DOCTOR OF PHILOSOPHY

Probing the impact of familiarity on effective connectivity within the Action Observation Network

Gardner, Thomas

Award date:
2016

Awarding institution:
Bangor University

Probing the impact of familiarity on effective connectivity within the Action Observation Network

Thomas Gardner

PRIFYSGOL
BANGOR
UNIVERSITY

Submitted for the degree of Doctor of Philosophy

2016

School Of Psychology

Bangor University

Supervisor: Dr Emily S. Cross
School of Psychology
Bangor University
UK

External Examiner: Dr James Kilner
Institute of Neurology
University College London
UK

Internal Examiner: Dr Patricia Bestelmeyer
School of Psychology
Bangor University
UK

Acknowledgements

First and foremost, I thank Emily Cross. Thank you for giving me this opportunity. Your enthusiasm and assurance have made this possible, as well as allowing me to aim big in all three of the empirical chapters (let alone the guitar creation).

I would also like to thank Richard Ramsey, my second supervisor, and Paul Downing, my PhD committee chair, for their valuable suggestions through the entire process.

My appreciation goes to Dave McKiernan for his patience with my demands for a fMRI safe guitar, as well as Paul Mullins and Andrew Fischer for their technical assistance in getting a signal out of the scanner room.

To my family, thank you for each providing unique assistance in all forms to allow me to pursue this career. In a similar vein, a thank you to James Chandler and Kristian Billimoria for their un-wielding belief in me.

Lastly, and by no means least, a huge thank you to my partner Zoe Oliver for her unconditional support in this process. From her reading of countless drafts, listening to my explanations of results and volunteering to be the guitar expert; I cannot fully express my gratitude for your help.

Contents

List Of Tables

List of Figures

Abstract

This thesis reports the results of three novel studies that used effective connectivity and region of interest methods on fMRI data to provide insights into how familiarity shapes action observation and execution.

Experiment 1 examined how observed movement familiarity modulates AON activity using dynamic causal modelling, a type of effective connectivity analysis. Participants viewed whole-body dance movements during scanning, and after scanning rated each movement on a measure of visual familiarity. These ratings were then applied as parametric modulators to the fMRI data, which revealed an attenuation of effective connectivity bidirectionally between parietal and temporal AON nodes when participants observed videos they rated as increasingly familiar. As such, the findings provide partial support for a predictive coding model of the AON, as well as illuminate how action familiarity manipulations can be used to explore simulation-based accounts of action understanding.

Experiment 2 examined the relationship between AON response amplitude and participants' familiarity with observed or executed actions. Specifically, this study examined whether increasing familiarity impacts AON engagement in a linear or quadratic manner. Using an elaborate guitar training intervention to probe the relationship between familiarity and AON engagement during action execution and action observation tasks, participants underwent fMRI scanning while executing one set of guitar sequences and observing a second set of sequences. Via region of interest analyses, linear, cubic and quadratic regression models were fitted to the data to match varying levels of familiarity. The data from the observation and execution conditions show mixed evidence for all types of models, suggesting that the response profile within key sensorimotor brain regions associated with the AON is not solely linear in nature in response to increasing familiarity. Moreover, by probing the objective and subjective nature of the prediction error signal, we show results that are consistent with a predictive coding account of AON engagement during action observation and execution.

The final study of this thesis aimed to test the assumptions of the predictive coding account using effective connectivity, when participants observed or executed familiar compared to unfamiliar actions. To test these predictions, we re-evaluated the same data collected for Study 2 wherein participants took part in an intensive guitar-training paradigm. Identifying core AON nodes from pre- and post-training scanning sessions, we then applied effective connectivity analyses to test whether changes in effective connectivity fit those hypothesised under the predictive coding account. We demonstrate that hypotheses derived from predictive coding that predict distinct patterns of modulation based on perceived or performed actions' familiarity are generally supported by the empirical data. These findings contribute valuable insights toward understanding the complex role played by familiarity in modulating action cognition.

The main empirical findings of this thesis show: 1) attenuation in connectivity within the AON when an action is perceived as more familiar; 2) the response profile of core AON regions to increasing familiarity (either objectively or subjectively defined) when performing or observing an action is not solely linear in nature; and 3) hypotheses derived from the predictive coding account concerning effective connectivity between core AON regions are largely supported when an intensive training paradigm is used to create a distinction between familiar and unfamiliar actions.

1 Chapter I

1.1 Aim of the thesis

The overarching aim of this thesis was to characterise how changes in familiarity shape connectivity within the Action Observation Network (AON). In regards to the organization of the chapters within this thesis, the three empirical chapters are written as standalone sections and are manuscripts that are either already published (Chapter III: Gardner, Goulden & Cross (2015) *Journal of Neuroscience)* or are currently under review (Chapter IV: Gardner, Aglinskas & Cross, under review at *Neuroimage*; Chapter V: Gardner & Cross, under review at *Journal of Neuroscience*). For this reason, some content is repeated between chapters.

In Chapter I, a brief outline of the Action Observation Network is presented, and the role played by familiarity in shaping AON engagement is considered. Prominent models (the direct matching hypothesis and the predictive coding account) are outlined with respect to how they might explain discrepant findings in the AON literature concerning the role of familiarity. Chapter II comprises a brief overview of fMRI, followed by a more extensive review of Dynamic Causal Modelling, the main analytical approach used in Chapters III and V of this thesis.

In Chapter III, the impact of subjectively-rated familiarity on the effective connectivity within the AON is examined. In Chapter IV, an intensive training paradigm is combined with region of interest analyses to examine the response profile of core AON regions and test predictions of two prominent models within the field: direct matching and predictive coding. In Chapter V, the same data from Chapter IV is reanalyzed using effective connectivity measures to test how varying levels of familiarity influence modulation between core AON regions. Finally, in Chapter VI, the results of the three empirical chapters are considered together, as are general limitations and future directions for this line of research.

1.2 Background

1.2.1 Linking Action with Perception

Humans are a highly social species in near-constant physical or visual contact with conspecifics. When we observe others moving around us, we are able to make many rich inferences about others' goals, intentions, and desires. For example, if we imagine sitting in a café and observing a friend reach out to grasp her coffee cup, from this action, we might infer that our friend is thirsty, or wishes to get warm by imbibing a hot beverage, or perhaps is hoping a dose of caffeine helps wake her up (Hamilton & Grafton, 2006; Hamilton, 2013; Ramsey & Hamilton, 2010). The goal of her reaching action would be to grasp the coffee cup, with the intention of lifting it up to take a drink from it. Integral to this concept is that we actively predict the actions as they unfold, allowing us to rapidly update predicted subsequent actions in an online fashion (Blakemore & Frith, 2005; Falck-Ytter et al., 2006). If, for example, our friend in the café appears to reach for her coffee cup but deviates towards the sugar, our understanding of her intention must change; no longer is the goal to grasp the coffee cup. In evolutionary terms, this rapid identification of goals from actions would be beneficial in attributing others' movements as threatening or friendly, and thus preparing us to respond accordingly (see Hauser & Wood (2010) for further discussion of this point). Our ability to obtain such rich information from others' actions likely stems from a particular type of neuron found within sensorimotor cortical regions of the primate brain: so-called "mirror neurons".

In 1992, di Pellegrino and colleagues aimed to investigate motor properties of the ventral premotor area F5 of the macaque monkey's brain (di Pellegrino et al, 1992). It was previously found that area F5, located rostrally within the inferior premotor cortex, responded when a monkey made arm movements (Rizzolatti et al, 1981; 1988). To test area F5's responsiveness to stimulus compared to movement related activity, single neuron recordings were taken. The task for the macaque involved passively observing the food (stimulus condition) or reaching out and grabbing the food (movement related condition). Incidentally, the authors noted that when the experimenter reached over to move the food back to its starting location, the

single neuron recording activity resembled that of when the monkey was performing the movement related condition. In subsequent papers by the same group of researchers, the incidental finding of di Pellegrino and colleagues (1992) was explicitly tested by use of the same methodology (Gallese et al, 1996; Rizzolatti et al, 1996; see also Umilta et al., 2001) and the finding was reliably replicated. These neurons in F5, which were found to respond to both observation and execution of the same or similar arm movements, were coined 'Mirror Neurons' (Gallese et al, 1996; Rizzolatti et al, 1996). Fogassi et al. (2005) further investigated the presence and attributes of mirror neurons in the macaque brain by extending focus to the inferior parietal lobule (IPL). Monkeys were required to observe grasping actions where the experimenter grasped to place or grasped to eat. The findings from the single neuron recordings showed different activations for grasping to eat compared to grasping to place. The researchers suggested that the action (in this case, grasping), is coded differently based on the final action goal. This conclusion can be drawn from the fact that the action was similar, however, the final outcome was different. This conclusion suggests that activity within IPL helps an observer to ascertain the intention of an observed agent (in this case, the experimenter). The finding of differential coding of initial action and final intention was later supported by Bonini and colleagues (2010). Bonini and colleagues aimed to investigate whether activity within ventral premotor area and IPL reflects the final action goal of a movement, which is embedded within the actions. The final action goal could be eating or placing, preceded by a grasping action (observed and executed by the macaque, allowing of the investigation of mirror neuron activity). Therefore, the kinematics of the actions was similar (grasp the object or food), yet the intention was only attainable at the latter stages of the action when the object or food was placed in the box or mouth, respectively for the conditions. Single unit recordings were taken from F5 (ventral premotor area) and area PFG of the IPL (an area that is cytoarchitecturally distinct within the IPL; Gregoriou et al., 2006). These two regions were used, as tight connections between them were previously shown (Cavada & Goldman-Rakic, 1989; Rozzi et al. 2006); therefore, functional coupling is plausible (Bonini et al., 2010). The results of this experiment were that the two regions encode different levels of motor abstraction. It was found that the mirror neuron activity found in F5 coded actions according to the final goal. However, compared to F5, PFG was preferentially stronger for the coding of

kinematics of the actions. Together, these findings indicate that this frontal-parietal circuit plays an important role in understanding the intentions of others.

1.2.2 From Mirror Neurons to the Action Observation Network

These seminal mirror neuron findings and others lent support to the notion that observed and executed actions could share representations within a similar system (Gallese & Goldman, 1998; Rizzolatti et al., 2001). Returning to the example used previously, when we observe our friend reaching for her coffee cup, we are able to understand the intention and goal of her action by simulating this same action within our own motor system, a concept that shall be returned to later in this chapter. However, it is important to note the findings described above relating to mirror neurons responding in a similar manner to observed and executed actions all came from investigations into the macaque brain. The fine-scale precision for measuring neural activity afforded by single-unit recordings is difficult to achieve in human participants, as single recordings in humans is very rare (but see Mukamel et al., 2010). However, a number of innovative attempts using human neuroimaging approaches, such as functional magnetic resonance imaging (fMRI), offer compelling evidence that similar neural mechanisms are at play within the human brain. For example, Kilner and colleagues (2009a) used a repetition suppression paradigm (where repeated presentation of a particular stimulus or feature results in reduced activity in a particular brain region of interest) to demonstrate that activity in the left inferior frontal gyrus (IFG) is consistent with what would be expected if this brain region contains mirror neurons. In contrast, Lingnau and colleagues (2009) showed that this finding did not emerge when the repeated stimuli were first executed and then observed. As such, Lingnau and colleagues (2009) suggest that their failure to find evidence for cross-modal adaptation for executed and observed motor acts calls into question core assumptions made by mirror neuron theories. When considered together, these findings suggest that identification of actual mirror *neurons* in humans will require further investigation (see Kilner et al., 2014 for applying adaption techniques in macaques). Whether these neurons do indeed exist in the human primate brain as they do in the non-human primate brain (c.f., Hickok, 2009; Hickok & Hauser, 2010; Dinstein et al, 2007; Molenberghs et al., 2012), will remain a rich area

for future work to further explore. However, what this work by Kilner and colleagues (2009a), Lingnau and colleagues (2009), and many others suggests is that at least *some* degree of correspondence clearly exist between executed and observed actions within sensorimotor circuits of the human brain. While some researchers remain committed to resolving the debate as to whether mirror neurons exist in the human brain, per se, others have expanded their focus beyond just inferior parietal and ventral premotor cortices to focus on the role played by sensorimotor cortical areas more broadly when engaging in action-related tasks.

When we observe others in action, a wealth of human neuroimaging studies have demonstrated widespread engagement of frontal, parietal, and occipitotemporal cortical regions, which have been collectively termed the Action Observation Network (AON) (Cross et al., 2009; Grafton, 2009; Keysers & Gazzola, 2009; Caspers et al., 2010). This extensive network includes the inferior frontal gyrus and inferior parietal lobe, homologues to area F5 and IPL in the macaque mirror neuron literature. In addition to these two regions, the posterior superior (and sometimes middle) temporal gyrus (STG/MTG) are also included as part of this network (Grafton, 2009; Cross et al., 2009), As these posterior temporal regions do not demonstrate motor properties, it is clear that the AON does not require strict correspondence between observed and executed actions, but is instead more broadly concerned with brain regions engaged when observing others in action (Iacoboni et al., 2001; Gallese et al., 2004; Schippers & Keysers, 2011). The main focus of this thesis concerns what happens to activity within these core AON regions when an observed or executed action is highly familiar vs. unfamiliar.

1.2.3 The Role of familiarity in Shaping the AON

1.2.3.1 Greater familiarity = Greater AON Engagement

Research into how familiarity shapes AON engagement when observing others in action has investigated familiarity in several domains, including an observed agent's form and motion (Press, 2011). Research investigating the role of familiarity

with an observed agent's form has shown greater activity within the AON when an observed agent is human compared to non-human (such as a dog or robot; Buccino et al., 2004; Tai et al., 2004; Shimada, 2010; Press, 2011). For example, Buccino and colleagues (2004) used fMRI to investigate AON activity while human participants viewing biting, speech reading, lip smacking and barking actions performed by humans, monkeys and dogs. The results showed that, within core AON regions, stronger engagement of the AON was elicited when the observed agent and actions were more familiar (i.e., more AON engagement when viewing speech reading and biting actions made by a human than when viewing barking actions made by a dog). As noted by the authors, this finding suggests that correspondence between an observer and observed agent's form and motor repertoire is crucial for engaging the AON during action observation. The finding from Buccino and colleagues (2004) was corroborated by Tai and colleagues (2004), who scanned participants' brains as they observed grasping actions performed by a human model or a robotic hand. They found premotor activity when participants observed grasping actions made by the human model, but not the robotic model. This finding was further supported by Shimada (2010) who found that compared to a congruent agent-action combinations (i.e., a human form performing an action with human kinematics), sensorimotor engagement attenuated within when the agent-action pairing was incongruent (human form performing a robotic movement). Taken together, these findings suggest that AON engagement is greater for when observing familiar compared to unfamiliar agents.

Support for greater AON engagement for familiar compared to less familiar kinematics has also been demonstrated by a large number of studies (Knoblich & Flach, 2001; Calvo-Merino et al., 2005; Cross et al., 2006; Macuga & Frey, 2011; Bischoff et al., 2012; Kirsch & Cross, 2015). For example, Calvo-Merino and colleagues (2005) recruited expert dancers from different dance backgrounds (classical ballet vs. capoeira) to participate in an fMRI study. When these dancers view videos of actions they had trained (e.g. ballet dancer watching ballet), greater activity emerged across the AON compared to when they observed actions not in their motor repertoire (e.g., ballet dancer watching capoeira). Corroborating and extending this finding, Cross and colleagues (2006) took professional dancers and trained them

on novel dance movements over a period of 5 weeks. The dancers were scanned once a week across the training period while they watched rehearsed movements and kinematically similar control movements. The authors showed that activity in the core AON regions was modulated as a function of the dancers' perceived ability to perform each dance action in that the better they could perform an observed action, the stronger the response was within PMv and IPL. Both these studies (Calvo-Merino et al., 2005 and Cross et al., 2006) suggest that even if an observed movement falls within an observer's general movement repertoire, activity within the AON is attuned based on the observer's ability to perform the action.

In another line of research aiming to explore the impact of observing familiar vs. unfamiliar actions on AON engagement, Liew and colleagues (2011) investigated perception of non-verbal communication of Caucasian and Chinese symbolic gestures by Caucasian and Chinese actors as Chinese participants underwent fMRI. The findings of this study were twofold. First, the authors reported greater AON activity when the actor was of the same race as the participant, supporting the notion that the AON responds more robustly to more familiar agents. However, the second finding of this study raises a challenge to the action familiarity findings reported in the preceding paragraph, in that Liew and colleagues (2011) found that the unfamiliar actions (i.e., Caucasian symbolic gestures observed by Chinese participants) elicited greater AON activity than familiar actions (e.g., Chinese symbolic gestures observed by Chinese participants). This contradictory finding of the role played by familiarity within the AON from the same study brings forth the point that our understanding of the role of familiarity remains unresolved.

1.2.3.2 Less familiarity = Greater AON Engagement?

A growing body of literature has added further fuel to this debate by showing that *unfamiliar* reliably elicit greater AON engagement (Gazzola et al., 2007; Liew et al., 2011; Cross et al., 2012). For example, Gazzola and colleagues (2007) found evidence for equivalent AON activation when participants observed grasping actions performed by a robotic or human hand, at odds with the findings reported by Tai et al. (2004). In this study by Gazzola and colleagues (2007), participants observed robotic

and human arms carry out goal directed or non-goal directed actions. They found that observing actions that were designed to achieve with a specific goal (kinematics aside), was sufficient to activate the AON. In addition, they report that regardless of agent, the repeated presentation of the identical stimuli in a block caused strong habituation of the AON (c.f. notion of repetition suppression presented above in section 1.2.2, also Hamilton & Grafton, 2006). These findings suggest that a linear relationship between increasing familiarity and increasing AON activity might be limited in its explanation of this network.

In a subsequent study, Cross and colleagues (2012) investigated how observing whole body dance movements impacts AON engagement. In this study, participants observed either a human dancer or an animated Lego robot perform actions that were either smooth and human-like in nature, or adhered to more rigid and robotic kinematics. Across two independent fMRI experiments, these authors found greater activity within core AON regions when participants observed the less familiar, rigid robotic movements compared to the more familiar human-like movements. These authors proposed a model whereby the relationship between AON engagement and familiarity was quadratic (U-shaped) rather than the linear assumptions of familiar actions and AON engagement. This model was further supported by Liew and colleagues (2013), in a study in which they investigated the role of familiarity in novices, experienced occupational therapists (OC) who had more extensive experience working with individuals with residual limbs, and subject CJ, who was born with a below-elbow residual limb. By studying these three groups of participants, the authors could investigate how varying levels of visual and physical familiarity with different types of intact and residual limb movements might modulate AON activity. While undergoing fMRI, subjects viewed both residual limb and hand actions. The authors found that for novices, greater AON activity emerged for residual limb viewings than hand actions. This was also shown to a lesser degree for OC participants, and in both cases, this activity attenuated with time. These findings provide further evidence that the observation of unfamiliar actions also elicits AON engagement. Additionally, AON engagement for the observation of residual limbs was greater than the hand actions for subject CJ, suggesting that extremely familiar

actions can also elicit robust AON engagement, a finding that is consistent with the U-shaped function proposed by Cross and colleagues (2012).

A fundamental question that remains the literature reviewed in both parts of section 1.2.3 is what, precisely, the role of familiarity is in shaping AON engagement. It is clear from the literature that AON engagement is not preferential for either familiar or unfamiliar actions, with evidence reviewed here to support both sides of this debate. However, one lesson we might take away from this debate is that continued use of magnitude-based approaches might have limited utility in shedding fresh light on this question (see Smith, 2012 for a discussion on the limitations of magnitude-based approaches). The model proposed by Cross et al. (2012) and Liew et al. (2013) frames the findings of the familiarity literature in a way in which one could start to appreciate and further explore the complex role played by familiarity in shaping AON responses, not just in terms of how much particular regions might respond to particular stimuli, but also how the core regions of the AON might exchange information between each other. In the following section, prominent action understanding models are considered in light of the debate of the role of familiarity, as are these models', abilities to address the fundamental questions posed by this thesis.

1.2.4 Relating Questions About Action Familiarity to Action Understanding Models

This section outlines two prominent models of action understanding: the direct matching hypothesis and the predictive coding account. While these are not the only two models of action understanding (e.g., see Oztop et al., 2013; Pezzulo et al., 2015; Csibra, 2005 for alternatives), these two models are well-developed and extensively researched, and both provide strong predictions about how fluctuations in familiarity might shape the exchange of information between core nodes of the AON.

1.2.4.1 Direct Matching

Following the literature that suggests greater AON engagement for familiar compared to unfamiliar actions, one could frame this evidence in light of the "like me" hypothesis (Meltzoff, 2007). This hypothesis states that if an action or actor is more "like me", or has been or could be performed by me, we would expect greater engagement of sensorimotor mechanisms that support action observation and execution (i.e., the AON). Returning to the coffee cup example from above, the fact that we can relate to this action because we have likely performed it many times ourselves, we should be able to understand the action in terms of the actor's goals and intentions by using our own experience to map on to our friend's action. This use of experience-driven simulation mechanisms (Sinigaglia, 2013) in understanding an action forms the foundations of the direct matching hypothesis.

The direct matching hypothesis of action understanding (Rizzolatti et al., 2001; Gallese & Goldman, 1998; Wolpert et al., 2003; Umilta et al., 2001; Kohler et al., 2002) suggests that an action's meaning is understood via the AON, supporting simulation of an observed action by matching the goal or intention of what that movement would be if performed by the observer. By use of forward models (Wolpert et al., 2003), the visual information of the observed action first enters the network through the higher order visual areas (for example, STS). This information then flows anteriorly through the AON, eventually arriving at motor and premotor cortices. Engagement of the motor system enables an observer to understand how the action should unfold, and thus supports an observer in reading the intentions and goals of an observed actor. In our example, as we have experience with reaching for coffee cups to drink from, we are able to simulate our friend's action and thus understand the intentions of our friend.

One criticism of this hypothesis is based on the name itself; *direct* matching. As noted by Csibra (1993), multiple means can be taken in order to obtain a given goal. For example, I can take a drink from my mug by either raising it to my face or lowering my face to it; the goal to drink remains the same. Thus, the direct matching implied by this hypothesis must remain resilient to the action and faithful to the goal.

As such, simply observing an action and mapping it onto one's own motor system, or a higher order understanding, must be present. This point was addressed by the authors of the original model (Rizzolatti & Sinigaglia, 2010) in an extension that is termed the goal-mirroring model (Steinhorst & Funke, 2014). In this variation, the authors propose that intention can be gauged by matching the goal of the observed action to an internal motor act with the same goal, regardless of whether the kinematics precisely match.

When we observe a familiar action, the direct matching hypothesis has a clear explanation for how the AON responds. As familiarity of actions increases, the simulation of how an action might unfold over time becomes more accurate and resonance between an observer's motor system and an observed action is maximized (c.f., Calvo-Merino et al., 2005; 2006). However, such an account would struggle to explain why an *unfamiliar* action that is *not* in the observer's repertoire could elicit greater AON activity. Therefore, the direct matching hypothesis provides a partial explanation for the role of familiarity within the AON, but stops short of being able to explain why some unfamiliar actions drive the AON more strongly than more familiar actions.

1.2.4.2 Predictive Coding

A model that extends the premise of the direct matching hypothesis is the predictive coding account. Predictive coding models of AON function (Keysers & Perrett, 2004; Kilner et al., 2007a,b; Gazzola & Keysers, 2009; Schippers & Keysers, 2011) are predicated on the use of perceptuomotor maps to predict and interpret observed actions (Lamm et al., 2007; Schubotz, 2007; Urgesi et al., 2010) and may potentially help to resolve seemingly discrepant findings concerning the relationship between familiarity of an observed movement and engagement of sensorimotor cortices.

This framework proposes a Bayesian comparison of predicted and observed actions (via use of forward and inverse models), creating a reciprocally modulated network comprising premotor, inferior parietal and posterior temporal cortices. This

network aims to minimise the differences between observed and predicted actions. When observing a less familiar action, predictions (feedback) are lacking or are under informed, and thus do not match the observed actions (feedforward), which equates to high prediction error. This should manifest as increased AON engagement for highly unfamiliar actions, as the influence of feedforward/perceptual activity is heavily relied upon. When viewing an action that is familiar, however, the predictions generated by the network are more precise, and thus minimise prediction error. The minimising of prediction error can also manifest as increased AON engagement, as predictions projected posteriorly are stronger than when prediction error was higher. The reciprocal nature of this framework allows for the explanation of higher AON engagement for both familiar and unfamiliar actions

Both direct matching and predictive coding remain two prominent models of action understanding. These two accounts both rely on forward models and higher order representations of actions, with the predictive coding account's use of inverse models enabling comparisons between observed and predicted actions. When we apply the question of familiarity to these models, the direct matching hypothesis cannot fully explain the findings in the literature of greater activity for the observation of both familiar and unfamiliar actions. On the other hand, predictive coding has potential to provide an explanation for these discrepant findings yet requires additional examination and empirical support.

1.2.5 Conclusions and Aims of Thesis

From the action observation literature described above, while it is clear that familiarity shapes AON engagement, it remains less clear *how* variations in an observed or executed action's familiarity are reflected in AON engagement, and *why* contradictory findings have been reported regarding the relationship between AON engagement and action familiarity. Models of action understanding postulate a pattern of connectivity between core AON nodes that might help resolve why we see greater AON engagement for both familiar compared to unfamiliar and unfamiliar compared to familiar actions across the literature. Through combining effective connectivity measures and ROI-based regression models with objective and subjective measures of

action familiarity and an intensive training intervention, this thesis systematically evaluates the relationship between familiarity and AON engagement by testing hypotheses derived from two prominent action understanding models: direct matching and predictive coding.

2 Chapter II

This chapter aims to provide a detailed overview of the neural connectivity measures currently applied to functional MRI (fMRI) data. Specifically, this chapter first focuses on the foundations of fMRI, its development and how it can be used to measure metabolic changes in the brain. Next, the concept of neural connectivity is introduced, and its relevance to understanding of complex cognitive functions is discussed. This leads to an in-depth review of dynamic causal modelling (DCM), in which key facets are explained, as are evidence for and limitations of this method.

2.1 Functional Magnetic Resonance Imagining (fMRI)

Since the end of the 20th century, magnetic resonance imaging has enabled clinicians and researchers to look inside the human body in a non-invasive manner and without the concurrent administration of harmful radiation or radioactive isotopes. Over the past several decades, the increasing sophistication of imaging techniques and analytical approaches has enabled researchers to investigate the functional operation of the human brain with ever more precision. This powerful tool owes its origin to our understanding of quantum physics.

2.1.1 MRI Physics

The foundations of Magnetic Resonance (MR) imaging originate within the field of quantum physics, and the spin characteristics of protons (Gerlach & Stern, 1922; Pauli, 1927; Dirac, 1928). It was shown that atoms possessed an intrinsic quantum state, or spin, and magnetic moment. In other words, atoms can be thought of as spinning tops with different spinning speeds and orientations (an analogy that will be used throughout this chapter). Core to MR imaging is the stimulation of nuclei, where they absorb the energy emitted to them from a transmitter, and emit the energy back until they reach a level of initial equilibrium. Returning to the spinning

top example, this principle would be similar to having a spinning top tethered to a taut string, flipping the top, and then measuring how fast it returns to its original state. This concept is known as nuclear magnetic resonance (NMR; Rabi et al., 1938). In order to perform the "flip" of the nuclei, a Radio Frequency (RF) pulse is used. The energy emitted by RF is absorbed by the nuclei, similar to the physics used for X-ray or CT scans, but without the need for ionizing radiation.

If one considers a small mass of tissue, the magnetic moment, or spin frequency, of the hydrogen atoms (the most abundant substance, which makes up around 80% of our bodies) within this mass are oscillating at different frequencies and are orientated randomly. In order to acquire an image of this tissue, the tissue has to be placed within a strong magnetic field. Purcell (1948) showed that by placing an electrical current within a strong magnetic field, the atoms would align with the direction of the magnetic field. If one thinks of the spinning tops within a small mass of tissue, the orientation of the spinning tops is not uniform in the natural state, and this equates to a net magnetisation of zero (which ensures we do not attract other ferromagnetic metals). However, by placing this small mass of tissue within a strong magnetic field, some of the atoms will align their orientation parallel to the direction of the magnetic field (from negative to positive), whilst others will align in the opposite direction (anti-parallel). Generally, the atoms will align with the direction of the magnetic field, called the low energy state. The amount of atoms that align with the magnetic field is directly proportional to the strength of the magnetic field. In a pivotal study, Bloch (1946) showed how an image of a solid could be taken using the physics described above. Bloch used a transmitter coil to emit electromagnetic energy and then record the unabsorbed energy in a detector coil. Another way to consider this experiment would be to take a stencil and apply spray paint over it; the stencil itself would absorb some of the paint, and the remaining paint would create a picture on the canvas. The experiment by Bloch (1946) identified that the quantum mechanical properties of atoms can be used in conjunction with a strong magnetic field to locate a single object in space. Before exploring the formulation of modern day 3D images, the excitation and relaxation principles are further explained.

Returning to the spinning tops example, the mass containing the spinning tops (atoms) is placed within a strong magnetic field. The frequency of these spins, or precess, is known as the Lamor Frequency. This frequency is different for every element; therefore, the correct frequency has to be used in order to flip them (in a 1.5 Tesla scanner, this is ~63MHz for hydrogen). In order to perform the "flip" of these atoms, the RF pulse needs to resonate with the hydrogen atoms. When the RF pulse is emitted, this causes excitation where the attuned atoms "flip". Note that the net magnetisation of the mass is aligned with the strong magnetic field (Z direction), thus a flip would cause this net magnetization to rotate perpendicular to this into the X-Y plane.

The absorbed energy by the atoms from the RF pulse causes a shift from a low energy state (aligned with the strong magnetic field – Z plane) to a high-energy state (due to the flip, along the X-Y plane). The atoms endeavor to return to the low energy state, therefore dissipate the absorbed energy that is recorded, allowing for the formation of an image. This relaxation process can be separated into two independent, and simultaneous processes which can be recorded: T1 and T2 relaxation. T1 relaxation, also known as spin-lattice relaxation due to the energy released into the surrounding tissue, describes the shift from high to low energy states in the Z plane. As the protons emit a RF wave, they return to equilibrium (where they were before administration of the RF pulse). T1 is defined as the time taken for the net magnetization in the Z plane (aligned with the strong magnetic field) to reach 63% of its original magnetization. Simultaneously, yet independently, T2 relaxation occurs on the X-Y plane. T2, or spin-spin relaxation time, concerns the spin coherence of the protons. To further explain how T2 works, our example of spinning tops in a tissue mass is useful. Before the RF pulse, the spinning tops are spinning with no coherence to one another. After the RF pulse, they are flipped onto the X-Y plane and in addition, their spins become coherent. This phase-coherence is subject to dephasing and T2 can be defined as when 37% of the original in-phase value is reached. T2 occurs faster than T1, and is also known as spin-spin relaxation. The distinction between these two relaxation processes is important when we consider that different tissues have different relaxation times (Damdian, 1971; Lauterbur, 1986), allowing for a high level of contrast between structures within a mass. By setting short RF

pulse repetitions (repetition time; TR), and short times in which the re-emitted RF waves are recorded (echo time; TE), T1 allows for the differentiation of anatomical structures as tissues with a high fat content appear brighter than those filled with water. T2 images provide the opposite contrast (with the use of long TR & TEs), and can provide value to pathological imaging.

The ability of MRI to generate spatial images was pioneered by Lauterbur (1973; 1974). He theorized that by changing the strength of the magnetic field at different spatial locations, the protons would have different Lamor frequencies. Returning to our spinning tops example, Lauterbur suggested that by changing the torque of the spinning tops in different locations, they would therefore only be flipped by the resonating radio frequency. By building up a series of these images, a 2D image would be able to be constructed. Lauterbur (1973) used these spatial gradients to create the first 2D image of water tubes. However, an early drawback of this method was the lengthy acquisition time. Building upon these principles, Mansfield (Mansfield & Maudsley, 1976; Mansfield, 1977) proposed and tested a method whereby all the spatial information of an object can be gathered at once by systemically changing the magnetic field gradients at a rapid speed, known as Echo Planar Imaging (EPI), which dramatically reduced acquisition time. By using an algorithm known as the Fourier transformation (Kumar, Welti & Ernst, 1975), the acquisitions could then be reconstructed into 3D images. For their contributions to this field, Lauterbur & Mansfield received the Nobel Prize in 2003, and by dropping the negative connotations associated with the word "nuclear", Magnetic Resonance Imaging (MRI) was founded.

2.1.2 Functional MRI (fMRI) and the BOLD Signal

The fact that MRI can be used to image the inside of the human body without exposure to radiation means that it provides many advantages over other widely-used non-invasive imaging approaches, such as X-rays and CT scans. Its use in identifying metabolic changes within the brain (which is known as functional Magnetic Resonance Imaging; fMRI), enables us to infer region-specific functionality. This process relies on our understanding of the hemodynamics of the brain (Buxton et al.,

2004) and how these relate to metabolic changes (Ogawa et al., 1990a, 1990b). Note, the hemodynamic model is described in more detail later in this chapter (*3.1.5*); therefore the Balloon model and cerebral blood volume are omitted from the current section.

The human brain comprises millions of neurons, each connected to thousands of others via axons (Kandel, Schwartz & Jessell, 2000). Signals are passed between neurons via action potentials, which involve a rapid rise and fall of electric membrane potential. This transferring of signals was found to relate to an alteration in blood flow (Roy & Sherrington; 1890). The hemodynamics of the brain have been subsequently described in more depth, and more contemporary research suggests a formation from stimulus presentation to metabolic changes (Buxton et al., 2004). This model suggests that upon stimulus presentation, a neural response is present in terms of action potentials between neurons. This activation requires the use of energy, subsequently replenished by oxygen from hemoglobin. The logical next step to be taken is to use MRI to identify areas of the brain involved in a task by examining oxygenated hemoglobin as it moves to replenish brain regions that were previously engaged. As noted by Pauling and Coryell (1936), oxygenated hemoglobin is diamagnetic and therefore has no magnetic moment; in other words, it does not absorb any RF pulse. On the other hand, deoxygenated hemoglobin is paramagnetic and has a magnetic moment, allowing for excitation via an RF pulse.

Ogawa and colleagues (Ogawa et al., 1990a, 1990b) were one of the first groups to investigate whether the principles of MRI could be applied to identifying changes in hemoglobin states. They manipulated the amount of oxygen breathed by a rodent and removed hemoglobin samples. Using a similar principled approach to Purcell (1948), they placed the samples within a strong magnetic field to test whether they caused any local field distortions, which would only be present if the substance aligned with the strong magnetic field. They found that deoxygenated hemoglobin did cause these distortions, thus revealing the possibility for an indirect measure of metabolic changes within the brain, which was named the Blood-Oxygenated-Level dependent (BOLD) response. Since this discovery, Belliveau and colleagues (Belliveau et al., 1991) were among the first to successfully apply this measure using

fMRI, and paved the way for the thousands of fMRI studies that have now been published to date.

2.1.3 Preprocessing

Once data are obtained from fMRI scanning, a number of pre-processing steps must be applied in order to determine how changes in neural activity are associated with a given cognitive, behavioural or motor task. The main pre-processing steps include slice time correction, realignment, registration and normalisation. In this thesis, all of these steps were performed using the SPM 12 toolbox (Welcome Department of Imaging Neuroscience, London, UK) running under MATLAB (2014 and 2015, The MathWorks, Natick, MA).

When acquiring an fMRI image, multiple slices are taken (c.f. Mansfield, 1997; Mansfield & Maudsley, 1976). These slices are taken across the duration of the TR, so that they cover the entire desired area before the next excitation pulse (RF). These slices of the brain can be altered in terms of thickness (in mm), number taken (dependent on thickness) and acquisition order (ascending, descending or interleaved). The sum of these slices is a volume, or 3D image of the brain, within the given TR (for example, if an experiment were to last 10 minutes and have a TR = 2000ms, then number of volumes collected would be 600/2 = 300). As different slices of a volume are measured at different times, collection of the actual BOLD response is staggered across time and space. Slice time correction is a common method that accounts for the order in which the slices are taken. By temporally interpolating the slices onto a reference slice (ordinarily the first slice taken), this discrepancy can be addressed (Schanze, 1995). When performing an fMRI experiment, participants are typically in the scanner for between one to two hours. Remaining still for this length of time can be difficult and head movement artifacts can occur ranging from muscle relaxation (or tension) to artifacts that arise as a result of other actions participants are required to do during the experimental task (e.g. button pressing). These movements can occur between slices or volumes, therefore, a realignment algorithm (Friston et al., 2005) using a 6 parameter (rigid body) spatial transformation and a least squares approach can be applied. This algorithm can account for gradual movement (e.g.

muscle relaxation). However, if the participant has coughed, for example, then the volumes should be removed. Additional steps taken in pre-processing are co-registering the individual subject's data to the structural scans and normalising the functional scans to a standardised template (MNI; Montreal Neurological Institute).

2.1.4 Strengths and weaknesses

The primary strength of fMRI is its relatively high spatial resolution. This resolution, in the order of mm, is due to the focus on the magnetic resonance of protons. Compared to a method such as electroencephalography (EEG), which relies on residual currents where the source is relatively unknown (however, see Phillips et al., 2002), fMRI is capable of identifying the *area* of activity. Further strengths of fMRI include the fact that it is non-invasive and is highly available to the scientific community (Glover, 2011).

One critical weakness of fMRI centers on its use of the BOLD signal (Logothetis & Wandell, 2004). Intersubject and inter-scanning session variability of the hemodynamic response and subsequent BOLD signal have been shown (McGonigle et al, 2000; Aguirre et al., 1998), implying that it is not a perfect model of brain function. Moreover, questions remain concerning whether BOLD activation is caused by spiking neurons (Heeger et al., 2000) or whether this activation is due to a mixture of synaptic activity and intracortical processing (Attwell & Iadecola, 2002). Additional limitations of fMRI include its low temporal resolution (this is a result of the BOLD signal peaking 6 seconds after neuronal activation; Buxton et al., 2004), as well as a lack of flexibility in tasks being conducted compared to methods such as EEG (Glover et al., 2011), due to the restrictions of participants lying supine and immobile in the scanner.

2.1.5 Conclusions

Functional Magnetic Resonance Imaging is a neuroimaging technique that has grown considerably in popularity and use over the past two and a half decades, and

relies on principled quantum physics and a hemodynamic model to derive task-driven metabolic changes in the brain. Although the underlying causes are still debated, the BOLD signal allows us to identify areas of the brain associated with performing particular tasks, and provides a useful tool for researchers to further develop our understanding of the human brain.

2.2 Introduction to Neural Connectivity

In the fields of cognitive neuroscience and neuropsychology, an overarching aim is to understand how the brain generates an understanding of the external world. Using non-invasive neuroimaging techniques such as fMRI, MEG and EEG, the traditional approach has been to identify brain regions that serve specific functions (an approach also known as functional segregation). Although functional segregation offers insights into the function-to-region relationship, it does not provide an understanding of nor enable investigation into the connections within systems. It is important to also examine these connections as they allow investigators to consider the brain as a connected, dynamic system, rather than a collection of separate, albeit functionally specific, nodes. For example, the identification of the primary motor area (M1) and supplementary motor area (SMA) as being crucial for action execution does not advance our understanding of how information might be exchanged between these two regions to subserve action execution. Interest in understanding the dynamic exchange of information between brain regions can be characterised as a shift from functional segregation to integration. The core aim of functional integration approaches is to investigate the interactions or relationships between anatomically distinct regions. The reason why this is possible is that functional integration methods facilitate exploration of statistical implications of activity (i.e. the BOLD response or Event Related Potentials), which consequently enable investigation into connections between regions and how these connections change with experimental manipulation (Friston, 2009a). For example, in the motor system example above, functional integration would be concerned with the presence of a connection between M1 and SMA and whether a particular experimental manipulation (e.g. kinematics of an action) might modulate or change the nature of this connection. It is important to note that functional segregation can be considered obsolete without integration and vice

versa (Friston, 2011). This is because identifying functionally specialised regions offers little to advance our understanding of a functionally-specific system; however, investigation into the modulatory factors that influence a connection can only be investigated after the identification of functionally-specialised regions.

In this section of this chapter, the paths of functional integration are discussed first. The core facets of the effective connectivity method used in this thesis (Dynamic Causal Modelling; Friston et al., 2003) are then described and the validity of this method debated.

2.3 Effective connectivity

There are two paths for investigating functional integration: functional connectivity and effective connectivity (see Frackowiak et al. (2003) for a comprehensive comparison of these). Functional connectivity is largely concerned with observable measures (e.g. BOLD signal in fMRI), which can have statistical dependencies and correlations (e.g., Granger Causality Modelling: Goebel et al., 2003; Valdes-Sosa, 2004). An example for fMRI would be to first observe activity (BOLD signal) in V1 and V2 when a participant views a rotating wheel. From the resulting BOLD signal, a time series of activity for each region when viewing the rotating wheel can be extracted. Functional connectivity can then address the question of whether the time series of activity for V1 correlates with the time series of activity for V2 under this manipulation. An important point to consider is that correlations can occur between many regions and for different reasons. This can make it difficult to be certain that a particular correlation has arisen from the manipulation of interest. In other words, it is not always clear whether a correlation between two regions is a result of the manipulation used or an alternative factor (Friston et al., 2003; 2011).

A prominent methodology that goes beyond the statistical dependencies between brain regions (i.e. correlations) is effective connectivity. This approach evaluates the directed (as opposed to correlated) influence one region exerts on another region at the neuronal population level (Friston et al., 2003). The key difference between functional and effective connectivity is that the latter aims to

create models of neural activity and quantifies this connectivity between regions in terms of the parameters of the connectivity model (Friston et al., 2009a). Moreover, effective connectivity generates feasible models of neural activity and then converts these into a time series, which can be compared to the actual time series, whilst also giving values to the connection strengths between brain regions of interest.

Effective connectivity is thought of as both dynamic and causal (Friston et al., 2009a). It is activity dependent, therefore dynamic, and also relies on a model of interactions or connections between brain regions (coupling), which allows the establishment of causality. For example, the activity in V2 relies on input from V1 (in the sense that it has "caused the activity"), yet this activity is subject to the fluctuations in activity of V1 – the relationship is dynamic. As models are generated, effective connectivity uses model comparison to test competing hypotheses (e.g. the presence or absence of a directed connection). Functional connectivity does not allow for this, as there is no generation of theoretical models of how data in one brain region 'causes' activity in another region. As such, functional connectivity is not stringent enough to compare two hypotheses in this manner (correlation versus no correlation). However, this is not to say functional connectivity approaches are not without their advantages, and such approaches are often adept at answering certain experimental questions of interest (Friston et al., 2011). For example, the hypothesis that two regions should be coupled by an experimental manipulation would be best suited to a functional connectivity approach. An effective connectivity method would not be best to address this question as it addresses how one region exerts influences on another.

