, 9, 68–81. https://doi.org/10.1016/j.dcn.2014.01.002

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

Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2002). Dynamic brain sources of visual evoked responses. *Science*, 295, 690–694. http://dx.doi.org/10 .1126/science.1066168

Mäkinen, V., Tiitinen, H., & May, P. (2005). Auditory event-related responses are generated independently of ongoing brain activity. *NeuroImage*, *24*(4), 961–968. https://doi.org/10.1016/j.neuroimage.2004.10.020

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Matti, S. & Olaf, H. (2013). Minimum-norm cortical source estimation in layered head models is robust against skull conductivity error. *NeuroImage*, 81, 265-272.

McDonald, J. H. (2014). *Handbook of Biological Statistics* (3rd ed.). Sparky House Publishing, Baltimore, Maryland.

Meister, I. G., Krings, T., Foltys, H., Boroojerdie, B., Muller, M., Topper, R., & Thron, A. (2004). Playing piano in the mind—an fMRI study on music imagery and performance in pianists. *Cognitive Brain Research, 19*, 219-228.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. Science, 198(4312), 75–78.Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, 50(4), 655–663. doi: 10.1016/j.neuron.2006.03.040

Michel, C. M., & Brunet, D. (2019). EEG Source Imaging: A Practical Review of the Analysis Steps. *Frontiers in Neurology*, *10*, 325. https://doi.org/10.3389/fneur.2019.00325

Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). The Link between Social Cognition and Self-referential Thought in the Medial Prefrontal Cortex. *Journal of Cognitive Neuroscience*, *17*(8), 1306–1315. https://doi.org/10.1162/0898929055002418

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable Medial Prefrontal Contributions to Judgments of Similar and Dissimilar Others. *Neuron*, *50*(4), 655–663. https://doi.org/10.1016/j.neuron.2006.03.040

Mobascher, A., Brinkmeyer, J., Warbrick, T., Musso, F., Wittsack, H.J., Stoermer, R., Saleh, A. Schnitzler, A. & Winterer, G. (2009). Fluctuations in electrodermal activity reveal variations in single trial brain responses to painful laser stimuli - A fMRI/EEG study. *NeuroImage, 44*(3), 1081-1092.

Molnar-Szakacs, I., & Uddin, L. Q. (2013). Self-Processing and the Default Mode Network: Interactions with the Mirror Neuron System. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00571

Moran, J. M., Lee, S. M., & Gabrieli, J. D. E. (2011). Dissociable neural systems supporting knowledge about human character and appearance in ourselves and others. *Journal of Cognitive Neuroscience*, 23(9), 2222–2230. doi: 10.1162/jocn.2010.21580

Moreau, N., Viallet, F., & Champagne-Lavau, M. (2013). Using memories to understand others: The role of episodic memory in theory of mind impairment in Alzheimer disease. *Ageing Research Reviews*, *12*(4), 833–839. https://doi.org/10.1016/j.arr.2013.06.005

Mu, Y., & Han, S. (2010). Neural oscillations involved in self-referential processing. *NeuroImage, 53*(2), 757–768. https://doi.org/10.1016/j.neuroimage.2010.07.008

Noel, J.-P., Cascio, C. J., Wallace, M. T., & Park, S. (2017). The spatial self in schizophrenia and autism spectrum disorder. *Schizophrenia Research*, *179*, 8–12. https://doi.org/10.1016/j.schres.2016.09.021

Nolte, G. (2003). The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Physics in Medicine and Biology*, *48*(22), 3637–3652. https://doi.org/10.1088/0031-9155/48/22/002

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *NeuroImage*, *31*(1), 440–457. https://doi.org/10.1016/j.neuroimage.2005.12.002

Ochsner, K. N., Beer, J. S., Robertson, E. R., Cooper, J. C., Gabrieli, J. D. E., Kihsltrom, J. F., & D'Esposito, M. (2005). The neural correlates of direct and reflected self-knowledge. *NeuroImage*, *28*(4), 797–814. https://doi.org/10.1016/j.neuroimage.2005.06.069

Olbrich, S., Mulert, C., Karch, S., Trenner, M., Leicht, G., Pogarell O., & Hegerl, U. (2009) EEG-vigilance and BOLD effect during simultaneous EEG/fMRI measurements, *NeuroImage, 45*, 319-332.

Olson, C. L. (1974). Comparative Robustness of Six Tests in Multivariate Analysis of Variance. *Journal of the American Statistical Association*, 69:348, 894-908.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, JM (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Computational Intelligence and Neuroscience, Volume 2011 (2011), Article ID 156869, doi:10.1155/2011/156869

Papo, D. (2015). How can we study reasoning in the brain? *Frontiers in Human Neuroscience,* 9. https://doi.org/10.3389/fnhum.2015.00222

Park, J., Kim, H., Sohn, J.-W., Choi, J., & Kim, S.-P. (2018). EEG Beta Oscillations in the Temporoparietal Area Related to the Accuracy in Estimating Others' Preference. *Frontiers in Human Neuroscience*, *12*, 43. https://doi.org/10.3389/fnhum.2018.00043

Pascual-Marqui, R.D.. (2002). Notes on some physics, mathematics, and statistics for time and frequency domain LORETA. December 8.

Pascual-Marqui, R. D., Faber, P., Kinoshita, T., Kochi, K., Milz, P., Nishida, K., & Yoshimura, M. (2018). *Comparing EEG/MEG neuroimaging methods based on localization error, false positive activity, and false positive connectivity* [Preprint]. https://doi.org/10.1101/269753

Peña, D., & Prieto, F. J. (2001). Multivariate outlier detection and robust covariance matrix estimation. *Technometrics*, 43(3), 286–310. https://doi.org/10.1198/004017001316975899

Perner, J. (1996). Simulation as explicitation of predication-implicit knowledge about the mind: arguments for a simulation-theory mix. In P. Carruthers & P. K. Smith (Eds.), *Theories of Theories of Mind* (pp. 90-104). Cambridge University Press.

Perner, J., Kloo, D., & Gornik, E. (2007). Episodic memory development: Theory of mind is part of re-experiencing experienced events. *Infant and Child Development*, *16*(5), 471–490. https://doi.org/10.1002/icd.517

Peterson, C., Wellman, H., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development*, 76(2), 502-517. doi: 10.1111/j.1467-8624.2005.00859.x

Pizzagalli, D., Pascual-Marqui, R. D., Nitschke, J. B., Oakes, T. R., Larson, C. L., Abercrombie, H. C., … Davidson, R. J. (2001). Anterior Cingulate Activity as a Predictor of Degree of Treatment Response in Major Depression: Evidence From Brain Electrical Tomography Analysis. *American Journal of Psychiatry*, *158*(3), 405–415. https://doi.org/10.1176/appi.ajp.158.3.405

Popov, T., Oostenveld, R., & Schoffelen, J. M. (2018). FieldTrip Made Easy: An Analysis Protocol for Group Analysis of the Auditory Steady State Brain Response in Time, Frequency, and Space. Frontiers in Neuroscience, 12, 711. https://doi.org/10.3389/fnins.2018.00711

Poulsen, J. and French, A. (2001). *Discriminant function analysis (DA)*. San Francisco State University, available at: http://userwww.sfsu.edu/efc/classes/biol710/discrim/discrim.pdf (accessed 3 December 2020).

Power, J. D., Fair, D. A., Schlaggar, B. L., & Petersen, S. E. (2010). The development of human functional brain networks. *Neuron*, 67(5), 735–748. doi: 10.1016/j.neuron.2010.08.017

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*(4), 741–763. https://doi.org/10.1037/0022-3514.67.4.741

Purkey, W. (1988). An overview of self-concept theory for counselors. Ann Arbor, MI: Clearinghouse on Assessment and Evaluation. http://ericae.net/edo/ed304630.htm

Rabin, J. S., & Rosenbaum, R. S. (2012). Familiarity modulates the functional relationship between theory of mind and autobiographical memory. *NeuroImage*, *62*(1), 520–529. https://doi.org/10.1016/j.neuroimage.2012.05.002

Redshaw, J., Nielsen, M., Slaughter, V., Kennedy-Costantini, S., Oostenbroek, J., Crimston, J., & Suddendorf, T. (2019). Individual differences in neonatal "imitation" fail to predict early social cognitive behavior. *Developmental Science.* doi: 10.1111/desc.12892

Rencher, A. C. (2002). *Methods of multivariate analysis* (2nd ed). J. Wiley.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Ann. Rev. Neurosci. 27,* 169-192.

Rosa, N. M., Deason, R. G., Budson, A. E., & Gutchess, A. H. (2015). Self-referencing and false memory in mild cognitive impairment due to Alzheimer's disease. *Neuropsychology*, *29*(5), 799–805. https://doi.org/10.1037/neu0000186

Rosen, J. B., Brand, M., Polzer, C., Ebersbach, G., & Kalbe, E. (2013). Moral decision-making and theory of mind in patients with idiopathic Parkinson's disease. *Neuropsychology*, 27(5), 562–572. doi: 10.1037/a0033595

Sabbagh, M.A. (2013). Chapter 8: Brain electrophysiological studies of theory of mind. In Baron-Cohen, S., Tager-Flusberg, H., & Lombardo, M.V. (Eds.), *Understanding Other Minds: Perspectives From Developmental Social Neuroscience,* 119-131. NY, NY: Oxford Press.

Saxe, R., Moran, J. M., Scholz, J., & Gabrieli, J. (2006). Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. *Social Cognitive and Affective Neuroscience*, *1*(3), 229–234. https://doi.org/10.1093/scan/nsl034

Scheeringa, R., Petersson, K. M., Kleinschmidt, A., Jensen, O., & Bastiaansen, M. C. M. (2012). EEG Alpha Power Modulation of fMRI Resting-State Connectivity. *Brain Connectivity*, *2*(5), 254–264. https://doi.org/10.1089/brain.2012.0088

Scheeringa, R., Petersson, K. M., Oostenveld, R., Norris, D. G., Hagoort, P., & Bastiaansen, M. C. M. (2009). Trial-by-trial coupling between EEG and BOLD identifies networks related to alpha and theta EEG power increases during working memory maintenance. *NeuroImage*, *44*(3), 1224–1238. https://doi.org/10.1016/j.neuroimage.2008.08.041

Schwoebel, J., Boronat, C. B., & Branch Coslett, H. (2002). The man who executed "imagined" movements: Evidence for dissociable components of the body schema. *Brain and Cognition*, *50*(1), 1–16. https://doi.org/10.1016/S0278-2626(02)00005-2

Shah, A. S., Bressler, S. L., Knuth, K. H., Ding, M., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2004). Neural dynamics and the fundamental mechanisms of event-related brain potentials. Cerebral Cortex, 14, 476– 483. http://dx.doi.org/10.1093/cercor/bhh009

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. Science, 303, 1157-1161.

Smetana, J. G., Jambon, M., Conry-Murray, C., & Sturge-Apple, M. L. (2012). Reciprocal associations between young children's developing moral judgments and theory of mind. *Developmental Psychology*, 48(4), 1144–1155. doi: 10.1037/a0025891

Spaulding, S. (2012). Mirror neurons are not evidence for the Simulation Theory. *Synthese*, *189*(3), 515–534. https://doi.org/10.1007/s11229-012-0086-y

Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science*, *28*(3), 303–333. https://doi.org/10.1207/s15516709cog2803_1

Song, J., Davey, C., Poulsen, C., Luu, P., Turovets, S., Anderson, E., Li, K., & Tucker, D. (2015). EEG source localization: Sensor density and head surface coverage. *Journal of Neuroscience Methods*, *256*, 9–21. https://doi.org/10.1016/j.jneumeth.2015.08.015

Sowndhararajan, K., Kim, M., Deepa, P., Park, S., & Kim, S. (2018). Application of the P300 Event-Related Potential in the Diagnosis of Epilepsy Disorder: A Review. *Scientia Pharmaceutica*, *86*(2), 10. https://doi.org/10.3390/scipharm86020010

Spreng, R. N., & Grady, C. L. (2010). Patterns of Brain Activity Supporting Autobiographical Memory, Prospection, and Theory of Mind, and Their Relationship to the Default Mode Network. *Journal of Cognitive Neuroscience*, *22*(6), 1112–1123. https://doi.org/10.1162/jocn.2009.21282

Subramaniam, K., Hinkley, L.B.N., Mizuiri, D., Kothare, H., Cai, C., Garrett, C., Findlay, A., Houde, J.F., & Nagarajan, S.S. (2019). Beta-band activity in medial prefrontal cortex predicts source memory encoding and retrieval accuracy. *Scientific Reports, 9*, article number 6814

Swick, D., Kutas, M., & Neville, H. J. (n.d.). *Localizing the Neural Generators of Event-Related Brain Potentials*. 49.

Taboga, Marco (2017). "Log-likelihood", Lectures on probability theory and mathematical statistics, Third edition. Kindle Direct Publishing. Online appendix. https://www.statlect.com/glossary/log-likelihood.

Tacikowski, P., & Nowicka, A. (2010). Allocation of attention to self-name and self-face: An ERP study. *Biological Psychology*, *84*(2), 318–324. https://doi.org/10.1016/j.biopsycho.2010.03.009

Thom, H. C. S., 1958: A note on the gamma distribution**.** *Mon. Wea. Rev.***,** 86**,** 117–122, https://doi.org/10.1175/1520-0493(1958)086<0117:ANOTGD>2.0.CO;2.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

Ujma, P.P., Simor, P., Steiger, A., Desler, M., & Bodizs, R. (2018). Individual slow wave morphology is a marker of aging. *Neurobiology of Aging*, 80, 71-82.

van Buuren, M., Vink, M., & Kahn, R. S. (2012). Default-mode network dysfunction and self-referential processing in healthy siblings of schizophrenia patients. *Schizophrenia Research*, *142*(1–3), 237–243. https://doi.org/10.1016/j.schres.2012.09.017

Verleger, R., Jaskowski, P., & Wascher, E. (2005). Evidence for an integrative role of P3b in linking reaction to perception. *Journal of Psychophsiology*, 20(2), 1-17.

Vogt, B.A., Finch, D.M., & Olson E.R. (1992). Functional heterogeneity in cingulate cortex: The anterior executive and posterior evaluative regions. *Cerebral Cortex*, 2(6), 435-443.

Voss, J. L., & Paller, K. A. (2017). Neural Substrates of Remembering: Event-Related Potential Studies. In *Learning and Memory: A Comprehensive Reference* (pp. 81-98).

Wang, J.-Z., Williamson, S. J., & Kaufman, L. (1992). Magnetic source images determined by a lead-field analysis: The unique minimum-norm least-squares estimation. *IEEE Transactions on Biomedical Engineering*, *39*(7), 665–675. https://doi.org/10.1109/10.142641

Wang, L., Saalmann, Y.B., Pinsk, M.A., Arcaro, M.J., & Kastner, S. (2012). Electrophysiological low-frequency coherence and cross-frequency coupling contribute to BOLD connectivity. *Neuron* 76, 1010–1020. doi:10.1016/j.neuron.2012.09.033

Wickens, C. D. & Carswell, C. M. (2006). Information processing. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

Wicker, B., Keysers, C., Plailly, J., Royet, J., Gallese, V., Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. Neuron, *40*, 655-664.

Yerys, B. E., Gordon, E. M., Abrams, D. N., Satterthwaite, T. D., Weinblatt, R., Jankowski, K. F., … Vaidya, C. J. (2015). Default mode network segregation and social deficits in autism spectrum disorder: Evidence from non-medicated children. *NeuroImage: Clinical*, *9*, 223–232. https://doi.org/10.1016/j.nicl.2015.07.018

Yu, C., Tu, S., Wang, T., & Qiu, J. (2010). The neural basis of self-evaluation processing in social judgment: *NeuroReport*, *21*(7), 497–501. https://doi.org/10.1097/WNR.0b013e3283383449

Yu, S.-Z. (2010). Hidden semi-Markov models. *Artificial Intelligence*, *174*(2), 215–243. https://doi.org/10.1016/j.artint.2009.11.011

Zeng, Y., Zhao, Y., Zhang, T., Zhao, D., Zhao, F., & Lu, E. (2020). A Brain-Inspired Model of Theory of Mind. *Frontiers in Neurorobotics*, *14*, 60. https://doi.org/10.3389/fnbot.2020.00060

Zhang, Q., van Vugt, M., Borst, J. P., & Anderson, J. R. (2018). Mapping working memory retrieval in space and in time: A combined electroencephalography and electrocorticography approach. *NeuroImage*, *174*, 472–484. https://doi.org/10.1016/j.neuroimage.2018.03.039

Appendix A: Timetable of Task Design

Table A1. Timetable of Task Design

Task Information	Method	Approximate Length of Time
Informed Consent	Paper	3 minutes
Demographic Questionnaire	Paper	2 minutes
EEG Net Application	Person	5 minutes
Impedance Check	Person	5 minutes
Resting EEG	Computer	10 minutes
Blocks 1-4	Computer	15 minutes
Impedance Check	Person	5 minutes
Blocks 4-8	Computer	15 minutes
Impedance Check	Person	5 minutes
Blocks 8-12	Computer	15 minutes
EEG Net Removal	Person	5 minutes
Debriefing Form	Paper	5 minutes
	Total	90 minutes

Appendix B: 64-Channel Montage and Electrode Selection
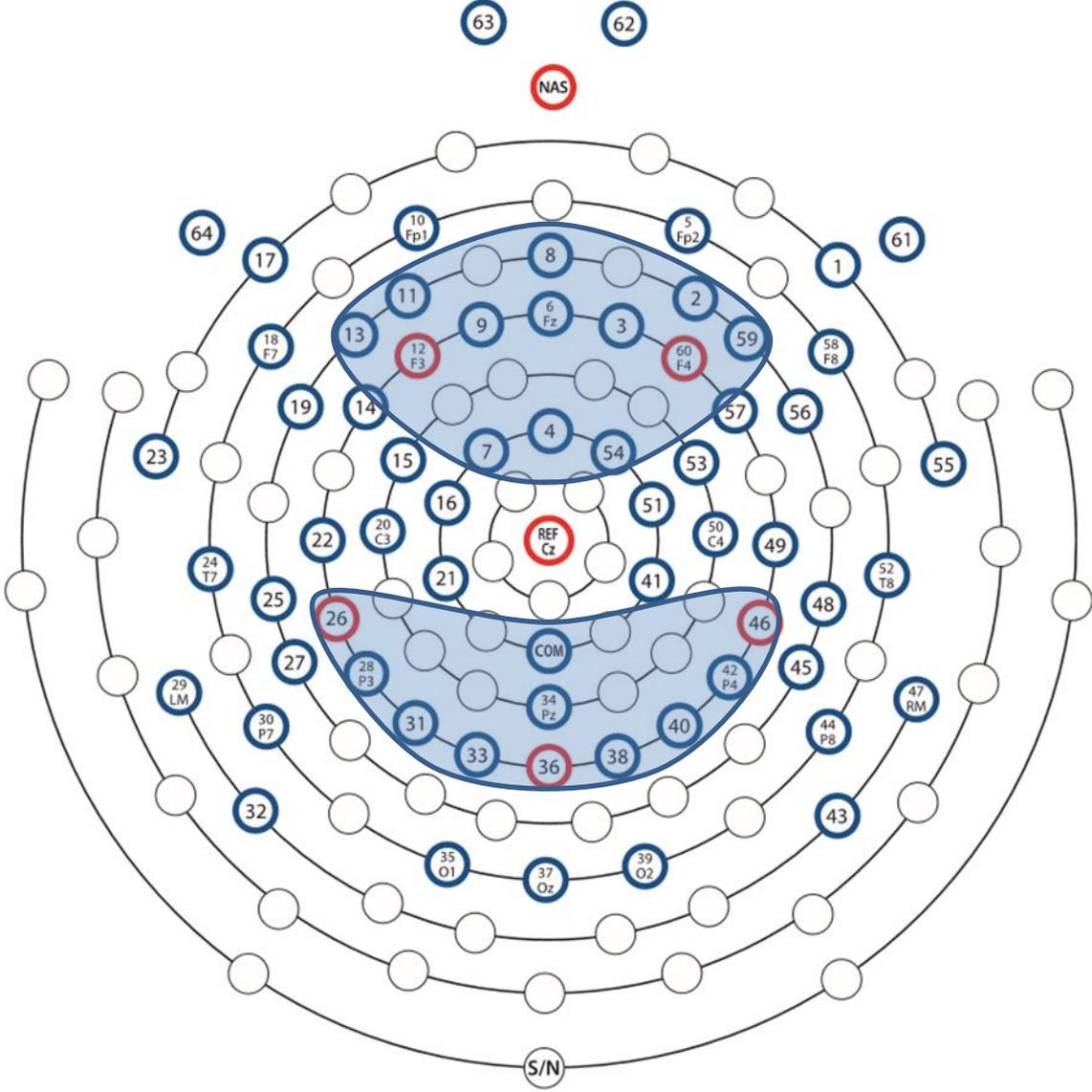
Figure B1. 64-Channel EEG montage. (a) frontal subset, (b) centroparietal subset.