As effective connectivity in fMRI relies on generating plausible models of brain function *a priori,* which are then used to evaluate the time series of the BOLD signal, this means that it is both hypothesis- and data-driven. This is an important distinction between functional and effective connectivity as we still do not fully understand the BOLD signal in terms of what is translated from the neuronal level (Attwell & Iadecola, 2002; Lauritzen, 2004; Logothetis et al., 2001; Logothetis & Wandell, 2004). To solely rely on the data-led approach, as functional connectivity does, can be risky, as a number of factors remain unknown concerning the link between BOLD signal and the neural dynamics that underpin its foundations. This is

also a shortcoming of some effective connectivity models that rely predominantly on the time series from the BOLD signal for the estimation of their coupling parameters (Structural Equation Modelling (SEM): McIntosh et al., 1994; Buchel & Friston, 1997; Horwitz et al., 1999; Autoregressive Models: Harrison et al., 2003; Roebroeck et al., 2005: psychophysiological interactions: Friston et al., 1997). This is because models of effective connectivity that ignore the transformation from neural activity to BOLD measurements can be difficult to interpret in terms of their validity due to the underlying physical premises (Gitelman et al., 2003). One way to overcome this issue is by combining models of neurodynamics with biophysical models to describe the activity measured. The result of this is a method that creates a model of an expected outcome for the signal and can then be compared to the actual signal – allowing for comparison of models to ascertain the most plausible model for the data set. One method of effective connectivity that takes this approach is Dynamic Causal Modelling (DCM) (Friston et al., 2003).

First proposed by Friston and colleagues (2003), Dynamic Causal Modelling is an input-state-output model for evaluating effective connectivity, which models how one brain region affects activity in another brain region. The brain is therefore treated as a deterministic, non-linear dynamic system that is subject to inputs and produces outputs. For example, if we are interested in whether V1 exerts influence on V2 (or vice versa) in a given task, we start by modelling the neuronal level predictions (or hidden activity as this is not directly observed from fMRI) that the activity in V1 exerts activity in V2, or vice versa. These neuronal level predictions can then be passed through the state change (a biophysically plausible model). The result of this gives us an output of the system, the predicted BOLD signal, which can be compared to the observed signal.

In relation to the other connectivity methods mentioned, DCM was developed to overcome a major shortcoming of its predecessors, e.g. biophysical and neurodynamic plausibility. Since its introduction, it has been refined and extended repeatedly (Friston et al., 2007, Kiebel et al., 2007, Marreiros et al., 2008, Stephan et al., 2008, 2007a) and implemented in EEG, MEG and local field potentials (LFPs) obtained from invasive recordings of both humans and animals; both in the time

domain (Daunizeau et al., 2009a; David et al., 2006; Kiebel et al., 2006) and frequency domain (Chen et al., 2008; Moran et al., 2007, 2008, 2009; Penny et al., 2009). fMRI will be the focus of this chapter, yet it is important to note that DCM can be implemented in multiple modalities. This modelling technique is considered causal as external inputs (experimental manipulations) cause changes in neuronal activity (the neuronal level predictions fed through a biophysically plausible model) that in turn cause changes in the predicted BOLD response

2.3.1 Dynamic Causal Modelling

The equations and illustrations presented in sections *2.3.1.1*, *2.3.1.2* and *2.3.1.6* derive from those reported by Friston et al. (2003). The equations in section *2.3.1.2.1* derive from Stephan et al. (2008) and the equations and illustrations in *2.3.1.7* follow those presented by Friston et al. (2003) and Penny et al. (2004).

2.3.1.1 Neural state equation

A key element of DCM that differentiates it from other forms of connectivity analysis is that it employs a plausible neurodynamic model. By passing predictions through a biophysically plausible model, a predicted BOLD response can be attained. This outcome can then be fitted to the observed time series (in Bayesian terms), and posterior calculations of connectivity can be obtained (Friston et al., 2003). This process will be described in due course, but it is important to begin with a model of how we believe the brain system we are examining will work.

In this chapter, a four-node system will be considered. The mathematics described work for any number of nodes. For the sake of clarity and illustration, here I focus on a four-node system. A four-node system could also accommodate a fully connected system where each node is interconnected with every other node; however, the example illustrated in *Figure 2.1* was intentionally kept simple for clarity.

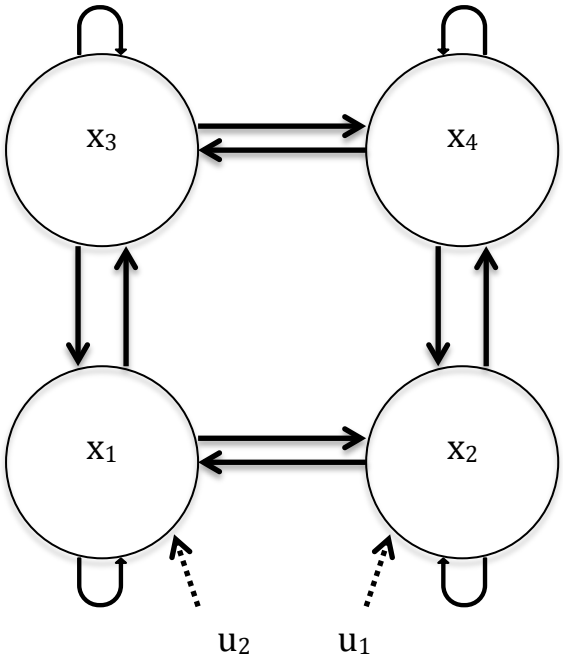


Figure 2.1. An example of a four-node system with hypothesised connectivity. x_i refers to the individual brain regions and the u-shaped arrows illustrate each region's influence on itself; u_i represents the experimental (sensory) input into the system.

In *Figure 2.1*, we assume that there is reciprocal intrinsic connectivity between regions (shown by straight, solid arrows). This system is theoretical, but it could represent a system where we are looking at inter-hemispheric connectivity, whereby x_1 and x_3 are part of one hemisphere and x_2 and x_4 are part of the other. The dotted arrows represent the experimentally manipulated input. This input can, and may, enter the system at any point or multiple points and it is key for the success of the model. If the input is incorrect then the system is not perturbed. In this situation, DCM will simply model a flat line BOLD response (Stephan et al., 2010). The curled arrows represent the self-connectivity. This is the influence the node has on itself (e.g. excitatory or inhibitory; discussed later in the one and two state sections).

With this theoretical and visual representation of a hypothetical system, DCM allows transformation of this into a neural state equation. The differential equation, which describes this first step, is:

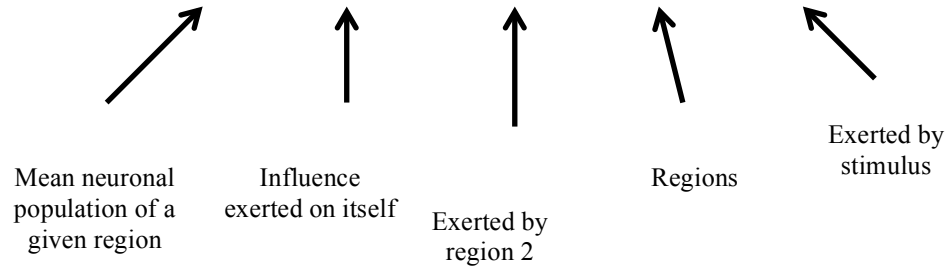
$$\frac{dy}{dx} = \mathrm{F}(\mathrm{x}, \mathrm{u}, \theta) \qquad (1)$$

where *F* is a non-linear function which describes the neurophysiological influences on a given region, (*x*), *u* is the experimental (or sensory) input, θ are the parameters which will be used later to make inference on. An example of how the mean activity in a given region can then be summarised as follows:

$$\dot{x}_1 = a_{11}\, x_1 + a_{12}\, x_2 + a_{13}\, x_3 + c_{12}\, u_2 \quad (2)$$
$$\dot{x}_2 = a_{21}\, x_1 + a_{22}\, x_2 + a_{24}\, x_4 + c_{21}\, u_1$$
$$\dot{x}_3 = a_{31}\, x_1 + a_{33}\, x_3 + a_{34}\, x_4$$
$$\dot{x}_4 = a_{42}\, x_2 + a_{43}\, x_3 + a_{44}\, x_4$$

where $\dot{x}$ is the derivative in time, equalling the mean population a region (acquired by extracting the region of interest from the GLM), *a* simply denotes that it belongs to the *a* matrix, or the intrinsic connectivity matrix. This non-linear expression shows all the regions and connections (including inputs) shown in *Figure 2.1*. To better understand what each element of this expression represents, we shall consider the linear expression of this and therefore just one region.

$$\dot{x}_1 = a_{11}\, x_1 + a_{12}\, x_2 + a_{13}\, x_3 + c_{12}\, u_2 \quad (3)$$



Equation 3 shows the connectivity within and to region x_1. Except from the time derivative, the other *x* values correspond to the other regions within the system. In the first part of the equation, the influence entered upon it corresponds to the curved arrow from *Figure 2.1*. The reason that the subscript numbers do not start at one is that they help us to understand what connection we are concerned with. The first number here denotes the connection to which region, and the second from what region. For example, a_{34} will correspond to the connection from x_3 from x_4. If we return to *Figure 2.1*, we can deduce that we believe connections to span from x_2 and x_3

(a_{12} and a_{13}) as well as that exerted by the stimulus (and therefore the *C* matrix). What is apparent is that the equation does not explain the influence this region exerts on others; however, if we expand this equation to include all regions then this is satisfied.

(4)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & 0 & a_{24} \\ a_{31} & 0 & a_{33} & a_{34} \\ 0 & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 & c_{12} \\ c_{21} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

State changes

Effective/endogenous activity

System state

Input parameters (how strongly the stimuli will effect)

External inputs (known as defined in GLM)

This expansion (*Equation 4*), can be reduced down to the form:

$$\dot{x} = A\dot{x} + Cu \tag{5}$$

Expressed in non-notational format, *Equation 5* states that the mean neuronal population activity in any given region is the result of all intrinsic connections to this region plus the influence of the experimental (sensory) input. This expression is the neural state equation for DCM. It describes the connectivity between the hypothesised regions, the effect the regions exert upon themselves and then influence exerted by the stimulus. The next step in the DCM process is then to identify whether this simplified equation should be modified into a bilinear or non-linear term.

2.3.1.2 Bilinear state equation

The bilinear state equation is an expansion of the linear state equation (or neural state equation), which investigates the coupling strength changes when the task is changed (Stephan et al., 2008). A visualisation of how this would look with our working example can be seen in *Figure 2.2*.

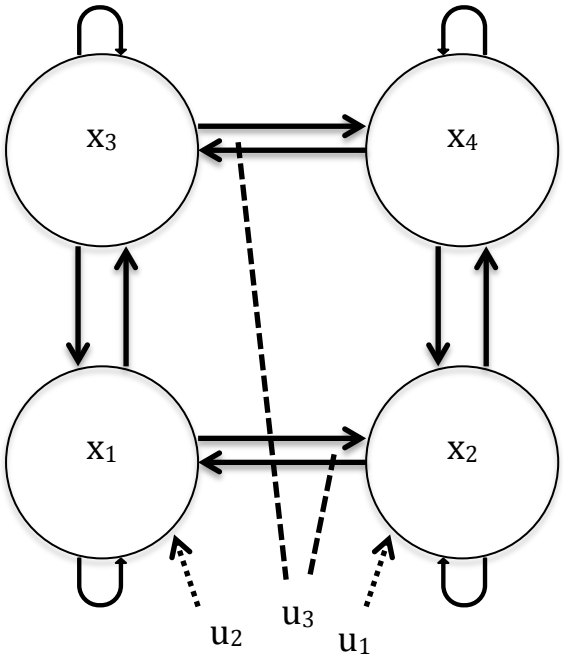


Figure 2.2. An expansion of *Figure 2.1* to include modulations by u_3 (context).

Figure 2.2 is identical to *Figure 2.1* but includes u_3, which is usually the context or experimental manipulation (task). What this additional parameter means is that the system is now allowing for the modulation in the intrinsic strength to be investigated when the context changes (or endogenous connectivity). Without this modulation, we cannot know how the coupling strengths change when the task changes, and instead we only learn how the system flows when only the input is considered. Most researchers using DCM are interested in how a system changes when the task/context changes. Use of the bilinear state equation allows for this. An example of how this could work is as follows: the connection between x_2 and x_1 may be 5 (an arbitrary number for this example), but when the task is changed to, for example, attention from non-attention, this coupling strength increases to 9. This example would suggest that attention modulates the connection between x_2 and x_1.

In order to make the original neural state equation bilinear, an additional matrix is added to take into account these modulations. The expansion involves including the matrix *b(j)* and gives:

$$\begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \\ \dot{X}_3 \\ \dot{X}_4 \end{bmatrix} = \left\{ \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & 0 & a_{24} \\ a_{31} & 0 & a_{33} & a_{34} \\ 0 & a_{42} & a_{43} & a_{44} \end{bmatrix} + u_3 \begin{bmatrix} 0 & b^{(3)}{}_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b^{(3)}{}_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 & c_{12} & 0 \\ c_{21} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

(6)

This bilinear state equation includes the *a* and *c* matrices discussed in the previous section as well as the modulation matrix *b*, located in the center of Equation 6. It is bilinear because it parameterizes the interactions between inputs and states. The external input u_3 is the context we wish to examine so is placed before the *b* matrix. If we wished to investigate multiple factors (e.g. attention and hemisphere) then we would include multiple *b* matrices. The *a* and *b* matrices are then combined to make a Jacobian Matrix of total connectivity.

If u_3 equals zero, then it is the same as the neural state equations and therefore no experimental modulation. If, however, u_3 equals 1 then the parameters from the *b* matrix are added to the strength of endogenous activity where this additive strength is a function of a controlled variable, or in our case, the context or task. We can collapse *Equation 6* and get the bilinear state equation:

$$\dot{x} = \left(A + \sum_{i=1}^{m} u_i B^{(i)} \right) x + Cu$$

(7)

In conclusion, DCM allows for the inclusion of modulatory variables that are used to show coupling strength changes between two interacting brain regions. This aspect of DCM is very important as the inversion and inference about the model relies on this stage of the process being precise and accurate.

2.3.1.2.1 Non-linear state equation

By modelling the modulations of the experimental manipulations via matrix b, we are able to investigate the effect that task or context has on the endogenous connections between brain regions. However, there are some limitations to this form of state equation. The main issues with this approach, as outlined by Stephan et al. (2008), are that it may not be the most appropriate, or adept, framework for modelling fast changes in effective connectivity, and that the neuronal origin of these modulatory effects are not specified.

Returning to our example illustrated in *Figure 2.2*, the connection between x_3 and x_1 may well be modulated by the context u_3. What we are unable to obtain from the bilinear approach is where this modulation originated. Of course, this is not necessarily a disadvantage, for there are many studies where the origin is unknown. However, knowing the origin makes the model better specified and more accurate. By using a non-linear approach, whereby the origin of the modulation is modeled, one can then investigate the fast changes in connectivity at the single neuron level. The reason this is attainable is because it eliminates the abstractness of suggesting there is a modulation and instead models that neurons in region x are modulating the connection between x_3 and x_1 (as in our example).

As highlighted in the original DCM paper, processes such as neuronal gain control, referring to the on/off switching of a connection and synaptic plasticity, can only be modeled if we were "…to go beyond bilinear approximations to allow for interactions among the states. This is important when trying to model modulatory or nonlinear connections such as those mediated by backward afferents that terminate predominantly in the supragranular layers and possibly on NMDA receptors" (Friston et al., 2003; pg. 1299). Such transparency by the original authors of DCM demonstrates some of the potential limitations of this method.

In order to obtain the non-linear equation, an additional matrix, d, is added to the bilinear form of the equation:

$$\dot{x} = \left(A + \sum_{i=1}^{m} u_i B^{(i)} + \sum_{j=1}^{n} x_j D^{(j)} \right) x + Cu \tag{8}$$

where matrix *d* models the input from a node within the system. To summarise, the use of a non-linear state equation allows for the modulation of connections by an additional node. Although eliminating abstractness of the origin of a modulation, strong hypothesis should exist of this origin for the use of this approach.

2.3.1.3 One state vs. two state

Within the *A* matrix, the diagonal cells represent the self connection of the nodes, for example, how region *1* has an effect on itself when stimulated by *u*. Marrieos and colleagues (2008) advocated an extension to the original DCM formulation whereby these self connections represent two states rather than the one state from the original formulation. In biological terms, this makes the models more plausible as it allows for both excitatory and inhibitory influences on the region. The use of two state DCM is more accurate (within reason), but time consuming. Moreover, where the nodes are not close in proximity, inhibitory coupling is not supported (Marrieos et al., 2008). This suggests that the use of two state DCM is preferable for some systems but not necessary. An additional benefit of using two state DCM over one state is that its use makes the overall system more stable when the predicted neuronal dynamics are converted into a predicted time series due to its increased biological plausibility.

2.3.1.4 Stochastic

The DCMs described thus far can be classified as deterministic as they are input-state-output models of effective connectivity – as characterised by Friston et al. (2003). An exciting new extension to DCM analyses allows for the modelling of stochastic effects (Li et al., 2011; Daunizeau et al., 2012). This development has taken place over many years (Riera et al., 2004; Penny et al., 2005; Daunizeau et al., 2009b)

with the end result being a system that allows for random fluctuations in the unobserved and physiological states – state noise. The reason such a model can exist is that the neural activity or hemodynamics of deoxyhemoglobin being modeled by both stochastic and deterministic DCM is allowed to be explained by the dynamics of the endogenous activity. To return to our example, we can say that x_2's activity is subject to x_1's; beyond whether or not the experimental manipulation (exogenous coupling) is characterised or not. This is highly applicable to studies modelling resting state data (Friston et al., 2014a; 2014b).

In order to obtain a stochastic neural state equation for DCM, the bilinear form of the equation is used with the addition of the parameters hidden causes; hemodynamic states and observational noise. This method adds a higher level of accuracy than the deterministic approach as it allows for the system to explain itself, rather than the user determining what the model will tell it, per se. One shortcoming of this approach is its time-consuming nature; deterministic models with four nodes would take an order of minutes whereas the same model where stochastic effects allowed would account for more variance, yet take an order of hours for each model to be estimated.

2.3.1.5 Hemodynamic model

Once the neural state equation is specified, it needs to be transformed into a BOLD signal (or estimated time series). The neural state equation simply specifies the priors of the model, in other words, the location and the strength of the connectivity. In order to incorporate this with fMRI data, the parameters obtained from this model need to be passed through a model that would then produce an expected BOLD signal. One model that does this is the hemodynamic model (Friston et al., 2002).

The hemodynamic model used in DCM is based on the Balloon Model. The Balloon Model (Buxton & Frank, 1997; Buxton et al., 1998) was developed to provide an input-state-output system model with the state variables volume (v) and deoxyhemoglobin content (q) which can explain the output of the system, the BOLD signal (y), based on the input of blood flow (f_{in}), which are all subject to time (t). The

model works by assuming that the increase in blood flow inflates a venous "Balloon" so that deoxygenated blood is diluted and then expelled at a faster rate. This clearance of deoxygenated blood then reduces the intravoxel dephasing and produces a greater signal (the increase in BOLD signal that fMRI analysis shows). Once the "balloon" is nearly filled to capacity, the flow of deoxygenated blood into the region increases into the venous compartment and this is expressed as an early dip in signal (this is what characterises the difference in explanation of the Balloon Model to the previous accounts (e.g. Ogawa et al., 1990b; Menon et al., 1995)). Once the balloon capacity has been filled, it then relaxes and the clearance and dilution is reduced, causing the undershoot observed post stimulus.

This model was extended by Friston et al. (2003) to include additional parameters to make it more biologically plausible for DCM: making the parameters of the Balloon Model signal decay, rate of flow, hemodynamic transit time, Grubb's exponent (see *Appendix 1; Glossary*) and the resting oxygen extraction fraction. The derivative of the neural state equation then serves as the input (the neuronal activity) and when fed through the model for each region, this results in a time series. Support for the use of the Balloon Model in DCM has been shown by Stephan and colleagues (2007b).

2.3.1.6 Bayesian model inversion

To summarise thus far, the fundamental underpinnings of DCM have been explained. The neural dynamics (and their subsequent equations) have been described as well as the method whereby the expected area specific BOLD signals are derived from a transformation of the neural dynamics by the hemodynamic forward model. In this next section, the principles of the Bayesian estimation scheme are described, leading to how this informs us how modeled BOLD signals are maximally similar to the experimentally measured BOLD signals.

Once the state equation is integrated and passed through the hemodynamic model, we can derive an estimated time series. This parameterised BOLD response is then inverted in a Bayesian fashion (Bishop, 2006; Gelman et al., 1995) to obtain two

estimates of quantities. The first quantity is the posterior distribution over all model parameters, which is used to make inferences about model parameters θ. The second quantity is the probability of the data given the model, in other words, a probability statistic describing how well the expected and observed fit. This is known as the model evidence and will be discussed in more detail later in this section.

The first quantity is the posterior distribution over all model parameters. This is an important aspect of DCM methodology as this approach creates parameter estimations for the connections within our model to identify the strength /existence of connections. The estimation scheme employed by DCM is performed within a Bayesian framework. In general Bayesian terms, the aim is to identify a quantity for an event, given prior constraints, and applying these to the observed in order to obtain the observed quantity. Formally, these terms are referred to as *priors* (or constraints), *likelihood* (or an observation model) and *posterior* (the resulting quantity). A schematic of the relationship of the variables can be seen in *Figure 2.3*.

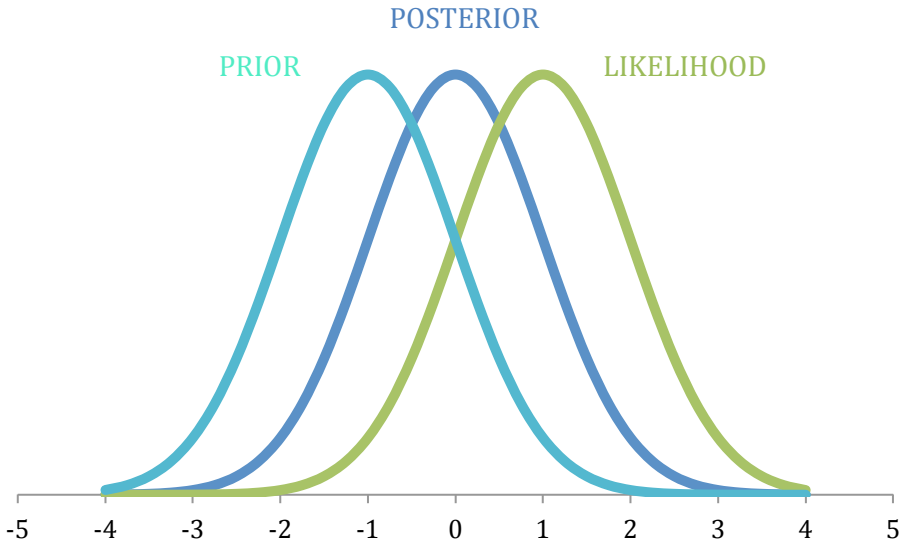


Figure 2.3. A schematic of the relationship between prior, likelihood and posterior quantities in a Bayesian approach.

For any model being estimated, the first goal is to estimate the parameters (or connection occurrence/strengths). The first part of this scheme is already conducted as

the model has been specified, both at the neuronal level and hemodynamic level which results in:

$$y = \lambda(x) \tag{9}$$

We then transform this modeled BOLD response in to an observation model by adding measurement error *e* and confounds *X* (e.g. drift). The reason we add these is comparable to that of general linear modelling, whereby the fitting of a function will not fit exactly to the data unless we account for variables that cannot be controlled. The resulting equation gives an observational model of:

$$y = h(u, \theta) + x\beta + e \tag{10}$$

In *Figure 2.3*, the likelihood term, $p(y|\theta)$, refers to the observational model whereby error and drift are added to our models of neural interactions and region specific hemodynamic. We also have a set of priors, $p(\theta)$, which include constraints of connections and hemodynamic parameters. The aim of the inversion scheme is then to make these points maximally similar to the point of convergence, resulting in a posterior parameter estimate for every parameter:

$$p(\theta|y) \propto p(y|\theta)p(\theta) \tag{11}$$

This equation states that the probability of the parameter given the data is proportional to the likelihood multiplied by the priors. This resulting posterior parameter distribution via the following Bayesian estimation framework is Gaussian (normally distributed), given by mean $\eta_{\theta|y}$ and covariance $C_{\theta|y}$.

Bayesian estimation in DCM is carried out using a Variational Bayes (Laplace) version of the Expectation-Maximisation algorithm (Dempster et al., 1977; Friston et al., 2002, Friston et al., 2003). In short, this algorithm consists of an *E* step (expectation) and an *M* step (maximisation). The goal of this algorithm is therefore to use an iterative scheme so that there is minimal discrepancy between the posterior (mean and covariance) and the likelihood quantities. Briefly, the *E* and *M* steps work

iteratively to approximate and then maximise the log-likelihood. A recent extension to this scheme in DCM is to use the free energy principle (Friston et al., 2008; 2009b). The free energy is the Kullback-Leibler divergence between the approximate and observed minus the log-evidence. The aim of this algorithm is then to minimise free-energy (Friston et al., 2007; Penny et al., 2004), giving what is known as log evidence.

2.3.1.7 Bayesian Model Selection (BMS)

In order to address competing hypotheses of the mechanistic nature of a system, the models created (and inverted by passing through a the hemodynamic model) have to be compared to each other in order to identify which is "best". This integral step to DCM has to be applied to both analyses concerned with inference on model parameters and structure alike. Once the optimal model has been selected, inferences on its parameters can be made (shown by Acs & Greenlee, 2008; Leff et al., 2008; Stephan et al., 2007b; Summerfield & Koechlin, 2008). The premise of Bayesian Model Selection is to compute an approximation to the model evidence, $p(y|m)$. Model evidence can be phrased as the probability of the data, y, given the model, m.

Dubbed the "Holy Grail" of model comparison (Stephan et al., 2010), model evidence has to satisfy two criteria: accuracy and complexity. These two properties define a good model but they also provide a trade-off. Consider the following example: one model is very complex (e.g. modulations on all connections by all experimental manipulations) and a second model which has the same intrinsic connectivity but one modulator by one manipulation on one connection between regions *A* and *B*. If we were to compare these two models solely on relative fit, we may infer that the more complex model best explains the data. However, a more complex model may not have high model evidence but is consistent across data sets. For example, a full model tested across different data sets may give model evidence relatively low, but it will be the same for each. On the other hand, the simple model may well be too simple. For example, a simple model might explain the data for one data set, but not be generalisable (see Gharamani (2004) for a schematic of this

relationship). In order for BMS to identify the best model out of those compared, a trade-off between accuracy and complexity must then occur, accounting for factors such as the number of free parameters (Pitt & Myung, 2002). It is important to note that the relationship between accuracy and complexity is monotonical, therefore the model will start to fit noise (otherwise known as "overfitting") which impacts generalisability of the subsequent conclusions.

In the first paragraph of this section, "best" was used to describe the model which ascribes the data, however, this optimal model assumption has to shift to address which model best represents the balance between accuracy and complexity. In BMS, the winning model is therefore the model that maximises the model evidence, mathematically noted as:

$$p(y \mid m) = \int p(y|\theta, m)p(\theta|m)\, d\theta \tag{12}$$

This integral cannot be solved analytically, therefore an approximation of model evidence has to be used. Individual models are fitted to the data and then this approximation is calculated using variational free-energy bound on the log evidence (Beal & Ghahramani, 2003; Friston et al., 2007). To articulate its aim, it provides an approximation to model evidence (in the form of logs) whereby the data are explained as simply as possible by using a minimal number of parameters that deviate from their priors to a minimal degree (Rosa et al., 2012). The use of free-energy in DCM was preceded by AIC and BIC yet free-energy has been shown to be a more accurate method in this context (Penny et al., 2004)

The resulting model evidence (or log model evidence) can be viewed as a measure of generalisability across data sets, much like cross validation (Pitt & Myung, 2002; MacKay, 1992). Moreover, we can comprehend (log) model evidence as the likelihood of the data, whilst taking into account the variability of its model parameters (Stephan et al., 2010). However, these probabilities do not necessarily provide us with a colloquial understanding on which model best describes the data. For example, if we take our two models, m_i and m_j, with posterior probabilities of 0.8

and 0.2 respectively. At first glance m_i is the optimal model here, however, the relationship between these two probabilities has to be quantified. In Bayesian statistics, the relative goodness of these two models is transformed into a Bayes Factor:

$$BF_{ij} = \frac{p(y|m_{i)}}{p(y|m_j} \approx \exp\left(F_i - F_j\right) \tag{13}$$

where BF is the Bayes Factor. This transformation allows one to compare the ratio of respective evidences (Kass & Raftery, 1995). These ratios can then be applied to guidelines of strength of difference suggested by Raftery (1995), which can be likened to that of *p* values in frequentists statistics (see also Stephan et al., 2008 for a comparison of alternative methods).

In summary, this essential component of DCM allows the comparison of two competing models. The method described above explains how models are compared, generalised to comparison of all models within the model space. Alternatively, one can group models into families (Penny et al., 2004). This approach allows for grouping of models that share attributes and then compare at a family level. For example, if one had a subset of model all with a modulated connection between two regions, and an equal number of models that do not have this attribute, one could use BMS and this approach to compare the relative evidences of these groups. This method is useful when there is no clear winner. Notably, BMS is a subjective method. Model evidence is defined for one particular data set and model space. For example, if one wishes to compare 10 models, then decides to add another 5 models to test; model space has changed and therefore the model evidence has changed. Moreover, adding an additional node changes the data and therefore model comparison. Similarly, model goodness is relative to the models being compared.

2.3.1.8 Fixed Effects vs. Random Effects

When conducting BMS, one must first identify the underlying cognitive mechanism that he or she is trying to evaluate. If we were to test whether an anatomical connection exists between two regions or a basic physiological mechanism (Stephan et al., 2009; Chen et al., 2009), then the most appropriate analysis to use is Fixed Effects (FFX) analysis. FFX analysis adopts the assumption that all optimal or winning models across individuals reflect that of the population (Stephan et al., 2010). For example, if testing the connections between visual areas V1 and V5, we can assume that all subjects will have this connection in order to process movement from the primary visual area of V1. FFX is the analysis method of choice when testing a hypothesis regarding the model structure (Stephan et al., 2010).

However, an important appeal of DCM is the ability to experimentally manipulate the context or task in order to test the effect this has on a system and its connections. FFX allows for confident testing of the structure of a system or physiological mechanism, but is not appropriate for testing models where the parameters (or experimental manipulations) are of interest. The reason for this is that cognitive tasks can be conducted in different ways, indicating that there may be heterogeneity across models (Stephan et al., 2010). Furthermore, if FFX is used on a group where the underlying mechanism is thought to be heterogenetic, outliers can easily influence the optimal model selection (Stephan et al., 2009). This can be illustrated when considering how each method computes the relative group statistic.

For FFX analysis, a Group Bayes Factor (GBF; Stephan et al., 2007c) or group log evidences are computed which rely on the relative evidence of one model to another. This relationship means that for FFX, if one subject has a different optimal model to the rest of the group then this outlier will give a false positive of which model is the winning model for the group. In contrast, the use of Random Effects (RFX) analysis accounts for outliers and computes posterior model probabilities and exceedance probabilities. This approach computes how likely that a given model would be the optimal model for a randomly chosen subject from its expected posterior model probability. From this, one can compare between models to gain an exceedance probability value to compare against other models within the model space (Stephan et al., 2009).

2.3.1.8.1 Bayesian Parameter Averaging

The use of BMS identifies an optimal model within the model space supplied. However, BMS alone is not sufficient to make conclusions about the parameter estimates or coupling strengths between the regions. Conceptually, this can be characterised in terms of identifying a model which best describes a system while not knowing the strength or influence (modulation or attenuation) of these connections. With DCM, different techniques are applied depending on whether Fixed Effects or Random Effects analyses were used.

For FFX analysis, Bayesian Parameter Averaging (BPA) is used to make inferences on parameter estimates. By taking the individual statistics for a given parameter, BPA combines these for each subject to obtain a joint posterior density where the posterior from one subject is used as the prior for the next (Garrido et al., 2007; Neuman & Lohmann, 2003). A crucial element of BPA is that the use of FFX is justified, i.e. the data will be skewed if this model is not optimal for every subject. Moreover, the use of these posterior covariances can also behave in a counter-intuitive way if there are high levels of signal to noise (Kasses et al., 2010). On the other hand, BPA has advantages such as producing a single posterior density that can then be easily interpreted with Bayesian Inference (Acs & Greenlee, 2008) and the fact that it is mathematically easy to calculate (Stephan et al., 2009; Stephan et al., 2010).

2.3.1.8.2 Bayesian Model Averaging

When using RFX analysis, the assumption that different cognitive approaches may be used to complete the task and therefore different subjects may have different models is accepted. This point can be clearly observed when viewing a BMS result where there is no outperforming model; characterised by matched model evidence across models. To make inferences on parameter estimates under these conditions, Bayesian Model Averaging can be used (BMA: Hoeting et al., 1999; Penny et al.,

2010). Unlike BPA, BMA refocuses from a particular model to that of the entire model space and calculates weighted averages of each model parameter where the posterior probability of a model is the weighting. This approach can be best conceptualised with using model family comparisons (Penny et al., 2010). Consider two families of models where each is grouped by a common connection that is not present in the other family. A simple BMS analysis of all models may reveal that there is not a clear winner, but the same analysis when the groupings are included reveals that one group is considerably more optimal. BMA allows one to remove this uncertainty about the structure of the model, as the models within these families may be all different in terms of fixed connectivity; BMA pools the information of these optimal models and produces parameter estimates (see Penny et al., 2010 for the mathematical calculations).

BMA gives parameter estimate averages for each connection; however, inferences concerning modulation strength cannot be made as it does not necessarily indicate that there is a significant connection. Using the parameter estimates from each subject, significant differences between connections and connection strengths can be calculated using frequentist statistical tests (e.g. *t*-test or ANOVA), making sure to correct for multiple comparisons.

2.3.2 Evidence and limitations

DCM combines realistic biological modelling with statistical techniques; these aspects, which separate it from other methods of functional and effective connectivity, can be viewed as "two sides of the DCM coin" (Daunizeau et al., 2011). That is, the innovations of this methodology can also be viewed as potential limitations when compared to regression-based measures of effective connectivity (those that do not necessarily rely on neurobiological plausibility or a Bayesian approach to inversion and inference, such as SEM (McIntosh et al., 1994; Buchel & Friston, 1997; Horwitz et al., 1999) or auto-regressive models (Harrison et al., 2003; Roebroeck et al., 2005)).

In the next section, the identified limitations of biological plausibility (or biomodelling) and statistical techniques are outlined (for an in-depth critical review, see Daunizeau et al., 2011). With reference to the biological plausibility of DCM, test, face and predictive reliability are also discussed in relation to SEM and GCM. General statistical inference concerns regarding the use of a Bayesian framework are then discussed, before outlining general design and subjective problems that DCM encounters.

2.3.2.1 Bio Modelling

As first shown in *Equation 1* of this chapter, DCM relies on the implementation of a neuronal model (which is later combined with a hemodynamic model and inverted within a Bayesian framework). This neuronal model is simplified so that causality of influence exerted can be examined between regions. However, the fact that this model is simplified (potentially too much so) may underestimate the importance of biological mechanisms. This concern has been outlined in a series of papers in the 'Comments and Controversies' special issue of *NeuroImage* dedicated to "The identification of interacting systems in the brain using fMRI: Model selection, causality and deconvolution" (Roebroeck et al., 2011a, 2011b; Valdes-Sosa et al., 2011; David, 2011; Friston, 2011; Bressler & Seth, 2011).

In addition, some have argued that the neural model of DCM may neglect the effects of inhibitory activity on the hemodynamic response (Sotero & Trujillo-Barreto, 2007; Shmuel et al., 2006). In other words, neurotransmitters with inhibitory effects may be involved in the system of question, yet DCM does not fully account for the effect of these when modelling the neural dynamics, which is therefore not accounted for in the hemodynamic forward model. This argument, as well as that of those highlighted in the 'Comments and Controversies' special issue of *NeuroImage* point to the fact that DCM requires further development and validation.

2.3.2.1.1 Test reliability

A fundamental aspect of any methodological approach, regardless of the discipline, is that it must satisfy test-retest reliability. This form of reliability presumes that when a technique is applied to the data on a second occasion, the results should be identical. In DCM, this re-test reliability should relate to both the model being selected and the parameter estimates of the given model.

Rowe et al. (2010) tested the reliability of the model selection procedure across multiple scanning sessions. The authors collected data from patients suffering from Parkinson's disease over two separate sessions (two separate time points). Critically, the authors found that the DCM analysis identified the same model across sessions, implying that the model selection was a reliable and robust measure between sessions. However, the authors of this paper do point out that the posterior estimates (coupling strengths) of these models were less reliable (i.e., similar, but not identical) across sessions. This finding contradicts other evidence that demonstrates the reliability of parameter estimates across sessions (Schuyler et al., 2010). One explanation for this discrepancy is that the models tested in Schuyler et al. (2010) were very simplistic and did not include clinical patients. Moreover, the time between sessions differed (weeks vs. minutes respectively) which could help explain why the parameter estimates differed in the Rowe et al. (2010) paper.

From these studies, we can see that model selection is reliable across sessions. This test-retest approach of validation allows us to be relatively confident in its accuracy. However, the conflicting findings concerning posterior estimates across sessions require further investigation to establish test-retest reliability.

2.3.2.1.2 Construct validity

Construct validity can be characterised as the extent to which a measure (*Z*) can explain variable *X*, where variable *Z* and an additional method (*Y*) are related (Cronbach & Meehl, 1955; Smith, 2012). In relation to connectivity analyses, this can be operationalised by testing two methods on the same data set. This is different to predictive validity (discussed in 3.1.3) as *X* is not fundamentally (physiologically in this case) grounded. Penny et al. (2010) aimed to test the construct validity of DCM

by comparing it to SEM by applying both connectivity methods approaches to the attention to motion data set supplied by Buchel & Friston (1997). This data set included three conditions: (1) photic; (2) motion with no attention; and (3) motion with attention.

Structural Equation Modelling (SEM) was developed in the field of econometrics and applied to imaging data by Mcintosh and Gozalez-Lima (1991). In short, with SEM, causal relationships are inferred from the data *a priori* (Pearl, 1998) where connection strengths relate to correlations. By minimising the discrepancy between then observed and assumed correlations, a model fit is calculated. This inference method of model fit by using likelihood ratio tests is analogous to Bayes factors in DCM (Penny et al., 2004). Penny and colleagues (2010) found that both methods led to the same conclusions with the data set investigated; reciprocal connectivity between regions and attention modulates connectivity between V1 and V5 (also shown in Friston et al., 2003). However, the advantages of DCM over SEM here relate to the fact that SEM makes no distinction between the neuronal and hemodynamic levels, as well as the fact that it also uses nested models. Firstly, the issue of biological plausibility of DCM can be seen as a strength in this study. For example, SEM makes assumptions of interactions between brain regions at the hemodynamic level (apparent as BOLD signal changes) however, as suggested by Gitelman et al. (2003), neuronal interactions do not necessarily lead to a detectable hemodynamic change. Penny et al. (2004) proceed to suggest that SEM is more applicable to PET where the neuronal-hemodynamic relationship is not so confounded. In addition, DCM is advantageous with its inference technique, whereby non-nested models can be compared (meaning smaller models do not necessarily attribute to a larger models), a feature not available to likelihood ratio tests. Penny et al. (2004) convey that the use of a biologically plausible model is an aspect of DCM that gives it strong construct validity.

Additional support for the construct validity of DCM can be seen in a study conducted by Lee et al. (2006). In this study, the authors aimed to test the validity of DCM using synthetic fMRI data, characterised by a neurobiological informed computation model (Horwitz & Tagamets, 1999). In brief, this simulated model

included inhibitory and excitatory synaptic activity (and random fluctuations) to best ascribe a neuronal system. By integrating the values, a hemodynamic fMRI time series for each region can be obtained (akin to the inversion scheme employed to DCM). Crucially, this method of spiking neural modelling allowed the authors to know the connectivity of the system, thereby validating DCM in its ability to model and make inference on which model best describes the system. Lee et al. (2006) deduced that with the use of large-scale spiking neuron models, DCM identified a well-characterised (in comparison to the known connectivity) and accurate model.

Considered together, these two studies offer support for the construct validity and partial face validity from Lee and colleagues (2006; see also Stephan et al., 2007b; 2008 for more simulated time series validation of DCM). By applying DCM to an informed neural system, DCM accurately identifies the optimal model. The use of a biologically plausible model is central to understanding its advantages. However, the key test of validity and reliability for any kind of modelling which aims to make inference on causality is to test how accurately it can identify this. In the next section, DCM shall be discussed in terms of its predictive validity.

2.3.2.1.3 Predictive validity

Predictive validity is the ability to model, or more specifically predict, future events. In connectivity approaches, this would relate to the ability of the methods to accurately model or predict the *actual* neural activity. This is attainable in cognitive neuroscience from rodent and primate studies as one can collect electrophysiological recording from single neurons, or specific sites that might make up a system. By using these data, one can test the predictive validity of models, such as DCM, which make assumptions and inferences at this level. In this section, Granger Causality Modelling will be briefly recapped and reviewed in relation to DCM (on a theoretical level). This leads to a discussion of a pivotal study by David et al. (2008), who directly compared the predictive validity of these two approaches of functional integration. This is followed by a brief discussion of additional studies demonstrating the predictive validity of DCM (more specifically, its bio-modelling element).

To recap, Granger Causality Modelling (GCM: Goebel et al., 2003; Valdes-Sosa, 2004)) is a form of functional connectivity, where connectivity is defined by statistical dependencies. In comparison to effective connectivity methods, no explanation as to how a connection is caused is modeled/included. Like DCM, GCM also aims to model time series data (Friston et al., 2012). The core principle with GCM is that X '*causes*' Y when X can help predict future events in Y (better than past events in Y can alone) (Granger, 1969; Seth, 2010). Although both methods appeal to notions of causality, GCM cannot draw conclusions that activity in one part of the brain *causes* activity in another. Unlike DCM, GCM does not use a forward model (neuronal dynamics). Furthermore, the model comparison technique employed in GCM is restricted by the use of statistical dependencies (Friston et al., 2012). Unlike DCM where multiple hypotheses can be compared (see BMS; 2.3.1.7), GCM tests the dependencies between regions against a null hypothesis.