Appendix C: Descriptive statistics and breakdown of within-subjects factors

Table C1. Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age	360	17	24	19.33	1.384
Gender	360	0	1	.67	.472
fArea	360	.02	4.05	1.07	.76
fLatency	360	271.31	949.33	511.37	82.27
cArea	360	.04	4.49	1.27	.82
cLatency	360	275.70	779.30	526.04	79.67
Inducedf_delta	360	52.79	68.31	60.94	3.01
Inducedf_theta	360	48.80	65.82	59.05	3.44
Inducedf_alpha	360	46.84	67.89	56.73	3.76
Inducedf_beta	360	42.49	58.22	50.33	3.05
Inducedf_lowgamma	360	38.77	49.78	44.30	2.48
Inducedf_highgamma	360	35.72	48.66	42.64	2.59
Inducedc_delta	360	-6.32	-.46	-2.41	1.04
Inducedc_theta	360	-8.48	-1.28	-4.40	1.53
Inducedc_alpha	360	-11.29	-1.06	-4.54	1.91
Inducedc_beta	360	-7.97	-1.15	-2.86	.91
Inducedc_lowgamma	360	-3.86	-1.31	-2.24	.38
Inducedc_highgamma	360	-3.27	-1.50	-2.20	.29
Evokedf_delta	360	48.59	68.46	61.05	3.23
Evokedf_theta	360	43.50	67.83	58.84	4.15
Evokedf_alpha	360	40.44	72.24	57.31	4.67
Evokedf_beta	360	37.64	60.44	50.15	3.28
Evokedf_lowgamma	360	34.57	49.75	43.70	2.40
Evokedf_highgamma	360	34.19	47.65	42.24	2.62
Evokedc_delta	360	-8.28	-.57	-2.49	1.25
Evokedc_theta	360	-10.42	-1.16	-4.25	1.88
Evokedc_alpha	360	-11.24	-.42	-4.35	2.14
Evokedc_beta	360	-6.87	-1.37	-2.87	.92
Evokedc_lowgamma	360	-3.92	-1.34	-2.25	.38
Evokedc_highgamma	360	-3.62	-1.56	-2.23	.32

F = frontal, c = centroparietal

Table C2. Within-Subjects Factors

		Value Label	N
Group	0	Character Self	45
	1	Character Fallon	45
	2	Character Mother	45
	3	Character Case	45
	4	Appearance Self	45
	5	Appearance Fallon	45
	6	Appearance Mother	45
	7	Appearance Case	45
Person	0	Self	90
	1	Fallon	90
	2	Mother	90
	3	Case	90
Word	0	Character	180
	1	Appearance	180

Appendix D: Reaction Time Output

Table D1. Mean reaction times by person and word condition

	Self	Mother	Fallon	Case
Character	1146.03(16)	1120.25(16)	1173.75(17)	697.43(15)
Appearance	1099.50(15)	1080.06(15)	1122.02(15)	696.73(14)

Note: Mean (standard error of the mean).

Table D2. Tests of within-subjects effects for reaction time

Source	Type III SS	df	Mean Square	F	Sig.	Partial η^2
Person	12382189.39	1.60	7708133.58	613.828	.000	.933
Error	887572.01	70.68	12557.49			
Word	108917.54	1.00	108917.54	59.160	.000	.573
Error	81006.48	44.00	1841.06			
Person*Word	36352.00	3.00	12117.33	11.311	.000	.204
Error	141410.55	132.00	1071.29			

Data is Huynh-Feldt corrected

Table D3. Post-hoc pairwise person and word comparisons for reaction time

(I) Person	(J) Person	Mean Difference (I-J)	Std. Error	Sig.[b]	95% CI for Difference[b] Lower Bound	Upper Bound
Self	Mother	22.610[*]	5.950	.003	6.218	39.001
	Fallon	-25.119[*]	7.742	.013	-46.445	-3.792
	Case	425.684[*]	15.485	.000	383.029	468.340
Mother	Self	-22.610[*]	5.950	.003	-39.001	-6.218
	Fallon	-47.728[*]	7.484	.000	-68.344	-27.112
	Case	403.075[*]	14.816	.000	362.261	443.888
Fallon	Self	25.119[*]	7.742	.013	3.792	46.445
	Mother	47.728[*]	7.484	.000	27.112	68.344
	Case	450.803[*]	16.908	.000	404.228	497.378
Case	Self	-425.684[*]	15.485	.000	-468.340	-383.029
	Mother	-403.075[*]	14.816	.000	-443.888	-362.261
	Fallon	-450.803[*]	16.908	.000	-497.378	-404.228

(I) Word	(J) Word					
Character	Appearance	34.788[*]	4.523	.000	25.673	43.903

Based on estimated marginal means
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Sidak.

Table D4. Post-hoc pairwise person*word comparisons for reaction time

Word	(I) Person	(J) Person	Mean Difference (I-J)	Std. Error	Sig.[b]	95% CI for Difference[b] Lower Bound	95% CI for Difference[b] Upper Bound
Character	Self	Mother	25.781*	7.994	.014	3.761	47.802
		Fallon	-27.719*	9.011	.021	-52.540	-2.898
		Case	448.596*	17.744	.000	399.716	497.476
	Mother	Self	-25.781*	7.994	.014	-47.802	-3.761
		Fallon	-53.501*	10.446	.000	-82.275	-24.726
		Case	422.815*	16.579	.000	377.145	468.484
	Fallon	Self	27.719*	9.011	.021	2.898	52.540
		Mother	53.501*	10.446	.000	24.726	82.275
		Case	476.315*	19.091	.000	423.728	528.903
	Case	Self	-448.596*	17.744	.000	-497.476	-399.716
		Mother	-422.815*	16.579	.000	-468.484	-377.145
		Fallon	-476.315*	19.091	.000	-528.903	-423.728
Appearance	Self	Mother	19.438*	6.842	.040	.591	38.284
		Fallon	-22.518	9.646	.137	-49.088	4.052
		Case	402.772*	14.328	.000	363.303	442.241
	Mother	Self	-19.438*	6.842	.040	-38.284	-.591
		Fallon	-41.956*	7.781	.000	-63.389	-20.522
		Case	383.335*	14.537	.000	343.290	423.379
	Fallon	Self	22.518	9.646	.137	-4.052	49.088
		Mother	41.956*	7.781	.000	20.522	63.389
		Case	425.290*	15.917	.000	381.445	469.136
	Case	Self	-402.772*	14.328	.000	-442.241	-363.303
		Mother	-383.335*	14.537	.000	-423.379	-343.290
		Fallon	-425.290*	15.917	.000	-469.136	-381.445

Person	(I) Word	(J) Word					
Self	Character	Appearance	46.529*	8.800	.000	28.794	64.265
Mother	Character	Appearance	40.186*	7.807	.000	24.452	55.919
Fallon	Character	Appearance	51.731*	8.454	.000	34.693	68.768
Case	Character	Appearance	.706	3.847	.855	-7.048	8.460

Based on estimated marginal means
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Sidak.

Appendix E: ERP Output

Table E1. Spearman Correlations between ERP DVs

		fArea	fLatency	cArea	cLatency
Spearman's rho	fArea	1.000	.205[**]	.801[**]	.214[**]
	fLatency	.205[**]	1.000	.145[**]	.502[**]
	cArea	.801[**]	.145[**]	1.000	.227[**]
	cLatency	.214[*]	.502[**]	.227[**]	1.000

*. Correlation significance at .05 level
**. Significance at the .001 level

Table E2. Multivariate within-subjects effects for ERP amplitude and latency

Within Subjects Effect		Value	F	Hypothesis df	Error df	Sig.	Partial η^2
Person	Pillai's Trace	.438	5.592	12.000	393.000	.000	.146
	Wilks' Lambda	.569	6.758	12.000	341.593	.000	.171
	Hotelling's Trace	.745	7.930	12.000	383.000	.000	.199
	Roy's Largest Root	.729	23.885	4.000	131.000	.000	.422
Word	Pillai's Trace	.333	5.113	4.000	41.000	.002	.333
	Wilks' Lambda	.667	5.113	4.000	41.000	.002	.333
	Hotelling's Trace	.499	5.113	4.000	41.000	.002	.333
	Roy's Largest Root	.499	5.113	4.000	41.000	.002	.333
Person * Word	Pillai's Trace	.125	1.426	12.000	393.000	.151	.042
	Wilks' Lambda	.877	1.444	12.000	341.593	.144	.043
	Hotelling's Trace	.137	1.457	12.000	383.000	.138	.044
	Roy's Largest Root	.112	3.657	4.000	131.000	.007	.100

Design: Intercept
Within Subjects Design: Person + Word + Person * Word
The statistic is an upper bound on F that yields a lower bound on the significance level.