In terms of biological modelling of the data, GCM is not concerned with what causes the pattern of data, and instead focuses on the temporal dependencies between regions (Harrison et al., 2003). Furthermore, the parameters of the GCM models have no biological meaning, just statistical dependencies (Friston et al., 2011). An important issue to arise from GCM is the fact that it relies on temporal statistical dependencies, for example, activity in region X is preceded by activity in region Y. As highlighted in the fMRI chapter of this thesis, fMRI suffers from temporal issues (slow dynamics) in relation to the hemodynamic response of a region (David et al., 2008). This is confounded for GCM as it relies on temporal statistical dependencies, for which it assumes that the signals are stationary (i.e. not necessarily a dynamic system; Roebroeck et al., 2005; Londei et al., 2006).

Although the use of GCM for modelling fMRI data can be questioned, this method has been shown to be useful in electrophysiological studies (for example, Broveli et al., 2004; Bosman et al., 2012). The reason for this is based on the fact that electrophysiological data measures the electrical discharge of neuronal activity. As no translation between neuronal and a measureable signal is required, GCM can accurately rely on the temporal correspondence of the signal in given regions, not accessible with fMRI.

A key study in the predictive validity of DCM, which also acted as a direct comparison to GCM, was conducted by David et al. (2008). In this study, the authors measured the brain responses of a rat to epileptic events, at both the source of the epileptic events and the connecting regions. This was a particularly empirically rich study due to the different sources of information gathered. The authors recorded EEG and fMRI activity during the seizure events, followed by intracranial signals in areas identified by EEG and fMRI. By combining methodological approaches in this manner, the authors were able to identify the true neural connectivity from the electrophysiological recordings (this is the benchmark procedure to test the forms of connectivity in terms of predictive validity), as well as a hemodynamic level of activity from the fMRI recording to apply DCM and GCM to. Applying connectivity measures to the fMRI time series for the seizure allowed for different conclusions can be drawn. The electrophysiological data showed that the driving spike, or epicenter of the seizures was located in the somatosensory cortex, a conclusion that was reached by DCM but not GCM. The authors also found substantial differences between the hemodynamic responses in different regions, an aspect that was only taken into account when hidden states (or neuronal discharge) were taken into account – as DCM does. David et al. (2008) conclude that the use of GCM is appropriate when brain states can be observed directly (as is the case with EEG), but is not when the data being recorded is a post hoc consequence of these states (BOLD). The results of this study potently point to the fact that DCM has predictive validity. It shows that, in comparison to GCM, the modelling of hidden states (in the form of separate regions, with different hemodynamic responses), can accurately predict the true electrophysiological model.

In addition to this influential study by David et al. (2008), the predictive validity of DCM has been investigated in the presence of strong physiological confounds (not modeled in the generative neuronal model of DCM; Reyt et al., 2010), changes in synaptic responses (Moran et al., 2008; 2009), excitatory and inhibitory synaptic processing (Moran et al., 2011; see also Brodersen et al., 2010 and Brodersen et al., 2011).

2.3.2.2 Statistical Inference Techniques and General Design Matters

The statistical inference techniques used for DCM (i.e. Bayesian comparisons) can be questioned by those accustomed to the frequentist approaches (for example *T*-tests and ANOVAs). The reason for this disparity is that DCM models cannot be formally tested; that is to say, they are not falsifiable. In other words, unlike frequentist methods, DCM models are not compared to a null hypothesis. Instead they are compared to an equally likely experimental hypothesis (indeed, this model can be a reductionist account of the experimental model). This use of relative falsification is embedded in the framework of Bayesian statistics, however. As such, one can question the generalisability of the subsequent results. Consider the following example: we tested the exogenous modulation of a system whereby we tested two models, one with a connection flowing anteriorly and one reciprocally. The model evidence showed that the anteriorly flowing model was the optimal model for this data set.

As noted by Lohmann and colleagues (2012), the tested models are actually tested on their plausibility, given the data. Therefore, when comparing the two models in our example, the anteriorly flowing model is more plausible than the alternative model. However, the alternative model may be inaccurate and the optimal model satisfies the accuracy trade-off referred to earlier in this chapter. Moreover, Lohmann et al. (2012) provided simulation data whereby they demonstrated that a winning model could not be found as there were other alternative models with greater model evidence than the true model (physiologically plausible). In response to this, Friston et al., (2012) reiterated the fact that a simpler explanation would have greater model explanation (accuracy vs. complexity trade-off). Indeed, Lohmann et al. (2012) included much simpler models in their simulations compared to the true model that was a fully connected and modulated model. What this debate clarifies is that the selection of model space has to be firmly grounded in theory (Stephan et al., 2009). This point becomes more apparent as a Bayesian approach is used where models are compared in a relative falsification procedure, and therefore, an ill-poised model space will lead to inaccurate results.

In addition to the realistically plausible biological model and statistical inference methods employed by DCM, there are also more general design issues that present potential limitations to its use. These issues are related to the size of model space and the specificity of the data set. The use of DCM is usually refined to the testing of two competing hypothesis (Friston et al., 2003). By only testing two alternative hypotheses, the size of model space is reduced, computational time is minimised and the interpretation of results is cleaner. However, Friston et al. (2011) presented an algorithm, which allowed one to test a number of models in the order of hundreds. This method relies on nested models, whereby connections are "turned off" to test the presence or absence of connections. With this approach, one cannot investigate the subtleties of a given model as each connection is "turned off" in turn. Lohmann et al. (2012) also criticised this approach for this reason, saying that DCM cannot be used as an exploratory tool. Indeed, Friston et al., (2012) suggest that looking at a large model space may seem appealing but this does not mean that the question of the study will be answered; just because it is possible to test many models does not mean that one should.

Finally, the following two limitations correspond to all measures of connectivity, functional or effective. The first limitation concerns the need for the researcher to select the ROIs upon which to perform the connectivity analysis. This approach presents inter-subject variability and care has to be taken into extracting consistent regions. Additionally, the extracted data and subsequent connectivity analysis is only specific to the given data set. This means that it is difficult to compare systems across studies, or where different ROIs or numbers of nodes have been used (Stephan et al., 2010).

2.3.3 Summary and conclusions

The approaches taken to understand neurocognitive function have shifted from identification of single areas associated with performing certain tasks to that of exploring the dynamics of interactions between a number of brain regions (or between systems) (Smith, 2012). DCM aids understanding of these systems by marrying biophysical with functional integration. As a method of effective connectivity, DCM

is not without its limitations, including those that make it unique, namely the biophysical model. This model is used to convert predicted neural connectivity into a time series that can be compared to the observed time series. The accuracy of this conversion is open to criticism with the use of DCM. A growing literature documents its validity for testing systems both in biological accuracy and model selection. Future extensions of DCM will take into account further understandings of the relationship between BOLD and neural function (Logothetis et al., 2004). Currently, however, it offers a powerful tool for systematically testing effective connectivity within the brain.

3 Chapter III

Dynamic modulation of the action observation network by movement familiarity

3.1 Abstract

When watching another person's actions, a network of sensorimotor brain regions, collectively termed the action observation network (AON), is engaged. Previous research suggests that the AON is more responsive when watching familiar compared with unfamiliar actions. However, most research into AON function is premised on comparisons of AON engagement during different types of task using univariate, magnitude-based approaches. To better understand the relationship between action familiarity and AON engagement, here we examine how observed movement familiarity modulates AON activity in humans using dynamic causal modelling, a type of effective connectivity analysis. Twenty-one subjects underwent fMRI scanning while viewing whole-body dance movements that varied in terms of their familiarity. Participants' task was to either predict the next posture the dancer's body would assume or to respond to a non– action-related attentional control question. To assess individuals' familiarity with each movement, participants rated each video on a measure of visual familiarity after being scanned. Parametric analyses showed more activity in left middle temporal gyrus, inferior parietal lobule, and inferior frontal gyrus as videos were rated as increasingly familiar. These clusters of activity formed the regions of interest for dynamic causal modelling analyses, which revealed attenuation of effective connectivity bidirectionally between parietal and temporal AON nodes when participants observed videos they rated as increasingly familiar. As such, the findings provide partial support for a predictive coding model of the AON, as well as illuminate how action familiarity manipulations can be used to explore simulation- based accounts of action understanding.

3.2 Introduction

Perceiving and interacting with others form the foundation of human social behavior. When watching others in action, we readily extract information about their goals and intentions (Hamilton & Grafton, 2006; Hamilton, 2013) and predict their subsequent behavior in a rapid, online fashion (Blakemore & Frith, 2005; Falck-Ytter et al., 2006). Neuroimaging studies identify frontal, parietal, and occipitotemporal regions collectively termed the Action Observation Network (AON) (Cross et al., 2009; Grafton, 2009; Keysers & Gazzola, 2009; Caspers et al., 2010) as critically involved in processing others' actions. Previous research demonstrates greater AON activity when watching visually or physically familiar actions (Buccino et al., 2004; Calvo- Merino et al., 2005; Cross et al., 2006; Vogt et al., 2007; Shimada, 2010; Press, 2011). These findings support a theory of action understanding that posits the AON is tuned to respond most to actions that are "like me" (Meltzoff, 2007) via experience-driven simulation mechanisms (Sinigaglia, 2013). However, most studies into how observers' prior action experience or familiarity impacts AON engagement have used magnitude-based approaches, investigating which brain regions show increased or decreased response amplitudes based on familiarity with an observed action. Some have suggested that such analyses might not be sufficient to construct a nuanced or complete picture of how sensorimotor brain regions support action understanding (Schippers & Keysers, 2011). One promising approach for further characterization of the relationship between familiarity and AON engagement is to examine functional connectivity between individual AON nodes during action observation, and ask how a particular task modulates these connections (Smith, 2012).

Prior theoretical work on the AON can inform predictions about how familiarity might modulate connectivity when observing others in action. Predictive coding models of AON function (Keysers & Perrett, 2004; Kilner et al., 2007a,b; Gazzola & Keysers, 2009; Schippers & Keysers, 2011) are predicated on the use of perceptuomotor maps to predict and interpret observed actions (Lamm et al., 2007; Schubotz, 2007; Urgesi et al., 2010). According to this account, forward models facilitate processing of familiar actions through use of stored action representations

and propagate information in a top-down manner from premotor to parietal to occipitotemporal regions. Unfamiliar actions are processed via a bottom-up, data-driven approach, where information propagates anteriorly from occipitotemporal to parietal to premotor cortices. To test whether and how action familiarity modulates effective connectivity among the component regions of the AON, we used dynamic causal modelling to test hypotheses under a predictive coding model of action observation. Our task involved observing whole-body movements followed by a short occlusion, after which participants chose which posture should follow in the movement sequence (similar to Stadler et al., 2011). By using an offline rating task, we examined how this noncategorical measure of familiarity impacts AON engagement (compare Cross et al., 2013; Liew et al., 2013). According to a predictive coding model, dynamic causal modelling (DCM) should demonstrate that increasingly familiar movements are associated with decreased feedforward influence from sensory/posterior nodes, as prediction errors are minimized when viewing more familiar actions, whereas feedback influence from anterior to posterior nodes should increase.

3.3 Materials and Methods

3.3.1 Participants

Twenty-one adult volunteers (mean age 21.95 years, SD 3.02 years) with normal or corrected to normal vision were recruited from the student population at Radboud University Nijmegen. Of this sample, 13 were female; 17 right-handed, 2 left-handed, and 2 ambidextrous based on the Edinburgh Handedness Questionnaire (Oldfield, 1971). All participants' data were used for the GLM analyses, and a subsample of 19 participants was used for the DCM analysis (see Definition of ROIs). All participants spoke English fluently and had no history of psychiatric or neurological disorders. All participants provided written informed consent, and the study procedures were approved by local ethics committees at both Bangor University and the Donders Centre for Cognitive Neuroimaging at Radboud University Nijmegen. Participants were reimbursed for their time with €15.

3.3.2 Experimental design and stimuli

The paradigm included two task manipulations that fell within a 2x2 factorial design. The two factors were familiarity (levels: familiar and unfamiliar) and task (levels: posture prediction or dot colour tracking). Action stimuli were consistent across the two tasks (only the instructions changed).

3.3.2.1 Stimuli construction and selection

Stimuli were created by filming a professional dancer performing a range of improvised and choreographed movement in a contemporary dance style that ranged from extremely simple (and thus relatively predictable) to much more complex (and much less predictable). To ensure these stimuli encapsulated a broad range of more and less familiar actions for dance-naive observers, we first piloted 157 video clips (each with a duration of 6 s) with a separate sample of 23 participants. From this pilot study, we selected the 30 movement stimuli rated as most familiar, and the 30 movement stimuli rated most unfamiliar by pilot participants. This split allowed us to make a stimuli set that should subjectively differ across the familiarity spectrum.

3.3.2.2 Prediction task

In the Prediction condition, participants were asked to monitor the movements of the agent in preparation for a two-way forced choice question asking them to choose which posture should follow on in the video clip after an occlusion of 0.6 s. The forced choice consisted of one coherent still frame (chosen from 0.6 s after when the video was occluded) and one that was plausible to the video but not chronological to the flow based on the occlusion duration (for an example, see *Figure 3.1*).

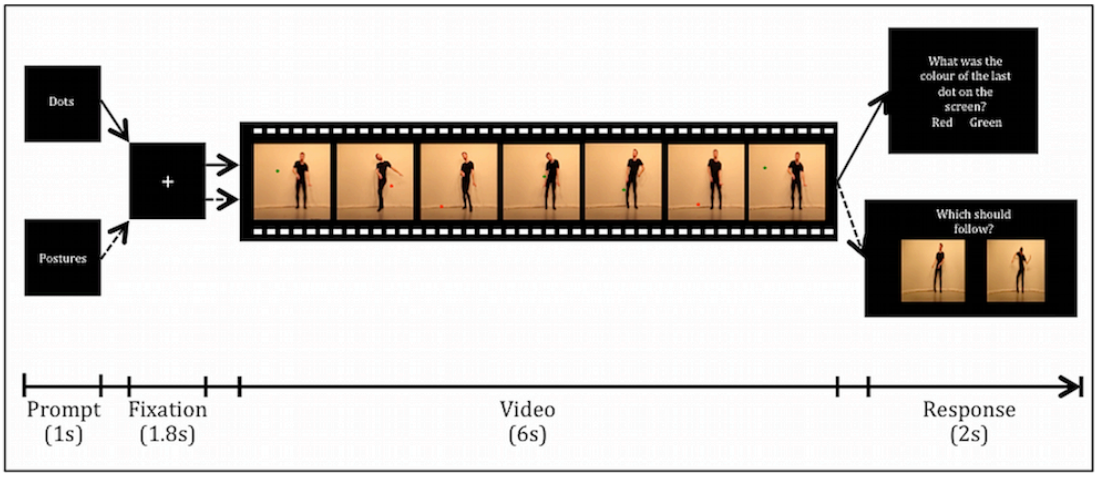


Figure 3.1. An illustration of the trials. The prompt screen was shown at the beginning of each block (10 trials) to orient the attention to the dots or postures. The fixation was shown at the start of each trial, followed by the video. This was then preceded by a question on the dots or a prediction question dependent on the block type. The response screen was displayed for 2 seconds and would display a blank black screen even if there was a response, until the 2-second window ceased.

3.3.2.3 Attentional control task

A control task was built into paradigm so that, for half the trials, participants were asked to monitor the colour of a dot that would randomly appear on the screen in different locations and in different colouurs (either red or green). The size of the dots was 5 mm and appeared at a rate of 1 per second (to clarify, coloured dots appeared during all video stimuli in the experiment, but participants were asked to attend to them during only half the trials). When participants were performing the attentional control task, they were asked to monitor the colour of the dot throughout the duration of the video clip. When the video clip ended, a question appeared asking participants to specify the colour of the last dot they saw on the screen.

3.3.2.4 Post scanning ratings

After the fMRI experiment, participants rated the videos on a Likert scale of 1–9 as to how familiar they found the movements within each video (0 very unfamiliar; 5 neutral; 9 very familiar). The dots were removed from these videos, and

the response window was not fixed. Still, participants were encouraged to answer as quickly and accurately as they could. The concept of familiarity was explained by the experimenter as how easily participants would be able to say what move should come next in each video, and participants were made explicitly aware that familiarity did not mean which videos had been recognized or remembered from the scanning experiment. Each video was displayed only once and in a randomized order.

To ensure that familiarity ratings and number or speed of movements were not confounded, we calculated the mean "motion energy" of each video stimulus based on a previously developed algorithm (Cross et al., 2012). Once we obtained an objective numeric score for how much pixel displacement each video contained, we computed a Pearson product- moment correlation coefficient to assess the relationship between the motion energy of each video and the average familiarity rating given by participants. The results revealed no correlation between the two variables (r 0.06, n 60, p 0.646). This suggests that the participants did not rate the stimuli based on how much movement each stimulus contained; instead, they were rating them on a more holistic, subjective view of familiarity.

3.3.3 fMRI design and procedure

Each participant completed one fMRI session that followed an event- related design. Participants completed two runs, lasting 13 min and containing 60 trials each. Trials were blocked into more familiar and less familiar stimuli (based on pilot data), and both kinds of stimuli were presented for both task conditions. At the beginning of the first run and the end of the second run, a 15 s rest period occurred. At the start of each block, a prompt (1 s) indicated which task participants were to perform for the upcoming block of trials (predicting postures or reporting on the dot colour). The ensuing blocks, consisting of 10 trials, were all from the same condition. At the start of each trial, a white fixation cross appeared on the center of the screen for 1.8 s, followed by a video clip (6 s). Based on the prompt at the start of the block, participants had 2 s to respond to the task and identify which still frame they thought

would follow or identify the colour of the last dot on the screen. This response period lasted for 2 s. If a button was pressed before 2 s had elapsed, the screen changed to a blank black screen until the 2 s time limit was reached. The order of the blocks and the video shown in each trial were pseudo-randomized so that each video was shown once in each of the conditions.

Stimulus presentation and response collection were performed using Psychophysics Toolbox (version 3) via MATLAB R2010a (MathWorks). The stimuli were projected onto a mirror above the head coil from a projector outside of the scanner. Participants made their responses with the forefinger and middle finger of the right hand, and responses were recorded from a custom-made MR-compatible button box.

Data acquisition was conducted at the Donders Centre for Cognitive Neuroimaging at Radboud University Nijmegen. Functional images were acquired on a 3.0T Siemens MRI scanner using a 32-channel head coil. Functional images were acquired covering the whole brain using an echo-planar imaging (EPI) sequence (35 axial slices, ascending slice acquisition, repetition time 2000 ms, echo time 30 ms, 90° flip angle, matrix 64x64, slice thickness: 3x3x3 mm, field of view (FOV): 224 mm). Before the functional run, 196 two-dimensional anatomical images (256x256 pixel matrix, T1-weighted) were obtained for normalization purposes.

3.3.4 fMRI data preprocessing and statistical analysis

A total of 338 volumes per participant per run were used in the analysis. Because of a technical error, two participants' data were not collected correctly at the start of the first functional run, resulting in a reduced number of volumes for these participants (615 volumes in total compared with 676 for all other participants). Data were analyzed using Statistical Parametric Mapping (SPM8: Wellcome Trust Centre for Neuroimaging, London) (Friston, 2007) implemented using MATLAB R2010a

(MathWorks). The data were first realigned and then slice-time corrected and preliminarily preorientated within standard stereotaxic space as defined by the MNI (Friston, 2007). This preorientation allowed for a better spatial normalization to the MNI template. Participants' EPI images were then coregistered to their T1 anatomical scans, which were then spatially normalized to standard stereotaxic space. The spatially normalized EPI images were filtered using a Gaussian kernel of 8 mm full-width at half maximum in the *x*, *y*, and *z* axes. A design matrix was fitted for each subject with a single regressor for all trials from the prediction task and a single regressor for all trials from the coloured dot-tracking task. A parametric regressor column was added to the design that included participants' individual ratings of each video from the posture prediction task (assigned outside the fMRI session). In modelling a single parametric effect of familiarity, we are effectively modelling the main effects of action observation and familiarity but not their interaction. The fixation, prompt, and response were fitted as noise regressors for each individual and combined into a single regressor of noninterest. Each trial was then modeled as a boxcar function for the duration of the video. For the GLM analyses, a cluster threshold of *k* 10 and a *p* value of 0.005 (uncorrected) was set to best observe the sensitive effects of the parametric analysis.

Although participants were encouraged to use the entire scale when rating the videos, many participants used a reduced range of the 9 point scale (for example, 2–8). Because we were interested in using each individual's ratings as a parametric regressor in the group fMRI model, it was important to equalize the relative rating scales across participants. To achieve this, participants' ratings were standardized via a *z* transformation.

3.3.4.1 Objectives of GLM neuroimaging analyses

Neural processes engaged during action prediction. The first analysis identified brain regions that responded to the task demands to predict movements by evaluating the task-based contrast of posture prediction attentional control task. This

contrast allowed us to explore the effects of online monitoring of an action while controlling for familiarity of the movements.

Parametric effects of increasing familiarity of an observed action. The second analysis allowed us to explore the sensitive measure of subject- specific (standardized) ratings of observed movements on AON activity. This parametric contrast was used to test the hypothesis that AON regions should show greater activity as a function of increasing familiarity. Unlike contrasts that would use segmentation of the stimuli into familiarity categories, our measure allows us to capitalize upon individual differences in familiarity ratings to identify brain regions that become increasingly or decreasingly active the more familiar a movement is rated by an observer.

3.3.5 DCM

Evaluation. DCM is a method of analyzing effective connectivity that uses a bilinear model of neural population dynamics, combined with a hemodynamic model, which aims to describe the neural activity in the measured BOLD response (Friston et al., 2003). By modelling feasible neuronal parameters, DCM aims to make a modeled BOLD response that is similar to the actual experimentally manipulated BOLD signal. The neural dynamics model created using DCM is then combined with a hemodynamic forward model that incorporates a balloon model (Buxton et al., 1998). The hemodynamic model provides a transformation of how the neural dynamics would propagate as a BOLD response, estimated via a standard Bayesian approach (variational Laplace). The fit of the combined neural model and hemodynamic forward model is estimated via a Bayesian approach, which uses conservative shrinkage priors for the coupling parameters.

A model is specified to have the fit to the data estimated by the previously mentioned routine. The input into the estimation procedure is three matrices (for bilinear DCMs, but see also Nonlinear DCMs) (Stephan et al., 2008). The first is the endogenous connection strengths (the *A* matrix), which represents the connectivity

between the regions of the model, sometimes called the fixed connectivity. The *B* and *C* matrices represent the experimentally manipulated conditions. Matrix *B* represents the modulation of an external input on a fixed connection, which describes the change in the value of the effective connectivity for a connection under a particular condition. The *C* matrix represents the input into the system, which describes the activity that is perturbing, or creating activity, in the system. This equation models the state changes by known inputs. By user-created variations of combinations of fixed connectivity, modulatory influences, and inputs, multiple DCMs can be created for the same set of regions and then compared with which fits the data best (while controlling for complexity).

Hypothesis. The experimental hypothesis tested in the current study is that differential connectivity will exist between the three core AON regions based on subjective familiarity ratings made by the participants. Moreover, we expected that more familiar actions would increase the influence of top-down connections (IFG-IPL-MTG) while decreasing the corresponding influence of bottom-up connections (MTG-IPL-IFG). Mechanistically, this corresponds to increasingly precise or confident top-down predictions afforded by familiar actions that, in our model, would be encoded by the modulatory effects of familiarity on the B parameters.

Preprocessing for DCM. To evaluate the effective connectivity of this network using DCM, volumes from the two separate runs were concatenated to form one single time series per participant. We repeated the GLM analysis in SPM12b to exploit recent developments in dynamic causal modelling (see below). The GLM was effectively the same as described above. However, in this case, we adjusted the data for the main effect of action observation. We then used the effect of stimulus movement as a driving input to the action observation network, whereas familiarity was used to modulate extrinsic (between node) connectivity. Effectively, this models the effect of familiarity in terms of context sensitive changes in coupling induced under action observation. Notice that we effectively removed responses to movement videos during the attend dots conditions, enabling us to focus on the effect of

familiarity during action observation. These will be subsequently referred to as the movement effect. The remaining conditions were modeled as a nuisance variable to ease DCM model specification (Stephan et al., 2010). These preprocessing steps allowed us to examine the effects of familiarity of actions on the effective connectivity of the system.

We allowed for stochastic effects within the model to more accurately model noise (Li et al., 2011). We also centered the input into the node. This gives the input a mean of zero and means that modulating parameters can increase and decrease the fixed connectivity, as opposed to simply increasing it as when the input is always positive. To allow for biological plausibility, we opted for two-state models that allow for both excitatory and inhibitory connectivity (Marreiros et al., 2008). Because we used a two-state DCM, we were able to enforce positivity constraints on the connection strengths and lend our interpretation a greater biological plausibility or validity. This is because all connection strengths and two-state DCMs are excitatory, where intrinsic (within node) excitatory connections activate inhibitory neurons to ensure stability of the modeled network. Crucially, the coupling strengths are log scale parameters. In other words, they represent the log of the scaling of an effective connection, such that a log of zero corresponds to a 100% scaling. This means that a negative log scaling parameter corresponds to a weaker connection and a positive parameter corresponds to a stronger connection. We will report the parameter estimates in log space (and perform *t*-tests on the log scaling parameters). This additional step allowed us to further examine the proposal by Keysers and Perrett (2004) that the MTG would exhibit inhibitory responses to familiar stimuli as well as help us understand how it relates to the predictive coding model proposed by Kilner et al. (2007b).

To identify the winning model (in this case, the model that best explains the system of interest), we used random-effects Bayesian model selection (BMS) to account for outliers (Stephan et al., 2009). This Bayesian approach used for DCM gives each model an exceedance probability, the probability that a model is more

likely than any other model tested (subject to a trade-off between model fit and complexity). The family comparison technique pools model evidence by the user-defined groups, where all models within a given group share a common feature (Penny et al., 2010).

3.3.6 System of interest

Definition of ROIs. The coordinates for the ROIs were identified from the parametric contrast of increasing familiarity (for coordinates, see *Table 3.1*). The time series for the ROIs were taken from the subject level *F*-contrast of movement implicit baseline as this was the most revealing contrast. This contrast revealed all regions that were active (two- tailed) when viewing movement, including the defined coordinates for MTG, IPL, and IFG in the left hemisphere.

Table 3.1. The coordinates used for the ROIs

X	Y	Z	Region	BA	Cluster Size
-48	-31	34	Inferior Parietal Lobule	40	141
-51	-58	-2	Middle Temporal Gyrus	39	60
-45	8	25	Inferior Frontal Gyrus	44	12

All ROIs were extracted by locating the nearest local maximum voxel to the coordinates of the group contrast of increasing familiarity. 19 of the 21 participants showed significant activity within the search radius of 16 mm. The 2 remaining participants were excluded from the connectivity analysis. The ROI time series for each region, for each participant, was extracted by using the eigenvariate (threshold of p 0.05), with sphere radius 8 mm and adjusted for effects of no interest.

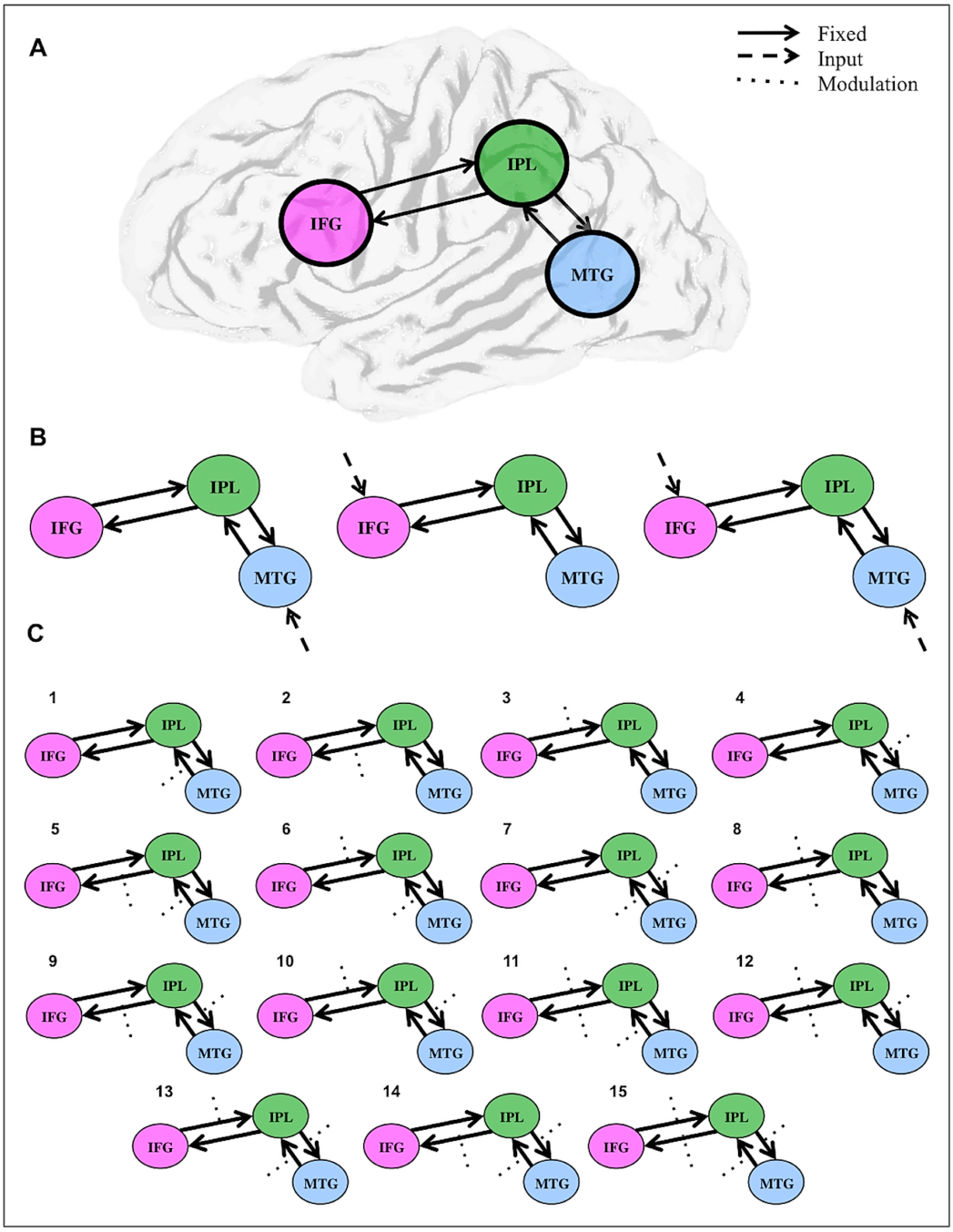


Figure 3.2. The model space. ***A***, The endogenous (fixed) connectivity between the three regions. ***B***, The three families of inputs (movement) tested: MTG, IFG and MTG, and IFG. For each of these three families, 15 models, shown in ***C***, were tested to identify the modulation of increasing familiarity, resulting in 45 models for each participant. Solid lines indicate fixed connectivity. Dashed arrows indicate input into the system by movement. Dotted arrows indicate modulations on the given connection by increasing familiarity.

Definition of network models. To test our hypothesis, all models were

manipulated with regard to their modulation and direct inputs but not endogenous connectivity. To counter the problem of expanding model space, the endogenous connectivity of all models was anatomically in- formed by the theoretical models from which our hypothesis was based (Keysers & Perrett, 2004; Kilner et al., 2007b). The connections between the nodes were reciprocal between MTG and IPL and reciprocal between IPL and IFG (*Figure 3.2*). This limiting of model space allowed us to examine the modulatory effects of increasing familiarity within a refined selection of models. These modulatory effects (B parameters) were tested on all extrinsic (between-node) excitatory connections and all combinations. For the direct inputs to the system, we modeled the movement effect. We tested the hypothesis of direct input into the system through MTG as well as direct input into IFG for all variants as well as direct input into both of these regions, resulting in 45 models per person, with a total of 855 models overall (*Figure 3.2*).

3.4 Results

3.4.1 Whole-brain GLM analyses

3.4.1.1 Neural processes engaged during action prediction

To address how action familiarity impacts AON activity using random-effects analyses, we report two contrasts. The first contrast identified brain regions more engaged when participants watched whole-body movements with the intention to predict which posture should follow after an occlusion, compared with watching the same videos but attending to the colour of a dot that was superimposed on top of the dancer (*Figure 3.3*).

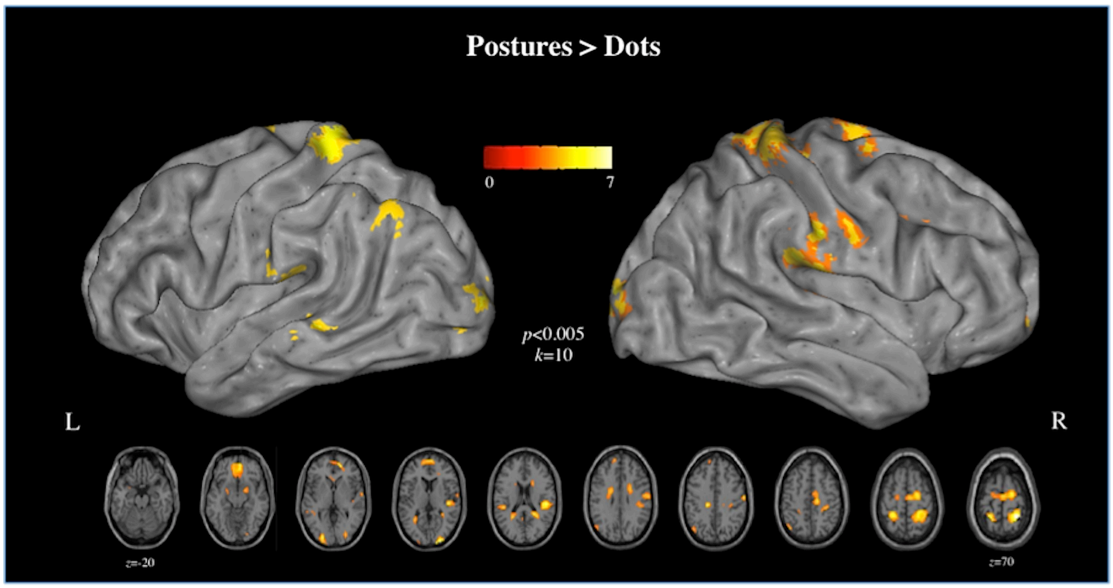


Figure 3.3. The GLM results of the main effect of performing the posture prediction task compared with the dot colour identification task. All *p* values 0.005; *k* 10 voxels.

This analysis revealed significant activity in M1 and the anterior regions of STG (for all peak coordinates from this contrast, see *Table 3.2*). This contrast shows the effects of the online updating of observed movements in anticipation of the decision-making period. The activity seen in the motor regions could perhaps represent the fact that, even though dance movements were observed in both conditions, the posture prediction task elicits more extensive engagement of motor cortical activity. Specifically, maintaining representations of actions so that they may be simulated during a period of occlusion may require more activity within primary motor regions (Kilner et al., 2009b; Stadler et al., 2011; Cross et al., 2013; Hari et al., 2014; compare mu suppression in this region during action observation). Interestingly, activity within higher-level occipital cortices (V4v and V3v) also emerged from this contrast. This pattern of activity could reflect increased visual attention demands of biological motion tracking, as opposed to what was required for the dot-tracking task.

Table 3.2. Localization of averaged BOLD response for the posture prediction versus dot tracking tasks

Anatomical region	BA	MNI coordinates			Putative functional name	*t* value	Cluster size	$P_{corrected}$ value
		x	*y*	*z*				
Postures vs. Dots								
R post central gyrus	1	27	-43	70	M1	7.25	326	0.032
R post central gyrus	2	27	-37	55	M1	5.13		0.140
R superior occipital gyrus	17	21	-100	7	V3v	6.77	90	0.032
R lingual gyrus	18	24	-88	-11	V4v	3.25		0.744
R rolandic operculum	41	48	-25	22	S2	6.70	345	0.032
R heschls gyrus	41/42	45	-22	10	PAC/STG	5.30		0.135
R postcentral gyrus	1	63	-13	34	M1	4.34		0.284
R mid orbital gyrus	10	6	50	-14	MPFC	5.72	415	0.129
R mid orbital gyrus	32/10	15	50	-2	ACC	5.40		0.129
R superior medial gyrus	10	9	59	10	MPFCv	4.34		0.284
L postcentral gyrus	2	-27	-40	58	M1	5.50	158	0.129
L postcentral gyrus	2	-21	-43	70	M1	5.10		0.140
L superior parietal lobule	31	-15	-25	43	PCC	5.46	32	0.129
R superior frontal gyrus	6	18	-10	67	PMd	5.16	382	0.140
R middle cingulate cortex	6	15	-22	46	PMd	5.01		0.145
R supplemental motor area	6	3	-16	67	SMAd	4.71		0.214
R parahippocampal gyrus	19/30	24	-46	16		4.98	102	0.145
L middle occipital gyrus	18	-24	-88	1	MOG	4.66	76	0.219
L middle temporal gyrus	22	-51	-40	-2	MTG	4.48	42	0.268
L medial temporal gyrus	22/37	-42	-46	1	MTG/STS	3.00		0.285
L inferior parietal lobule	19	-30	-55	13	IPL	4.47	91	0.915
L superior temporal gyrus	42	-45	-31	19	pSTG	4.30	45	0.285
L superior temporal gyrus	42/43	-60	-28	16	STG	2.96		0.344
R caudate nucleus	23	24	2	28		4.24	96	0.720
L caudate nucleus	23	-18	-1	28		4.02	82	0.344
L caudate nucleus	23	-18	14	22		3.29		0.462
L inferior parietal cortex	40/7	-42	-67	49	IPC	4.02	51	0.462
L amygdala	25	-48	-73	34	AMYG	3.77		0.529
R amygdala	25	-21	-1	-14	AMYG	3.75	34	0.529
R amygdala	25	21	-1	-11	AMYG	3.61	28	0.627
L superior frontal gyrus	9	-12	53	31	SFG	3.49	25	0.285
L posterior cingulate cortex	31	-6	-52	25	PCC	3.46	43	0.915

MNI coordinates of peaks of relative activation within regions responding to the main effects of task: postures vs. dots. Results were calculated at $p_{\text{uncorrected}} < 0.005$, k = 10 voxels. Up to 3 local maxima are listed when a cluster has multiple peaks more than 8mm apart. Abbreviations for brain regions: V4v = ventral visual area 4; S2 = sensorimotor area; PAC = primary auditory cortex; STG = superior temporal gyrus; M1 = primary motor cortex; MPFC = medial prefrontal cortex; ACC = anterior cingulate cortex; MPFCv = ventral medial prefrontal cortex; PCC = posterior cingulate cortex; PMd = dorsal premotor cortex; SMAd = dorsal supplemental motor area; MOG = middle occipital gyrus; MTG = middle temporal gyrus; STS = superior temporal sulcus; pSTG = posterior superior temporal gyrus; AMYG =

amygdala; IPL = inferior parietal lobule; SFG = superior frontal gyrus; IPC = inferior parietal cortex.

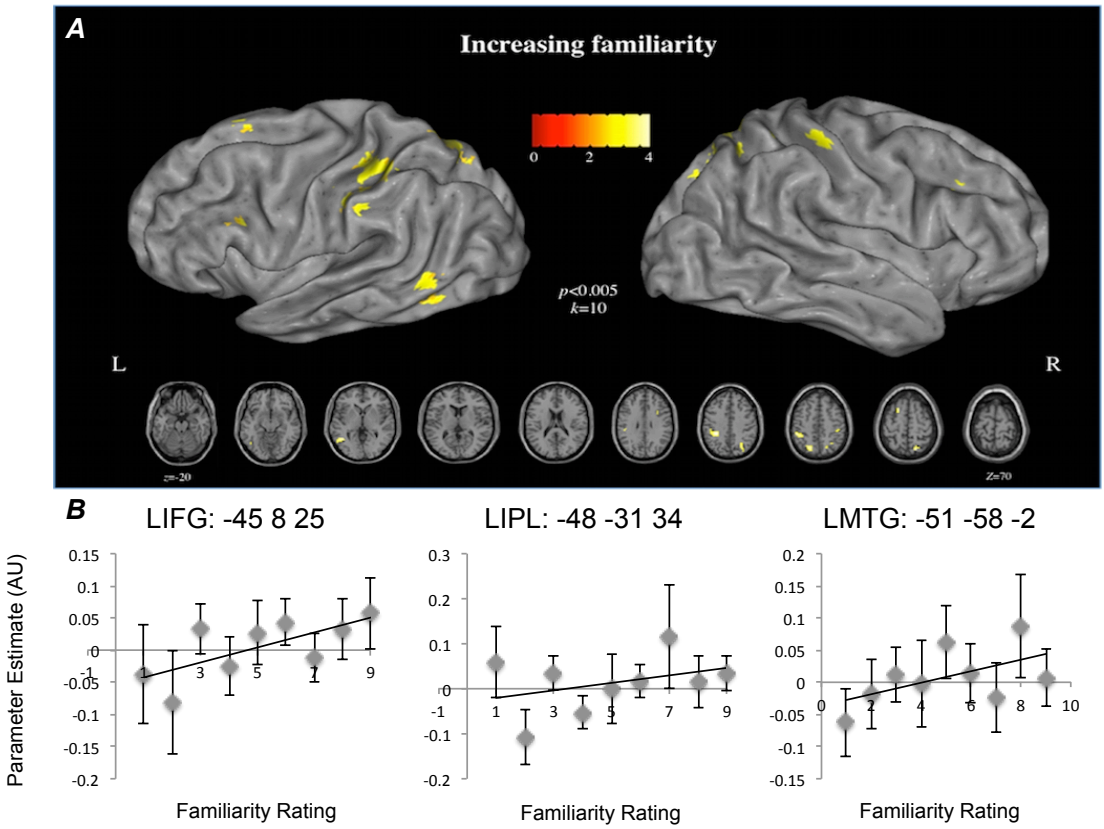


Figure 3.4. A The GLM results of the parametric rating contrast for the posture prediction task. All *p* values 0.005; *k* 10 voxels. ***B*** Parameter estimates for core AON regions plotted against familiarity rating. These show a linear relationship between famialirty rating and parameter estimate, error bars represent SEM.

Table 3.3. Localization of averaged BOLD response for observation of actions rated as increasingly and decreasingly familiar

Anatomical region	BA	MNI coordinates			Putative functional name	*t* value	Cluster size	$P_{corrected}$ value
		x	*y*	*z*				
Increasing familiarity								
Left SupraMarginal Gyrus	40	-48	-31	34	SmG	4.02	141	0.837
Left Inferior Parietal Lobule	7/40	-33	-37	40	IPL	3.99		0.837
Left Inferior Parietal	7/40	-48	-37	52	IPL	3.80		0.837

Lobule								
Left Middle Temporal Gyrus	39	-51	-58	-2	MTG	3.96	60	0.837
Left Superior Frontal Gyrus	6	-18	11	61	SFG	3.74	10	0.837
Left Superior Parietal Lobule	7	-21	-67	46	SPL	3.74	83	0.837
Left Superior Parietal Lobule	7	-15	-67	52	SPL	3.63		0.837
Right Superior Parietal Lobule	7	18	-67	61	SPL	3.68	96	0.837
Right Angular Gyrus	7/19	30	-70	37	AG	3.49		0.881
Right Angular Gyrus	7	27	-61	43	AG	3.27		0.897
Right Inferior Frontal Gyrus (p. Opercularis))	44	33	11	28	IFG	3.63	13	0.837
Right Postcentral Gyrus	2	48	-28	52	M1	3.41	25	0.881
Right Inferior Parietal Lobule	40	36	-43	55	IPL	3.02		0.993
Left Inferior Frontal Gyrus (p. Opercularis)	44	-45	8	25	IFG	3.24		
Decreasing familiarity								
Right Superior Occipital Gyrus	19	36	-76	4	SOG	4.88	105	0.157
Right Superior Occipital Gyrus	19	24	-82	16	SOG	3.76		
Right Superior Temporal Gyrus	22	48	-34	22	STG	4.68	16	0.556
Left Anterior Cingulate Gyrus	24	-3	20	28	AC	4.30	36	0.513
Cerebellum		0	-61	-2	CERR	4.26	106	0.157
Right Middle Cingulate Gyrus	32	9	-22	43	SPL	4.15	62	0.340
Right Middle Cingulate Gyrus	32	12	-31	46	SPL	3.53		
Right Middle Cingulate Gyrus	32	3	-10	40	MCG	3.43		
Right Cuneus	17	18	-73	34	CU	4.07	25	0.513
Left Insula	46	-39	-13	-2	INS	3.87	25	0.513
Left Lingual Gyrus	18	-12	-91	-11	V3	3.83	123	0.157
Left Lingual Gyrus	18	-21	-85	-2	V3	3.76		
Left Middle Occipital Gyrus	19	-36	-76	4	V5	3.59		
Left Cuneus	17	-15	-82	34		3.53	23	0.513
Left Middle Occipital Gyrus	18	-21	-94	10	V3v	3.37	14	0.556
Right Lingual Gyrus	18	12	-37	-8	CERR/ HIPP	3.06	14	0.556

MNI coordinates of peaks of relative activation within regions responding to the parametric effects of increasing familiarity. Results were calculated at $p_{uncorrected} < 0.005$, k = 10 voxels. Up to 3 local maxima are listed when a cluster has multiple peaks more than 8mm apart.

Abbreviations for brain regions: SmG = SupraMarginal Gyrus; IPL = inferior parietal lobule; MTG = middle temporal gyrus; SFG = superior frontal gyrus; SPL = superior parietal lobule; AG = angular gyrus; IFG = inferior frontal gyrus; M1 = primary motor cortex; SOG = Superior Occipital Gyrus; STG = Superior Temporal Gyrus; AC = Anterior Cingulate; CERR = Cerebellum; HIPP = Hippocampus; MCG = Middle Cingulate Gyrus; CU = Cuneus; INS = Insula.

3.4.1.2 Evaluated regional activations sensitive to increasing familiarity

The second random-effects analysis evaluated brain regions sensitive to the increasing familiarity of a complex whole-body movement (for all peak coordinates from this contrast, see *Table 3.3*). This contrast allows us to locate the regions of interest for the DCM analysis as well as assess general effects of perceived action familiarity based on participants' individual subjective ratings of each action stimulus. As can be seen in *Figure 3.4A*, the parametric contrast of increasing familiarity reveals activity within bilateral premotor, parietal cortices, and left temporal cortices. The reverse of this contrast, brain regions sensitive to decreasing familiarity, is also reported in *Table 3.3*. This did not reveal any regions of the core AON, with the main sensitivity seen in visual regions. The increasing familiarity contrast suggests that, as the movements are rated as increasingly familiar, the stronger the response becomes within core AON regions. This relationship is also shown in *Figure 3.4B*, whereby the relationship between parameter estimates and rating is linear. These results add support to the view that the more familiar an observed action is, the stronger activity is within the AON (Buccino et al., 2004; Calvo-Merino et al., 2005; Cross et al., 2006). The three nodes in the left hemisphere form the ROIs for the DCM analysis, which evaluates their effective connectivity during action observation and how increasing familiarity modulates connections within this triad of brain regions.

3.4.2 Effective connectivity analyses

The DCMs were created using the three ROIs listed in *Table 3.1*: MTG, IPL, and IFG (all in the left hemisphere). The input(s) into the system was movement, and the modulations were increasing familiarity. First, we identify where there is input into the system using a family comparison analysis. Next, we report the BMS analysis

of the models within the winning family. Finally, we used Bayesian model averaging (BMA) to present the parameter estimates of the winning family of models.

3.4.2.1 Family level inference

As described in *Materials and Methods*, we classified our 45 models into three families based on input into the system. To recap, the families were MTG input, IFG input, and input into both MTG and IFG. The results of the family comparison of inputs showed that there is unequivocally strong evidence for the family with input into both regions of the system. The exceedance probability of this family was (almost) 1, implying that there is a near certainty that this class of input explains the network. This identification of the inputs into the system allowed us to concentrate our model comparison on the 15 models with MTG and IFG input into the system.

3.4.2.2 BMS and BMA within optimal family

After assessing the existence of input parameters with family comparisons, we were then able to perform a BMS analysis on the 15 models within the MTG and IFG input family. As shown in *Figure 3.5A*, there is a clear optimal model, model 15.

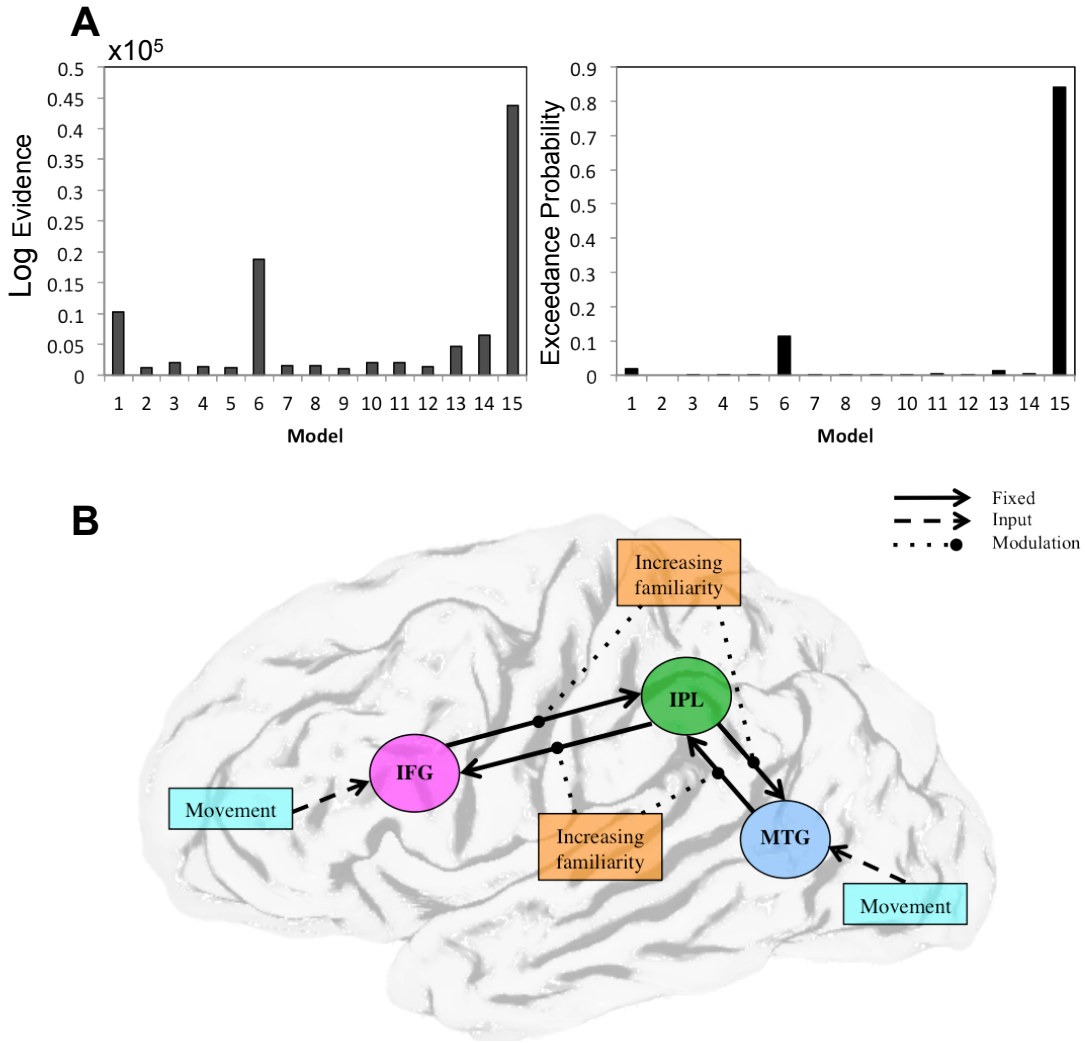


Figure 3.5. BMS on the 15 models in the MTG and IFG input family. ***A***, Log Evidence and Exceedance probabilities for all models. There is one clear winner, model 15, which had the posterior probability of 0.72. ***B***, Model 15, the winning model, had modulation of increasing familiarity on all fixed connections.

The optimal model was the model that had modulation by increasing familiarity on all fixed connections. The probability of this model was 0.72 (*Figure. 3.5B*), which is a high probability. The next best model, model 6, had an exceedance probability of 0.11. In relation to our hypothesis that there would be greater top- down modulation for increasing familiarity, a winning model that had fully modulated connections would neither confirm nor refute this hypothesis. Although the posterior probability of the winning model was 0.72, this does not provide definitive evidence for this and only this model. Therefore, we used BMA to obtain estimates of effective

connectivity (and their modulation) that accommodate uncertainty about models (Stephan et al., 2010). It is not always possible to find sufficient evidence for one model being the optimal model. Depending on which model is declared the best fitting, different inferences will be made regarding the behavior of a system. BMA helps to resolve this ambiguity by averaging over the models, accounting for the model evidence in the averaging; therefore, models with a low probability will not contribute very much to the BMA. This BMA also incorporates Bayesian parameter averaging over subjects to provide robust estimates of quantitative changes in coupling that are weighted by our relative confidence in the 15 models considered. The results of the BMA for the IFG and MTG input family are shown in *Figure 3.6*.

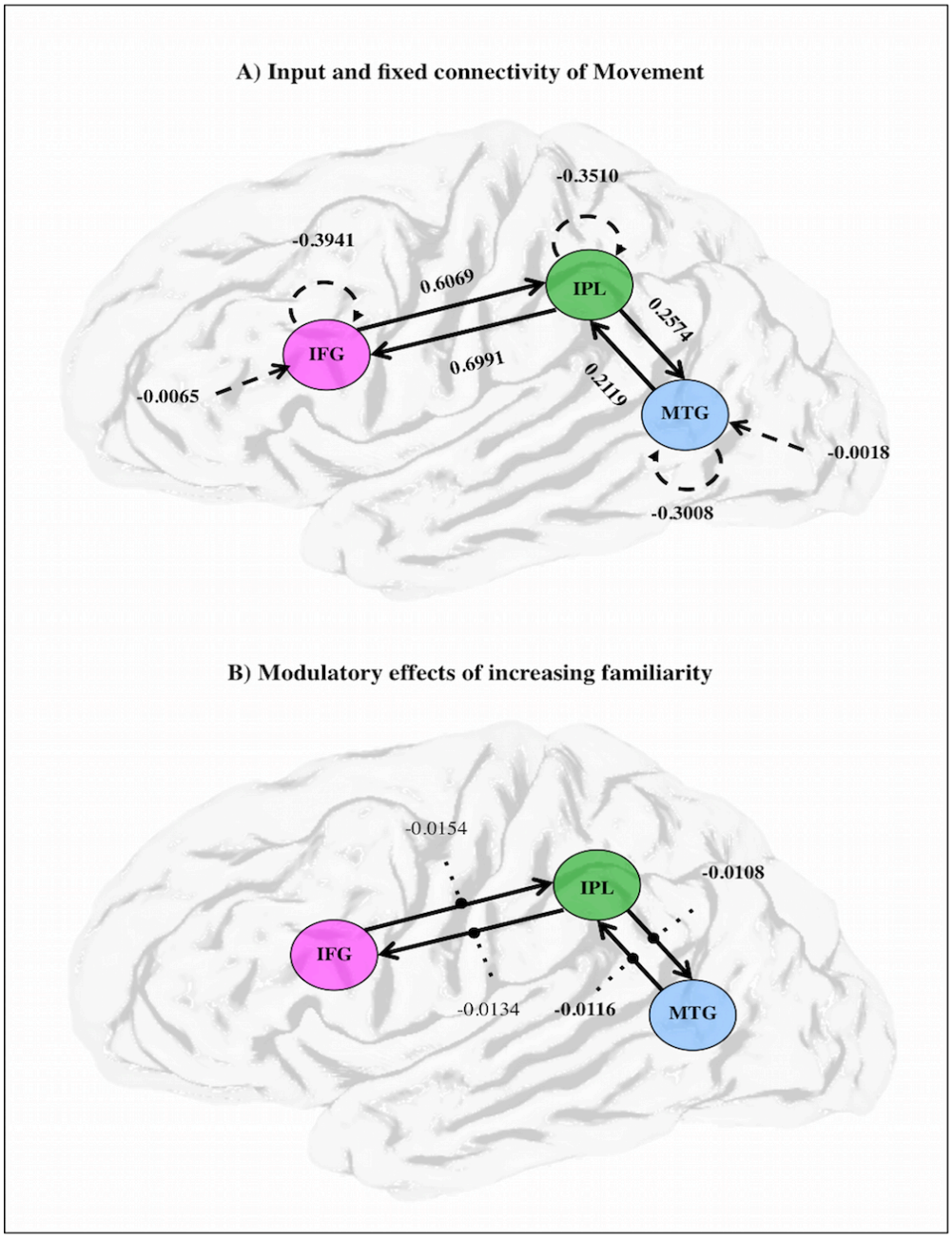


Figure 3.6. The resulting BMA of the winning model in *Figure 3.5*. ***A***, The input and fixed connectivity of the model, subject to the input of movement. The dashed semicircled arrows indicate self-connections. Solid arrows indicate the endogenous connectivity. Straight dashed lines indicate the input into the system. ***B***, The modulatory effects of increasing familiarity on the fixed connections. In all cases, the numerical value represents the connection strength/modulation in log scaling parameters.

Table 3.4. Results of a one sample *t* test for coupling parameters of endogenous activity of movement: endogenous connectivity

From			
	IFG	**IPL**	**MTG**
To			
IFG	-0.3941 (0.0267)* (p<0.0001)	0.6991 (0.0107)* (p<0.0001)	_
IPL	0.6069 (0.0108)* (p<0.0001)	-0.3510(0.0285)* (p<0.0001)	0.2119 (0.0141)* (p<0.0001)
MTG	_	0.2574(0.0100)* (p<0.0001)	-0.3008(0.0220)* (p<0.0001)

Data are the mean (SEM) for each connection. —, Not investigated. The threshold was set at p 0.007 (corresponding to an FDR-corrected threshold of p 0.05 for multiple comparisons).*Significant connection.

Table 3.5. Results of a one sample *t* test for coupling parameters of modulatory activity of movement: modulatory

From			
	IFG	**IPL**	**MTG**
To			
IFG	_	-0.013 (0.005) (p=0.017)	_
IPL	-0.608 (0.0520) (p=0.207)	_	-0.012 (0.003)* (p=0.001)
MTG	_	-0.011(0.004)* (p=0.005)	_

Data are the mean (SEM) for each connection. —, Not investigated. The threshold was set at p 0.0125 (corresponding to an FDR-corrected threshold of p 0.05 for multiple comparisons).*Significant connection.

Upon examination of the effects of movement on the network, the BMA results demonstrate that all the connections are significantly 0, supported by a one-sample *t* test, shown in *Table 3.4* (the use of *t* tests provides a way of scoring the standardized effects sizes in relation to intersubject variability). Beginning with the inputs into the network, the coupling strengths are relatively weak and inhibitory. The

reason for this may be that we allowed for stochastic effects, which means that the endogenous connectivity within the system is not reliant on a deterministic input into the network (Li et al., 2011). The inclusion of stochastic effects accounts for an unmodeled node modulating the input, meaning that the endogenous connectivity is not affected by weak inputs.

The endogenous connections between the nodes show us that the anatomical basis of our models was correct; movement reveals connections. The reciprocal connections between IPL and IFG appear stronger than the connections between MTG and IPL. It should be noted that the movement input contained videos that were rated as both familiar and unfamiliar, so these coupling parameters can be considered the average of the two polarities. Therefore, to understand these coupling strengths, we next investigated the modulatory effects of increasing familiarity.

The impact on the network's modulatory connectivity from increasing familiarity can be seen in *Figure 3.6B*. These modulatory connections were then placed into a one-sample *t* test that revealed that not all significantly differed from zero (*Table 3.5*). Only the modulatory influences between MTG-IPL and between IPL-MTG were significantly different from zero across individuals.

The modulatory effect of increasing familiarity on the fixed connection reveals several noteworthy findings described here and reconciled in the Discussion. The first is that, when a video is rated as increasingly familiar, the connection from MTG to IPL is attenuated. The sign change indicates that the bottom-up connection is attenuated by more familiar movements, thus supporting our hypothesis and assumptions made by a predictive coding model. The second finding is an attenuation of the connection from IPL to MTG. This finding was not predicted and, as such, does not clearly support our hypothesis, as we would have predicted significantly augmented modulation in this top-down direction. However, as discussed below, this finding nonetheless informs our understanding of a predictive coding account of

action perception.

3.5 Discussion

Our aim was to characterize how subjective familiarity of an observed action impacts AON engagement through complementary use of parametric GLM and effective connectivity analyses. We found that, as observed movements were rated as more familiar, BOLD signal in left MTG, IPL, and IFG increased. We then used DCM to explore effective connectivity between these regions to evaluate the hypothesis that, as movements are rated as increasingly familiar, the anterior influence (MTG ➔ IPL ➔ IFG), representing bottom-up, feedforward action coding, attenuates, whereas connectivity for posterior influence (IFG ➔ IPL ➔ MTG), representing top-down, feedback action coding, is up-regulated. DCM provided partial support of this hypothesis by demonstrating attenuated influence from connections leading to and originating from MTG. In the following, we consider these findings in terms of a predictive coding model of AON function and how they advance our understanding of the impact of familiarity on action perception.

3.5.1 The impact of familiarity on the AON and the "like me" hypothesis

The main GLM findings show that primary motor and AON regions are recruited during perception and prediction of movements and that increasingly familiar movements are associated with concomitant increases in activity. The first result corroborates research demonstrating sensorimotor engagement during tasks that explicitly tap into prediction processes (Blakemore & Frith, 2005; Falck-Ytter et al., 2006; Stadler et al., 2011; Cross et al., 2013). The parametric analyses reveal that increasingly familiar movements preferentially engage core AON regions, supporting the general premise of a "like me" hypothesis of AON function (Buccino et al., 2004; Calvo-Merino et al., 2005; Cross et al., 2006; Meltzoff, 2007; Liew et al., 2011; Press, 2011). The parametric contrast allowed us to examine the relationship between AON activity and familiarity in a nonbinary fashion, thus advancing knowledge from

previous studies. This approach subsequently enabled examination of causal connectivity between core AON regions to further characterize AON contributions to action understanding.

3.5.2 Effective connectivity

Our first effective connectivity aim was to create a feasible model space based on AON regions proposed previously (Keysers and Perrett, 2004). By tailoring this model space, we could address where movement inputted to the system and examine the models within this family (Penny et al., 2010). We showed inputs to both IFG and MTG in a reciprocally connected network. Next, we sought to identify the optimal model of modulation/modulatory activity (B matrix) based on these fixed connections. The BMS results revealed the optimal model as one where increasing familiarity modulates all fixed connections. By use of BMA, we found that increasing movement familiarity attenuates anterior influence between MTG and IPL while also attenuating posterior influence between IPL and MTG.

The parameter estimates (from the BMA) suggest that increasing familiarity causes diffuse and small reductions in effective connectivity among the three areas studied. The small effect sizes deserve some comment; in this analysis, we used stochastic DCM. In stochastic DCM, one estimates both the neuronal activity and effective connectivity that best explain observed responses. Generally, this provides smaller estimates of changes in coupling because condition-specific effects can also be modeled by differences in neuronal activity. Although small, our Bayesian model comparison suggests that familiarity-related changes in effective connectivity are evident in the data. The effect sizes are expressed in terms of log-scale parameters and can be interpreted (approximately) as proportional changes. In other words, familiarity induces a change of 1% of the underlying connectivity strengths. Furthermore, as predicted, ascending connections to IPL show a familiarity-related decrease greater in magnitude than the decrease in descending connections from IPL.

3.5.3 Predictive coding models revisited

Our findings support several assumptions made by models of action perception that rely on notions of predictive coding (Keysers & Perrett, 2004; Kilner et al., 2007a,b). These models suggest greater influence from visual to motor regions when an observed movement is unfamiliar. Such activity is hypothesized to be indicative of action representations being built in a perceptually driven, bottom-up manner. When an observed movement is familiar, top-down predictions should have a greater influence on perceptual processing at lower levels. These predictions are formally identical to corollary discharge (i.e., the expected consequences of an intended movement). This connects predictive coding accounts of action perception to machine learning accounts of motor control (Wolpert et al., 2003) that appeal to notions of corollary discharge. In the predictive coding account of action observation (Kilner et al., 2007a,b), the more familiar a movement, the more precise top down predictions become. Effectively, increasing familiarity should increase influence of top down predictions relative to bottom up information (which, in this model, originates in MTG). We therefore anticipated that more familiar actions should be associated with increasing influence from connections spanning posteriorly from IFG to IPL and IPL to MTG, and decreasing influence from connections spanning anteriorly from MTG to IPL and IPL to IFG.

DCM analyses partially support these hypotheses. We show that, with increasing familiarity, MTG exerts an attenuated influence on IPL in an anterior direction, as hypothesized. A similar dampening of influence from IPL to IFG with increasing familiarity is also observed, although this finding did not survive corrections for multiple comparisons. Our hypotheses concerning the influence of familiarity on posterior-projecting connections were not supported. Increasing action familiarity did not significantly up-regulate influence from IFG to IPL, as was expected. One possible explanation for this could be that the influence IFG exerts on IPL is impervious to variations in familiarity examined in the present study. Most crucially, and contrary to our original hypothesis, a significant attenuation of influence was observed in posterior connections from IPL to MTG.

At first glance, this finding might appear to contradict a predictive coding account of action observation. However, our initial hypotheses might have been overly simplistic, as they did not take into account the range of excitatory and inhibitory influences between nodes within a predictive coding framework (Keysers & Gazzola, 2014). In keeping with our original hypotheses, the reason why effective connectivity between MTG and IPL in both directions attenuates with increasing familiarity could be that increasingly accurate predictions decrease demand on incoming or ongoing perceptual processing. As such, the present findings illustrate a version of predictive coding in which outputs from, and inputs to, MTG attenuate with increasing familiarity, whereas reciprocal influence between parietal and premotor areas is less impacted by increasing action familiarity. To an extent, these findings support a predictive coding framework of the AON in a nuanced and subtle manner. The general idea that increased familiarity is associated with decreased prediction error, and a dampening of sensory inputs is supported by our data. However, the present findings do not yield clear support for more specific predictions concerning the concurrent upregulation of posterior-projecting connections (IFG to IPL and IPL to MTG) and downregulation of anterior-projecting connections (MTG to IPL and IPL to MTG).

An alternative framing of these DCM results is that, with increasing familiarity, localized representations become sufficient to perform the task at hand, resulting in decreased cross-node coordination within the AON (as this coordination becomes less necessary). Such an interpretation would be consistent with the univariate findings reported in the present study and others (Buccino et al., 2004; Calvo-Merino et al., 2005; Cross et al., 2006; Vogt et al., 2007; Shimada, 2010; Liew et al., 2011), which show increased activity within the individual AON nodes during familiar action observation. This interpretation remains speculative at this stage. To test this possible explanation, follow-up research could investigate observation of actions that cover a broader range of familiarity to determine whether the degree of familiarity drives the connection strengths between AON nodes, such that observation of highly unfamiliar actions is associated with stronger AON connections, whereas

observation of actions rated as highly familiar is associated with attenuated connections. Were this alternative theoretical framing to be supported by follow-up work, it would call for an updating and reconsideration of a predictive coding account of action perception.

3.6 Conclusions

In conclusion, the present findings provide partial support for a predictive coding account of action observation, as well as offer novel insights into how familiarity modulates effective connectivity within the AON. By using participants' individual, subjective ratings of familiarity, we show greater AON activity when movements were perceived as increasingly familiar. Through the use of effective connectivity analyses, we demonstrate attenuation of reciprocal connections between IPL and MTG with increasing familiarity and presumed decreasing prediction error. Although this finding is broadly in line with a predictive coding account of action perception (Kilner et al., 2007b), the data does not provide clear or complete support for this theoretical model for how familiarity should impact influence between AON nodes. An alternative possibility is that increased familiarity of an observed action results in a shift away from coordinated activity between AON nodes and toward more individuated processing within individual nodes. A challenge for future work is to further explore how familiarity modulates influence between AON regions. Of particular importance for future investigation is how different levels of familiarity impact parietal and premotor regions, as our findings less clearly demonstrate marked changes in effective connectivity between these two regions with increasing familiarity. As a whole, the present study demonstrates the utility of effective connectivity analyses to explore simulationist accounts of social information processing.

4 Chapter IV

Probing the Action Observation Network response to varying levels of familiarity

4.1 Abstract

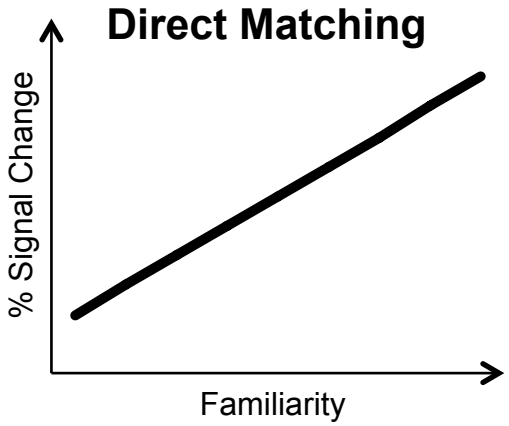
Watching other people move elicits engagement of a collection of sensorimotor brain regions collectively termed the Action Observation Network (AON). A large body of previous literature suggests that the AON responds more robustly when observing familiar compared to unfamiliar actions, and that the amplitude of AON response positively correlates with an observer's familiarity with an observed movement. On the other hand, a growing number of studies document patterns of AON activity counter to these findings, whereby in some circumstances, unfamiliar actions lead to more AON engagement than familiar actions. In an attempt to reconcile these conflicting findings, some have proposed that the relationship between AON response amplitude and observed action familiarity is nonlinear, or quadratic, in nature. In the present study, we used an elaborate training intervention to probe the relationship between movement familiarity and AON engagement during action execution and action observation tasks. Participants underwent fMRI scanning while executing one set of guitar sequences and observing a second set of sequences. Participants then acquired further physical practice or observational experience with half these stimuli outside the scanner across 3 days. Participants then returned for a second scanning session identical to the first, wherein they executed and observed an equal number of familiar (trained) and unfamiliar (untrained) guitar sequences. Via region of interest analyses, we extracted activity within AON regions engaged during both scanning sessions, and then fit linear, quadratic and cubic regression models to these data. The data show evidence for linear and quadratic models, suggesting that the response profile within key sensorimotor brain regions associated with the AON is not solely linear in nature in response to familiarity. Moreover, by probing the subjective nature

of the prediction error signal, we show results that are consistent with a predictive coding account of AON engagement during action observation and execution.

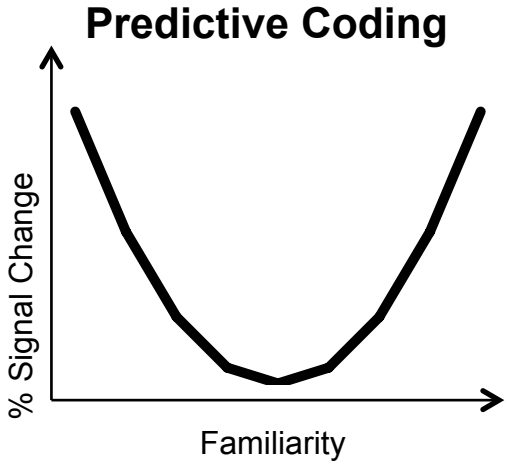
4.2 Introduction

Watching others in action provides important information about other people's goals, intentions and desires. When we observe others moving around us, we can predict how their current and future actions might unfold, thus enabling us to respond appropriately to people we encounter in a social world (Blakemore & Frith, 2005). The observation of actions elicits activity in a network of sensorimotor brain regions collectively termed the Action Observation Network (AON; Cross et al., 2009; Grafton, 2009; Keysers & Gazzola, 2009; Caspers et al., 2010). The core brain regions that compose the AON include occipitotemporal cortical regions associated with observing bodies in motion, as well as the premotor cortex and inferior parietal lobule. These latter two brain regions have been shown to contain so-called mirror neurons in the non-human primate brain (Di Pellegrino et al., 1992; Gallese et al., 1996; Rizzolatti et al., 2001; Umilta et al., 2001), and demonstrate a similar response profile of engagement to observed and executed actions in the human brain (Gazzola & Keysers, 2009; for a review see Molenberghs, 2012). Previous literature demonstrates that the more familiar an observed action is, the stronger the response is within these core AON regions (Buccino et al., 2004; Calvo-Merino et al., 2005; Cross et al., 2006; Vogt et al., 2007; Shimada, 2010). Moreover, we recently demonstrated that complex whole body movements that participants rated as more familiar were associated with greater AON activity compared to those movements rated as less familiar (Gardner et al., 2015). These magnitude-based approaches support experience-driven simulation accounts of action perception (Sinigaglia, 2013), which form the foundation of the direct matching hypothesis of action understanding (Rizzolatti et al., 2001; Gallese & Goldman, 1998; Wolpert et al., 2003; see Csibra, 2005 for a debate on this account). This account of action understanding suggests that an action's meaning is understood via the AON supporting simulation of an observed action, by matching the goal or intention of what that movement would be if performed by the observer. In terms of familiarity, a linear relationship between magnitude of AON activity and familiarity would be consistent with this hypothesis: as familiarity increases, the simulation of how an

action might unfold over time becomes more accurate and resonance between an observer's motor system and an observed action is maximised. This relationship can be seen in *Figure 4.1**A***.



(Rizzolatti et al., 2001; Gallese & Goldman, 1998; Wolpert et al., 2003)



(Cross et al., 2012; Liew et al., 2013; Keysers & Perrett, 2004; Kilner et al, 2007a)

Figure 4.1. The hypothesised relationship between familiarity and % signal change (BOLD signal) for both *A* direct matching and *B* predictive coding.

On the other hand, an increasing number of studies are reporting findings demonstrating that AON activity does not necessarily follow this linear trend of increasing engagement with increasing familiarity (Gazzola et al., 2007; Liew et al., 2013; Cross et al., 2012). These studies demonstrate greater AON activity when participants observe actions that are unfamiliar (compared to more familiar actions), a

finding that appears to be at odds with a simulation-based account of AON function at first glance. The findings from these studies suggest that a linear relationship between AON activity and familiarity is likely too simplistic. In terms of the direct matching hypothesis, this theory would struggle to explain why an *unfamiliar* action that is *not* in the observer's repertoire would elicit greater AON activity (*Figure 4.1**A***). Predictive coding models of AON function (Keysers & Perrett, 2004; Kilner et al., 2007a,b; Gazzola & Keysers, 2009; Schippers & Keysers, 2011), predicated on the use of perceptuomotor maps to predict and interpret observed actions (Lamm et al., 2007; Schubotz, 2007; Urgesi et al., 2010) may potentially help to resolve these seemingly discrepant findings concerning the relationship between familiarity of an observed movement and engagement of sensorimotor cortices. This framework proposes a Bayesian comparison of predicted and observed actions, creating a reciprocally modulated network comprising premotor, inferior parietal and posterior temporal cortices. This network aims to minimise the differences between observed and predicted actions. When observing a less familiar action, predictions (feedback) are lacking or are under informed, and thus do not match the observed actions (feedforward), which equates to high prediction error. This should manifest as increased AON engagement for highly unfamiliar actions, as the influence of feedforward/perceptual activity is heavily relied upon. When viewing an action that is familiar, however, the predictions generated by the network are more precise, and thus minimises prediction error. The minimising of prediction error can also manifest as increased AON engagement, as predictions projected posteriorly are stronger than when prediction error is higher. The reciprocal nature of this framework allows for the explanation of higher AON engagement for both familiar and unfamiliar actions (illustrated in *Figure 4.1**B***). One possibility that arises from this explanation is that this predictive framework is adaptive and dynamic. During the trough of the quadratic function (moderate familiarity), the AON would be both successfully and unsuccessfully attempting to minimise prediction error while it attempts to understand and predict the action, thus a smooth curve would not best explain this relationship. A cubic function would better fit this potential explanation, while still supporting the predictive coding account. To our knowledge, these two theoretical framings of how the AON is modulated by familiarity have not yet been directly compared with empirical evidence.

In the current study, we aimed to address two distinct questions relating to the relationship between familiarity and AON engagement. First, we aimed to compare the direct matching hypothesis (linear model) with the predictive coding account (non linear or quadratic model), in terms of which model of AON engagement best explains the impact of varying levels of familiarity. To test whether the response of the AON to varying levels of familiarity is either linear or quadratic, we combined an intensive training intervention, pre- and post-training fMRI scans, and a region of interest-led analytical approach (similar in methodology to that of Mattavelli et al., 2012). Our task involved two types of action-related task: action observation and action execution. In the observation condition, participants observed an expert musician playing short musical sequences on a bass guitar, after which participants responded to an attentional control question. In the execution condition, the participants played a different set of short musical sequences in the scanner on a scanner-safe bass guitar. The use of a training paradigm enabled us to establish a clear distinction between familiar and unfamiliar stimuli, and the use of execution and observation conditions facilitates closer comparison and scrutiny of how two kinds of experience shape AON responses. The regions of interest were identified from an action observation and action execution vs. implicit baseline contrast. For both conditions, these regions were taken from both days of scanning, for all blocks, to which a linear regression model, a quadratic regression model and a cubic regression model were fitted for each region, for each run and for every participant. According to the direct matching hypothesis (Rizzolatti et al., 2001; Gallese & Goldman, 1998; Wolpert et al., 2003), we should see that when the blocks were ordered by familiarity, the response in these regions should increase linearly in magnitude (*Figure 4.1**A***). Alternatively, a quadratic or cubic relationship between familiarity and BOLD signal would provide support for the predictive coding account (*Figure 4.1**B***).

Our second aim was to evaluate the internal consistency of our findings. We recently demonstrated that ratings of familiarity provide a sensitive measure of familiarity (Gardner et al., 2015). Therefore, we took ratings of familiarity by each participant and used these as the independent variable in the regression models, as an

alternative to simply the number of times participants were exposed to each musical sequence. This approach enabled us to establish the extent to which findings from our first approach (using objective measures of exposure/familiarity) are replicated via subjective ratings of familiarity reported by each participant for each musical sequence.

4.3 Materials and methods

4.3.1 Participants

Twenty-two healthy young adult volunteers recruited from the local University community took part in the experiment and received £20 in exchange for their time. Two volunteers were excluded from the final sample due to excessive head motion during scanning. The final sample comprised of 20 volunteers (9 males, *M_age* = 20.60 years, *SD* = 1.73). All participants had normal or corrected to normal vision with no history of neurological illness. All participants were right-handed and required to play a right-handed guitar in the scanner. Participants were brought to the lab prior to scanning to ensure that they could play the instrument in the manner required inside the scanner bore, also to ensure that all were guitar novices. The study was approved and conducted following the guidelines of the Ethics Committee of the School of Psychology at Bangor University and the Bangor Imaging Unit. All participants provided written informed consent prior to their participation.

4.3.2 Stimuli & Apparatus

4.3.2.1 Sequences

We chose 16 sequences from the computer game Rocksmith® (Ubisoft, 2014), lasting an average of 15.8s (*SD* = 2.37s). These sequences were excerpts taken from songs initially chosen due to their lack of lyrical content. This restriction was selected so participants would not associate any particular action sequence with lyrics.

The sequences were also matched on the difficulty level assigned by Rocksmith®. Rocksmith® assigns difficulty level via an algorithm that assesses song speed and the number of notes to be played within the time window (the difficulty level is visible at the top of *Figure 4.2*). To ensure that difficulty levels were matched as closely as possible across all stimuli, the number of notes to be played and the length of the sequences were matched, as were beats per minute for the individual song excerpts. In addition, the inclusion criteria for these sequences required that notes fell within a fret range of 1-7 and string range of 1-3. This restriction ensured that participants would not have to move their heads to identify frets during scanning, while at the same time maintaining a level of difficulty that would challenge participants throughout the training period. Furthermore, we also excluded technical guitar playing movements such as "hammering" and "sliding" so that the actions required would be accessible to novices. Finally, we matched the sequences on mean amount of motion energy displayed in each video (see Bobick, 1997; Schippers et al., 2010; Cross et al., 2012), to ensure gross differences in the amount of motion displayed in individual sequence videos did not contribute to basic visual differences between stimuli or training conditions.

4.3.2.2 Rocksmith® Guitar Task

Shown in *Figure 4.2* is a screenshot of the gameplay from Rocksmith® for the sequences that were physically executed (Panel A) and those that were observed (Panel B). The horizontal coloured lines at the foot of the screen correspond to the strings of the guitar being played. We used a bass guitar for this study as it has fewer strings than a standard electrical guitar (4 vs. 6) and at a basic level is generally considered easier to learn to play. These coloured strings are illustrated from a first person view, as if one is looking through the back of the fretboard of the guitar (i.e., when holding the guitar, the red string corresponds to the top string, green the second string from the top, and so on).

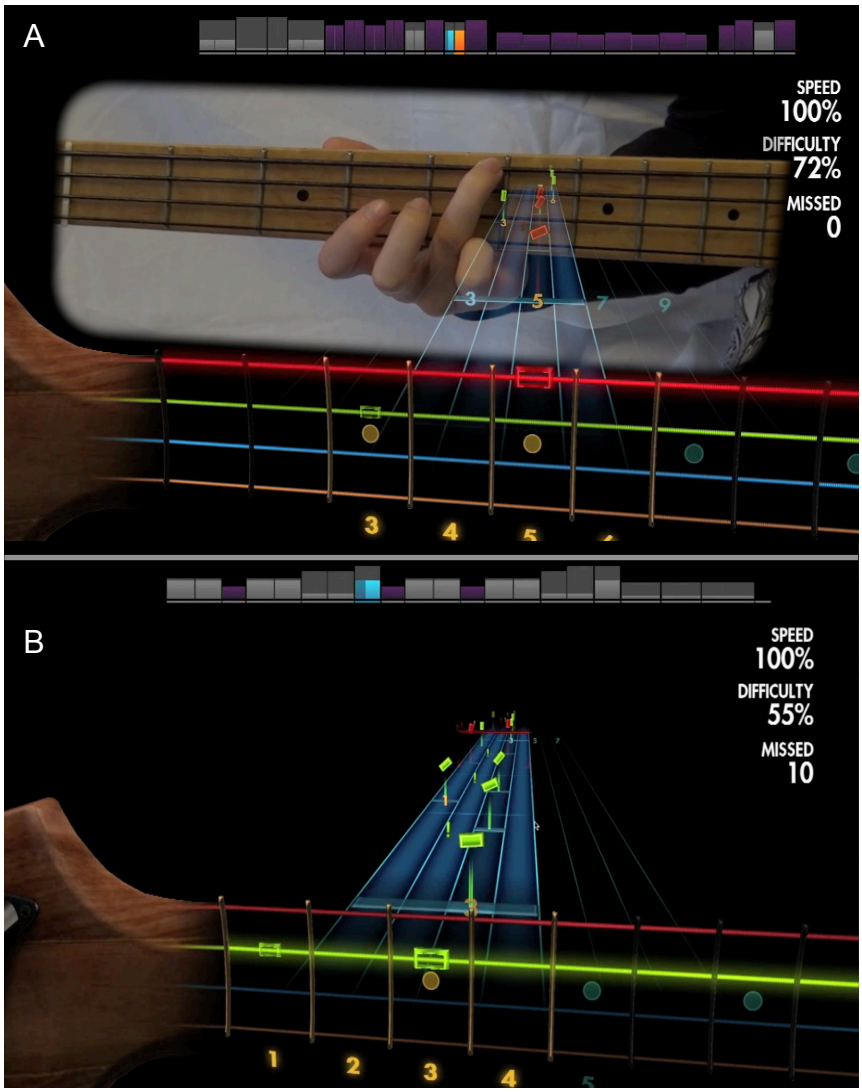


Figure 4.2. A still video frame example of the stimuli used in the observation condition ***A*** and execution condition ***B***. The hand and fretboard of the musician playing the guitar in the centre of ***A*** is superimposed on the actual game play.

The second aspect of the gameplay to be explained is the translucent blue "conveyor belt" of notes seen in the centre of the screen. The numbers on this conveyor correspond to the fret number on the guitar itself. The coloured rectangles are critical for the participant to attend to in order to make the correct movement. The colour of the rectangle corresponds to the string to be played with the right hand, and the number fret it appears on corresponds to which fret should be pressed down with the left hand. As the rectangles and numbers on the conveyor move towards the fore of the screen, the rectangles on the conveyor rotate 90 degrees from vertical to horizontal. When they reach the horizontal position, they contact the strings at the

bottom of the screen, and this is when the participant must play the appropriate note. If correctly executed, the rectangle illuminates slightly but if missed, the word “miss” appears on the string and fret that should have been played and the counter located at the upper right corner of the screen is adjusted accordingly. To obtain a perfect score, the correct fret on the correct string had to be plucked within a ± 250ms window of the onset of the note.

4.3.2.3 Observation Task

Figure 4.3 shows a schematic of the observation task. Participants first viewed a brief fixation cross, followed by a video of an expert musician playing one song excerpt. The musician performs all of the songs without fault, thus providing a precise template for participants to observe.

Figure 4.3. A schematic of one observation trial. The task was to attend to the hand and respond to whether there was a palming of the strings.