Table E3. Univariate tests of within-subjects effects for ERP amplitude and latency

Source	Measure	Type III SS	df	Mean Square	F	Sig.	Partial η^2
Person	fArea	4.159	2.434	1.709	7.887	.000	.152
	fLatency	125443.921	1.525	82241.174	5.552	.011	.112
	cArea	3.823	2.233	1.712	6.161	.002	.123
	cLatency	30331.819	2.304	13162.836	1.975	.137	.043
Word	fArea	.543	1.000	.543	3.311	.076	.070
	fLatency	31836.152	1.000	31836.152	12.906	.001	.227
	cArea	.128	1.000	.128	.997	.323	.022
	cLatency	38079.772	1.000	38079.772	13.359	.001	.233
Person * Word	fArea	.296	2.864	.103	1.146	.332	.025
	fLatency	4656.660	2.615	1780.965	.695	.538	.016
	cArea	1.023	2.793	.366	4.019	.011	.084
	cLatency	9940.223	3.000	3313.408	1.594	.194	.035

Data is Huynh-Feldt corrected.

Table E4. Post-hoc pairwise person comparisons for ERP amplitude and latency

Measure	(I) Person	(J) Person	Mean Difference (I-J)	Std. Error	Sig.	Lower CI	Upper CI
fArea	Self	Mother	-.050	.051	.912	-.191	.091
		Fallon	-.092	.043	.206	-.210	.026
		Case	.189	.074	.078	-.013	.392
	Mother	Self	.050	.051	.912	-.091	.191
		Fallon	-.042	.053	.966	-.188	.104
		Case	.239	.074	.014	.034	.444
	Fallon	Self	.092	.043	.206	-.026	.210
		Mother	.042	.053	.966	-.104	.188
		Case	.281	.072	.002	.082	.481
	Case	Self	-.189	.074	.078	-.392	.013
		Mother	-.239	.074	.014	-.444	-.034
		Fallon	-.281	.072	.002	-.481	-.082
fLatency	Self	Mother	1.945	6.337	1.000	-15.511	19.401
		Fallon	-4.411	8.062	.995	-26.618	17.795
		Case	-43.602	16.043	.055	-87.794	.589
	Mother	Self	-1.945	6.337	1.000	-19.401	15.511
		Fallon	-6.357	6.790	.928	-25.059	12.346
		Case	-45.547	16.507	.049	-91.019	-.076
	Fallon	Self	4.411	8.062	.995	-17.795	26.618
		Mother	6.357	6.790	.928	-12.346	25.059
		Case	-39.191	17.975	.191	-88.706	10.324
	Case	Self	43.602	16.043	.055	-.589	87.794
		Mother	45.547	16.507	.049	.076	91.019
		Fallon	39.191	17.975	.191	-10.324	88.706
cArea	Self	Mother	-.023	.055	.999	-.173	.128
		Fallon	-.034	.055	.990	-.184	.117
		Case	.217	.088	.101	-.025	.460
	Mother	Self	.023	.055	.999	-.128	.173
		Fallon	-.011	.045	1.000	-.135	.112
		Case	.240	.080	.025	.021	.459
	Fallon	Self	.034	.055	.990	-.117	.184
		Mother	.011	.045	1.000	-.112	.135
		Case	.251	.074	.009	.048	.455
	Case	Self	-.217	.088	.101	-.460	.025
		Mother	-.240	.080	.025	-.459	-.021
		Fallon	-.251	.074	.009	-.455	-.048
cLatency	Self	Mother	-.384	8.376	1.000	-23.458	22.690
		Fallon	1.116	7.450	1.000	-19.407	21.639
		Case	21.404	12.865	.480	-14.035	56.843
	Mother	Self	.384	8.376	1.000	-22.690	23.458
		Fallon	1.501	8.942	1.000	-23.131	26.132
		Case	21.788	11.639	.344	-10.274	53.850
	Fallon	Self	-1.116	7.450	1.000	-21.639	19.407
		Mother	-1.501	8.942	1.000	-26.132	23.131
		Case	20.288	13.268	.576	-16.262	56.837
	Case	Self	-21.404	12.865	.480	-56.843	14.035
		Mother	-21.788	11.639	.344	-53.850	10.274
		Fallon	-20.288	13.268	.576	-56.837	16.262

Table E5. Post-hoc pairwise word comparisons for ERP amplitude and latency

Measure	(I) Word	(J) Word	Mean Difference (I-J)	Std. Error	Sig.[a]	95% CI for Difference[a] Lower Bound	Upper Bound
fArea	Character	Appearance	-.078	.043	.076	-.164	.008
fLatency	Character	Appearance	-18.808	5.235	.001	-29.359	-8.257
cArea	Character	Appearance	-.038	.038	.323	-.114	.038
cLatency	Character	Appearance	-20.570	5.628	.001	-31.912	-9.228

Based on estimated marginal means

a. Adjustment for multiple comparisons: Sidak.

Table E6. Post-hoc pairwise person*word comparisons for ERP amplitude and latency

	Person	Person	fArea	fLatency	cArea	cLatency
Character	Self	Mother	0.022	9.961	0.113	10.459
		Fallon	-0.033	-0.459	0.038	2.698
		Case	.255*	-44.727	0.232	18.077
	Mother	Self	-0.022	-9.961	-0.113	-10.459
		Fallon	-0.054	-10.420	-0.074	-7.761
		Case	0.233	-54.688	0.120	7.618
	Fallon	Self	0.033	0.459	-0.038	-2.698
		Mother	0.054	10.420	0.074	7.761
		Case	.287*	-44.267	0.194	15.379
	Case	Self	-.255*	44.727	-0.232	-18.077
		Mother	-0.233	54.688	-0.120	-7.618
		Fallon	-.287*	44.267	-0.194	-15.379
Appearance	Self	Mother	-0.122	-6.071	-0.158	-11.227
		Fallon	-0.151	-8.364	-0.106	-0.466
		Case	0.124	-42.478	0.203	24.730
	Mother	Self	0.122	6.071	0.158	11.227
		Fallon	-0.030	-2.293	0.052	10.762
		Case	.246*	-36.407	.360*	35.958*
	Fallon	Self	0.151	8.364	0.106	0.466
		Mother	0.030	2.293	-0.052	-10.762
		Case	.275*	-34.114	.309*	25.196
	Case	Self	-0.124	42.478	-0.203	-24.730
		Mother	-.246*	36.407	-.360*	-35.958*
		Fallon	-.275*	34.114	-.309*	-25.196

Measure	Person	(I) Word	(J) Word	Mean Difference (I-J)	Std. Error	Sig.[b]	95% CI for Difference[b] Lower Bound	Upper Bound
fArea	Self	Character	Appearance	.020	.052	.693	-.084	.125
	Mother	Character	Appearance	-.123	.066	.069	-.256	.010
	Fallon	Character	Appearance	-.098	.080	.229	-.260	.064
	Case	Character	Appearance	-.110	.073	.137	-.257	.037
fLatency	Self	Character	Appearance	-13.386	12.166	.277	-37.906	11.134
	Mother	Character	Appearance	-29.418	9.757	.004	-49.082	-9.753
	Fallon	Character	Appearance	-21.290	7.054	.004	-35.506	-7.075
	Case	Character	Appearance	-11.137	10.707	.304	-32.716	10.442
cArea	Self	Character	Appearance	.073	.050	.148	-.027	.174
	Mother	Character	Appearance	-.197	.065	.004	-.329	-.065
	Fallon	Character	Appearance	-.071	.079	.372	-.230	.088
	Case	Character	Appearance	.044	.064	.495	-.084	.172
cLatency	Self	Character	Appearance	-16.020	11.189	.159	-38.571	6.530
	Mother	Character	Appearance	-37.707	11.042	.001	-59.961	-15.453
	Fallon	Character	Appearance	-19.184	8.550	.030	-36.416	-1.952
	Case	Character	Appearance	-9.367	9.142	.311	-27.792	9.058

Based on estimated marginal means
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Sidak.
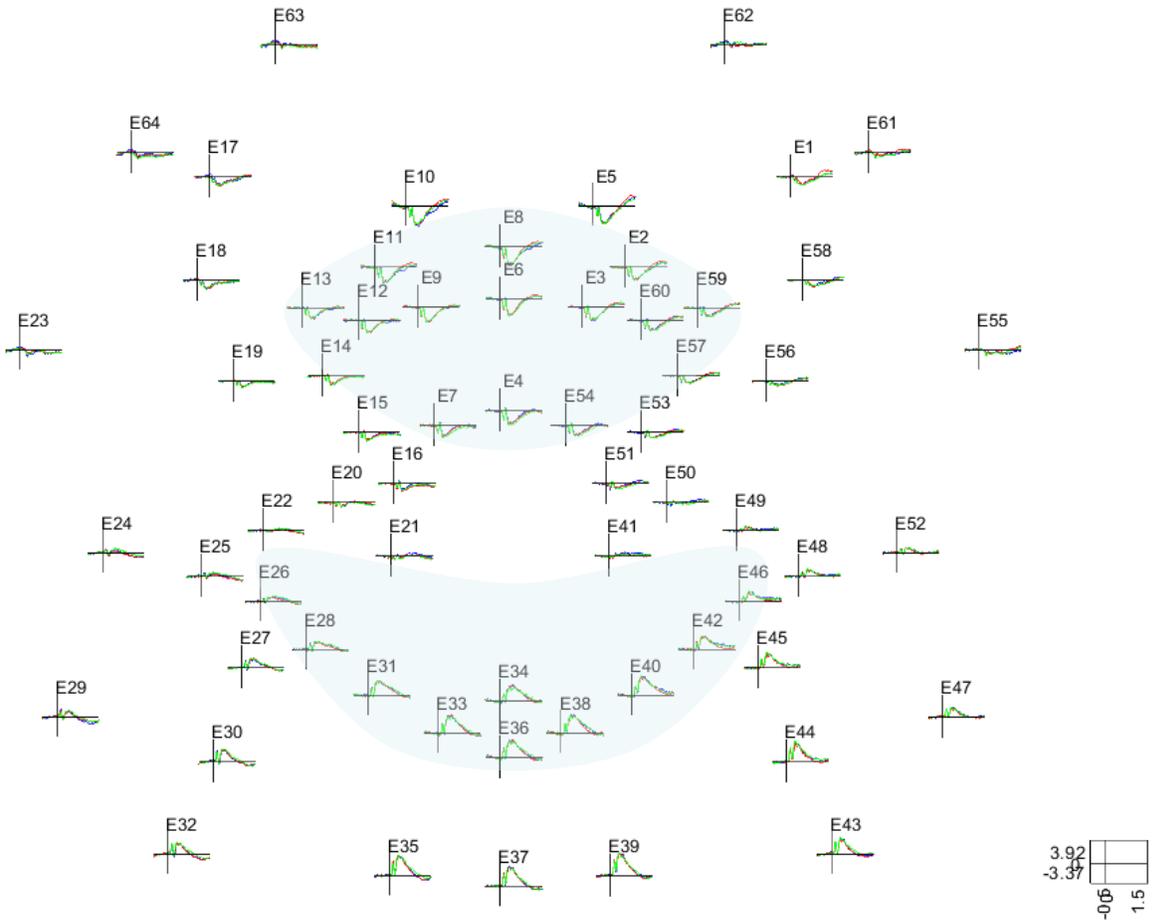
Figure E1. Multiplot of all grand average ERPs for Self, Mother, and Fallon character responses
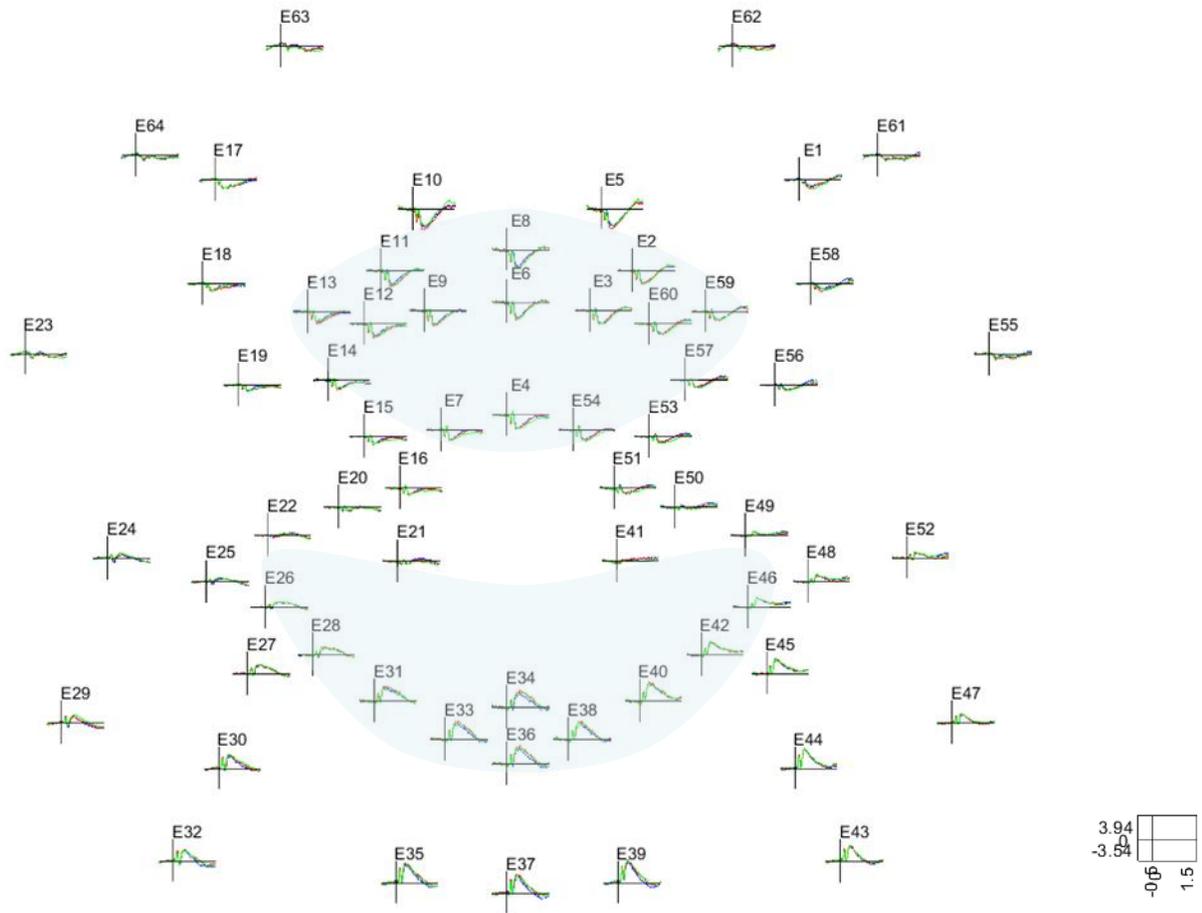
Figure E2. Multiplot of all grand average ERPs for Self, Mother, and Fallon appearance responses

Appendix F: Evoked and Induced Power Output

Table F1. Spearman correlations for induced and evoked power

Induced Power		Frontal						Centroparietal					
		Delta	Theta	Alpha	Beta	Lowγ	Highγ	Delta	Theta	Alpha	Beta	Lowγ	Highγ
Frontal	Delta	1.000	.855**	.710**	.791**	.565**	.374**	.827**	.766**	.653**	.724**	.297**	0.069
	Theta	.855**	1.000	.828**	.786**	.399**	.167**	.800**	.881**	.756**	.757**	.185**	-0.050
	Alpha	.710**	.828**	1.000	.831**	.355**	.164**	.682**	.760**	.856**	.778**	.163**	-0.065
	Beta	.791**	.786**	.831**	1.000	.597**	.383**	.695**	.704**	.711**	.861**	.325**	.098*
	LowGamma	.565**	.399**	.355**	.597**	1.000	.872**	.435**	.356**	.314**	.449**	.670**	.549**
	HighGamma	.374**	.167**	.164**	.383**	.872**	1.000	.240**	.146**	.132**	.246**	.639**	.678**
Centroparietal	Delta	.827**	.800**	.682**	.695**	.435**	.240**	1.000	.887**	.742**	.780**	.356**	.106*
	Theta	.766**	.881**	.760**	.704**	.356**	.146**	.887**	1.000	.839**	.783**	.293**	0.043
	Alpha	.653**	.756**	.856**	.711**	.314**	.132**	.742**	.839**	1.000	.825**	.251**	0.005
	Beta	.724**	.757**	.778**	.861**	.449**	.246**	.780**	.783**	.825**	1.000	.420**	.166**
	LowGamma	.297**	.185**	.163**	.325**	.670**	.639**	.356**	.293**	.251**	.420**	1.000	.891**
	HighGamma	0.069	-0.050	-0.065	.098*	.549**	.678**	.106*	0.043	0.005	.166**	.891**	1.000

Evoked Power													
Frontal	Delta	1.000	.498**	.127**	.157**	0.050	0.024	.565**	.404**	.153**	.150**	0.037	0.049
	Theta	.498**	1.000	.488**	.226**	-0.038	-0.052	.408**	.629**	.409**	.302**	0.001	0.054
	Alpha	.127**	.488**	1.000	.502**	-0.043	-0.057	.166**	.379**	.710**	.411**	0.018	0.033
	Beta	.157**	.226**	.502**	1.000	0.017	-0.079	.127**	.242**	.408**	.590**	-0.016	-0.003
	LowGamma	0.050	-0.038	-0.043	0.017	1.000	.161**	0.075	0.072	0.017	-0.078	.195**	-0.005
	HighGamma	0.024	-0.052	-0.057	-0.079	.161**	1.000	0.066	0.008	-0.044	-0.013	0.065	.261**
Centroparietal	Delta	.565**	.408**	.166**	.127**	0.075	0.066	1.000	.618**	.255**	.172**	-0.001	-0.011
	Theta	.404**	.629**	.379**	.242**	0.072	0.008	.618**	1.000	.577**	.342**	-0.012	0.013
	Alpha	.153**	.409**	.710**	.411**	0.017	-0.044	.255**	.577**	1.000	.525**	-0.018	0.058
	Beta	.150**	.302**	.411**	.590**	-0.078	-0.013	.172**	.342**	.525**	1.000	-0.074	-0.023
	LowGamma	0.037	0.001	0.018	-0.016	.195**	0.065	-0.001	-0.012	-0.018	-0.074	1.000	0.078
	HighGamma	0.049	0.054	0.033	-0.003	-0.005	.261**	-0.011	0.013	0.058	-0.023	0.078	1.000

*p < .05; **p < .001

Table F2. Multivariate within-subjects tests for evoked and induced power

Induced Power: Within Subjects Effects		Value	F	Hypothesis df	Error df	Sig.	Partial η²
Person	Pillai's Trace	.584	2.479	36.000	369.000	.000	.195
	Wilks' Lambda	.505	2.586	36.000	358.236	.000	.204
	Hotelling's Trace	.809	2.689	36.000	359.000	.000	.212
	Roy's Largest Root	.525	5.382	12.000	123.000	.000	.344
Word	Pillai's Trace	.229	.818	12.000	33.000	.631	.229
	Wilks' Lambda	.771	.818	12.000	33.000	.631	.229
	Hotelling's Trace	.297	.818	12.000	33.000	.631	.229
	Roy's Largest Root	.297	.818	12.000	33.000	.631	.229
Person * Word	Pillai's Trace	.375	1.466	36.000	369.000	.045	.125
	Wilks' Lambda	.668	1.456	36.000	358.236	.048	.126
	Hotelling's Trace	.435	1.445	36.000	359.000	.052	.127
	Roy's Largest Root	.192	1.968	12.000	123.000	.033	.161

Evoked Power: Within Subjects Effects		Value	F	Hypothesis df	Error df	Sig.	Partial η²
Person	Pillai's Trace	.385	1.510	36.000	369.000	.034	.128
	Wilks' Lambda	.656	1.527	36.000	358.236	.030	.131
	Hotelling's Trace	.463	1.541	36.000	359.000	.028	.134
	Roy's Largest Root	.254	2.603	12.000	123.000	.004	.203
Word	Pillai's Trace	.168	.556	12.000	33.000	.861	.168
	Wilks' Lambda	.832	.556	12.000	33.000	.861	.168
	Hotelling's Trace	.202	.556	12.000	33.000	.861	.168
	Roy's Largest Root	.202	.556	12.000	33.000	.861	.168
Person * Word	Pillai's Trace	.341	1.314	36.000	369.000	.113	.114
	Wilks' Lambda	.694	1.307	36.000	358.236	.117	.114
	Hotelling's Trace	.391	1.300	36.000	359.000	.122	.115
	Roy's Largest Root	.195	1.998	12.000	123.000	.030	.163

a. Design: Intercept
 Within Subjects Design: Person + Word + Person * Word
b. Tests are based on averaged variables.
c. The statistic is an upper bound on F that yields a lower bound on the significance level.
d. Exact statistic