In order to ensure that participants paid close attention to the actions performed by the expert musician in the observation condition, an attentional control task was implemented whereby participants had to identify whether the musician palmed the strings during each stimulus video. A palming of the strings occurs when the musician removes her fingers from the fretboard, extends them vertically, and places the palm of her hand over the frets. This action was performed quickly as the musician then immediately continued to play the correct notes without stopping and without error. At the end of each trial, participants were asked whether a palming action was seen in the last sequence, and were required to make their response by plucking one of two strings with the right index finger. To ensure there was no confusion about what was required of participants during the observation task, the concept of palming the strings was explained and demonstrated before the experiment began.

4.3.2.4 Execution Task

In the execution condition, *Figure 4.2 (Panel B)*, participants were instructed to play along with the song to the best of their ability. The instruction was also given for participants to simply move on to the next note if they missed one note, to ensure participants did not move excessively when trying to compensate for an error. After each sequence was played, the gameplay presented a count of how many notes were missed, providing feedback on performance.

4.3.3 Training procedure

The study began with all participants taking part in an fMRI scanning session (the first column of *Figure 4.4*), wherein they observed eight sequences and executed eight different sequences. Participants played or watched each of the eight sequences from the different training conditions twice per run, and completed two runs of both kinds of training on each day of scanning. All the stimuli on the first day are labelled unfamiliar, as they were all novel to the participants. The order of the stimuli, which set was observed and executed, and the order of conditions were counterbalanced across participants.

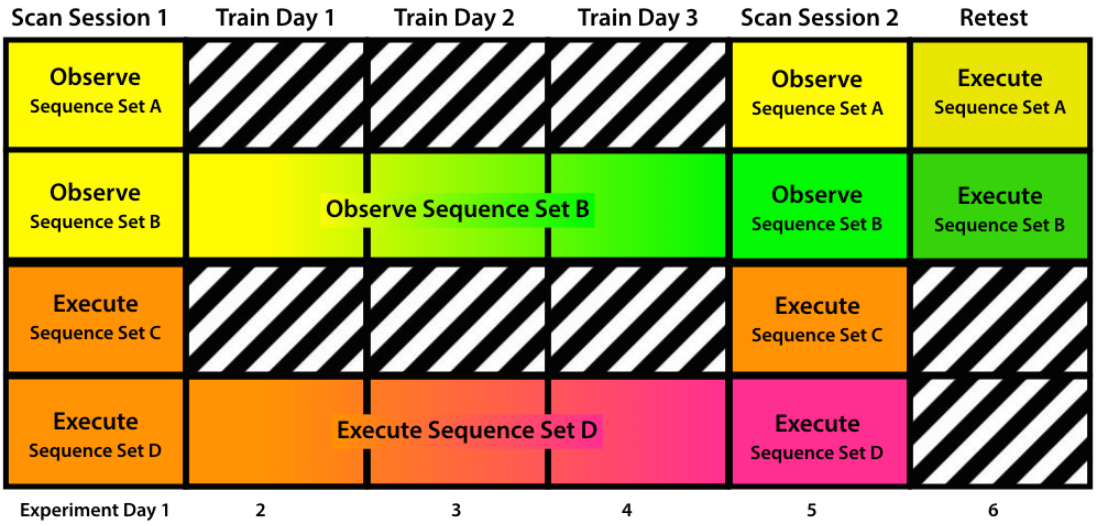


Figure 4.4. The timeline of both training and test days for both conditions. The gradient of colour illustrates the change in familiarity due to training. The Retest (far right column) always occurred after Scan Session 2, although not the day immediately after (between 14 and 21 days post Scan Session 2).

Days 2 through 4 of testing were the training phase (green boxes in *Figure 4.4*). During these days, participants were invited to the lab and asked to perform the same tasks as those they completed in the scanner, on precisely half of the execution sequences and half of the observation sequences. The four execution sequences and four observation sequences were performed or observed four times per training day, and the order of practice was counterbalanced across condition and sequences. The training set up in the laboratory was designed to match that of the scanner as closely as possible. Due to the fragile nature of the scanner safe guitar, a standard 4-stringed bass guitar was used in the training sessions to eliminate risk of damage to the equipment. Participants were required to lay on a table with the guitar placed over their midriff – similar to how the guitar was positioned in the scanner, and participants viewed the 24-inch iMac screen through prism glasses (this can be seen in *Figure 4.5*). Stimulus presentation and response collection were performed using Psychophysics Toolbox (v3) via MATLAB R2015a (MathWorks). The instruction was also given to keep movement to a minimum (for example, no tapping the guitar to the beat of the song), and the researchers monitored this. During the observation condition, participants were asked to rest their left hand over the frets so that there was no possibility that they could move their hand along with the musician's,

ensuring that any learning was due to observation alone. During the execution condition, participants were asked to play the songs as accurately as they could.

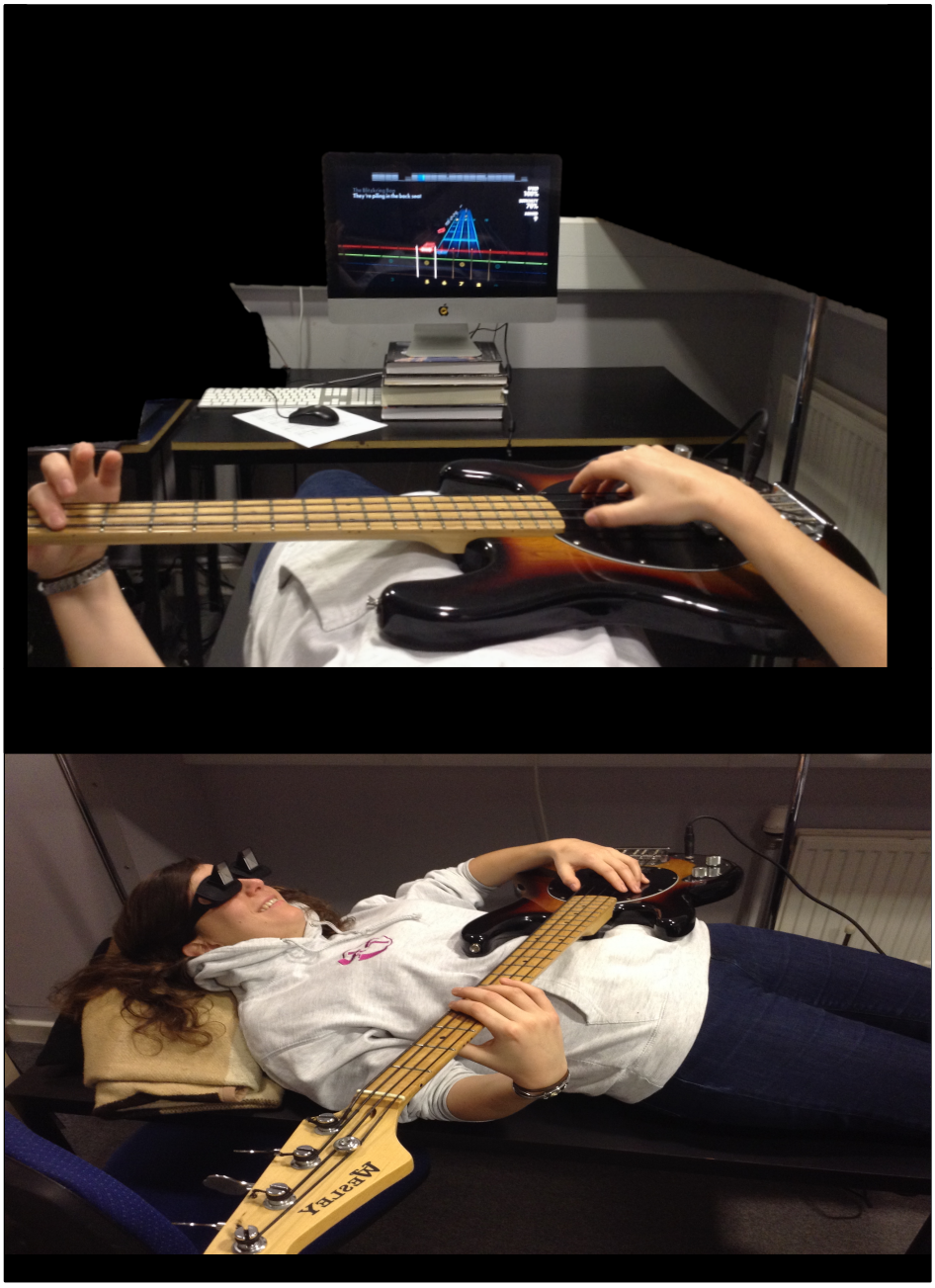


Figure 4.5. The setup for the training period. This setup was designed so that it matched the fMRI setup as accurately as possible. *N.B*. The guitar was held like a double bass, with the neck of the guitar at 30 degrees, the guitar is held at 90 degrees above for illustrative purposes only.

On the fifth day of testing, participants returned to the scanner where they completed an identical scanning protocol to Day 1, performing or observing both

trained and untrained sequences. Following the second scanning session, participants were invited back to the laboratory for a sixth and final day of testing, wherein they were asked to perform all the songs from the observation condition (including the four observationally trained and the four untrained songs; Retest, *Figure 4.4*). Due to scheduling complications, not all participants were able to attend this final testing session, and we thus report data from this final day of testing for the 15 participants who were able to return for this final session.

4.3.4 Familiarity rating

After the second scanning session, participants came out of the scanner and performed a rating task on the stimuli. Participants were asked to observe videos of the expert guitarist playing each sequence and to rate on a Likert scale of 1-9 on how familiar they were with each sequence (with anchors 1 = highly unfamiliar and 9 = highly familiar; the identical scale to that used by Gardner et al., 2015). Participants were asked to use the whole scale and to respond as quickly as they could.

4.3.5 Neuroimaging procedure & parameters

Each participant completed two identical fMRI sessions on days 1 and 5 of the experiment that followed an event related design. Each scanning session featured two action execution runs and two action observation runs, presented in a counterbalanced order across participants and across scanning sessions. All eight excerpts were experienced twice per run. Each observation run lasted an average of 5 minutes (range = 5.07 - 5.70 minutes), and each action execution run lasted 11 minutes (range = 10.97 - 11.80 minutes). This difference occurred due to the buffering varying lengths of the different sequences and loading times, two factors that were not modelled within a trial nor used in matching sequence length. For the observation trials (shown in *Figure 4.3*), at the start a fixation cross was shown for 1.8s. This was followed by the video clip of the agent playing along with the excerpt (audio was included here and during the training period). After each clip, a black screen was presented with the question “did the musician seen in the video make a palming action over the strings?” The question screen was displayed for 2s before moving on to the next trial.

Participants were required to respond to the question within that 2s window by plucking the appropriate string to denote their answer. During the action execution runs, participants first saw a brief interval where the song was being loaded. This aspect was unavoidable as we wanted to gain actual response accuracy via the game, so had to load each song as if it were selected by the user (the transition between menu and sequence to play was automated via a MATLAB script, and the entire load time before each sequence ranged between 19.02 and 41.45 seconds). Once the appropriate sequence was selected, there was a buffer supplied via the game so that there was an adequate amount of time before the participant began performing the sequence, allowing for finger position adjustment before each execution sequence began. After playing along with each sequence, participants' accuracy scores were displayed for participants to see before returning to the menu screen to begin the next trial.

Stimulus presentation and response collection were performed using Psychophysics Toolbox (version 3) using MATLAB R2015a (MathWorks) run via a MacBook Pro laptop computer. The stimuli were presented on a 24-inch LCD BOLDScreen (Cambridge Research Systems), which was visible to the participant via a mirror mounted on the head coil. Participants listened to the song excerpts through Phillips MR-compatible headphones.

Participants were given a MR-compatible bass guitar to make their responses during the execution and observation runs. The guitar was a full-length bass guitar, which presented some challenges for participants to manage whilst in the scanner. Participants were positioned into the scanner bore slowly and shown the best way to hold the bass guitar so as not to damage the guitar or the head coil. The guitar worked via a piezo-pickup embedded in the head of the guitar, under the strings (which were made of nylon). The guitar's tuning pegs were manufactured via 3D printing using a glass/plastic alloy. The output of the guitar was passed along a fibre optic cable from the scanner room to the control room where it was amplified and fed into the MacBook Pro running MATLAB. Offline, the responses for the observation condition

were filtered to remove any RF interference created by the scanner. The gameplay applied filtering for the execution condition so that the note being played could be heard by the participant.

Data acquisition was conducted at the Bangor Imaging Unit at Bangor University, Wales. Functional images were acquired on a 3.0T Phillips MRI scanner using a SENSE phased-array 32-channel head coil. Functional images were acquired covering the whole brain using an echo-planar imaging (EPI) sequence (35 axial slices, ascending slice acquisition, repetition time = 2000 ms, echo time = 30ms, 90° flip angle, matrix = 64 × 64, slice thickness: 3 × 3 × 3 mm, field of view (FOV): 224 mm). Before the functional run, 196 two-dimensional anatomical images (256 × 256 pixel matrix, T1-weighted) were obtained for normalization ROI selection and manipulation

4.3.6 fMRI data analysis

The total number of functional scans collected for the observation runs ranged between 156 and 178, and between 316 and 340 scans for each execution run. The number of scans for each subject was identical across scanning days. The data was analyzed using Statistical Parametric Mapping (SPM12: Wellcome Trust Centre for Neuroimaging, London; Friston, 2007) and implemented using MATLAB R2015a (MathWorks). The data was first realigned and then slice-time corrected and preliminarily preorientated within standard stereotaxic space as defined by the MNI (Friston, 2007). This preorientation allowed for better spatial normalization to the MNI template. Participants' EPI images were then coregistered to their T1 anatomical scans, which were then spatially normalized to standard stereotaxic space. The spatially normalized EPI images were filtered using a Gaussian kernel of 8 mm full-width at half maximum in the *x*, *y*, and *z* axes. For the observational runs, the design matrix was fitted for each subject with a single regressor for the familiar stimuli, a single regressor for the unfamiliar stimuli and a single regressor for the fixation and responses. For the execution runs, this setup was the same with the inclusion of a

loading period rather than the fixation and response regressor. The 4 runs (2 observation, 2 execution) were placed into the same design matrix, enabling us to create a total of two design matrices per participant; one pre- and one post-training. A whole brain analysis was performed for all analyses, $p < 0.001$, $k = 10$. Only clusters that survived FWE correction were considered for further ROI analysis. All brain regions that emerged from analyses were identified via the Anatomy Toolbox (Eickhoff et al., 2007).

The main neuroimaging analyses were designed to achieve two distinct objectives:

4.3.7 Imaging Objective 1: Evaluate shape of regression function within ROIs based on objective measure of familiarity

The first imaging objective was designed to compare the direct matching hypothesis with the predictive coding account, in terms of which model of AON engagement encapsulates varying levels of familiarity best for our task. The steps in this process involved first identifying ROIs and then fitting the regression models.

4.3.7.1 Identification of ROIs

ROIs were identified and extracted from the final observation and execution runs from the post-training scan session. The contrast used to identify these ROIs was observed and executed sequences > implicit baseline. This enabled us to identify regions that were active when viewing both familiar and unfamiliar actions (as the final run contained trained and untrained sequences for both the observation and execution conditions), as well as active for both action observation and execution. The clusters were identified with a threshold of $p < 0.001$ (uncorrected) and $k = 10$ voxels at the group level. The same threshold was then used to identify regions for each participant for each condition and for each run of the remaining three runs (two runs from the pre-training session and the first run from the post-training session; split by familiarity, creating six runs). The MARSBAR toolbox (Brett et al., 2002) was

then used to extract the time series for each region and for each subject, which was then transformed into percent signal change. ROI analyses for the observation condition and execution condition were performed independently. The resulting percent signal change values for each run and condition were then used to address the main hypotheses in the study.

4.3.7.2 Fitting regression models across varying levels of familiarity

The first aim of this study was to address whether the AON's response to varying levels of familiarity is best captured by a linear (direct matching) or quadratic/cubic (predictive coding) regression model. To address this, we fitted linear, quadratic and cubic regression models to the percent signal change within each ROI, following the procedures reported by Mattavelli et al. (2012).

The runs were first ordered in terms of familiarity. At the pre-training scan session, all stimuli were unfamiliar, yet for clarity, we label them here as stimuli that will or will not be practiced. The 6 blocks were therefore ordered as follows, moving from least familiar to most: Session1_Un_Run1 (first scan session, sequences to remain untrained, first run), Session1_Fam_Run1 (first scan session, sequences to be trained, first run), Session1_Un_Run2 (first scan session, sequences to remain untrained, second run), Session1_Fam_Run2 (first scan session, sequences to be trained, second run), Session2_Un_Run1, Session2_Fam_Run1. Appropriate weighting was applied to the stimuli so that the differences between them were comparable to that of the amount of training (to be) undertaken. For example, the difference between Session2_Un_Run1 and Session2_Fam_Run1 was 12 as the former was only viewed three times over the course of the experiment, whereas the latter was viewed 12. This weighting allowed for a better approximation of the level of familiarity of the runs, thus facilitating a more accurate fit of the regression equations. For each participant, for each training condition, a linear regression and a quadratic regression curve were fitted to each region. For each region, and each training condition, three *R*-squared values were taken: one indicating the fit of the linear regression model, the second indicating the fit of the quadratic regression model and the third the fit of the cubic regression model.

One potential limitation of this approach is that the complex models would explain more variance. To address this, the model comparison technique of Akaike Information Criteria (AIC: Akaike, 1987) was adopted. This method penalizes the more complex models, therefore controlling for overfitting. To calculate AIC with *R*-squared values, the following formula was used (Kaps & Lamberson, 2004);

$$AIC = n \log(SS\ RES / n) + 2\ p$$

where SS RES is 1- *R*-squared, *n* is the number of observations (6), and *p* is the number of parameters (linear = 1, quadratic = 2, cubic = 3). As a small sample was used, AICc was calculated whereby;

$$AICc = AIC + ((2p(p+1)/n-p-1))$$

From this, a Akaike weights (w_i; Burnham & Anderson, 2002) were calculated, showing the probability that a given model is the best model given the data set;

$$\Delta_i\ AICc = AIC_i - min\ AIC$$

$$W_i = (\exp(0.5\Delta_i\ AICc)/ (\Sigma^{m}_{m=1}(\exp(0.5\Delta_i\ AICc)$$

where *m* is a given model. With this statistics, we are able to see the probability that each model is the best model, for each region and for each subject. We then implemented paired t-tests to compare the probability of each model being the best model in a stepwise manner, for each region (correcting for multiple comparisons). Therefore, the differences between linear and quadratic and quadratic and cubic were investigated. The difference between linear and cubic models was not compared as we wished to see the cumulative influence of each parameter. As these models were nested, we assume that the null hypothesis is that the relationship is linear if no significance difference between the models is found.

4.3.8 Imaging Objective 2: Evaluate shape of regression function within ROIs based on subjective measure of familiarity

The aim of the second set of imaging analyses was to evaluate the internal consistency of our findings concerning objective and subjective measures of familiarity. These analyses were performed on the first run of the second scan session only. The rationale behind this approach was that the subjective ratings should only be meaningful from the second scanning session, once participants have undergone the training period. The ROIs were extracted as stated above and participants' individual subjective ratings for each sequence were used as the independent variable in the regression models, rather than exposure. Sequences were then pooled across ratings. One participant was excluded from the execution condition due to only having used three rating values (out of the 1 – 9 rating scale, see *section 4.3.4*).

4.4 Results

4.4.1 Behavioural results

For the observation condition, the accuracy score related to percent correct in terms of whether there was a palming of the strings or not (coded by the experimenter). For the execution condition, Rocksmith® provided a count of how many notes were missed. This was then transformed into a percentage of notes correctly played and used for analysis. The same protocol was used for the training and scanning sessions.

4.4.1.1 Observation Condition

The observation responses for the training sessions (see *Figure 4.7**A***) showed that for all training days, the responses were significantly greater than chance (50%): training day 1 ($M = 82.29$, $SE = 3.20$); $t(19) = 10.07$, $p < 0.001$; training day 2 ($M = 81.77$, $SE = 2.31$); $t(19) = 13.75$, $p < 0.001$; and training day 3 ($M = 85.21$, $SE = 2.29$); $t(19) = 15.36$, $p < 0.001$). These results suggest that on each day of training,

participants could accurately identify whether there was palming of the strings, confirming close attention was paid to the hand actions performed in the stimuli during the observation condition.

To assess the responses made to the observation task *(n.b.* df reflects the fact that not all data was recoverable) during the scanning sessions (see *Figure 4.7**B***), a 2x2 ANOVA was run with factors Scan Session (Session 1/pre-training vs. Session 2/post-training) and Familiarity (familiar vs. unfamiliar). A main effect of familiarity emerged, $F(1, 16) = 114.79$, $p < 0.001$, such that the average accuracy of the familiar stimuli ($M = 80.94$, $SE = 3.03$) was greater than the unfamiliar stimuli ($M = 63.06$, $SE = 3.98$). No main effect of scanning session was found, however, there was an interaction between scanning session and familiarity, $F(1, 16) = 114.79$, $p = 0.044$. Post hoc tests revealed that this interaction arises due to the significant differences between the scanning sessions, with a greater difference on accuracy for the observation task between familiar and unfamiliar sequences at scanning session 2 ($p = 0.047$); suggesting that the time spent observationally training on a subset of sequences resulted in superior performance for recognising the palming action for the trained stimuli only during the post-training scan session.

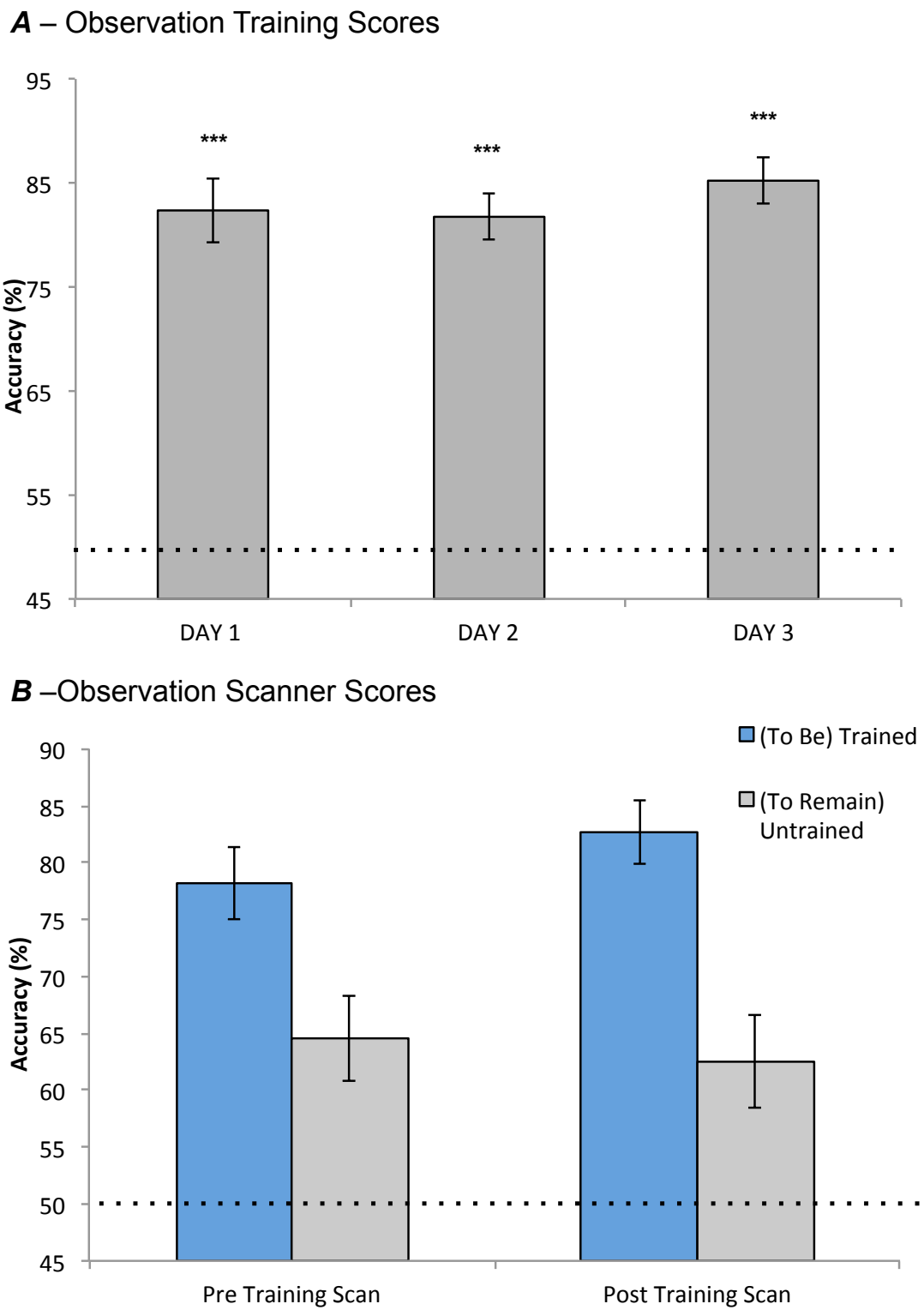


Figure 4.7. Accuracy scores (%) for the observation condition for *A*, the training period and *B*, the scanning sessions. The dotted line denotes chance (50%). *** $p < 0.001$. Bonferroni correction for multiple comparisons applied; error bars represent the standard error of the mean.

A subsample of 15 participants returned to the laboratory after completing both fMRI sessions and the training sessions to perform those stimuli which were only observed, never executed, in order to get an objective measure of how much participants actually learned via observation (results illustrated in *Figure 4.8*). This follow-up test indicated that participants could play the observed sequences significantly better than those that remained untrained; $t(14) = 5.782$, $p < 0.001$. This result suggests that participants did indeed learn to some degree how to perform those sequences that were only observed during the training period.

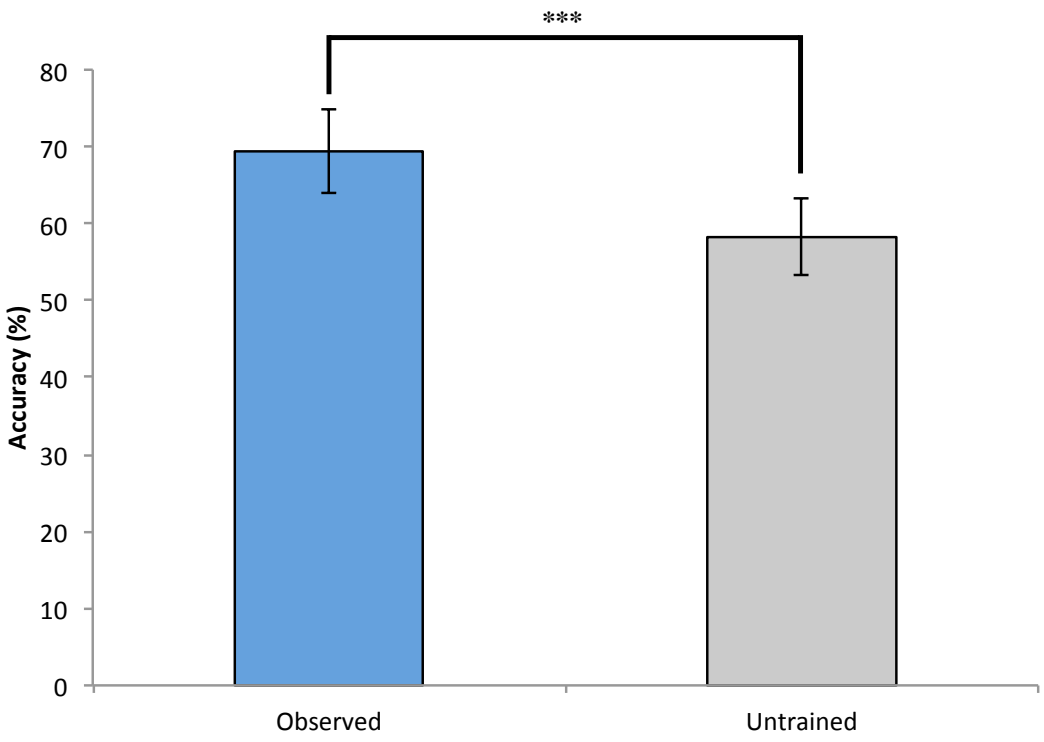


Figure 4.8. Retest results of the guitar sequences that were observed throughout the training process and untrained sequences from the observation condition. Accuracy is the % of notes hit. **p < 0.001; error bars represent the standard error of the mean.

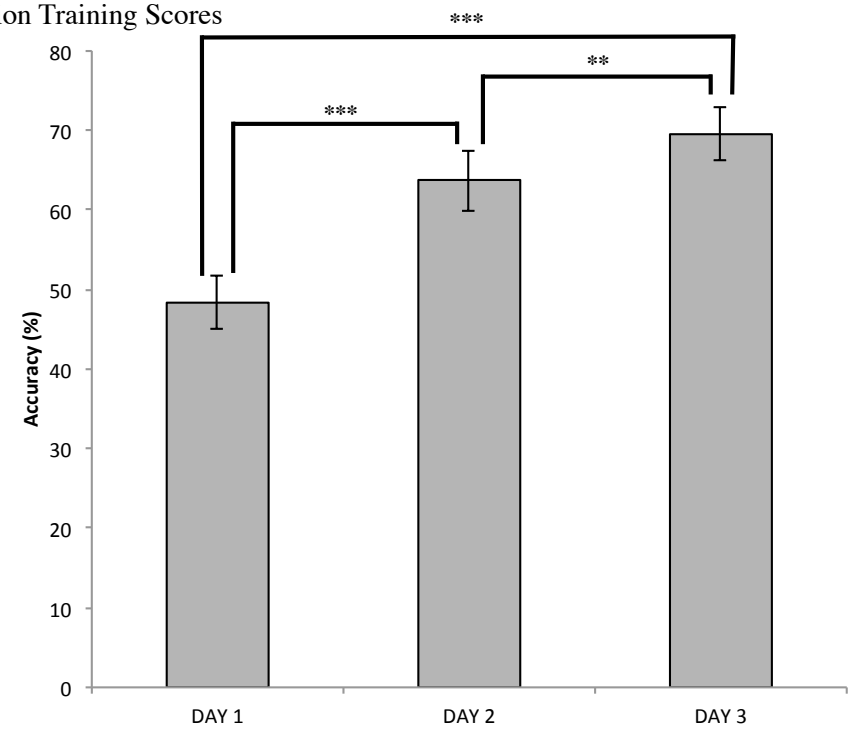
4.4.1.2 Execution Condition

For the execution condition, a repeated measures ANOVA revealed a significant difference in participants' ability to play the guitar riffs across training days; F (2,42) = 40.00, $p < 0.001$ (*Figure 4.9**A***). Further analysis revealed that a significant increase in accuracy between Day 1 (M = 48.01) and Day 2 ($M = 62.05$; $p < 0.001$), Day 1 and Day 3 ($M = 69.58$; $p < 0.001$), and between Day 2 and Day 3 ($p =$

0.022). These differences indicate that accuracy significantly improved across training days, thus demonstrating clear learning induced by the RockSmith® guitar playing task.

The execution scores achieved during scanning are illustrated in *Figure 4.9**B***. These scores were subjected to a 2x2x2 ANOVA with factors Scan Session (Session 1/pre-training vs. Session 2/post-training), Run (Run 1 vs. Run 2), and Familiarity (familiar vs. unfamiliar). A main effect of scan session emerged, $F(1, 21) = 68.78$, $p < 0.001$, such that the average performance accuracy at scan session 2 ($M = 55.96$, $SE = 2.81$) was greater than that at scan session 1 ($M = 33.12$, $SE = 2.27$). There was also a main effect of run, $F(1, 21)= 20.02$, $p < 0.001$, indicating that accuracy in the second run of each scan session ($M = 47.37$, $SE = 2.33$) was better than the first run ($M = 41.70$, $SE = 2.14$). A main effect of familiarity also emerged, $F(1, 21) = 30.41$, $p < 0.001$, indicating that accuracy was better for familiar (to-be-trained/trained; $M = 48.25$, $SE = 1.99$) compared to unfamiliar excerpts (to remain untrained/untrained; $M = 40.82$, $SE = 2.49$).

A – Execution Training Scores



B – Execution Scanner Scores

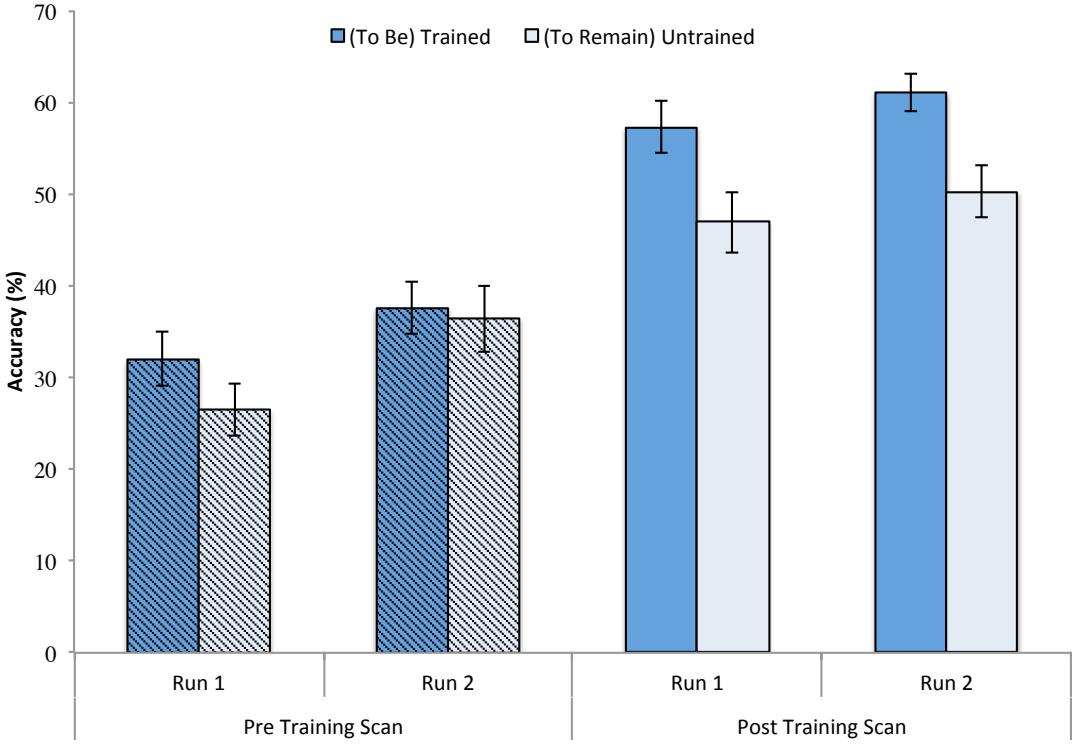


Figure 4.9. Accuracy scores (%) for the execution condition for ***A***, the training period and ***B***, the scanning sessions. Gradient bars indicate parentheses in legend. *** $p < 0.001$, ** $p < 0.01$. Bonferroni correction for multiple comparisons applied; error bars represent the standard errors of the mean.

In terms of interactions, a significant interaction emerged between familiarity and scan session, $F(1, 21) = 22.79$, $p < 0.001$, indicating that differences in performance accuracy for unfamiliar compared to familiar excerpts increased as a function of training. A significant interaction between familiarity and run was also present, $F(1, 21) = 5.02$, $p = 0.036$, suggesting that some learning occurred across the runs (regardless of scanning session), driven by the unfamiliar excerpts showing more marked improvements in performance in the second run compared to the first run of each scanning session. No significant interaction between session and run was found, nor was a 3-way interaction between session, run, and familiarity; all p values > 0.05.

4.4.2 fMRI Results

4.4.2.1 ROI Identification

In order to identify regions of interest with the action observation network, we report a single random effects contrast from the second run of the post-training scan session: familiar and unfamiliar execution and observation vs. implicit baseline.

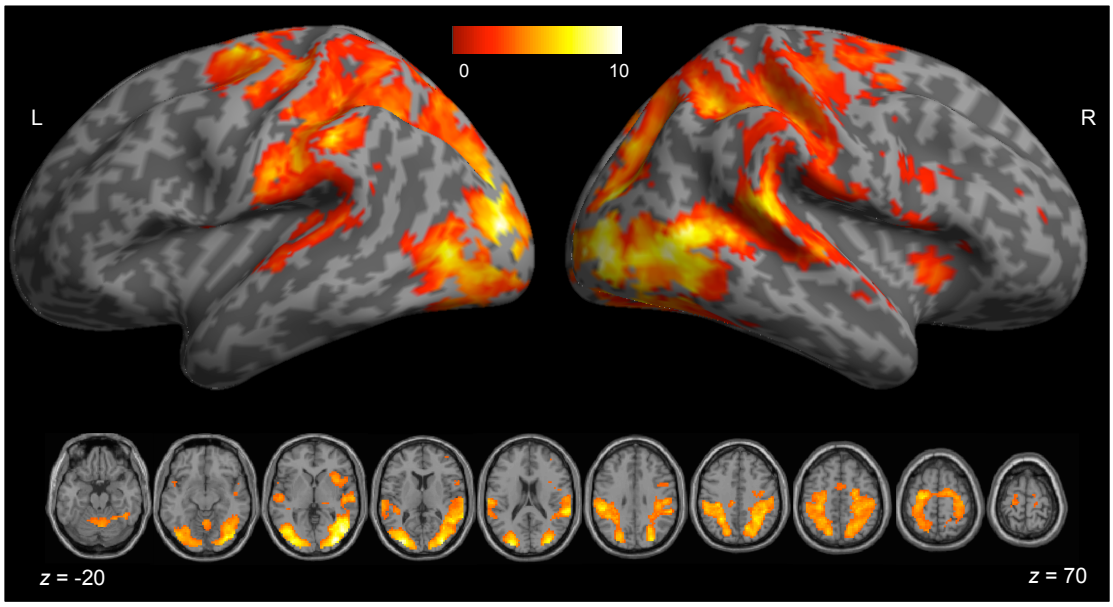


Figure 4.10. Whole brain group analysis Session 2 Run 2 familiar and unfamiliar execution and observation vs. implicit baseline. All p values < 0.0001 (uncorrected), K = 10 voxels.

The ROI contrast of execution and observation vs. implicit baseline revealed widespread engagement of AON regions (See *Table 4.1* for a full list of regions). Only cluster corrected regions (denoted in bold in the table) were used for later analysis. The threshold of p < 0.001 (uncorrected) revealed two cluster-corrected regions, R IFG & R MOG (see supplementary materials). Further examination revealed that the R MOG cluster extended over 8000 voxels, therefore the threshold was elevated to p < 0.05 *(FWE*-corrected) to enable investigation of cluster corrected regions within this larger cluster. From the p < 0.05 *(FWE*-corrected) threshold, Bilateral MOG, Left SmG, R STG, L PMd and R AG were identified. In addition to these six regions, R IFG was also considered for all subsequent ROI analyses.

Table 4.1. Regions associated with observation & execution activity.

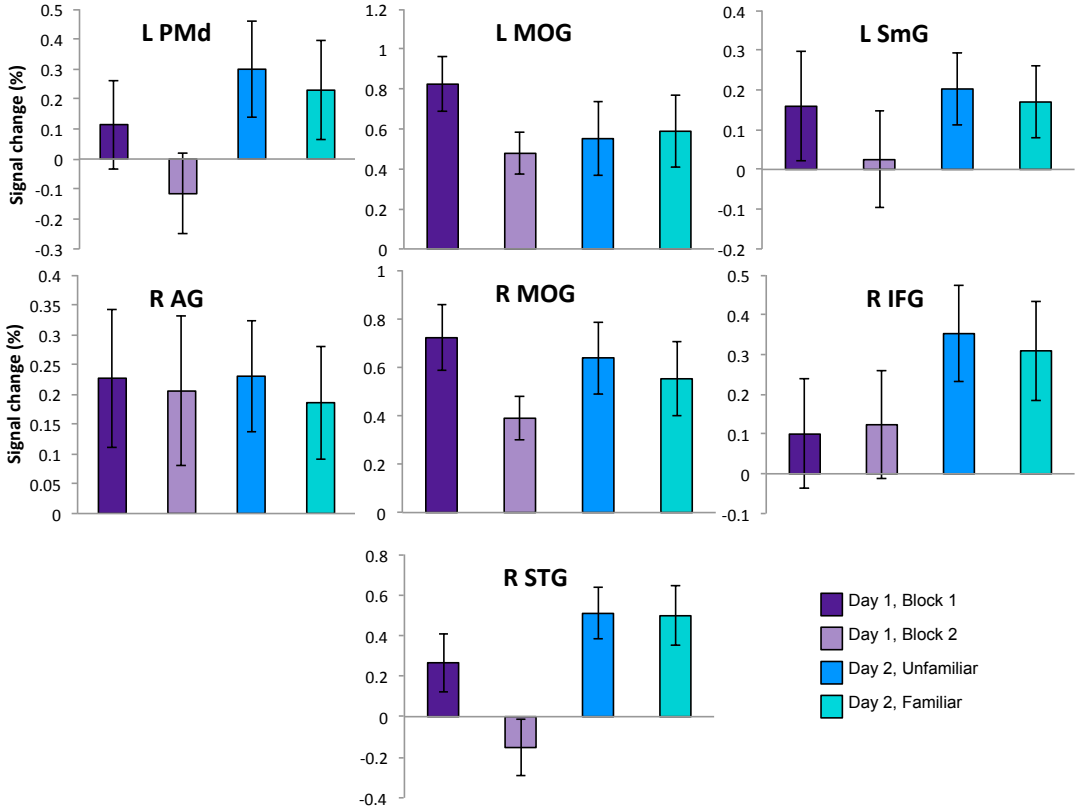
Anatomical region	BA	MNI Coordinates X	Y	Z	Putative functional name	T value	Cluster size	P_{FWE} corrected
P ***p < 0.05*** **$_{FWE}$ corrected**								
R Middle Occipital Gyrus	**18**	**36**	**-85**	**1**	**MOG**	**10.91**	**468**	**0.000**
R Middle temporal Gyrus	37	48	-70	1	MTG	10.04		
R Middle Occipital Gyrus	19	33	-85	10	pFG	9.69		
L Middle Occipital Gyrus	**18**	**-30**	**-85**	**7**	**MOG**	**10.71**	**237**	**0.000**
L Middle Occipital Gyrus	19	-27	-88	16	pFG	9.01		
L Middle Occipital Gyrus	19	-39	-73	4	EBA	8.85		
L SupraMarginal Gyrus	**40**	**-51**	**-34**	**34**	**SmG**	**8.65**	**33**	**0.000**
L Inferior Parietal Cortex	40/1	-54	-22	25	IPC	7.44		
R Superior Temporal Gyrus	**22**	**63**	**-34**	**16**	**STG**	**8.41**	**50**	**0.000**
L Precentral Gyrus	**6**	**-24**	**-10**	**64**	**PMd**	**7.89**	**10**	**0.001**
R Angular Gyrus	**39/7**	**30**	**-49**	**37**	**AG**	**7.17**	**14**	**0.000**
R Angular Gyrus	39	33	-58	40	AG	7.02		

Bold indicates cluster corrected regions. MOG = Middle Occipital Gyrus, MTG = Middle Temporal Gyrus, pFG = Posterior Fusiform Gyrus, EBA = Extrastriate Body Area, SmG = SupraMarginal Gyrus, IPC = Inferior Parietal Cortex, STG = Superior Temporal Gyrus, PMd = Dorsal Premotor Cortex, AG = Angular Gyrus.