Table F3. Univariate within-subjects tests for induced and evoked power

Source	Induced Power Measure	Type III SS	df	Mean Square	F	Sig.	Partial η²
Person	fDelta	13.975	2.047	6.828	3.510	.033	.074
	fTheta	5.728	2.816	2.034	1.607	.194	.035
	fAlpha	6.657	3.000	2.219	1.240	.298	.027
	fBeta	1.342	2.808	.478	.825	.476	.018
	fLowGamma	5.534	2.086	2.653	4.743	.010	.097
	fHighGamma	6.362	1.927	3.301	4.689	.013	.096
	cDelta	38.589	2.149	17.960	8.310	.000	.159
	cTheta	11.884	2.560	4.643	2.153	.107	.047
	cAlpha	1.596	2.941	.543	.303	.819	.007
	cBeta	1.711	2.436	.702	1.230	.300	.027
	cLowGamma	17.386	1.943	8.947	9.982	.000	.185
	cHighGamma	16.637	1.846	9.012	13.390	.000	.233
Person * Word	fDelta	1.821	2.888	.630	.781	.503	.017
	fTheta	4.161	2.984	1.394	1.180	.320	.026
	fAlpha	9.959	3.000	3.320	2.124	.100	.046
	fBeta	1.700	3.000	.567	1.219	.305	.027
	fLowGamma	.417	2.877	.145	.614	.601	.014
	fHighGamma	2.598	3.000	.866	4.450	.005	.092
	cDelta	5.344	2.997	1.783	1.853	.141	.040
	cTheta	1.874	3.000	.625	.568	.637	.013
	cAlpha	4.249	3.000	1.416	.921	.433	.021
	cBeta	3.788	3.000	1.263	2.065	.108	.045
	cLowGamma	1.063	2.827	.376	1.272	.287	.028
	cHighGamma	.161	2.794	.058	.240	.626	.005

Source	Evoked Power Measure	Type III SS	df	Mean Square	F	Sig.	Partial η²
Person	fDelta	7.379	2.718	2.715	2.638	0.058	0.057
	fTheta	7.790	2.881	2.704	2.177	0.096	0.047
	fAlpha	1.861	3.000	0.620	0.341	0.796	0.008
	fBeta	2.271	2.801	0.811	1.531	0.212	0.034
	fLowGamma	0.919	2.852	0.322	2.331	0.081	0.050
	fHighGamma	0.294	3.000	0.098	1.138	0.336	0.025
	cDelta	13.941	2.350	5.932	4.426	0.010	0.091
	cTheta	1.632	2.892	0.564	0.369	0.768	0.008
	cAlpha	5.359	3.000	1.786	0.979	0.405	0.022
	cBeta	2.032	2.589	0.785	1.776	0.163	0.039
	cLowGamma	0.807	3.000	0.269	1.811	0.148	0.040
	cHighGamma	0.370	3.000	0.123	1.075	0.362	0.024

Data is Huynh-Feldt corrected.

Table F4. Post-hoc person comparisons for evoked and induced power

Induced Power		fDelta	fTheta	fAlpha	fBeta	fLow Gamma	fHigh Gamma	cDelta	cTheta	cAlpha	cBeta	cLow Gamma	cHigh Gamma
Self	Mother	0.170	0.029	-0.057	-0.108	.193*	0.150	0.205	-0.043	0.066	-0.109	0.001	0.099
	Fallon	-0.330	-0.230	-0.017	-0.168	-0.073	-0.103	-0.221	-0.147	0.006	-0.190	0.024	-0.032
	Case	0.134	0.110	-0.335	-0.117	0.217	0.241	.666*	0.337	-0.118	-0.064	.515*	.506*
Mother	Self	-0.170	-0.029	0.057	0.108	-.193*	-0.150	-0.205	0.043	-0.066	0.109	-0.001	-0.099
	Fallon	-.499*	-0.259	0.040	-0.059	-.266*	-.253*	-.427*	-0.104	-0.060	-0.081	0.023	-0.131
	Case	-0.035	0.081	-0.278	-0.009	0.024	0.091	0.460	0.381	-0.184	0.044	.514*	.407*
Fallon	Self	0.330	0.230	0.017	0.168	0.073	0.103	0.221	0.147	-0.006	0.190	-0.024	0.032
	Mother	.499*	0.259	-0.040	0.059	.266*	.253*	.427*	0.104	0.060	0.081	-0.023	0.131
	Case	0.464	0.340	-0.318	0.050	0.290	0.345	.887*	0.485	-0.123	0.125	.491*	.538*
Case	Self	-0.134	-0.110	0.335	0.117	-0.217	-0.241	-.666*	-0.337	0.118	0.064	-.515*	-.506*
	Mother	0.035	-0.081	0.278	0.009	-0.024	-0.091	-0.460	-0.381	0.184	-0.044	-.514*	-.407*
	Fallon	-0.464	-0.340	0.318	-0.050	-0.290	-0.345	-.887*	-0.485	0.123	-0.125	-.491*	-.538*

| Evoked Power		fDelta	fTheta	fAlpha	fBeta	fLow Gamma	fHigh Gamma	cDelta	cTheta	cAlpha	cBeta	cLow Gamma	cHigh Gamma
Self	Mother	-0.184	-0.096	-0.008	0.083	-0.089	-0.043	-0.194	-0.029	-0.169	0.098	0.112	-0.016
	Fallon	0.211	0.094	-0.155	0.017	0.050	0.036	0.187	-0.043	-0.208	0.078	0.027	0.037
	Case	-0.063	-0.303	-0.139	-0.135	-0.036	-0.016	-0.334	-0.175	-0.342	-0.093	-0.006	-0.052
Mother	Self	0.184	0.096	0.008	-0.083	0.089	0.043	0.194	0.029	0.169	-0.098	-0.112	0.016
	Fallon	.395*	0.190	-0.147	-0.066	0.138	0.079	.381*	-0.013	-0.040	-0.020	-0.085	0.053
	Case	0.121	-0.207	-0.131	-0.219	0.052	0.027	-0.140	-0.146	-0.173	-0.191	-0.119	-0.036
Fallon	Self	-0.211	-0.094	0.155	-0.017	-0.050	-0.036	-0.187	0.043	0.208	-0.078	-0.027	-0.037
	Mother	-.395*	-0.190	0.147	0.066	-0.138	-0.079	-.381*	0.013	0.040	0.020	0.085	-0.053
	Case	-0.274	-0.397	0.016	-0.152	-0.086	-0.052	-0.521	-0.133	-0.133	-0.171	-0.033	-0.089
Case	Self	0.063	0.303	0.139	0.135	0.036	0.016	0.334	0.175	0.342	0.093	0.006	0.052
	Mother	-0.121	0.207	0.131	0.219	-0.052	-0.027	0.140	0.146	0.173	0.191	0.119	0.036
	Fallon	0.274	0.397	-0.016	0.152	0.086	0.052	0.521	0.133	0.133	0.171	0.033	0.089

Based on estimated marginal means
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Sidak.

Table F5. Post-hoc pairwise person*word comparisons for induced power

Induced Power (I)	(J)			Frontal						Centroparietal			
		Delta	Theta	Alpha	Beta	Low Gamma	High Gamma	Delta	Theta	Alpha	Beta	Low Gamma	High Gamma
Character Self	Mother	0.260	0.017	0.336	0.026	0.259	.289*	0.296	-0.122	0.371	-0.157	0.016	0.094
	Fallon	-0.171	-0.351	-0.022	-0.003	0.019	0.102	0.077	-0.034	0.123	-0.067	0.143	0.009
	Case	0.117	-0.158	-0.296	-0.085	0.255	0.260	.924*	0.297	0.034	-0.228	.494*	.492*
Mother Self	-0.260	-0.017	-0.336	-0.026	-0.259	-.289*	-0.296	0.122	-0.371	0.157	-0.016	-0.094
	Fallon	-0.431	-0.368	-0.359	-0.029	-0.241	-0.187	-0.219	0.089	-0.247	0.090	0.128	-0.085
	Case	-0.143	-0.175	-0.633	-0.112	-0.004	-0.029	0.629	0.420	-0.337	-0.071	.478*	0.399
Fallon Self	0.171	0.351	0.022	0.003	-0.019	-0.102	-0.077	0.034	-0.123	0.067	-0.143	-0.009
	Mother	0.431	0.368	0.359	0.029	0.241	0.187	0.219	-0.089	0.247	-0.090	-0.128	0.085
	Case	0.288	0.194	-0.274	-0.083	0.236	0.158	.847*	0.331	-0.090	-0.161	0.350	.483*
Case Self	-0.117	0.158	0.296	0.085	-0.255	-0.260	-.924*	-0.297	-0.034	0.228	-.494*	-.492*
	Mother	0.143	0.175	0.633	0.112	0.004	0.029	-0.629	-0.420	0.337	0.071	-.478*	-0.399
	Fallon	-0.288	-0.194	0.274	0.083	-0.236	-0.158	-.847*	-0.331	0.090	0.161	-0.350	-.483*
Appearance Self	Mother	0.079	0.041	-0.450	-0.243	0.128	0.011	0.115	0.036	-0.238	-0.060	-0.014	0.105
	Fallon	-0.488	-0.108	-0.012	-0.332	-0.165	-.309*	-.520*	-0.261	-0.112	-0.312	-0.096	-0.073
	Case	0.151	0.377	-0.374	-0.149	0.180	0.222	0.407	0.378	-0.269	0.099	.537*	.520*
Mother Self	-0.079	-0.041	0.450	0.243	-0.128	-0.011	-0.115	-0.036	0.238	0.060	0.014	-0.105
	Fallon	-.568*	-0.149	0.438	-0.090	-.292*	-.319*	-.635*	-0.297	0.126	-0.252	-0.082	-0.178
	Case	0.072	0.336	0.076	0.094	0.052	0.212	0.292	0.342	-0.031	0.159	.550*	.415*
Fallon Self	0.488	0.108	0.012	0.332	0.165	.309*	.520*	0.261	0.112	0.312	0.096	0.073
	Mother	.568*	0.149	-0.438	0.090	.292*	.319*	.635*	0.297	-0.126	0.252	0.082	0.178
	Case	0.640	0.486	-0.362	0.184	0.345	.531*	.927*	0.639	-0.157	0.411	.633*	.593*
Case Self	-0.151	-0.377	0.374	0.149	-0.180	-0.222	-0.407	-0.378	0.269	-0.099	-.537*	-.520*
	Mother	-0.072	-0.336	-0.076	-0.094	-0.052	-0.212	-0.292	-0.342	0.031	-0.159	-.550*	-.415*
	Fallon	-0.640	-0.486	0.362	-0.184	-0.345	-.531*	-.927*	-0.639	0.157	-0.411	-.633*	-.593*
Self	Character-Appearance	-0.005	-0.214	0.027	0.142	0.017	0.115	0.115	-0.197	0.093	-0.045	0.037	-0.054
Mother	Character-Appearance	-0.186	-0.190	-.760*	-0.127	-0.114	-.164*	-0.065	-0.039	-.516*	0.052	0.008	-0.043
Fallon	Character-Appearance	-0.322	0.029	0.037	-0.188	-0.166	-.296*	-.482*	-0.425	-0.143	-0.290	-0.202	-0.137
Case	Character-Appearance	0.029	0.321	-0.051	0.079	-0.057	0.077	-0.403	-0.117	-0.210	0.282	0.080	-0.027

Values are differences (I-J) between estimated marginal means.
*. The mean difference is significant at the .05 level.
Adjustment for multiple comparisons: Sidak.
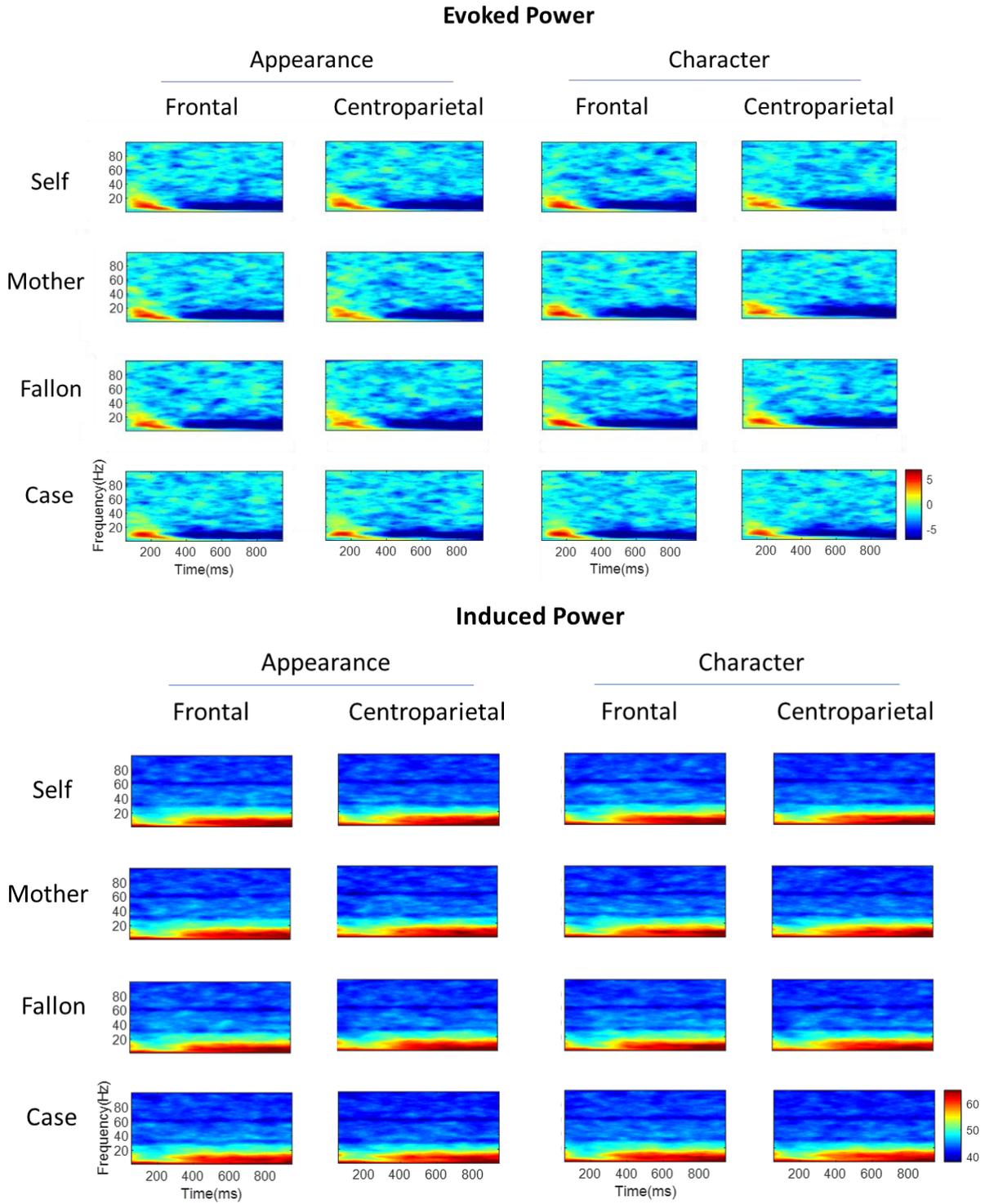
Figure F1. Evoked and induced power (μV^2) by person, word, and region

Appendix G: Source Analysis Output
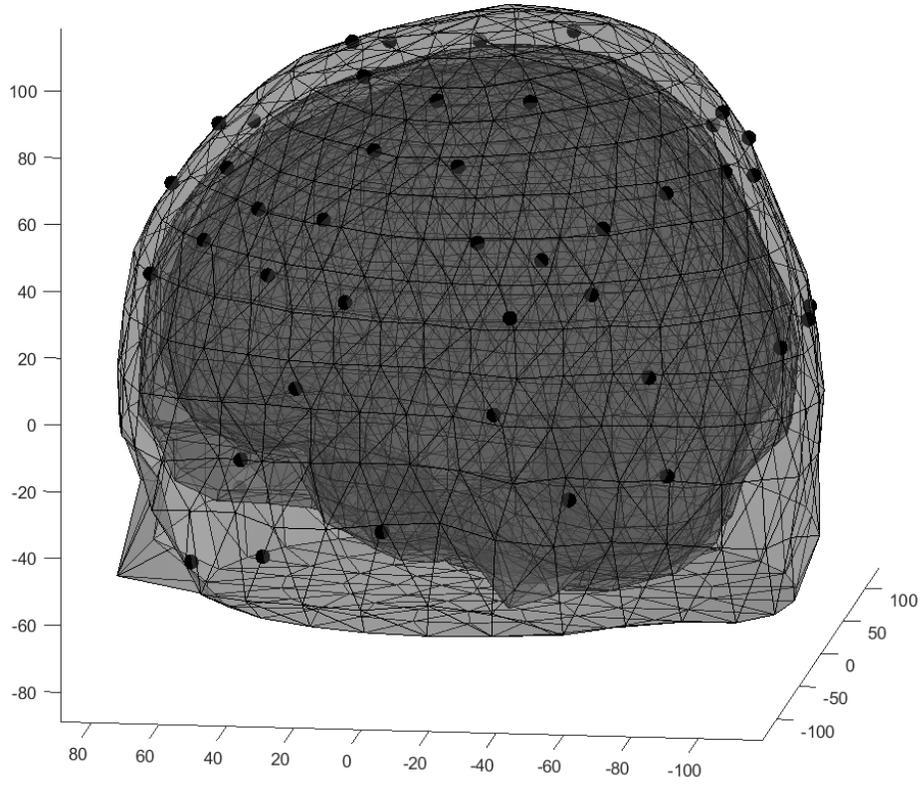
**Figure G1.** Electrode alignment to the standard BEM head model
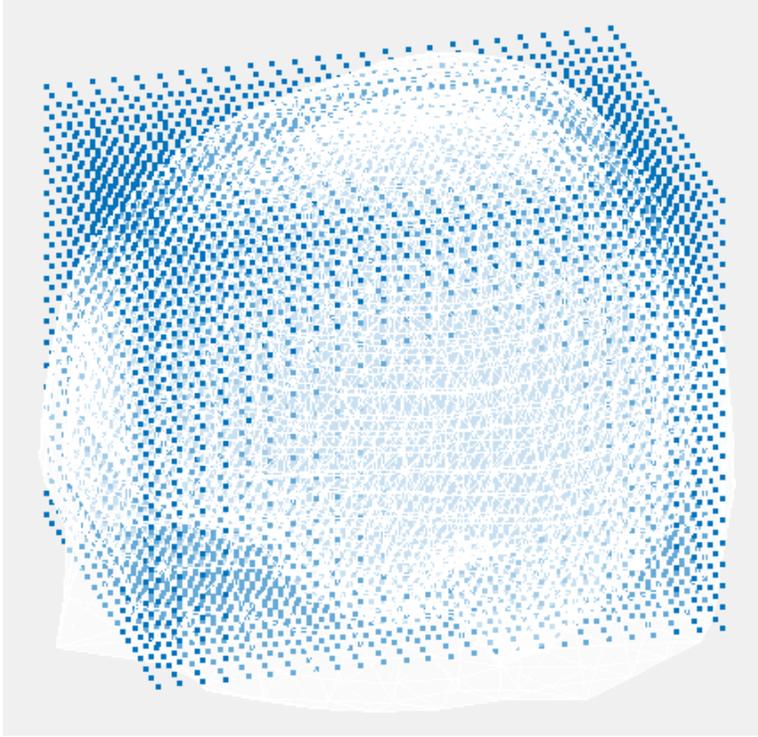
**Figure G2.** Leadfield alignment with the standard BEM head model
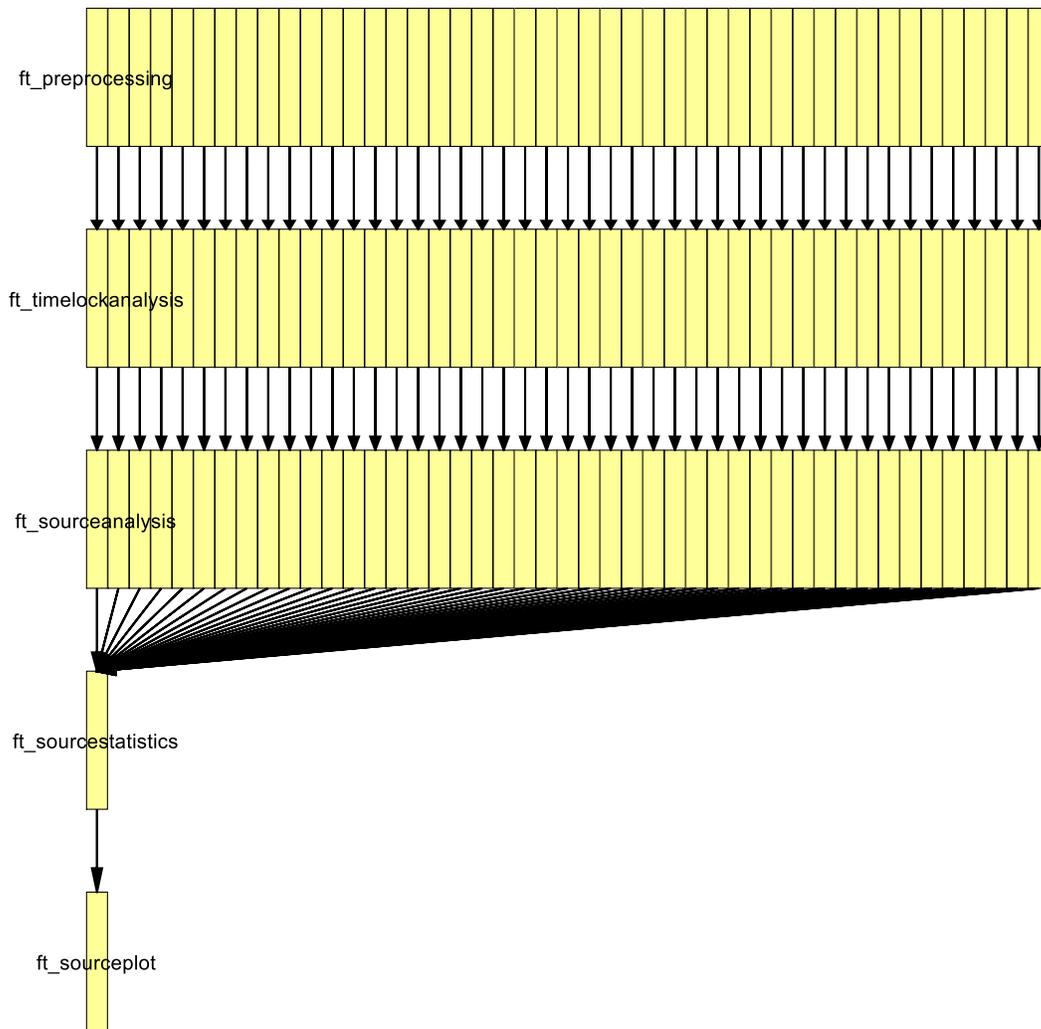
Figure G3. Source plot procedure pipeline