4.4.2.2 AON response profile to varying levels of familiarity: Testing the linear (direct matching) vs. quadratic/cubic (predictive coding) accounts

Figure 4.11 shows the mean percent signal change for each region of interest. At the group level, quadratic and cubic patterns of response amplitude in response to familiarity generally emerge across the regions. To assess which pattern is present in all participants, a linear regression, quadratic regression and cubic regression were fitted to each region for each participant. The average *R*-squared values are shown in *Figure 4.12*.

A – Observation Condition



B – Execution Condition

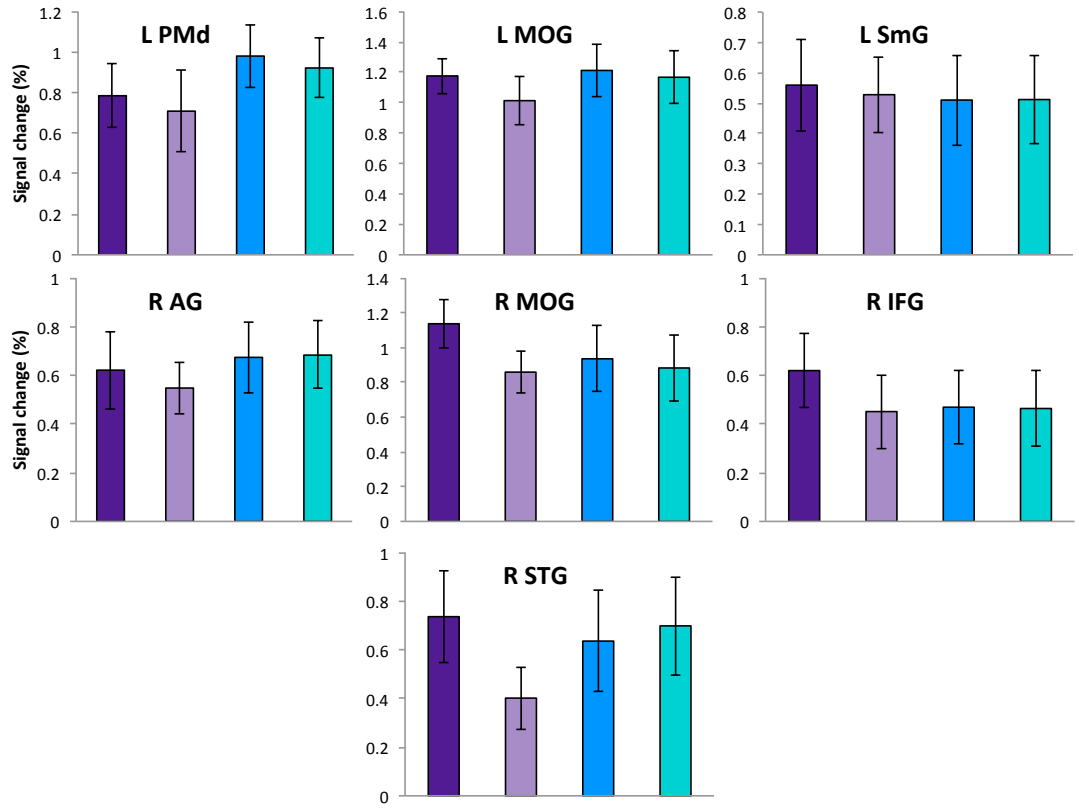


Figure 4.11. Response profiles within each ROI, expressed as percent signal change over familiarity. ***A*** illustrates the results of the observation condition and ***B*** the execution; error bars represent the standard errors of the mean. MOG = Middle Occipital Gyrus, PMd = Dorsal Premotor Cortex, SmG = SupraMarginal Gyrus, AG = Angular Gyrus, STG = Superior Temporal Gyrus, IFG = Inferior Frontal Gyrus.

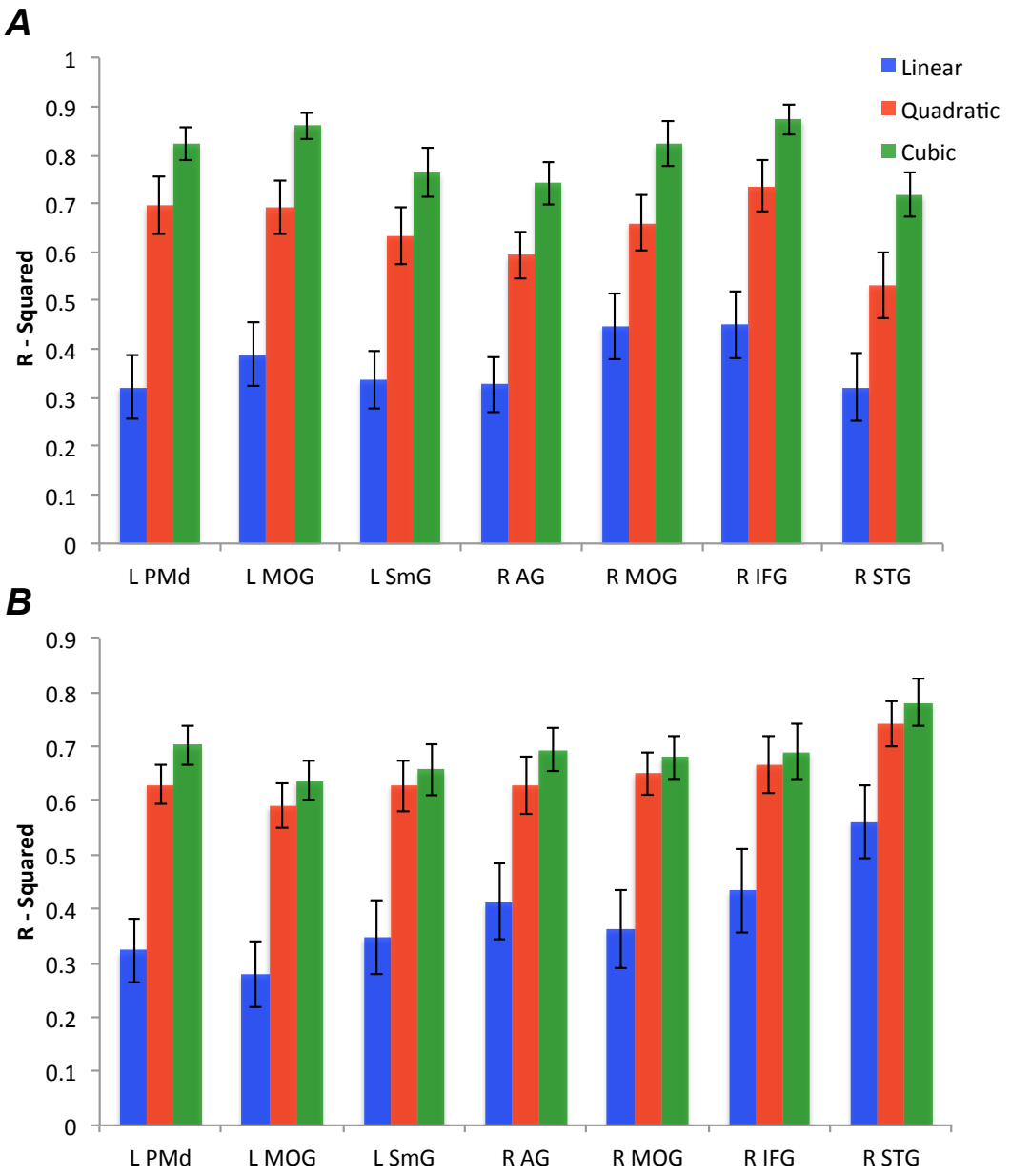


Figure 4.12. *R*-squared values for the linear, quadratic and cubic regressions for each ROI, for both observation ***A***, and execution ***B***. Error bars represent the standard error of the mean. MOG = Middle Occipital Gyrus, PMd = Dorsal Premotor Cortex, SmG = SupraMarginal Gyrus, AG = Angular Gyrus, STG = Superior Temporal Gyrus, IFG = Inferior Frontal Gyrus.

As shown in *Figure 4.12* ***A***, quadratic and cubic regression models appear to fit the data better than a linear model, for all seven ROI's for the observation condition. This finding is also confirmed in the execution condition, shown in *Figure 4.12* B. The limitation of more complex models overfitting the data can be seen in both conditions with the R STG. We see here that the linear model has an *R*-squared value of .057, which is objectively large. However, we see that the quadratic and cubic models both better fit the data, yet this could be that the additional parameters fit the data slightly better. This finding supports the use of a model comparison technique, which penalizes more complex models.

Individual *R*-squared values were then taken and a weighted AIC value (W_i) was calculated for each model, for each ROI. The resulting W_i (shown in *Figure 4.13*) for each model was compared via paired t-tests. These paired t-tests were evaluated with a Bonferroni adjusted alpha level of $p < 0.025$.

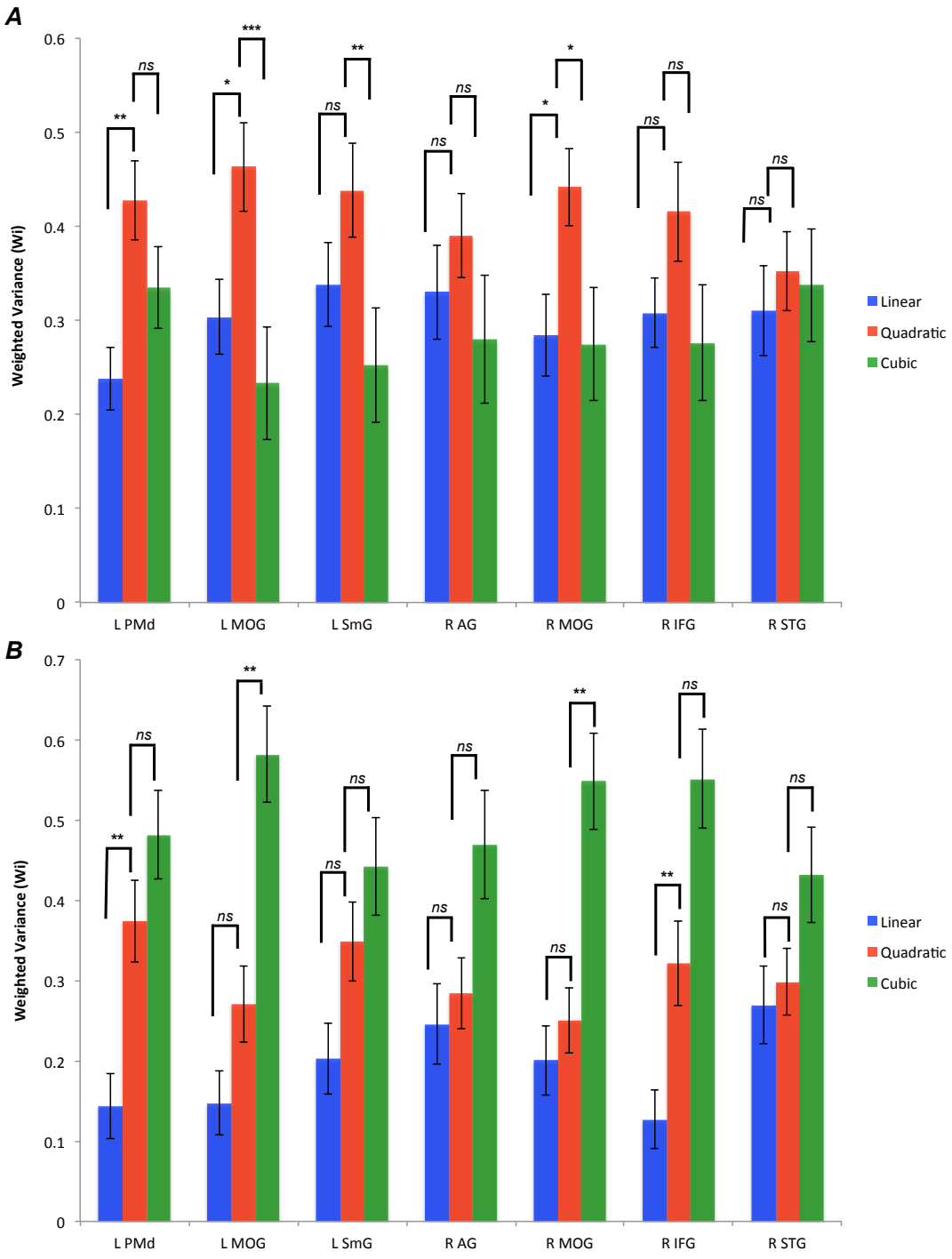


Figure 4.13. *Wi* values for the linear, quadratic and cubic regressions for each ROI, for both observation ***A***, and execution ***B***. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Error bars represent the standard error of the mean. MOG = Middle Occipital Gyrus, PMd = Dorsal Premotor Cortex, SmG = SupraMarginal Gyrus, AG = Angular Gyrus, STG = Superior Temporal Gyrus, IFG = Inferior Frontal Gyrus.

As shown in *Figure 4.13* **A**, a clear trend exists when observing actions: the more complex the model, the greater W_i. In L SmG, there is not significant difference between linear and quadratic models; $t(19) = -1.957$, $p=0.065$, nor between quadratic and cubic models, $t(19) = -0.921$, $p=0.369$. In R AG, there is not significant difference between linear and quadratic models; $t(19) = -0.569$, $p=0.576$, nor between quadratic and cubic models, $t(19) = -1.804$, $p=0.087$. In R STG, there is not significant difference between linear and quadratic models; $t(19) = -0.427$, $p=0.674$, nor between quadratic and cubic models, $t(19) = -1.455$, $p=0.162$. These results suggest that there is no difference between the models for these regions, suggesting that in L SmG, R AG and R STG, the relationship between familiarity and AON engagement is linear. In L MOG, there is not significant difference between linear and quadratic models; $t(19) = -1.931$, $p=0.069$, there is a difference between quadratic and cubic models, $t(19) = -3.109$, $p=0.006$. In R MOG, there is not significant difference between linear and quadratic models; $t(19) = -0.846$, $p=0.408$, there is a difference between quadratic and cubic models, $t(19) = -3.203$, $p=0.005$. These results suggest that there is no difference linear and quadratic models, however, there is a difference between quadratic and cubic terms. This suggests that for L & R MOG, the relationship between familiarity and AON engagement is cubic. In L PMd, there is a difference between linear and quadratic models; $t(19) = -3.118$, $p=0.006$, yet there is not a significant difference between quadratic and cubic models; $t(19) = -1.088$, $p=0.290$. In R IFG, there is a difference between linear and quadratic models; $t(19) = -2.898$, $p=0.009$, yet there is not a significant difference between quadratic and cubic models; $t(19) = -2.117$, $p=0.048$. These results suggest that there is a difference between linear and quadratic models, yet no difference between quadratic and cubic models. This suggests that the cubic function does not significantly explain than the quadratic model, which is significantly more accurate than the linear model; indicating that for L PMd & R IFG, there is a quadratic relationship between familiarity and AON engagement for the observation of actions.

As shown in *Figure 4.13* **B**, a clear trend exists when executing actions: the quadratic model has greater W_i than linear and cubic models. In R AG, there is not significant difference between linear and quadratic models; $t(19) = -1.021$, $p=0.320$, nor between quadratic and cubic models, $t(19) = 1.432$, $p=0.168$. In R IFG, there is

not a difference between linear and quadratic models; $t(19) = -1.621$, $p=0.121$, yet there is not a significant difference between quadratic and cubic models; $t(19) = 1.862$, $p=0.078$. In R STG, there is not significant difference between linear and quadratic models; $t(19) = -0.730$, $p=0.474$, nor between quadratic and cubic models, $t(19) = 0.177$, $p=0.861$. These results suggest that there is no difference between the models for these regions, suggesting that in R AG, R IFG and R STG, the relationship between familiarity and AON engagement for the execution of actions is linear. Interestingly, in L MOG, there is not significant difference between linear and quadratic models; $t(19) = -2.383$, $p=0.028$, there is a difference between quadratic and cubic models, $t(19) = 1.098$, $p<0.001$. In L SmG, there is not significant difference between linear and quadratic models; $t(19) = -1.234$, $p=0.232$, there is a difference between quadratic and cubic models, $t(19) = 3.625$, $p=0.002$. For both these regions, the cubic model is significantly less probable than then quadratic model, which is not significantly greater than the linear model. This suggests that in L SmG and L MOG, this relationship is linear. In R MOG, there is significant difference between linear and quadratic models; $t(19) = -2.792$, $p=0.012$, also between quadratic and cubic models, $t(19) = 2.800$, $p=0.011$. This shows that this region is therefore cubic with each parameter significantly explaining more variance, thus more probable. In L PMd, there is a difference between linear and quadratic models; $t(19) = -3.068$, $p=0.006$, yet there is not a significant difference between quadratic and cubic models; $t(19) = 1.184$, $p=0.251$. This suggests that for PMd, there is a quadratic relationship between familiarity and execution of actions.

4.4.2.3 Internal Consistency

Thus far, we have shown that for the observation condition, there is a linear relationship between familiarity and engagement in temporal and parietal ROIs, quadratic and frontal AON regions and cubic in occipital ROIs; when using exposure as the independent variable. In the execution condition, similar findings exist for L PMd, L SmG, R AG, R MOG & R STG. The differences arise in L MOG, which is linear in the execution condition (cubic in the observation) and in R IFG, which is linear in the execution condition (quadratic in observation). To test the extent to which an objective measure of action familiarity (i.e., number of times any given

sequence has been practiced or observed) is consistent with subjective measure of familiarity, we extended the same analytical approach described above and this time used participants' own subjective ratings as the independent variable for setting up the regression models. With this approach, we could evaluate whether comparable findings emerge when using this more sensitive/subject-specific measure of familiarity. The subject ratings were taken from the 1 – 9 Likert scale collected post scanning. For each of the stimuli, the subject's individual rating was assigned to it, becoming the independent variable in the regression model (no weighting was applied as there was equal distances between the rating points, e.g. the differences between a rating of 1 and was equal to 2 - 3). Visualization cannot be provided for this data as not all rating points were used across the participants. Linear, quadratic and cubic regression models were fitted to the data and the average *R*-squared values for each model, for each ROI is shown in *Figure 4.14*

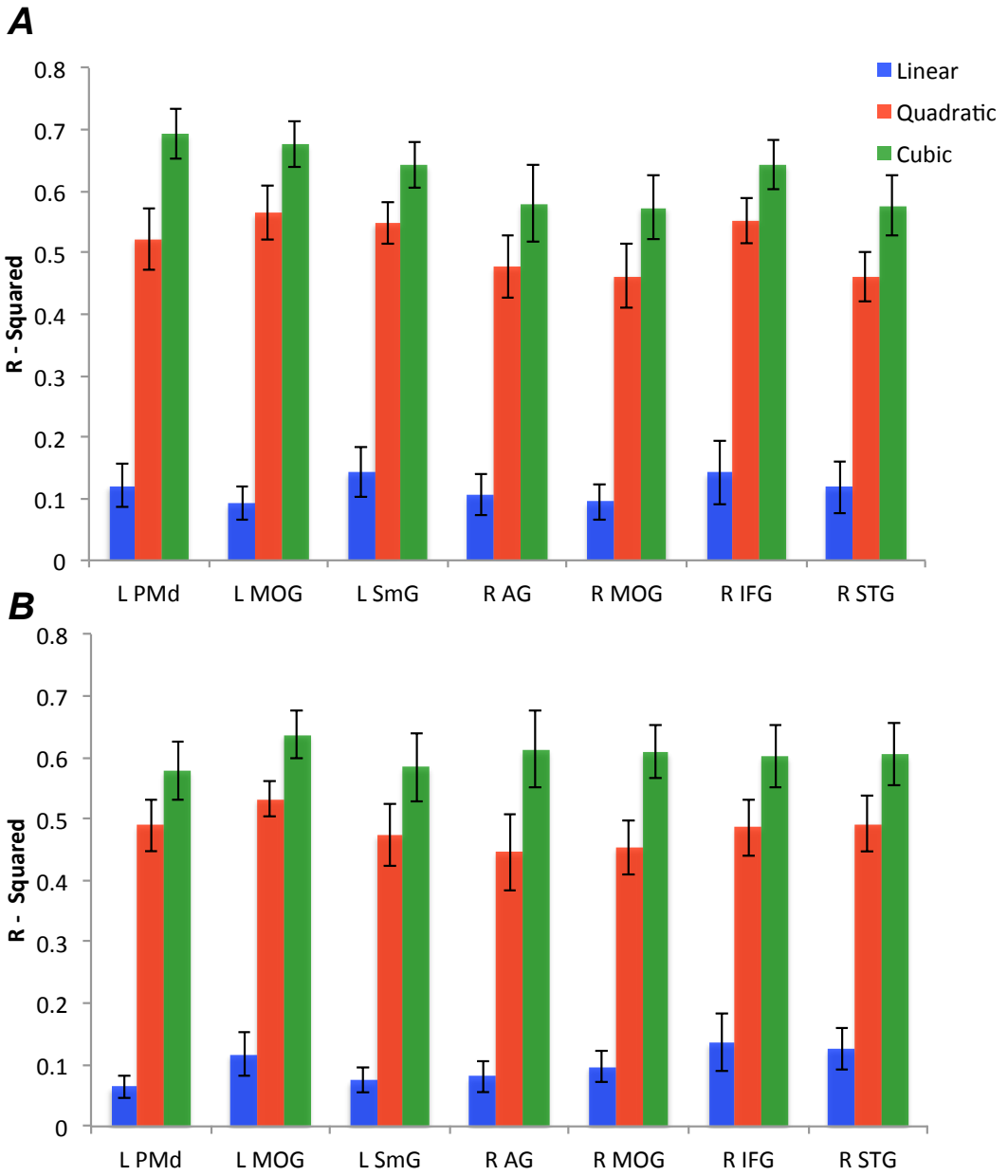


Figure 4.14. *R*-squared values for the linear, quadratic and cubic regressions for each ROI, for both observation ***A***, and execution ***B***. Error bars represent the standard error of the mean. MOG = Middle Occipital Gyrus, PMd = Dorsal Premotor Cortex, SmG = SupraMarginal Gyrus, AG = Angular Gyrus, STG = Superior Temporal Gyrus, IFG = Inferior Frontal Gyrus.

As shown in *Figure 4.14*, the *R*-squared values are very similar to those when using exposure as the independent variable, with the quadratic and cubic models explaining more variance. Following the procedure used above, the *R*-squared values for each subject, for each model, for each region were used to calculate a W_i value.

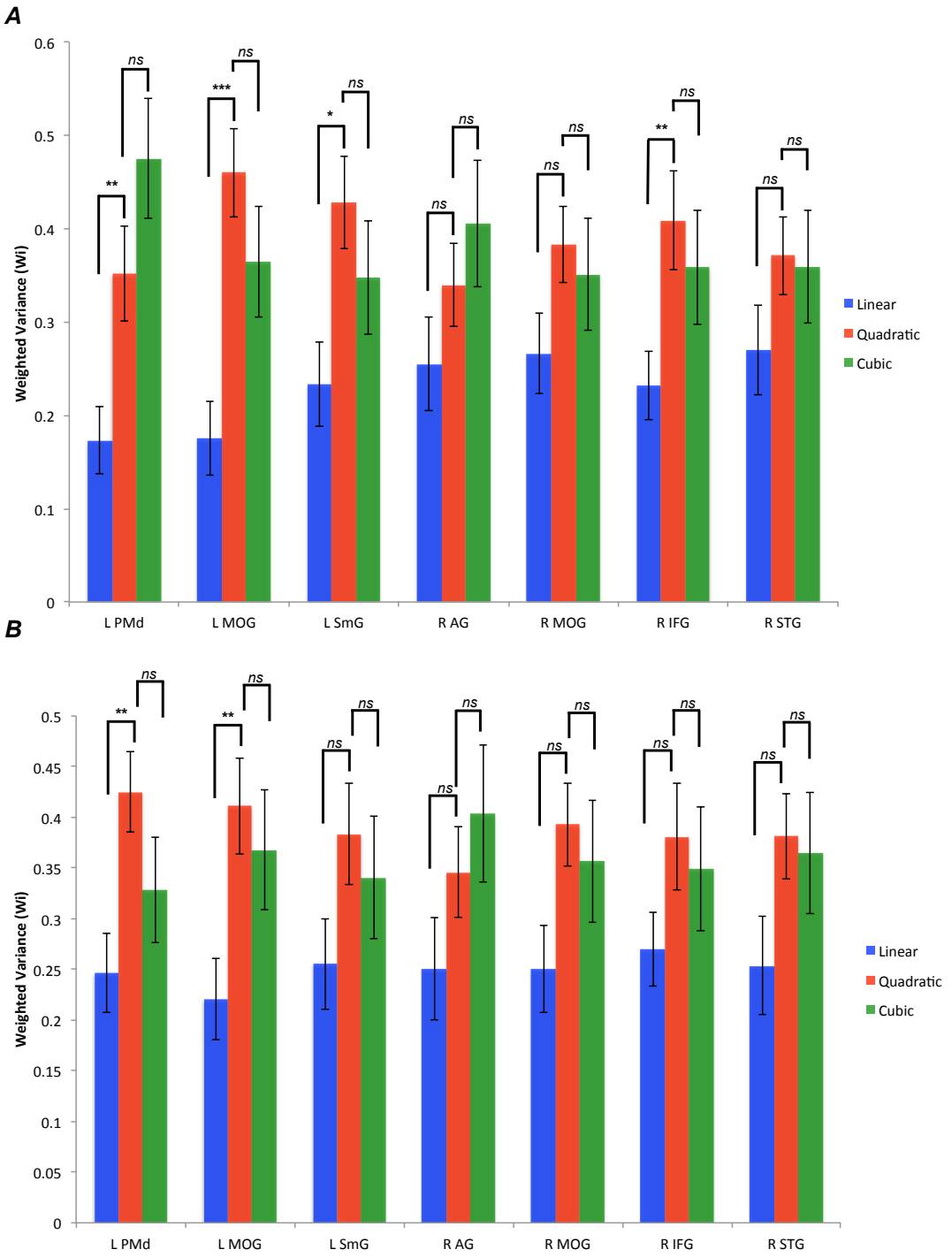


Figure 4.14. *Wi* values for the linear, quadratic and cubic regressions for each ROI, for both observation ***A***, and execution ***B***. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Error bars represent the standard error of the mean. MOG = Middle Occipital Gyrus, PMd = Dorsal Premotor Cortex, SmG = SupraMarginal Gyrus, AG = Angular Gyrus, STG = Superior Temporal Gyrus, IFG = Inferior Frontal Gyrus.

As shown in *Figure 4.14* **A**, when observing actions the more complex the model, the greater W_i. In R AG, there is not significant difference between linear and quadratic models; $t(19) = -1.315$, $p=0.204$, nor between quadratic and cubic models, $t(19) = -0.571$, $p=0.575$. In R MOG, there is not significant difference between linear and quadratic models; $t(19) = -2.080$, $p=0.051$, nor between quadratic and cubic models, $t(19) = 0.371$, $p=0.715$. In R STG, there is not significant difference between linear and quadratic models; $t(19) = -2.162$, $p=0.044$, nor between quadratic and cubic models, $t(19) = 0.128$, $p=0.900$.These results suggest that there is no difference between the models for these regions, suggesting that in R MOG, R AG and R STG, the relationship between familiarity and AON engagement is linear. In L SmG, there is significant difference between linear and quadratic models; $t(19) = -2.526$, $p=0.021$, yet there is not a difference between quadratic and cubic models, $t(19) = 0.924$, $p=0.367$. In L MOG, there is a significant difference between linear and quadratic models; $t(19) = -4.693$, $p<0.001$, there is not a difference between quadratic and cubic models, $t(19) = 1.098$, $p=0.286$. In L PMd, there is a difference between linear and quadratic models; $t(19) = -3.118$, $p=0.009$, yet there is not a significant difference between quadratic and cubic models; $t(19) = -1.088$, $p=0.274$. In R IFG, there is a difference between linear and quadratic models; $t(19) = -2.939$, $p=0.008$, yet there is not a significant difference between quadratic and cubic models; $t(19) = 0.548$, $p=0.590$. These results suggest that there is a difference between linear and quadratic models, yet no difference between quadratic and cubic models. This suggests that the cubic function does not significantly explain than the quadratic model, which is significantly more accurate than the linear model; indicating that for LPMd, RIFGL MOG and L SmG, there is a quadratic relationship between familiarity and AON engagement for the observation of actions.

As shown in *Figure 4.14* **B**, when executing actions that the more complex the model, the greater W_i. In R AG, there is no significant difference between linear and quadratic models; $t(19) = -1.308$, $p=0.207$, nor between quadratic and cubic models, $t(19) = -0.522$, $p=0.608$. In R IFG, there is not difference between linear and quadratic models; $t(19) = -1.750$, $p=0.096$, nor between quadratic and cubic models; $t(19) = 0.345$, $p=0.734$. In R MOG, there is no significant difference between linear and quadratic models; $t(19) = -2.114$, $p=0.048$, nor between quadratic and cubic

models, $t(19) = 0.373$, $p=0.713$. In R STG, there is not a significant difference between linear and quadratic models; $t(19) = -1.948$, $p=0.066$, nor between quadratic and cubic models, $t(19) = 0.160$, $p=0.875$. In L SmG, there is not a significant difference between linear and quadratic models; $t(19) = -1.911$, $p=0.071$, nor between quadratic and cubic models, $t(19) = 0.481$, $p=0.636$.These results suggest that there is no difference between the models for these regions, suggesting that in L SmG, R MOG, R IFG, R AG and R STG, the relationship between familiarity ratings and AON engagement is linear. In L PMd, there is a significant difference between linear and quadratic models; $t(19) = -3.060$, $p=0.006$, there is not a difference between quadratic and cubic models, $t(19) = 0.482$, $p=0.635$. In L MOG, there is a difference between linear and quadratic models; $t(19) = -3.600$, $p=0.006$, yet there is not a significant difference between quadratic and cubic models; $t(19) = 1.146$, $p=0.266$. These results suggest that there is a difference between linear and quadratic models, yet no difference between quadratic and cubic models. This suggests that the cubic function does not significantly explain than the quadratic model, which is significantly more accurate than the linear model; indicating that for L PMd and L MOG, there is a quadratic relationship between familiarity rating and AON engagement for the execution of actions.

4.5 Discussion

The aim of the present study was to investigate how varying levels of familiarity modulate engagement of the action observation network (AON) during the observation or execution of guitar playing sequences. Specifically, we were interested in testing whether a direct matching account (Rizzolatti et al., 2001; Gallese & Goldman, 1998; Wolpert et al., 2003) or predictive coding account (Keysers & Perrett, 2004; Kilner et al., 2007a,b; Gazzola & Keysers, 2009; Schippers and Keysers, 2011) better explained the impact of increasing familiarity on AON engagement. To address this, we asked participants to take part in identical fMRI sessions where they observed and executed different guitar playing sequences, and in between the sessions, trained (via observational or physical practice) on half the sequences. By performing a region of interest analysis on areas that exhibited activity

during both observation and execution of actions, were able to address two distinct questions: (1) does the shape of response within core AON regions correspond more to a direct matching or predictive coding profile?: and (2) how do subjective measures of familiarity relate to the response profile within core AON regions?

4.5.1 AON response to varying levels of familiarity – Evaluating the direct matching and predictive coding accounts

The first question we addressed was whether the relationship between increasing familiarity and AON response was better captured by a linear response profile (more in keeping with a direct matching account; Rizzolatti et al., 2001; Gallese & Goldman, 1998; Wolpert et al., 2003) or a quadratic/cubic response profile (which would be more consistent with a predictive coding account; Keysers & Perrett, 2004; Kilner et al., 2007a,b; Gazzola & Keysers, 2009; Schippers & Keysers, 2011). Through use of pre- and post-training scanning sessions on either side of an intensive training intervention, we were able to manipulate the degree of familiarity participants had with specific guitar riffs. After identifying regions of interest within the AON that respond to executing or observing guitar playing actions, we fitted linear, quadratic and regression models to each region for each subject.

We found that for the observation condition, the ROIs in the parietal and temporal lobes were best captured with a linear regression model. In the execution condition, all ROIs except R MOG and L PMd are also best captured by a linear regression model. These findings are in keeping with a direct matching account (Rizzolatti et al., 2001; Gallese & Goldman, 1998; Wolpert et al., 2003), which suggests that as familiarity increases, the engagement of the AON increases. For the observation condition, the occipital and frontal AON regions exhibited cubic and quadratic responses respectively. For the execution condition, R MOG and L PMd exhibit cubic and quadratic responses respectively. These findings suggest that at each end of the familiarity spectrum, the response of the AON is similar, and is overall stronger than the response observed at intermediate levels. This finding falls in line with theoretical framings that suggest a quadratic function of AON engagement based on increasing familiarity (Cross et al., 2012; Liew et al., 2013; Diersch et al., 2013).

Generation of linear and quadratic regression models enables us to directly compare two dominant models of AON function: the direct matching framework (Rizzolatti et al., 2001; Gallese & Goldman, 1998; Wolpert et al., 2003) and the predictive coding framework (Keysers & Perrett, 2004; Kilner et al., 2007a,b; Gazzola & Keysers, 2009; Schippers & Keysers, 2011). Briefly, the predictive coding framework suggests reciprocal modulation between the nodes of the AON, which provides feedback predictions, and feedforward updates with the aim to minimise error. A quadratic function of familiarity would therefore fit this assumption, as at the highly familiar end of this scale, greater familiarity with an observed or executed action should result in greater/more accurate prediction; while at the highly unfamiliar end of the scale, less familiar actions should be associated with more on-going updating. In contrast, according to a direct matching framework (Rizzolatti et al., 2001; Gallese & Goldman, 1998; Wolpert et al., 2003) as familiarity increases, greater AON activity should emerge in a linear fashion. The present study is the first to directly test these two accounts by examining the shape of the response profile within key AON ROIs, based on varying levels of familiarity. Our data suggest that, opposed to the direct matching hypothesis, the response profile of all the core AON regions is not linear. These findings provide partial support for the predictive coding account as consensus across regions is not present, however, this suggests that the relationship between familiarity and AON engagement is complex, and in certain regions, nonlinear.

In addition, it is of interest that the results of L PMd are remarkably similar between the observation and execution conditions (as well as using familiarity ratings), whereby the response is quadratic. This result corroborates one of the main findings reported by neurophysiological investigations of mirror neurons in non-human primates: namely, that the response profile of some cells within parietal and premotor cortices during observation is comparable to those same cells' response profile during action execution (Di Pellegrino et al., 1992; Gallese et al., 1996; Rizzolatti et al., 2001; Umilta et al., 2001). Naturally, we cannot conclude that the ROIs examined in the present study are actually coding information in the same way during observation and execution. Nonetheless, the fact that quadratic functions best capture the response profiles of the PMd during both execution and observation support the notion that LPMd (or F5 in the macaque), play a critical role in action

observation and execution, but that they respond to changes in action familiarity in a similar manner as well.

4.5.2 Consistency between objective and subjective familiarity ratings

By evaluating a separate analysis wherein we used subjective ratings of familiarity as an independent variable within our regression model, we could investigate whether objective and subjective measures of familiarity align or differ within core AON regions. Participants' subjective familiarity ratings were obtained after the last scanning session and were assigned to the corresponding sequence within the GLM. Both of the approaches reliably show that 4/7 regions for both conditions replicate indicate that subjective ratings offer a sensitive and subject-specific measure of experience. Such individual ratings can add value when used in conjunction with time-consuming training interventions (Cross, Hamilton & Grafton, 2006; Kirsch, Dawson & Cross, 2015; Casile & Giese, 2006; Läppchen et al., 2015), or when used in isolation when taking physical measures of performance are not feasible or possible (e.g., Gardner et al., 2015; Cross et al., 2011; Press & Kilner, 2013; Kawabata & Zeki, 2004).

4.5.3 Limitations and Future Directions

One potential limitation of the current study concerns the use of a weighted regression model. By using the number of times exposed to the stimuli as the objective familiarity parameter, the independent variable in the regression model was consequently not evenly distributed, potentially skewing the regression model. However, this limitation is less of a concern due to the fact that we were able to by and large replicate our findings when using subjective ratings of familiarity as the independent variable in the regression models. This suggests that any potential skew introduced to the regression model by the not even distribution of the independent variable is ungrounded. An extension of current work would then be to improve the regression model whereby more time points (and more evenly distributed time points)

are examined, thus providing a better representation of the curve (c.f. Braams et al., 2015).

An additional potential limitation of this study that warrants consideration is the use of different training scenarios for the observation and execution conditions. Specifically, in the execution condition, participants received online real-time feedback about their performance, whereas no such feedback is possible when observing. This limitation would be of greater concern if we were interested in drawing direct comparisons about the effects of both kinds of training on performance. However, our aims in the current study focused on how physical and observational experience impact AON engagement independently, but within the same group of participants. As such, the differences in training experience are both necessary and warranted in this case. A recent study sheds further light on using rich training interventions with multisensory training experience (Kirsch & Cross, 2015). In this study, the authors showed that by use of a multisensory training paradigm that layering auditory, visual and physical experience has a cumulative effect in shaping engagement of the premotor cortex during action observation (Kirsch & Cross, 2015). In the current study, we used visual plus auditory stimuli in the observation condition, and multimodal stimuli that involved motor, visual and auditory systems in the physical training condition. As such, one could argue that we have effectively "stacked the deck" in our favour (in terms of maximally engaging the AON) by using rich, multisensory stimuli to investigate ROI responses. This was a deliberate decision, as past research also shows that auditory experience in addition to visual cues enables participants to reconcile the timing of the movements in accordance with the auditory feedback (see Lotze, 2013 for insights on the importance of motor, somatosensory, auditory and visual aspects in musician imagery; with particular focus on the audiomotor loop). It would be valuable for future work, however, to investigate in finer detail the impact of unimodal vs. multimodal experience as it relates to theoretical models of AON engagement and familiarity.

4.6 Conclusions

In conclusion, the present findings address core questions concerning how familiarity shapes action observation and execution-related processing within the AON. Via a region of interest analysis, we directly tested which of two key theoretical models of the AON better accounted for the impact of varying levels of familiarity: the direct matching and predictive coding accounts. The findings indicate, both for objective and subjective familiarity, that the nature of the AON is complex and not exclusively linear, providing support for the predictive coding account.

5 Chapter V

Complex Action Learning Shapes Dynamic Modulation of the Action Observation Network

5.1 Abstract

Humans spend considerable time observing and executing actions. Many of these actions are highly familiar and well-practiced, such as tying one's shoes, while others that are much less familiar, such as watching a somersaulting gymnast. In the human action cognition literature, debate exists on the role of familiarity and subsequent engagement of a network of temporal, parietal, and frontal regions; collectively termed the Action Observation Network (AON). The predictive coding model, a prominent theoretical model of AON function, holds potential for reconciling this debate. This model features reciprocal connections spanning the core AON nodes, which perform Bayesian comparisons of expected and observed actions. According to this model, familiar and unfamiliar actions might activate core AON nodes to a similar degree, but how information is exchanged between the nodes should differ based on an individual's familiarity with an observed or executed action. To test these predictions, we used an intensive guitar-training paradigm, enabling rigorous manipulation of familiarity in visual and motor domains. Identifying core AON nodes from pre- and post-training scanning sessions, we then applied effective connectivity analyses to test the hypothesised modulations of the predictive coding account. We demonstrate that predictive coding hypotheses of distinct patterns of modulation based on perceived or performed actions' familiarity are generally supported by the empirical data, and contribute valuable insights toward understanding the complex role played by familiarity in modulating action cognition.

5.2 Introduction

Observing others in action elicits activity across a network of occipitotemporal, premotor and parietal cortices, which are collectively termed the Action Observation Network (AON; Cross et al., 2009; Grafton, 2009; Keysers & Gazzola, 2009; Caspers et al., 2010). While broad consensus exists that action perception and execution likely share common neural underpinnings (Gazzola & Keysers, 2009; Kilner et al., 2009a; Hommel et al., 2001; Rizzolatti & Craighero 2004; Rizzolatti et al., 2014), many questions remain concerning how experience shapes responses within these regions, including the role played by familiarity in the visual or motor domains. Some literature demonstrates that the more physically familiar an observed action is, the stronger the response is within core AON regions (Buccino et al., 2004; Calvo-Merino et al., 2005; Cross et al., 2006; Vogt et al., 2007; Shimada, 2010). However, other evidence shows that the AON's response profile does not necessarily follow this linear trend of increasing engagement with increasing familiarity (Gazzola et al., 2007; Liew et al., 2013; Cross et al., 2012), and instead responds more robustly when observing unfamiliar compared to familiar actions. In order to reconcile this debate and to better understand how experience shapes AON engagement, it is imperative to move from magnitude-based neuroimaging analyses to more sophisticated multivariate approaches that enable closer examination of how familiarity dynamically modulates sensorimotor brain regions (Smith, 2012; Gardner et al., 2015). Crucially, such multivariate approaches also enable sensitive testing of a theoretical model that holds significant promise for advancing understanding of the role played by experience in shaping AON engagement: predictive coding

Predictive coding accounts of AON function (Keysers & Perrett, 2004; Kilner et al., 2007a,b; Gazzola & Keysers, 2009) propose a Bayesian comparison between predicted and observed actions, across reciprocally connected nodes of the AON. This Bayesian comparison aims to minimise differences between actual and predicted actions. This should manifest as increased AON engagement for highly unfamiliar actions, as the influence of feedforward connections from posterior perceptual regions is heavily relied upon. When viewing or performing an action that is familiar,

however, the network's predictions for how a movement should unfold in real time are more precise, thus minimising prediction error. This minimised prediction error is manifest as robust AON engagement, driven by an upregulation of posterior projections originating from the anterior premotor node feeding back through the network. When evaluating past studies that demonstrate more AON engagement for familiar vs. unfamiliar actions (e.g., Buccino et al., 2004; Cross et al., 2006), and those showing more AON engagement for unfamiliar vs. familiar actions (e.g., Liew et al. 2013; Cross et al., 2012), predictive coding accounts begin to add some clarity as to why such conflicting findings might emerge when using univariate analytical approaches.

Here, we systematically test the impact of instilling physical and visual familiarity with novel actions on AON engagement by combining intensive training procedures, pre- and post-training fMRI sessions, and measures of effective connectivity. This approach enables us to evaluate how changes in familiarity induced by practice modulate the exchange of influence between core nodes of the AON. By questioning how effective connectivity *changes* with increasing experience, we can build upon recent partial support for a predictive coding account of AON engagement provided by effective connectivity (Gardner et al., 2015; Thioux & Keyser, 2015). Using a guitar-playing videogame paradigm, we studied guitar-naïve participants as they learned to play simple riffs via physical or visual experience. To test how well a predictive coding account of AON function explains variations in familiarity, we used Dynamic Causal Modelling (Friston et al., 2003), a form of effective connectivity that allows investigation of experimentally manipulated modulators. Based on the predictive coding account, we expect to see anterior-projecting modulation of the AON when observed or executed actions are unfamiliar. In contrast, as these actions become familiar, predictive coding hypothesizes increased precision of predictions for executed or observed movements, which should manifest as the presence of posterior modulation.

5.3 Materials and methods

5.3.1 Participants

Twenty-two young adult volunteers recruited from the Bangor University community took part in the experiment. Two volunteers were excluded from the final sample due to excessive head motion during scanning (> 6 mm). The final sample comprised 20 volunteers (9 males, *M_age* = 20.60 years, *SD*=1.73). All participants had normal or corrected to normal vision with no history of neurological illness. All participants were right-handed and required to play a right-handed bass guitar in the scanner. Participants were brought to the lab prior to scanning to ensure they could play the instrument in the manner required inside the scanner bore, and also to ensure that all were guitar novices. The study was approved and conducted following the guidelines of the Ethics Committee of the School of Psychology at Bangor University and the Bangor Imaging Unit. All participants provided written informed consent prior to participation.

5.3.2 Stimuli & Apparatus

5.3.2.1 Sequences

We chose 16 sequences from the computer game Rocksmith® (Ubisoft, 2014), lasting an average of 15.8s (SD = 2.37s). These sequences were excerpts taken from songs initially chosen due to their lack of lyrical content. This restriction was selected so participants would not associate any particular action sequence with lyrics. The sequences were also matched on the difficulty level assigned by Rocksmith®. Rocksmith® assigns difficulty level via an algorithm that assesses song speed and the number of notes to be played within the time window (the difficulty level is visible at the top of *Figure 5.1*). To ensure that difficulty levels were matched as closely as possible across all stimuli, the number of notes to be played and the length of the sequences were matched, as were beats per minute for the individual song excerpts. In addition, the inclusion criteria for these sequences required that notes fell within a fret

range of 1-7 and string range of 1-3. This restriction ensured that participants would not have to move their heads to identify frets during scanning, while at the same time maintaining a level of difficulty that would challenge participants throughout the training period. Furthermore, we also excluded technical guitar playing movements such as “hammering” and “sliding” so that the actions required would be accessible to novices. Finally, we matched the sequences on mean amount of motion energy displayed in each video (see Bobick, 1997; Schippers et al., 2010; Cross et al., 2012), to ensure gross differences in the amount of motion displayed in individual sequence videos did not contribute to basic visual differences between stimuli or training conditions.

5.3.2.2 Rocksmith® Guitar Task

Shown in *Figure 5.1* is a screenshot of the gameplay from Rocksmith® for the sequences that were physically executed (Panel A) and those that were observed (Panel B). The horizontal coloured lines at the foot of the screen correspond to the strings of the guitar being played. We used a bass guitar for this study as it has fewer strings than a standard electrical guitar (4 vs. 6) and at a basic level is generally considered easier to learn to play than a traditional guitar. These coloured strings are illustrated from a first person view, as if one is looking through the back of the fretboard of the guitar (i.e., when holding the guitar, the red string corresponds to the top string, green the second string from the top, and so on).

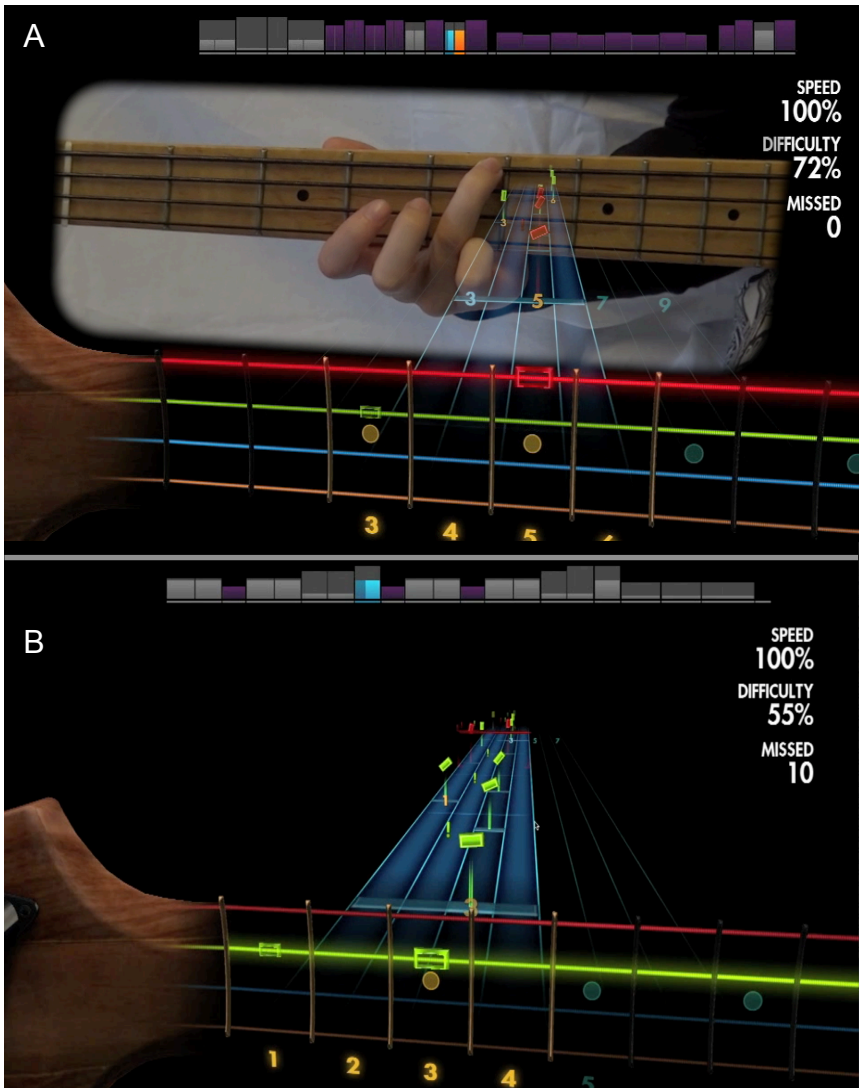


Figure 5.1. A still video frame example of the stimuli used in the observation condition ***A*** and execution condition ***B***. The hand and fretboard of the musician playing the guitar in the centre of (A) is superimposed on the actual game play.

The second aspect of the gameplay to be explained is the translucent blue "conveyor belt" of notes seen in the centre of the screen. The numbers on this conveyor correspond to the fret number on the guitar itself. The coloured rectangles are critical for the participant to attend to in order to make the correct movement. The colour of the rectangle corresponds to the string to be played with the right hand, and the number fret it appears on corresponds to which fret should be pressed down with the left hand. As the rectangles and numbers on the conveyor move towards the fore of the screen, the rectangles on the conveyor rotate 90 degrees from vertical to horizontal. When they reach the horizontal position, they contact the strings at the

bottom of the screen, and this is when the participant must play the appropriate note. If correctly executed, the rectangle illuminates slightly, but if missed the word “miss” appears on the string and fret that should have been played and the counter located at the upper right corner of the screen is adjusted accordingly. To obtain a perfect score, the correct fret on the correct string had to be plucked within a ±250ms window of the onset of the note.

5.3.2.3 Observation Task

Figure 5.2 shows a schematic of the observation task. Participants first viewed a brief fixation cross, followed by a video of an expert musician playing one song excerpt. The musician performs all of the songs without fault, thus providing a precise template for participants to observe.

Figure 5.2. A schematic of one observation trial. Participants' task was to attend to the hand and respond to whether the musician performed a palming action of the strings.

In order to ensure participants paid close attention to the actions performed by the expert musician in the observation condition, an attentional control task was implemented whereby participants were required to identify whether the musician palmed the strings during each stimulus video. A palming of the strings occurs when the musician removes her fingers from the fretboard, extends them vertically, and places the palm of her hand over the frets. This action was performed quickly as the musician then immediately continued to play the correct notes without stopping and without error. At the end of each trial, participants were asked whether a palming action was seen in the last sequence, and were required to make their response by plucking one of two strings with the right index finger. To ensure no confusion about what was required of participants during the observation task, the concept of palming the strings was explained and demonstrated before the experiment began.

5.3.2.4 Execution Task

In the execution condition, *Figure 5.1 (Panel B)*, participants were instructed to play along with the song to the best of their ability. The instruction was also given for participants to simply move on to the next note if they missed one note, to ensure participants did not move excessively when trying to compensate for an error. After each sequence was played, the gameplay presented a count of how many notes were missed, providing feedback on performance.

5.3.3 Training procedure

The study began with all participants taking part in an fMRI scanning session (far left column of *Figure 5.3*), wherein they observed eight sequences and executed eight different sequences. Participants played or watched each of the eight sequences from the different training conditions twice per run, and completed two runs of both kinds of training on each day of scanning. All the stimuli on the first day are labeled unfamiliar, as they were all novel to the participants. The order of the stimuli, which set was observed and executed, and the order of conditions were counterbalanced across participants.

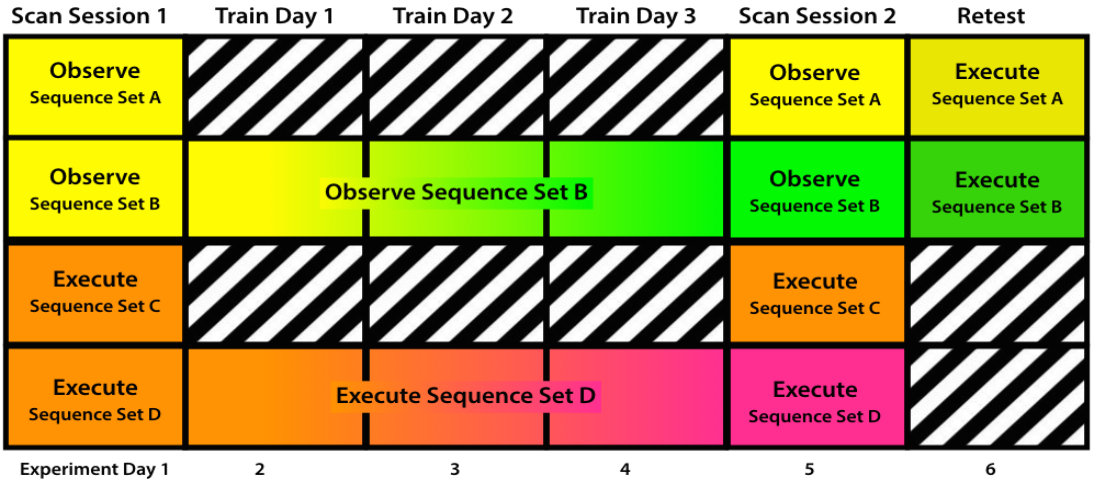


Figure 5.3. The timeline of both training and test days for both conditions. The gradient of colour illustrates the change in familiarity due to training. The Retest occurred after scan session 2, not necessarily the day immediately after.

Days 2 through 4 of testing were the training phase (yellow to green and orange to pink boxes in *Figure 5.3*). During these days, participants were invited to the lab and asked to perform the same tasks as those they completed in the scanner, on precisely half of the execution sequences and half of the observation sequences. The four execution sequences and four observation sequences were performed or observed four times per training day, and the order of practice was counterbalanced across condition and sequences. The training set up in the laboratory was designed to match that of the scanner as closely as possible. Due to the fragile nature of the scanner-safe guitar, a standard 4-stringed bass guitar was used in the training sessions to eliminate risk of damage to the equipment. Participants were required to lay on a table with the guitar placed over their midriff – similar to how the guitar was positioned in the scanner, and participants viewed the 24-inch iMac screen through prism glasses (this can be seen in *Figure 5.4*). Stimulus presentation and response collection were performed using Psychophysics Toolbox (v3) via MATLAB R2015a (MathWorks). The instruction was also given to keep movement to a minimum (for example, no tapping the guitar to the beat of the song), and the researchers monitored this. During the observation condition, participants were asked to rest their left hand over the frets so that there was no possibility that they could move their hand along with the musician's, ensuring that any learning was due to observation alone. During the execution condition, participants were asked to play the songs as accurately as they could.

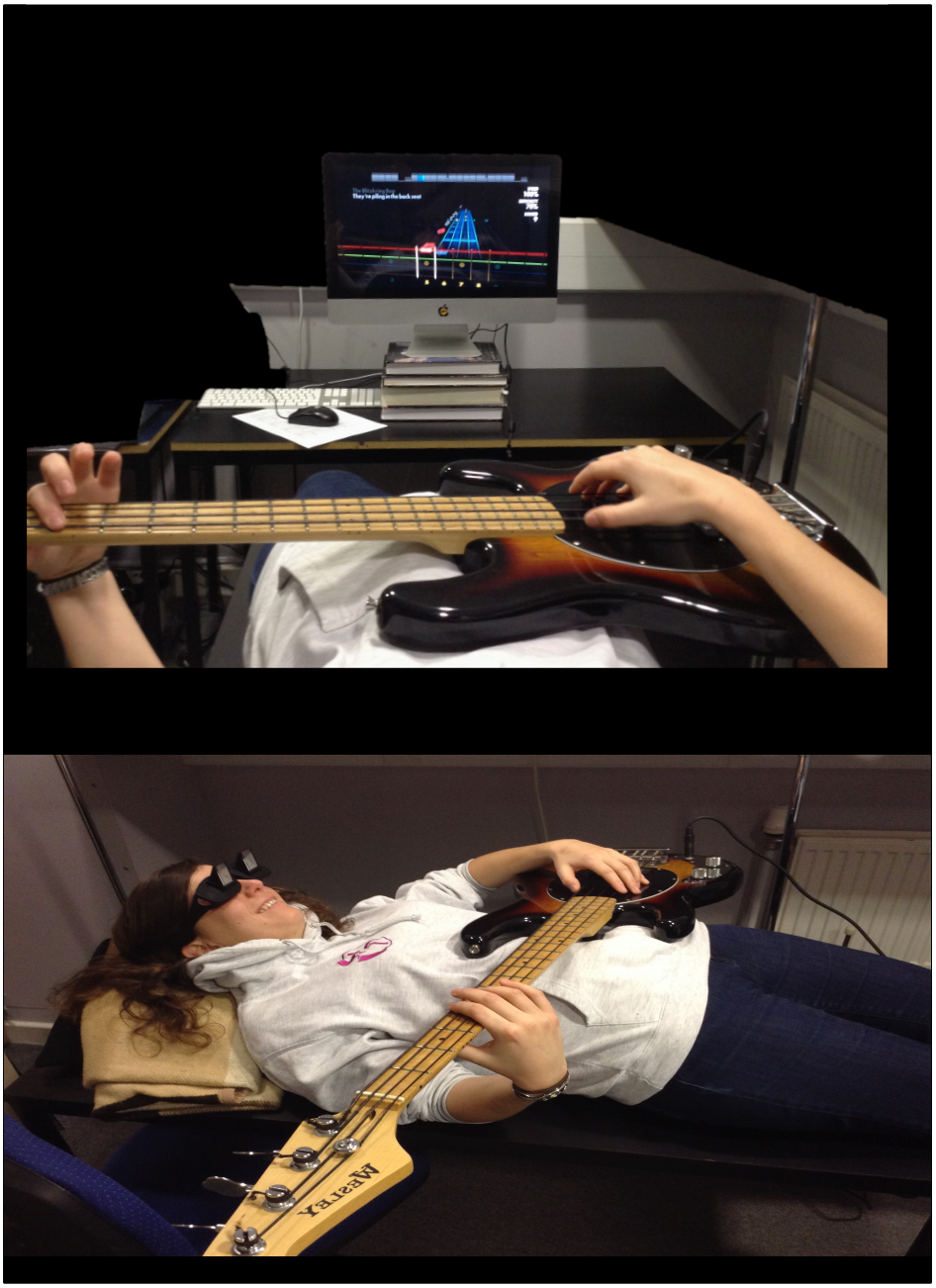


Figure 5.4. The setup for the training period. This setup was designed so that it matched the fMRI setup as accurately as possible. *N.B*. The guitar was held like a double bass, with the neck of the guitar at 30 degrees, the guitar is held at 90 degrees above for illustrative purposes only.

On the fifth day of testing, participants returned to the scanner where they completed an identical scanning protocol to Day 1, performing or observing both trained and untrained sequences. Following the second scanning session, participants were invited back to the laboratory for a sixth and final day of testing, wherein they were asked to perform all the songs from the observation condition (including the four

observationally trained and the four untrained songs; Retest, *Figure 5.3*). Due to scheduling complications, only 14 participants were able to attend this final testing session.

5.3.4 Neuroimaging procedure & parameters

Each participant completed two identical fMRI sessions on days 1 and 5 of the experiment that followed an event related design. Each scanning session featured two action execution runs and two action observation runs, presented in a counterbalanced order across participants and across scanning sessions. All eight excerpts were experienced twice per run. Each observation run lasted an average of 5 minutes (range = 5.07-5.70 minutes), and each action execution run lasted 11 minutes (range = 10.97-11.80 minutes). This difference occurred due to varying lengths of buffering between the different sequences and loading times, two factors that were not modelled within a trial nor used in matching sequence length. For the observation trials (shown in *Figure 5.2*), at the start a fixation cross was shown for 1.8s. This was followed by the video clip of the agent playing along with the excerpt (audio was included here and during the training period). After each clip, a black screen was presented with the question "did the musician seen in the video make a palming action over the strings?" The question screen was displayed for 2s before moving on to the next trial. Participants were required to respond to the question within that 2s window by plucking the appropriate string to denote their answer. During the action execution runs, participants first saw a brief interval where the song was being loaded. This aspect was unavoidable as we wanted to gain actual response accuracy via the game, so had to load each song as if it were selected by the user (the transition between menu and sequence to play was automated via a MATLAB script, and the entire load time before each sequence ranged between 19.02 and 41.45 seconds). Once the appropriate sequence was selected, there was a buffer supplied via the game so that there was an adequate amount of time before the participant began performing the sequence, allowing for finger position adjustment before each execution sequence began. After playing along with each sequence, participants' accuracy scores were

displayed for participants to see before returning to the menu screen to begin the next trial.

Stimulus presentation and response collection were performed using Psychophysics Toolbox (version 3) through MATLAB R2015a (MathWorks) run via a MacBook Pro laptop computer. The stimuli were presented on a 24-inch LCD BOLDScreen (Cambridge Research Systems), which was visible to the participant via a mirror mounted on the head coil. Participants listened to the song excerpts through Phillips MR-compatible headphones.

Participants were given a MR-compatible bass guitar to make their responses during the execution and observation runs. The guitar was a full-length bass guitar. Participants were positioned into the scanner bore slowly and shown the best way to hold the bass guitar so as not to damage the guitar or the head coil. The guitar worked via a piezo-pickup embedded in the head of the guitar, under the strings (which were made of nylon). The guitar's tuning pegs were manufactured via 3D printing using a glass/plastic alloy. The output of the guitar was passed along a fibre optic cable from the scanner room to the control room where it was amplified and fed into the MacBook Pro running MATLAB. Offline, the responses for the observation condition were filtered to remove any RF interference created by the scanner. The gameplay applied filtering for the execution condition so that the note being played could be heard by the participant.

Data acquisition was conducted at the Bangor Imaging Unit at Bangor University, Wales. Functional images were acquired on a 3.0T Phillips MRI scanner using a SENSE phased-array 32-channel head coil. Functional images were acquired covering the whole brain using an echo-planar imaging (EPI) sequence (35 axial slices, ascending slice acquisition, repetition time = 2000 ms, echo time = 30ms, 90° flip angle, matrix = 64 × 64, slice thickness: 3 × 3 × 3 mm, field of view (FOV): 224 mm). Before the functional run, 196 two-dimensional anatomical images (256 × 256

pixel matrix, T1-weighted) were obtained for normalization ROI selection and manipulation

5.3.5 fMRI data analysis

The total number of functional scans collected for the observation runs ranged between 156 and 178, and between 316 and 340 scans for each execution run. This variation arose due to varying loading times of the stimuli between the counterbalanced groups that were beyond experimental control. The number of scans collected for each subject was identical across scanning days. Data were analyzed using Statistical Parametric Mapping (SPM12: Wellcome Trust Centre for Neuroimaging, London; Friston, 2007), implemented using MATLAB R2015a (MathWorks). The data were first realigned and then slice-time corrected and preliminarily preorientated within standard stereotaxic space as defined by the MNI (Friston, 2007). This preorientation allowed for better spatial normalization to the MNI template. Participants' EPI images were then coregistered to their T1 anatomical scans, which were then spatially normalized to standard stereotaxic space. The spatially normalized EPI images were filtered using a Gaussian kernel of 8 mm full-width at half maximum in the *x*, *y*, and *z* axes. For the observational runs, the design matrix was fitted for each subject with a single regressor for the familiar stimuli, a single regressor for the unfamiliar stimuli and a single regressor for the fixation and responses. For the execution runs, this setup was the same with the inclusion of a loading period rather than the fixation and response regressor. The 4 runs (2 observation, 2 execution) were placed into the same design matrix, enabling us to create a total of two design matrices per participant; one pre- and one post-training. All brain regions that emerged from analyses were identified via the Anatomy Toolbox (Eickhoff et al., 2007).

5.3.5.1 Dynamic Causal Modelling

Evaluation. Dynamic causal modelling (DCM) is a method of analyzing effective connectivity that uses a bilinear model of neural population dynamics,

combined with a hemodynamic model, which aims to describe the neural activity in the measured BOLD response (Friston et al., 2003). By modelling feasible neuronal parameters, DCM aims to make a modeled BOLD response that is similar to the actual experimentally manipulated BOLD signal. The neural dynamics model created using DCM is then combined with a hemodynamic forward model that incorporates a balloon model (Buxton et al., 1998). The hemodynamic model provides a transformation of how the neural dynamics would propagate as a BOLD response, estimated via a standard Bayesian approach (variational Laplace). The fit of the combined neural model and hemodynamic forward model is estimated via a Bayesian approach, which uses conservative shrinkage priors for the coupling parameters (Friston et al., 2003).

A model is specified to fit the data estimated by the previously mentioned routine. The input into the estimation procedure is three matrices (for bilinear DCMs, but see also Nonlinear DCMs; Stephan et al., 2008). The first is the endogenous connection strengths (the A matrix), which represents the connectivity between the regions of the model, and is sometimes referred to as the fixed connectivity. The B and C matrices represent the experimentally manipulated conditions. Matrix B represents the modulation of an external input on a fixed connection, which describes the change in the value of the effective connectivity for a connection under a particular condition. The C matrix represents the input into the system, which describes the activity that is perturbing, or creating activity in, the system. This equation models the state changes by known inputs.

Hypothesis: The experimental hypothesis tested in the current study is that connectivity will change between the core regions of the AON as a function of familiarity over time, in directions predicted by the predictive coding model of AON function (Keysers & Perrett, 2004; Kilner et al., 2007a,b). Specifically, when the actions are unfamiliar, we expect to see the presence of modulation in an anterior direction (Temporal➔Parietal➔ Premotor; a pattern of connectivity that would reflect a high level of prediction error or imprecise/underspecified predictions about how an action should unfold in real time). As sequences become familiar through physical or observational experience, we expect modulations to become more precise

(reflecting more accurate predictions for the observed or executed actions), as demonstrated by increased modulation in the posterior direction (Premotor➔Parietal ➔Temporal).

5.3.5.1.1 Definition of Regions of Interest (ROIs)

The ROIs were identified from data collected during the first run of the second scanning session using the contrast (familiar and unfamiliar execution AND familiar and unfamiliar observation) > (implicit baseline). The rationale behind this was that this contrast would encapsulate the activity for both familiar and unfamiliar sequences as this run followed the training period. The use of the second run within this scanning session would be less of a true test of brain activity when observing or executing unfamiliar guitar sequences, as the unfamiliar sequences would have been practiced during the first run, and would thus be (slightly) more familiar. Therefore, drawing ROIs from the first run of the post-training scan session should best reflect the polarities of familiarity within our design.

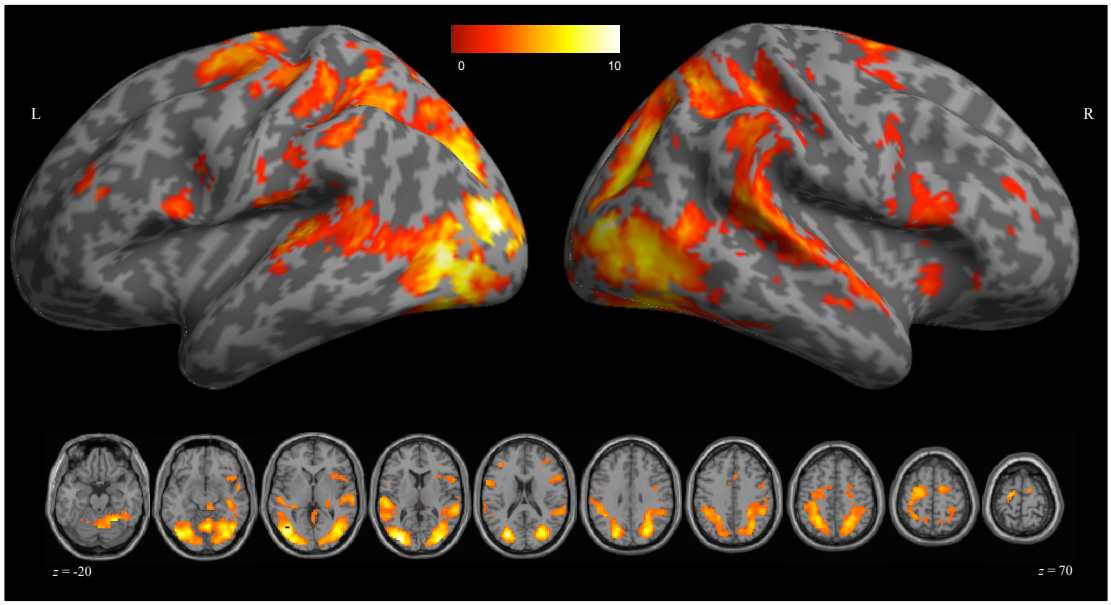


Figure 5.5. Whole brain group analysis Scan Session 2, Run 1, familiar and unfamiliar execution and observation trials vs. implicit baseline. All *p* values <0.0001 (uncorrected), *K*=10 voxels.

Table 5.1. Regions associated with ROI contrast.

Anatomical Region	MNI Coordinates			Putative Functional Name	T Value	Cluster Size	P_{FWE} Corrected
	X	Y	Z				
L Middle Occipital Gyrus	**-30**	**-82**	**10**	**MOG**	**11.28**	**1732**	**< 0.001**
L Middle Occipital Gyrus	-42	-70	4	MOG	9.91		
L Middle Occipital Gyrus	-27	-70	-14	MOG	8.74		
L Inferior Parietal Lobule	-36	-52	46	IPL	6.46		
R Middle Occipital Gyrus	**33**	**-85**	**10**	**MOG**	**9.93**	**1712**	**< 0.001**
R Superior Parietal Lobule	33	-67	34	SPL	8.99		
R Middle Occipital Gyrus	39	-73	1	V5	8.78		
L Superior Temporal Gyrus	**-48**	**-28**	**7**	**STG**	**7.65**	**109**	**< 0.001**
L Superior Temporal Gyrus	-54	-22	10	STG	7.03		
L Inferior Parietal Cortex	-30	-31	7	IPC	5.54		
R Superior Temporal Gyrus	**57**	**-40**	**13**	**STG**	**7.06**	**158**	**< 0.001**
R Superior Temporal Gyrus	66	-25	4	STG	6.36		
R SupraMarginal Gyrus	57	-40	22	SmG	5.82		
R Supplementary Motor Area	**15**	**-4**	**61**	**SMA**	**6.43**	**34**	**0.005**
L Hippocampus	**42**	**-25**	**-11**	**HIPP**	**6.05**	**19**	**0.032**
L Middle Frontal Gyrus	-36	35	22	MFG	6.01	12	0.083
R Superior Temporal Gyrus	**57**	**-7**	**-5**	**STG**	**5.88**	**20**	**0.028**
R Superior Temporal Gyrus	57	8	-8	**STG**	5.19		
R Inferior Parietal Lobule	**54**	**-40**	**34**	**IPL**	**5.82**	**56**	**0.001**
R SupraMarginal Gyrus	45	-40	43	SmG	5.79		
L Inferior Frontal Gyrus	**-57**	**8**	**19**	**IFG**	**5.53**	**25**	**0.015**
L Inferior Frontal	-45	2	25	IFG	5.01		

Gyrus							
R Insula	**45**	**17**	**-8**	**INS**	**5.51**	**23**	**0.019**
R Cerebellum	9	-31	-11	CERR	5.23	11	0.096
R Fusiform Gyrus	48	-34	-14	FG	5.23	12	0.083
R Inferior Frontal Gyrus	51	14	16	IFG	4.94	11	0.096

Bold indicates cluster corrected regions at the threshold of $p_{FWE\text{-}corrected} < 0.05$.

This contrast (illustrated in *Figure 5.5*) revealed widespread, bilateral engagement of the AON (see *Table 5.1* for a full list of regions). Initially, our search was focused on the left hemisphere (as reported by Gardner et al., 2015), however upon examination of the results from this whole-brain analysis, we found bilateral AON engagement in the core nodes of this network. Specifically, we found bilateral activity within IFG, IPL, and STG[1]. Therefore, these 6 regions served as the ROIs for subsequent DCM analyses (see *Table 5.2* for coordinates of ROIs). The time series were taken from the subject level t-contrasts from all experimental runs vs. implicit baseline. This contrast revealed all regions active across the runs (within the scanning session), including the 6 ROIs noted above.

Table 5.2. Coordinates of all 6 ROIs. Mean Size of ROI (in voxels) is reported for both days.

				Mean ROI size (SEM)	
X	***Y***	***Z***	***Region***	***Day 1***	***Day 2***
-48	-28	7	L STG	21 *(4)*	33 *(5)*
57	-40	13	R STG	30 *(5)*	49 *(5)*
-36	-52	46	L IPL	35 *(5)*	62 *(4)*
54	-40	34	R IPL	33 *(6)*	57 *(6)*
-57	8	19	L IFG	28 *(4)*	31 *(5)*
51	14	16	R IFG	37 *(5)*	49 *(6)*

[1] In a previous study examining the impact of self-reported familiarity on the exchange of information within core AON nodes, we focused on MTG rather than STG (Gardner et al., 2015). However, in the current study, MTG emerged within the left hemisphere only, while activity within STG, a brain region originally proposed to be involved in the predictive coding account of the AON (Kilner et al., 2007b), emerged bilaterally. For this reason, we focus subsequent connectivity analyses on STG as the temporal node of the AON in the present study.

All ROIs were extracted by locating the nearest local maximum voxel to the group contrast of Session 2 Run 1 familiar and unfamiliar, execution and observation vs. implicit baseline. As we aimed to identify the ROIs across scanning sessions, special steps were taken. First, a 20mm search radius was used to identify each ROI. Of the 20 subjects, 3 were removed from the analysis, as they did not exhibit significant activity within the stated search radius. The next step was to ensure that the coordinates across scanning sessions fell within the same functional and anatomical regions. When looking across scanning sessions, all but 3 of the remaining 17 participants had ROIs that emerged within the specified search radius of the given coordinates, and whose coordinates remained consistent (functionally and anatomically) over time. The ROI time series for each region, for each run, for each day and for each of the 14 participants, was extracted using the eigenvariate (threshold of $p < 0.05$) with a sphere radius of 8mm and adjusted for effects of no interest (the average size of the ROIs can be found in *Table 5.2*).

5.3.5.1.2 Definition of network models

First, we set the endogenous connectivity to follow the proposed flow of information suggested in the action observation literature (Keyser & Perrett, 2004; Kilner et al., 2007a,b), so that reciprocal connectivity was established between STG-IPL and IPL-IFG (for both hemispheres). Next, we defined the parameters of the models. We allowed for stochastic effects within the model to more accurately model noise (Li et al., 2011). We also centered the input into the node giving the input a mean of zero, which means that modulating parameters can increase and decrease the fixed connectivity, as opposed to simply increasing it as when the input is always positive. To allow for biological plausibility, we opted for two- state models that allow for both excitatory and inhibitory connectivity (Marreiros et al., 2008). Because we used a two-state DCM, we were able to enforce positivity constraints on the connection strengths to lend our interpretation greater validity and biological plausibility.

Finally, we created the models of interest for Scanning Session 1 (pre-training scan) and Scanning Session 2 (post-training scan). For Scanning Session 1, the to-be-

trained and to- remain-untrained stimuli were collapsed into one modulator as all the stimuli were unfamiliar/untrained at this stage. The input into the system was modelled so that unfamiliar stimuli would have a direct effect on the neuronal activity in STG (bilaterally). In accordance with the predictive coding account, we modelled modulations by the unfamiliar stimuli on the anterior flow between STG ➔ IPL and IPL ➔IFG. For Scanning Session 2, we modelled the unfamiliar (untrained stimuli) modulator to the same connections as Scanning Session 1 (STG ➔ IPL and IPL ➔IFG; input into bilateral STG). Furthermore, we allowed for the modulation of the connections in the posterior direction (IPL ➔STG and IFG ➔IPL) by the familiar (trained) stimuli. The direct effect of this modulator was modelled to assess its influence on IFG (bilaterally). These models were created for both the left and right hemisphere networks, for both observation and execution. Separate models were created for each Run, averaged across within each scanning session for the analysis. Posterior estimates (in Hz) were extracted for each participant for intrinsic, modulatory (familiar and unfamiliar) and direct input (C matrix). To test the presence of connections/modulations, one sample t-tests were performed on the posterior estimates. Bonferonni correction for multiple comparisons was applied to the *p* values ($p = .05/4$ for the intrinsic connections and context dependent modulations; $p = .05/2$ and $p = .05/4$ for the direct influence on Scanning Sessions 1 and 2 respectively).

5.4 Results

The DCMs were created using the six regions of interest identified in *Table 5.2*: bilateral STG, IPL and IFG. The models tested the influence of unfamiliar stimuli on the anterior flow between STG ➔ IPL and IPL ➔ IFG, and familiar stimuli in the posterior direction (IPL ➔ STG and IFG ➔ IPL). Here, we report the observation condition followed by the execution condition. Tables reporting the exogenous effects and direct influence (modulatory input) are included in the main text.

5.4.1 Action Observation

Dynamic causal modelling (DCM) was conducted to explore changes in effective connectivity within the AON when visual familiarity is systemically varied. The 3-node model (within each hemisphere) is illustrated in *Figure 5.6*. The posterior estimates (in Hz) for the intrinsic connections were significant for both scanning sessions, showing significant reciprocal connectivity between STG & IPL and IPL & IFG for the observation of actions. *Table 5.3* gives the posterior estimates for the direct and context dependent modulations by the different levels of familiarity.

Table 5.3. Input and connection modulations for the observation condition

	Modulator	Connection/ Input node	Mean (SEM)	T-value	*p* - value
Scanning Session 1 Left Hemisphere	**Un**	**STG**	**-0.0079 (0.0027)**	**-3.07**	**.009**
	Un	**STG ➔ IPL**	**0.0182 (0.0072)**	**2.54**	**.025**
	Un	IPL ➔ IFG	0.0111 (0.0086)	1.28	.222
Scanning Session 1 Right Hemisphere	Un	STG	-0.0042 (0.0024)	-1.75	.104
	Un	**STG ➔ IPL**	**0.0192 (0.0070)**	**2.88**	**.013**
	Un	IPL ➔ IFG	0.0239 (0.0104)	2.30	.039
Scanning Session 2 Left Hemisphere	**Fam**	**IFG**	**-0.0104 (0.0032)**	**-3.24**	**.006**
	Un	STG	-0.0023 (0.0048)	-.48	.637
	Fam	IPL ➔ STG	0.0450 (0.0177)	2.54	.025
	Fam	**IFG ➔ IPL**	**0.0492 (0.0167)**	**2.95**	**.011**
	Un	**STG ➔ IPL**	**0.0396 (0.0105)**	**3.78**	**.002**
	Un	**IPL ➔ IFG**	**0.0569 (0.0171)**	**3.33**	**.005**
Scanning Session 2 Right Hemisphere	**Fam**	**IFG**	**-0.0112 (0.0034)**	**-3.26**	**.006**
	Un	STG	-0.0018 (0.0041)	-.44	.665
	Fam	**IPL ➔ STG**	**0.0555 (0.0149)**	**3.73**	**.003**

	Fam	IFG ➔ IPL	0.0382 (0.0143)	2.66	.020
	Un	**STG➔ IPL**	**0.0518 (0.0136)**	**3.80**	**.002**
	Un	**IPL ➔ IFG**	**0.0646 (0.0179)**	**3.62**	**.003**

Modulations which survive corrections for multiple comparisons are highlighted in bold.

For Scanning Session 1, we modelled the modulation by the observation of unfamiliar actions on anterior connections (STG **è** IPL & IPL ➔ IFG), which corresponds to the predictions of the predictive coding account (Keysers & Perrett, 2004; Kilner et al., 2007a,b; Schippers & Keysers, 2011). In the left hemisphere, the direct input into the system through STG is present. Of the tested connections, we see that there is modulation on the STG ➔ IPL connection. In the right hemisphere, the same connection is modulated by unfamiliar actions. The direct input into STG is not consistent across participants in the right hemisphere.

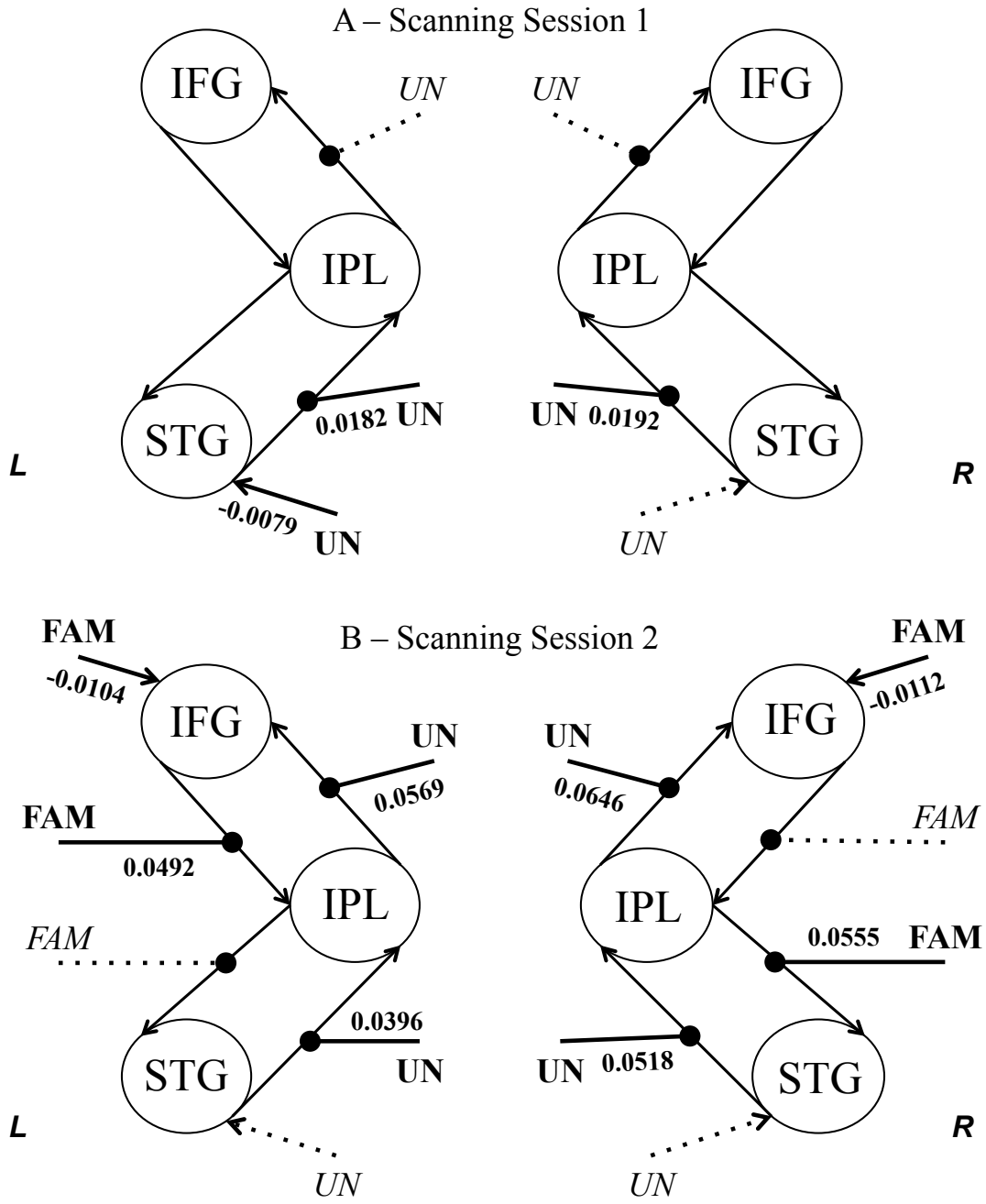


Figure 5.6. The DCM results for the observation condition; Scan Sessions 1 & 2, for both hemispheres. Arrows indicate direct input/ intrinsic modulations. Dotted lines indicate connections which are not significant. Statistics shown are for the significant modulations in mean posterior estimates (in Hz). Superior Temporal Gyrus, STG; Inferior Parietal Lobule, IPL; Inferior Frontal Gyrus, IFG.

Between the scanning sessions, an intensive training period occurred where a subset of stimuli were observed over three days (overall, each trained stimulus was viewed a total of 12 times). As such, when participants returned for the second

scanning session, we had systematically created a distinction between stimuli that remained untrained (familiar) and those that were only seen in the first scanning session and not trained (unfamiliar). The unfamiliar stimuli were modelled to modulate the anterior connections, identical to the first scanning session. Both bilateral anterior connections (STG ➔ IPL & IPL ➔ IFG) were modulated by the observation of unfamiliar actions, showing full support for the predictive coding account. It is of note that the direct input into STG was not present in either hemisphere. The familiar actions showed direct input into bilateral IFG. In the left hemisphere, we find modulation by familiar stimuli within the parietal-premotor loop between IFG ➔ IPL. The connection between IPL ➔ STG is not significant in the left hemisphere, suggesting inconsistency across participants. In the right hemisphere, the IPL ➔ STG connection is significantly modulated, whereas the IFG ➔ IPL connection does not survive correction for multiple comparisons. To summarise, the observation of unfamiliar actions modulates some anterior connections within the AON, as predicted. When an observed action is most unfamiliar (scanning session 1), anterior connections from STG ➔ IPL are present. As the task become more familiar in the second scanning session, all anterior connections are modulated when observing the less familiar/untrained sequences, providing clear support for the predictive coding account. After training, when participants watch familiar actions they observed 12 times across three days of training, evidence emerges for modulation in the posterior direction within the premotor-parietal loop in the left hemisphere and within the parietal – temporal loop in the right hemisphere. Together, these findings provide positive (albeit not full) support for a predictive coding account of the role played by familiarity in shaping AON responses.