Appendix H: HSMM-MVPA Output

Table H1. Pairwise comparisons of stage durations for all subjects in 3-bump HSMM-MVPA models of self, mother, and Fallon for character and appearance

Character

Stage	(I) Person	(J) Person	Mean Difference (I-J)	Std. Error	Sig.
1	Self	Mother	-8.44	4.82	.238
		Fallon	99.98*	6.53	.000
	Mother	Self	8.44	4.82	.238
		Fallon	108.42*	7.77	.000
2	Self	Mother	5.29*	1.59	.005
		Fallon	-74.68*	4.08	.000
	Mother	Self	-5.29*	1.59	.005
		Fallon	-79.97*	4.26	.000
3	Self	Mother	11.63*	3.87	.013
		Fallon	-2.59	3.64	.859
	Mother	Self	-11.63*	3.87	.013
		Fallon	-14.22*	4.78	.014
4	Self	Mother	-8.48	3.83	.093
		Fallon	-22.71*	3.79	.000
	Mother	Self	8.48	3.83	.093
		Fallon	-14.23*	5.07	.022

Appearance

Stage	(I) Person	(J) Person	Mean Difference (I-J)	Std. Error	Sig.
1	Self	Mother	-6.381	4.269	.369
		Fallon	106.149	6.444	.000
	Mother	Self	6.381	4.269	.369
		Fallon	112.530	6.736	.000
2	Self	Mother	-.114	2.102	1.000
		Fallon	-79.730	4.145	.000
	Mother	Self	.114	2.102	1.000
		Fallon	-79.616	4.503	.000
3	Self	Mother	4.691	3.216	.390
		Fallon	-4.351	3.162	.440
	Mother	Self	-4.691	3.216	.390
		Fallon	-9.042	3.605	.047
4	Self	Mother	1.804	4.226	.965
		Fallon	-22.068	3.600	.000
	Mother	Self	-1.804	4.226	.965
		Fallon	-23.872	4.904	.000

Sidak corrected estimated marginal means. Extra Fallon rows excluded to save space. No information was lost.

Table H2. Pairwise comparisons of bump magnitudes for all subjects in 3-bump HSMM-MVPA models of self, mother, and Fallon collapsed across character and appearance

Bump	(I) Person	(J) Person	Mean Difference (I-J)	Std. Error	Sig.[b]
1	Self	Mother	.093[*]	.001	.000
		Fallon	.266[*]	.001	.000
	Mother	Self	-.093[*]	.001	.000
		Fallon	.173[*]	.001	.000
	Fallon	Self	-.266[*]	.001	.000
		Mother	-.173[*]	.001	.000
2	Self	Mother	.040[*]	.001	.000
		Fallon	.098[*]	.001	.000
	Mother	Self	-.040[*]	.001	.000
		Fallon	.058[*]	.001	.000
	Fallon	Self	-.098[*]	.001	.000
		Mother	-.058[*]	.001	.000
3	Self	Mother	-.027[*]	.000	.000
		Fallon	-.195[*]	.001	.000
	Mother	Self	.027[*]	.000	.000
		Fallon	-.168[*]	.001	.000
	Fallon	Self	.195[*]	.001	.000
		Mother	.168[*]	.001	.000

Sidak corrected estimated marginal means.
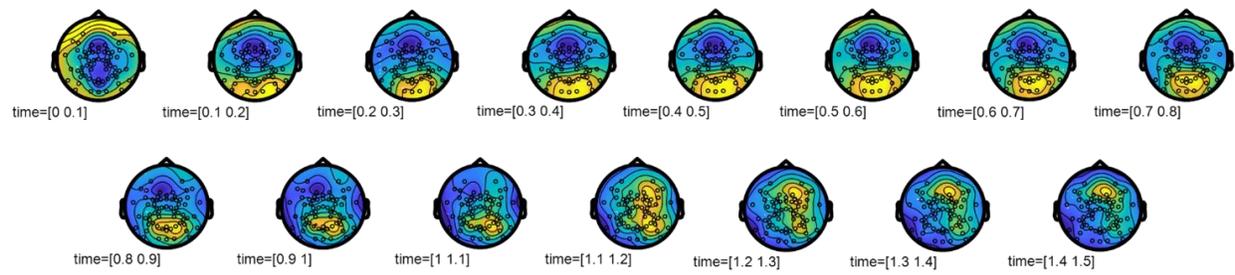
Figure H1. Topoplots of averaged ERP data from the self-character condition from 0 to 1.5 seconds using Fieldtrip. All other conditions showed similar topographies but on slightly different timescales, especially for case.

4-Bump Model

Self
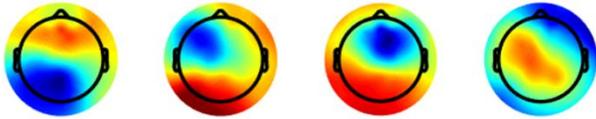
Mother
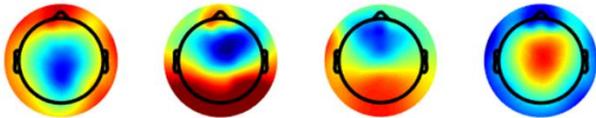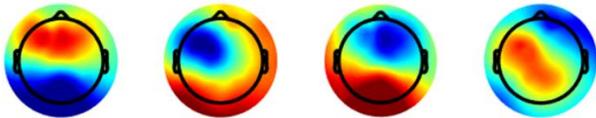
Fallon

5-Bump Model

Self
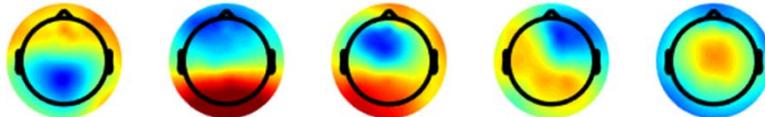
Mother
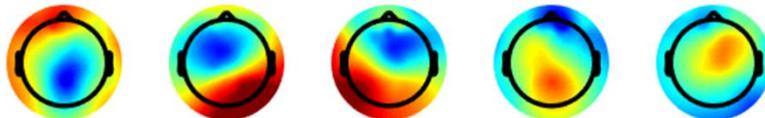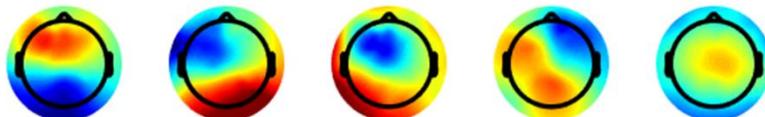
Fallon

Figure H2. Topographies of self, mother, and Fallon for 4 and 5-bump HSMM-MVPA models. Like the 3-bump model, these model has considerable similarity to the topoplots in Figure H1.