5.4.2 Action Execution

The execution condition required participants to actively play the Rocksmith ™ game whilst in the scanner. The Dynamic Causal Modelling (DCM) procedures conducted on imaging data collected during the execution condition were identical to those run on the observation condition data, in terms of examining connections within the 3-node model (within each hemisphere). The posterior estimates (in Hz) for the intrinsic connections were all significant for both scanning sessions, showing

significant reciprocal connectivity between STG & IPL and IPL & IFG for the execution of actions. *Table 5.4* gives the posterior estimates for the direct and context dependent modulations by the different levels of familiarity.

Table 5.4. Input and connection modulations for the execution condition

	Modulator	Connection/ Input node	Mean (SEM)	T-value	*p* - value
Scanning Session 1 Left Hemisphere	**Un**	**STG**	**-0.0086 (0.0019)**	**-4.57**	**.001**
	Un	**STG➔IPL**	**0.0379 (0.0239)**	**5.93**	**<.001**
	Un	**IPL ➔ IFG**	**0.0626 (0.0154)**	**4.05**	**.001**
Scanning Session 1 Right Hemisphere	**Un**	**STG**	**-0.0073 (0.0065)**	**-4.21**	**.001**
	Un	**STG ➔ IPL**	**0.0460 (0.0078)**	**5.87**	**.001**
	Un	**IPL ➔ IFG**	**0.0606 (0.0138)**	**4.39**	**<.001**
Scanning Session 2 Left Hemisphere	Fam	IFG	-0.0064 (0.0032)	-2.00	.067
	Un	STG	-0.0039 (0.0027)	-1.45	.171
	Fam	**IPL ➔ STG**	**0.0936 (0.0155)**	**6.05**	**<.001**
	Fam	**IFG ➔ IPL**	**0.0910 (0.0128)**	**7.10**	**<.001**
	Un	**STG ➔ IPL**	**0.0455 (0.0094)**	**4.82**	**<.001**
	Un	**IPL ➔ IFG**	**0.0982 (0.0100)**	**9.83**	**<.001**
Scanning Session 2 Right Hemisphere	**Fam**	**IFG**	**-0.0085 (0.0020)**	**-4.29**	**.001**
	Un	STG	-0.0045 (0.0022)	-2.05	.061
	Fam	**IPL ➔ STG**	**0.1006 (0.0204)**	**4.94**	**<.001**
	Fam	**IFG ➔ IPL**	**0.0694 (0.0129)**	**5.39**	**<.001**
	Un	**STG ➔ IPL**	**0.0539 (0.0103)**	**5.24**	**<.001**
	Un	**IPL ➔ IFG**	**0.0949**	**6.43**	**<.001**

	(0.0148)

Modulations which survive corrections for multiple comparisons are highlighted in bold.

During the first scanning session, the execution of unfamiliar actions modulates anterior connections (STG ➔ IPL & IPL ➔ IFG), a finding that corresponds to the predictions made by the predictive coding account (see also active inference; Friston et al., 2005). Unlike the observation conditions, all the tested modulations and direct inputs are consistent across participants, indicating that a full anterior modulation of this network is recruited when executing actions that are unfamiliar.

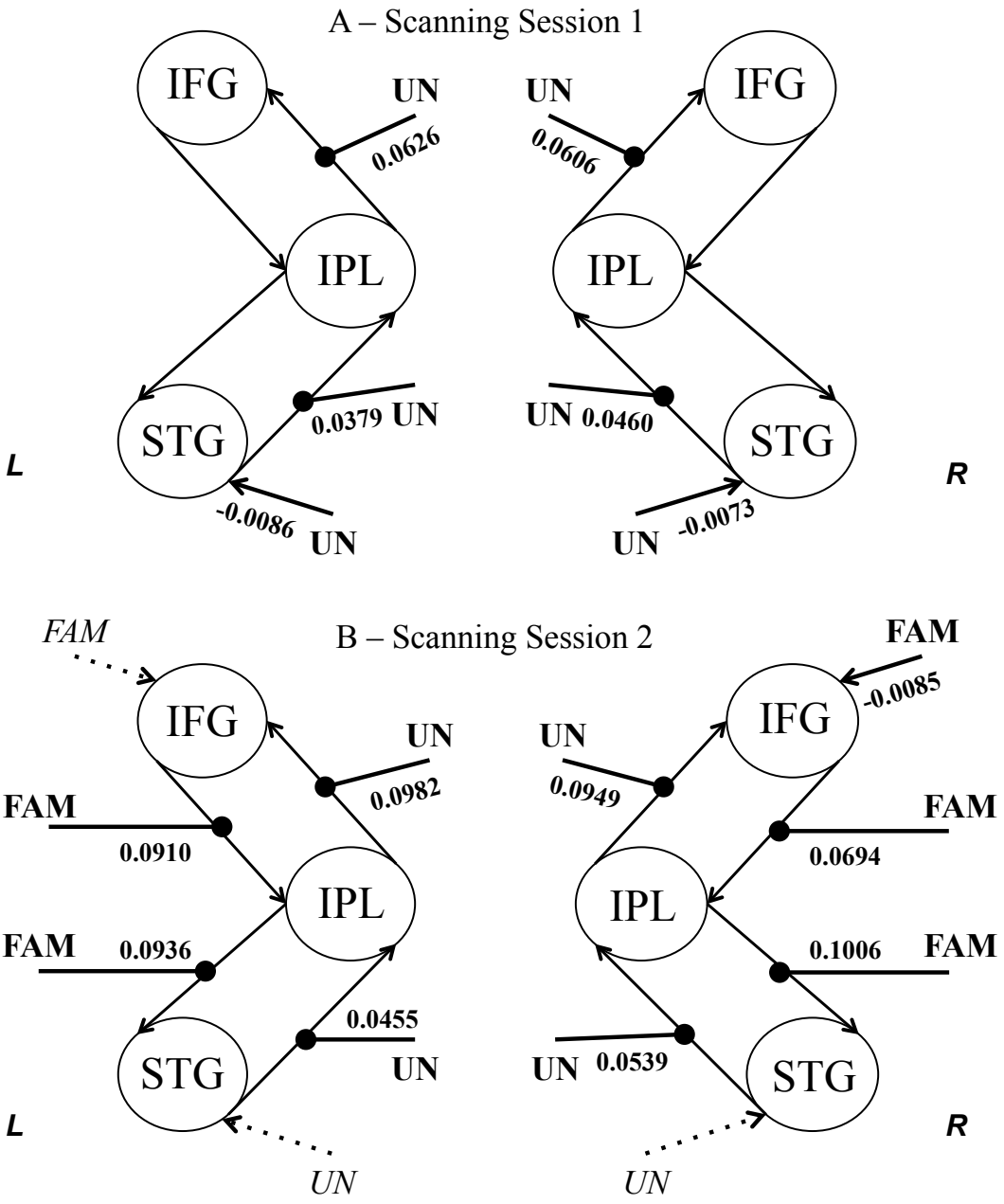


Figure 5.7. The DCM results for the execution condition; Scan Sessions 1 & 2, for both hemispheres. Arrows indicate direct input/ intrinsic modulations. Dotted lines indicate connections that are not significant. Statistics shown are for the significant modulations in mean posterior estimates (in Hz). Superior Temporal Gyrus, STG; Inferior Parietal Lobule, IPL; Inferior Frontal Gyrus, IFG.

During training, participants practiced performing a subset of guitar sequences over the course of three days, establishing the distinction between familiar and unfamiliar actions. The DCM analysis of the second scanning session data revealed that the unfamiliar stimuli modulated the anterior connections as in scanning session 1

(shown in *Figure 5.7*). The direct input into the network by unfamiliar actions no longer modulates bilateral STG. Both bilateral posterior connections (IPL➔STG & IFG➔IPL) were modulated by the familiar actions, showing full support for the predictive coding account. The direct input by the familiar actions to IFG is only present in the right hemisphere. To summarise, the execution of unfamiliar actions modulates the anterior connections within this network. When an action becomes familiar after intensive practice, modulation of the posterior connections (IPL➔STG and IFG➔IPL) within the network emerges. These findings provide strong evidence in support of the predictive coding account of how familiarity shapes AON engagement during action execution.

5.5 Discussion

By systemically controlling familiarity via a guitar-training paradigm, we tested predictions generated from the predictive coding account regarding how changes in familiarity shape AON engagement during action observation and execution. We hypothesised that when observed or executed actions were unfamiliar, anterior modulation would be present between AON nodes (shown by reliable modulation from STG➔IPL and from IPL➔IFG). As actions became familiar, we expected an anterior to posterior shift in modulation within the AON, manifested as increased posterior modulation between IPL➔STG and IFG➔IPL. The overall pattern of findings provides compelling, although not complete, support for the predictive coding account for how changes in familiarity modulate AON engagement.

5.5.1 Synthesis with the Predictive Coding Account: Action Observation

Analyses of effective connectivity between core AON nodes when observing unfamiliar actions revealed significant modulation from STG➔IPL. According to the predictive coding account (Keysers & Perrett, 2004; Kilner et al., 2007a,b; Gazzola & Keysers, 2009; Schippers & Keysers, 2011), upregulation of this connection relates to increased prediction error (i.e., the observed action does not precisely match the predicted action), creating a need for more information to be fed into the network

from the posterior (visual) node (Kilner et al., 2007a,b). This modulation was also present when participants observed unfamiliar sequences during the post-training scan session, suggesting untrained stimuli remained unfamiliar. At odds with the predictive coding account is the absence of modulation from IPL➔IFG when observing unfamiliar actions during the first scanning session (however, this modulation does emerge by the second scanning session). This discrepancy might be explained by a Hebbian learning principle (Keyser & Perrett, 2004; Keyser & Gazzola, 2014), which suggests that when learning, consistency and causality of firing between neurons increases the efficiency of their connection (Hebb, 1949; Keyser & Gazzola, 2014). The predictive coding account, on a basic level, has foundations within this framework and when applied to our data, we see anterior modulation of AON nodes becoming more complete as the *task* becomes more familiar (across scanning sessions), and as differences between familiar and unfamiliar sequences become established. When observing others in action, STG provides visual input to IPL, which in turn modulates IFG, an area proposed to play a role in higher-order representations of others' actions (Press et al., 2012; Kilner et al., 2009a). Although the data from the present study are collected across two separate scanning sessions, when considered as a whole, the finding of increased anterior modulation within the AON when observing unfamiliar action sequences is broadly consistent with the predictive coding account.

While rarely the focus of AON studies (c.f., Caspers et al., 2010), hemispheric differences that emerge in the present study warrant closer consideration. Greater familiarity increased modulation between IFG➔IPL in the left hemisphere, and from IPL➔STG in the right hemisphere during action observation. According to the predictive coding account, both modulations should be present (ostensibly in both hemispheres, though the primary predictive coding literature does not specify this). As such, our findings provide only partial support for a predictive coding account predicated on similar functioning of the AON across both hemispheres. Previous findings offer some insight as to why increased modulation from IPL➔STG in the left hemisphere might be absent. Although predictive coding hypothesises upregulation of this connection with increasing familiarity, we recently showed this connection to *attenuate* with increasing familiarity in the left hemisphere during

action observation (Gardner et al., 2015). Considering the current findings together with that of Gardner et al. (2015), effective connectivity analyses reveal mixed support for the observation of increasingly familiar actions to lead to increased influence in a posterior/feedback direction within the AON. A challenging task for future work will be to further characterise the nature of the prediction error signal (see Rao & Ballard, 1999; Friston, 2005; Brodski et al., 2015), and to more closely scrutinise hemispheric similarities and differences in predictive coding.

5.5.2 Synthesis with the Predictive Coding Account – Action Execution

It is of note that the predictions derived from the predictive coding account for action execution (formally the active inference principle; Friston et al., 2005) are by and large supported during pre- and post-training scanning sessions, and across both hemispheres. In other words, we find evidence for anterior modulation by unfamiliar actions throughout the AON and posterior modulation for familiar actions. According to active inference, the execution of a goal-directed action should be reflected in a minimising of free energy (Friston et al., 2010). This concept is echoed by minimisation of prediction error within the predictive coding framework (Keysers and Perrett, 2004; Kilner et al., 2007a,b). Within this framework, forward models (Wolpert et al., 1995, 2003; Wolpert & Miall, 1996) predict the sensory consequences of actions, which leads to fine-tuning of movements (Blakemore et al., 2003). This notion is reflected in our results by modulation in an anterior direction by unfamiliar actions. Once the motor codes required to execute an action become familiar (after three days of intensive training within our design), the network should maintain a homeostatic state (Pezzulo et al., 2015). This state is created by minimising free energy, which results in accurate motor commands to carry out the desired action (through inverse models). Thus, once an action is familiar, the posterior modulation by familiarity seen in our results reflects (more) accurate comparison between predicted and actual sensory consequences, resulting in minimised free energy (Friston et al., 2010).

5.5.3 Limitations and future directions

One potential limitation concerns our focus on a single model for evaluating our hypotheses particularly when a strength of DCM is its ability to compare competing models (Friston et al., 2003, Stephan et al., 2010). Here we focus on a single model because of strong predictions about how connectivity (anterior and posterior modulations) should be shaped by changes in familiarity, based on prior theoretical and empirical accounts of predictive coding (Keysers & Perrett, 2004; Kilner et al., 2007a,b; Gazzola & Keysers, 2009; Schippers & Keysers, 2011; Gardner et al., 2015). This method of examining one model with DCM, also used by others to examine AON function (c.f. Thioux & Keyser, 2015), respects the recommendation within the DCM literature to keep models simple (Stephan et al., 2010). This is also reflected in our use of a 3-node model that closely adheres to the AON's neurophysiology (Kilner et al., 2007a,b). While meta-analyses reveal additional regions active during action observation and execution (Caspers et al., 2010 & Rizzolatti et al., 2001, 2014), by focusing on the three core AON nodes, and evaluating specific predictions about modulation by familiarity, we are able to constrain the exponential expansion of model space (see Stephan et al., 2010), and thus perform more targeted, hypothesis-driven investigations of our theoretical model.

The present findings raise several important questions for the AON literature more broadly. While we found strong support for the predictive coding framework during action execution, support was more mixed during action observation. Further work could directly examine how changes in familiarity shape AON engagement during action observation compared to execution, as well as more clearly define how prediction error occurs (see Rao & Ballard, 1999; Friston, 2005; Brodski et al., 2015 for debates on the nature of this signal). While the present study underscores the importance of exploring how the core nodes of the AON influence each other when observing or executing familiar compared to unfamiliar actions, our data provide only a snapshot of how this connectivity unfolds. We report partial confirmation for predictions derived from a predictive coding model, but the lack of full support indicates that more work is required to either update the theoretical models or, more precisely, specify how experience-driven plasticity influences reciprocal

communication between core AON nodes. A further challenge remains to more precisely delineate the active role played by each node within this network, perhaps by combining effective connectivity analyses with other multivariate, data-driven approaches, such as MVPA (Abrams et al., 2013; Coutanche & Thompson-Schill, 2013). A final question worth exploring in terms of understanding the relationship between experience and AON engagement concerns what happens if familiar stimuli become unfamiliar again (through lack of practice and/or across time). It would be illuminating to determine whether anterior modulation across the network re-emerges as lack of practice causes posterior modulations to fade.

5.6 Conclusions

The present findings generally support a predictive coding account of AON function, as well as offer insights into how systemically controlled familiarity modulates effective connectivity within this network. By using an intensive training paradigm, enabling tight control of visual and physical familiarity, we found distinct patterns of posterior and anterior modulation of the AON to emerge when observed or executed actions were more or less familiar, respectively. Our findings highlight many avenues for future work, including examining how familiarity modulates activity within individual AON regions, as well as how dynamic fluctuations in familiarity shape effective connectivity over longer time periods. As a whole, the present study demonstrates the empirical utility of combining a training paradigm with effective connectivity analyses to explore the complexities of how we perceive and perform actions in a social world.

6 Chapter VI

General Discussion

The empirical chapters presented in this thesis aimed to examine the role of familiarity on action observation network (AON) engagement during action observation and action execution. The work generated as part of this thesis explored how varying levels of familiarity modulate connectivity between the core nodes of this network using functional neuroimaging approaches. In the first empirical chapter (Chapter III), we found attenuation of influence between posterior nodes of the AON with increasing familiarity of observed actions (Gardner, Goulden & Cross, 2015). In the second empirical chapter (Chapter IV), we used regression analyses to explore AON regions of interest and found that both linear and nonlinear functions capture these regions' response profiles to increasing familiarity during action observation *and* execution (Gardner, Aglinskas & Cross, under review). In the final empirical chapter (Chapter V), we coupled an intensive training paradigm and multiple scanning sessions with effective connectivity to test a predictive coding account of how the AON is shaped by familiarity, and found compelling (though incomplete) support for this account (Gardner and Cross, under review). Below, the main findings from each chapter are briefly summarised. This is followed by consideration of the broader implications of these findings on the action observation literature, with a special focus on the relationship between familiarity and AON engagement, particularly how the present findings inform action understanding models, and the utility of rich, ecologically-valid action stimuli and training paradigms in this line of research. Methodological implications arising from the empirical work of this thesis are then discussed, and the thesis concludes with suggestions for future work.

6.1 Summary of findings

The study reported in Chapter III examined how participants' familiarity with observed movements modulates effective connectivity (measured by dynamic causal modelling; DCM) between core nodes of the AON. During scanning, participants viewed whole-body dance movements, and then rated each dance movement on a measure of visual familiarity after scanning. The main GLM result of this study was that videos rated as increasingly familiar were associated with more robust engagement within core regions of the AON. Participants' ratings were also used as parametric modulators in the DCM models, which revealed an attenuation of effective connectivity bidirectionally between parietal and temporal AON nodes when participants observed videos they rated as increasingly familiar. As such, the findings provide partial support for a predictive coding model of the AON, as well as illuminate how action familiarity manipulations might be used to explore simulation-based accounts of action understanding.

In Chapter IV, a study was conducted to examine the relationship between AON response amplitude and objective and subjective measures of familiarity with observed and executed actions. The region of interest (ROI)-based approach featured in Chapter IV enabled direct testing of prominent models of action understanding, whereby a linear relationship between familiarity and AON amplitude would support the direct matching hypothesis and a quadratic/cubic relationship between familiarity and AON amplitude would be more consistent with the predictive coding account. Using an elaborate guitar training intervention and pre- and post-training scanning sessions, participants executed one set of guitar sequences and observed a second set of sequences during scanning. Via ROI analyses, linear, quadratic & cubic regression models were fitted to varying levels of familiarity (determined by the number of exposures participants had for each sequence). From both observation and execution conditions, the data show evidence for linear and quadratic models, suggesting that the response profile within key sensorimotor brain regions associated with the AON is not solely linear in nature in response to familiarity. Moreover, by probing the subjective nature of the prediction error signal, we show results that are consistent

with a predictive coding account of AON engagement during action observation and execution.

Finally, Chapter V used effective connectivity analyses to test predictions derived from the predictive coding account of AON function when participants observed or executed familiar and unfamiliar actions. To test these predictions, we used a targeted DCM approach to the same data collected for Chapter IV, which used an intensive guitar-training paradigm to enable rigorous manipulation of familiarity in visual and motor domains. After identifying core AON nodes from pre- and post-training scanning sessions, we then analysed effective connectivity to test the hypothesised modulations between AON nodes made by the predictive coding account. We demonstrated that predictive coding hypotheses of distinct patterns of modulation based on perceived or performed actions' familiarity are generally supported by the empirical data, and contribute valuable insights toward understanding the complex role played by familiarity in modulating action cognition.

In sum, the results from all three empirical chapters suggest the relationship between familiarity and AON engagement is, primarily, nonlinear in nature. Furthermore, how varying levels of familiarity impact connectivity between core regions of the AON generally fits with the predictive coding account of action understanding.

6.2 Implications for the role of familiarity plays within the AON

6.2.1 Mapping the shape of increases in familiarity and AON engagement

In Chapter III, the main GLM result was that when using participants' offline ratings of familiarity, dance movements that are rated as increasingly familiar were associated with greater AON engagement. In relation to the AON literature, this finding relates to the action being within the physical repertoire of the observer (Calvo- Merino et al., 2005; Cross et al., 2006; Cross et al., 2009), and is therefore consistent with the notion that greater AON engagement occurs when observing an

action that is more "like me" (Meltzoff, 2007). However, this finding appears to contradict what we report in the ROI analyses of Chapter IV. By using an objective measure of familiarity (in this case, the number of exposures), in this subsequent training study, we found the relationship between familiarity and AON response was not solely linear. Specifically, watching or performing actions that were highly unfamiliar or highly familiar resulted in more robust AON engagement than actions that were an intermediate level of familiarity in certain regions, most notably PMd. Moreover, and more directly comparable to the result of Chapter III, when using participants' subjective ratings as the predictor within the regression model, the quadratic relationship between familiarity and AON engagement remained. Taken together, these findings represent a microcosm of the complex debate concerning familiarity and AON engagement when using univariate approaches.

A key focus of the action understanding literature concerns shared representations between observed and executed actions (Gallese & Goldman, 1998; Rizzolatti et al., 2001). In Chapter IV, we were able to use a complex task that could be performed in the fMRI scanner, allowing us to identify regions that were engaged during action observation or action execution. The response profile within these regions (which spanned bilateral sensorimotor regions of the AON) for both conditions was both linear and nonlinear across regions, partially showing greater engagement for familiar and unfamiliar actions, and less engagement for actions that were somewhere in between. Returning to the mirror neuron literature, this pattern of findings has high face validity, in that if particular neurons fire both during action observation and execution in a similar manner, then it seems logical that these same neurons might be shaped by variations in familiarity with executed or observed actions in a similar fashion. Comparable models of observation and execution, for example predictive coding and active inference, also support this conclusion (Kilner et al., 2007a;b; Friston et al., 2005)

The GLM findings reported in Chapter III do not fit the proposed U-shaped (or quadratic) function of AON engagement and familiarity (Cross et al., 2012; Liew et al., 2013), whereby the highest AON activity emerges when observing a highly familiar *or* highly unfamiliar action. The direct comparison between linear and

nonlinear response profiles reported from the training study in Chapter IV revealed mixed evidence across ROIs, suggesting that both linear and nonlinear relationships are present. If we were to reconcile the data from this thesis with these proposed models, it could be that some mix between linear and quadratic models best defines the data (for example, a skewed or "J-shaped" quadratic model or cubic model, with greater BOLD engagement for highly familiar actions than for unfamiliar actions). Such a modified model should capture the positive quadratic function as a function of familiarity, yet the profile of this function would not be fully symmetrical. Rather, the finding of Chapter III would suggest that the "curve" on the more familiar side of the function would be steeper than the less familiar side (similar to an inverted version of the model proposed by Diersch et al. (2013)).

Further investigation into the complex nature of the relationship between AON engagement and familiarity is required, specifically, how best to characterise or capture the response profile to increasing familiarity. The extant data from the literature combined with the new data generated by this does not yet enable us to achieve this goal, due to the fact that not all points on the curve are represented. To this end, future work would be advised to examine more levels of familiarity (thus filling in more points on the curve/line), to generate a more accurate and complete picture of this profile when moving from unfamiliar to familiar actions (c.f., Kirsch & Cross, 2015). Furthermore, the complexity of the relationship between familiarity and AON engagement also supports the use of multivariate, connectivity measures to further investigate the interplay between regions, as magnitude-based approaches (such as those reviewed in this section) might only get us so far.

6.2.2 Synthesis with models of action understanding

To briefly recap, the prominent models of action understanding investigated within this thesis were the direct matching hypothesis and the predictive coding account. The direct matching hypothesis (Rizzolatti et al., 2001; Gallese & Goldman, 1998; Wolpert et al., 2003; Umilta et al., 2001; Kohler et al., 2002) suggests that an action's meaning is understood via the AON, which supports simulation of an observed action by matching the goal or intention of an observed movement with

what the goal or intention of that same movement would be if performed by the observer. Predictive coding models of AON function (Keysers & Perrett, 2004; Kilner et al., 2007a, b; Gazzola & Keysers, 2009; Schippers & Keysers, 2011), predicated on the use of perceptuomotor maps to predict and interpret observed actions (Lamm et al., 2007; Schubotz, 2007; Urgesi et al., 2010), suggest a Bayesian comparison of observed (or executed) and expected actions; manifested as exchange of information in an anterior direction for unfamiliar actions and in a posterior direction for familiar actions. In the following, the findings of this thesis are discussed in relation to these models of action understanding.

The findings of Chapters III and V provide partial support for the predictive coding account of action understanding. In Chapter III, the connectivity analysis results showed that when an action was rated as more familiar, MTG exerts an attenuated influence on IPL in an anterior direction (precisely as would be predicted under predictive coding), and an attenuation of influence was also observed in the posterior connection from IPL to MTG (not explicitly predicted under predictive coding). Together, these findings illustrate a version of the predictive coding account in which outputs from, and inputs to, MTG attenuate with increasing familiarity, whereas reciprocal influence between parietal and premotor areas is less impacted by increasing action familiarity. An alternative explanation for this finding of attenuation of the influence of the posterior node of the AON is that as observed actions become more familiar, the intentions are more easily understood, allowing for more precise predictions and less error. This would also support the view of Hamilton and Grafton (2006) and other studies using repetition suppression paradigms to study AON function (Dinstein et al., 2007; Hamilton & Grafton, 2008; Kilner et al., 2009a; Ramsey & Hamilton 2010), whereby repeated presentation of action-relevant stimuli causes habituation of the BOLD response within the AON. In addition, this finding could be explained by considering the range of excitatory and inhibitory influences between nodes within a predictive coding framework (Keysers and Gazzola, 2014). The hypothesis of the predictive coding framework suggests that there is anterior modulation for unfamiliar actions and a posterior modulation for familiar within the core regions of the AON (Kilner et al., 2007b). However, the nature of the signals is

not described, for example, it is not clear that a high prediction error necessarily equates to a greater influence between nodes.

According to the predictive coding theory (Rao & Ballard, 1999), a mismatch between actual and expected action outcomes, which results in high prediction error, should equate to increased neural activity. The findings of Chapter III generally support this claim (between MTG and IPL), as more familiar actions should create less prediction error, demonstrated by attenuation of connections (c.f., Mumford, 1992; Rao & Ballard, 1999; Friston, 2005). However, the ROI's which show a quadratic or cubic relationship between familiarity and AON engagement reported in Chapter IV challenges this point, as robust AON engagement was seen for both highly unfamiliar and familiar actions. These findings from Chapter IV support the view of signal suppression due to increased prediction error (Grossberg, 2007, 2013; Carpenter & Grossberg, 2010). A question remains, however, concerning how there can be increased activity for both high and low prediction error.

Based on the data presented in this thesis, we might speculate that a familiar action should generate less prediction error (as shown in Chapter III), but as an action reaches a high level of familiarity, this dynamic system is able to evaluate predicted and actual actions at an even finer level of detail, in an online fashion (Blakemore & Frith, 2005; Falck-Ytter et al., 2006; Cross et al., 2012). This possibility requires further investigation to validate. If validated, it would support the quadratic framework proposed by Cross et al. (2012) and Liew et al. (2013), and would also fit well with broader considerations of the brain as a dynamic, predictive organ (Bubic, 2010). This idea can be further explored by considering the coffee cup example introduced at the start of this thesis. If we observe our friend reach out to grasp the coffee cup, an action we are familiar with, our predictions of this action, which we have performed countless times ourselves, should be strong and accurate, thus creating little prediction error. However, as the AON is effectively simulating our friend's action in an online manner, and comparing this to the actual observed action, the simulations must have some flexibility built into their predictions, allowing for the intention of the action to change. If our friend reached toward the cup and then continued her reach towards the sugar, the mechanisms responsible for predicting this

action must remain agile enough to reconfigure the predictions, otherwise, our "predictive" brains would be running a finite number of simulations to gain the intention, not adaptive to what is actually being observed. Furthermore, as reaching for a coffee cup is a highly familiar action, any deviation in the act, for example a small hesitation or tremor, this would create differences between the observed and expected, equating to more robust activity. Therefore, these assumptions could explain why there is a quadratic function between familiarity and AON activity; yet, this presumption would need further investigation.

In Chapter IV, we set out to directly compare the direct matching hypothesis with the predictive coding account of action understanding, in terms of how both theories could account for variations in action familiarity. As described above, we found positive evidence for the predictive coding account over the direct matching hypothesis as there was not consensus linear activity across ROIs. The findings of this study have two critical implications. First, we showed via a simple regression analyses on the ROIs both linear and nonlinear relationships between familiarity (in terms of number of times participants were exposed to an executed or observed action sequence) and amplitude of AON response. Second, and importantly for internal consistency, we showed that when using a subjective rating scale (identical to that used in Chapter III), this pattern remained. This second point validates the methodological approach (and, consequently, findings) from Chapter III.

To precisely test the predictions about familiarity derived from the predictive coding account, we used a highly theoretically-motivated DCM model in Chapter V. Within this model, familiar actions were set to modulate posterior connections and unfamiliar actions were set to modulate anterior connections. On the first scanning session (pre-training), when all actions were novel and highly unfamiliar, the anterior connection between STG and IPL was modulated by unfamiliar actions, yet the IPL to IFG modulation did not reliably emerge across participants. After training, the STG to IPL connection remained modulated by unfamiliar actions, and now the IPL to IFG modulation emerged when watching unfamiliar actions (supporting predictions from the predictive coding account). For the familiar (trained) stimuli, the posterior connection between IFG and IPL was also significantly modulated, whereas the IPL

to STG connection was not, providing partial support for the predictive coding account. From our findings, the main question to arise here is why we fail to find full support for the predictive coding account. Moreover, we might ask whether the predictive coding account is able to provide a complete (and accurate) framework for understanding the role of familiarity with the AON.

To begin to address this question, we can revisit particular findings from Chapter V. First, we can consider unfamiliar actions when they are most unfamiliar (i.e., the pre-training scanning session). As outlined in the discussion of Chapter V, the absence of an IPL to IFG modulation could be explained by a Hebbian learning principle (Hebb, 1949; Keyser & Perrett, 2004; Keyser & Gazzola, 2014). As an observed action is extremely novel, a generative model is required in order to create a high order representation of the action (Kilner et al., 2007b). By the second scanning session, this representation is beginning to be formed, and even as the action remains relatively unfamiliar (compared to the trained actions), the IPL to IFG connection emerges. After training, the familiar stimuli modulate the IFG to IPL connection, as predicted by predictive coding. In Chapter III, it was shown that this connection was not subject to modulation by increasing familiarity, while the findings from Chapter V show this connection to be modulated when an action *becomes* highly familiar. Therefore, the predictions of the network, once formed through experience, appear to be projected posteriorly from IFG. In terms of the absence of a modulation between IPL to STG, again we can reference the findings of Chapter III, where this connection *was* modulated by increasing familiarity (attenuated). When an action reaches a level where the prediction of the network successfully matches the observed, the modulatory influence from IPL to IFG attenuates, and in Chapter V, this is indirectly shown as an absence of a significantly positive connection across participants.

To summarise, the evidence presented in this thesis provides general, if not complete, support for the predictive coding account. Although the evidence does not perfectly support the account as proposed in the literature (Kilner et al., 2007b; Friston et al., 2005), it provides a valuable insights into how this framework might be applied to empirical data and using complex stimuli and training manipulations, as

well as an elaborated understanding of how the prediction error signal is shaped by modulations in familiarity.

6.3 Methodological considerations: complex actions, training paradigms and objective and subjective ratings of familiarity

Throughout this thesis, a number of complex actions and experimental manipulations were used to explore AON functionality. The reason we used whole body movements in the observation task featured in Chapter III is that such movements provide rich form and motion information (Cross et al., 2009; 2012). Moreover, the actions we observe others perform in our daily lives are rarely limited to single body parts (i.e., just hands or just feet); therefore the use of whole body movements provides ecological validity and social relevance. Clearly, a challenge of using such a rich source of action information is that it draws into question where participants are fixating when observing the action. Laβberg and colleagues (2014) investigated the role of gaze when viewing whole body movements. Using eye tracking while participants observed gymnasts somersault, they found tight coupling between participants' eye gaze and the gymnast's eye, head and spine in anticipation of actions. Thus, when observing a complex action, participants were aiming to understand the kinematics and goal of the actor via spinal orientation and the gymnast's gaze. Moreover, Green and colleagues (2014) also showed that adults (but not infants) used anticipatory gaze when observing goal directed actions. Applied to the use of whole body movements in this thesis, we might assume that participants were predicting the actions of the dancer to anticipate how he might next arrange his limbs. Furthermore, the use of complex, whole-body movements performed by a professional dancer allowed for a wide range of movement familiarity and complexity (controlled for by motion energy). This variety in complexity helped create a robust measure of familiarity to investigate.

The reason we shifted to studying hand movements instead of whole body movements in Chapters IV and V was that we were interested in both action observation and execution in these chapters, and thus an action that could be performed in the scanner was required. Rather than simple reaching actions, which

are often used in the action observation field (c.f. Tai et al., 2004; Gazzola et al., 2007; Liew et al., 2013) we opted for a more complex category of hand action: guitar playing. The use of a guitar playing actions has been used in AON studies previously (Buccino et al., 2004; Vogt et al., 2007) and, like dance, it also offers a rich form of action information. Playing the guitar requires participants to perform a sequence of simple finger actions, coupled with appropriate timing and forward planning. Such complex actions composed of simple sub-actions have been shown in other domains, for example, tennis playing (Basler et al., 2014) or golf (Beilock & Gonso, 2008; Kim & Cruz, 2011). Crucial to these types of complex actions is that a novice can learn to perform aspects of the action in isolation, even if he or she cannot necessarily perform a unified version or sequence of the full action in a coherent manner. This was important for the use of guitar playing in this thesis as we wished to use complete guitar-playing novices, for whom the skill we wished them to master was attainable (not to a professional level but an adequate level of familiarity) within a short training window.

In the wider picture of the thesis and its findings, one could ask: if the tasks were switched between chapters, would we expect to find the same results? The short answer to this question is yes. Kirsch and Cross (2015) used a training paradigm (discussed in more detail below) where participants learned to dance to specific sequences, showing that with increased experience, there was greater AON engagement. In relation to the connectivity findings, the predictive coding account would suggest that regardless of action being observed, a Bayesian comparison between observed and expected should occur, manifesting as differential anterior and posterior modulation.

Returning to the point of training paradigms, in Chapters IV and V an intensive training paradigm was employed to establish systemic variation in familiarity in the visual and motor domains. During the training period, participants observed one set of actions and spent an equivalent amount of time executing another set of actions. One question that could be raised with this approach is the extent to which mere action observation induces any training benefits in participants. As shown by Kirsch and Cross (2015), the layering of different levels of experience (visual,

auditory and physical) has an additive value in terms of AON engagement. In other words, the more layers to the training experience, the more value it provides to the trainee in terms of both behavioural performance and AON engagement during action observation. Based on the behavioural findings reported in Chapter IV (from both the training period and retest), this same pattern of findings also emerges within the present study. For those guitar sequences that were observed throughout the training week, participants' ability to play these sequences at the retest stage was lower than their ability to play the physically practiced sequences after training (however, the retest results do show that participants could play the observed sequences significantly better than the untrained sequences). The considerable differences that emerge in participants' ability to play sequences experienced during physical practice compared to visual experience is a valid criticism of the use of this kind of training paradigm (see also Kirsch & Cross, 2015; Cross et al., 2009). However, to counter this point, our aim was not to directly compare physical and visual experience in this study *per se*. By keeping these two elements separate throughout the design (and analysis), this allowed us to systematically vary familiarity separately for action observation and execution, which fits with our original aim.

Finally, the use of familiarity ratings within this thesis merits consideration. If objective, physical measures of performance ability are not feasible or possible (Cross et al., 2011; Press & Kilner, 2013; Kawabata & Zeki, 2004), alternative measures that rely on participants' self-report can enable further investigation of familiarity or perceived physical ability. By asking subjects to rate how familiar the actions observed during scanning were to them in an offline task in Chapter III, this allowed us to tap into a subjective reflection of the familiarity variable. One positive consequence of this approach is that it allows participants to express a subjective, overall feeling about each stimulus, which is not strictly related to how many times they have performed or seen the action before. For example, if we consider the guitar actions of Chapters IV and V, it could be the case that even though certain sequences were trained, participants may felt very familiar and competent with some of the trained actions, and less familiar or competent with other trained actions (and their ability to successfully execute these sequences may reflect this). To address this potential issue, we also took subjective ratings of familiarity in Chapter IV (as we did

in Chapter III), to enable closer comparison between the two approaches. When instructions were given to participants to rate the stimuli on familiarity, for the studies in both Chapter III and Chapter IV, participants were asked to use the entire scale. In Chapter IV, this instruction appeared to work and the whole scale of 1-9 was used across participants. For Chapter III, this instruction was not as effective, so a *z* standardization of the ratings was used instead of absolute ratings. This transformation allowed for the range of ratings used by each subject represented of the entire scale, allowing for further GLM and connectivity analysis to be more representative of the data.

To conclude this section, we used different types of actions within this thesis in order to address complementary questions about the role of familiarity in shaping AON responses. Familiarity was manipulated in Chapters IV and V through use of an intensive training paradigm, and all empirical chapters of the thesis also used a sensitive self-report rating measure of familiarity. Together, these methods enable us to manipulate and measure action familiarity to further understand experience-dependent plasticity within the AON.

6.4 Methodological implications: Effective connectivity and ROIs

In concert with training manipulations and measures of familiarity, this thesis also featured the use of effective connectivity measures to probe the exchange of influence between core nodes of the AON. We used the effective connectivity technique of dynamic causal modelling (Friston et al., 2003; see Chapter II for a discussion on its use), as it provides a measure of the *causal* relationship between the nodes of a given network. Compared to other methods such as Granger causality modelling (GCM; Goebel et al., 2003; Valdes-Sosa, 2004), DCM allows us to look beyond correlations within the observed data. By using a method such as GCM, the understanding of *causal* modulation between nodes of the network is not attainable, and the principles of the predictive coding account of action understanding would not be as clearly testable. For example, if a correlation between IPL and IFG is found, this would not illuminate whether familiarity drives the exchange of influence from IPL to IFG; rather, we would conclude that the activity between the regions is coupled when

familiarity varies. Therefore, the use of DCM to investigate the questions of this thesis is justified.

In addition, the use of different DCM approaches in this thesis warrants discussion. In Chapter III, we used a classical approach to DCM analysis, whereby all models containing permutations of connections were created and then compared in relation to model fit on both accuracy and complexity (Friston et al., 2003; Stephan et al., 2010). In Chapter V, a single model was constructed for each condition at each scanning session, which was then tested for consistency of connections across participants. A question that might arise here is why was the same method not used for both studies? The reason for using an alternative approach in Chapter V, as supported by the DCM literature, is that it is key to this kind of analysis to keep the model space as simple and hypothesis-driven as possible (Stephan et al., 2010). In Chapter III, we used the hypothesis of the presence of increasing familiarity on all connections between the core regions of the AON. This allowed for a more exploratory method and therefore we examined multiple models. However, in Chapter V, we tested strong predictions of the predictive coding account. Here, we also have multiple time points making the potential amount of models to evaluate very large. However, by using a strong hypothesis, we could define a more targeted set of models to directly test a hypothesis derived from the literature. Therefore, the use of different model spaces between the chapters was necessary and allowed us to explore the core questions of this thesis in the most targeted and appropriate way possible.

The use of ROIs, both in the DCM analysis of Chapter V and regression analysis of Chapter IV provide a strong position for this thesis. When investigating the profile of response amplitude within AON regions, a regression model was fitted to changes in amplitude as a function of familiarity (both number of exposures and participants' subjective ratings). This method, as described by Mattavelli et al. (2012), highlights certain limitations within our design. When number of exposures was used as a measure of familiarity, the increments between the points were not uniformly distributed. We adjusted these values to represent the exposure to the stimuli, but in future work more even increments (and indeed, more increments in general) should be used. However, within the same study, we aimed to complement the analysis of

objective measures of familiarity with the complementary use of subjective familiarity ratings as predictors in the regression model. These results provide support for the findings of a quadratic function with the objective exposure data, and more generally provide a clean measure of internal consistency between objective and subjective familiarity measures within the thesis.

Furthermore, the ROI approach used in Chapter V provided a powerful tool for testing familiarity over time. Here, we used identical regions for both pre- and post-training scanning sessions, allowing us to investigate how connectivity between these core regions changes as actions transition from novel and unfamiliar to trained and familiar. However, one issue with using this powerful approach is that a considerable number of participants were excluded from the final analysis (sample dropped from 21 to 14 participants for final analysis). These participants were required to be excluded, as they did not show activity in all the regions over both time points. This exclusion could have potentially limited the power of the subsequent analysis on the connection strengths (as $N = 14$ is a far from ideal sample size). Here, however, we trade power within the tests for a complete and robust within-subjects approach, which we argue ultimately allows for better understanding of how the system processes different levels of familiarity across a rigorously controlled training paradigm.

6.5 Future research

Many opportunities exist for future work to build upon the findings of this thesis. Here, we showed support for the predictive coding account using fMRI and DCM. An extension of this would be to investigate this connectivity with other neuroimaging techniques. Although fMRI provides evidence for connections between the core nodes of the AON, a matter that remains outstanding is how these reciprocally-connected regions process familiarity within the time domain. For example, we can conclude that increasing familiarity feeds posteriorly through the network. However, an unanswered question concerns whether the IFG first sends a prediction or whether the anterior connections first relay a signal about the observed action (c.f. Lewis et al., 2016 for an example of how the predictive coding accounts

dynamics can be applied to language). For example, using simultaneous fMRI and EEG, one could investigate the connectivity within this network as well as exploring the temporal aspects of this processing. EEG is often used to investigate the AON (c.f. Cochin et al., 1999; Grafton & Tipper, 2012; Hari et al., 2014; Kilner et al., 2004; Lepage & Theoret, 2006; Southgate et al., 2009; Press et al., 2011), and coupled with a powerful tool of connectivity analysis such as DCM (c.f. Kiebel et al., 2008 for the use of DCM in EEG), would provide a richer understanding of this network.

In order to further test the role of familiarity within the AON and, more specifically, test the predictions of the predictive coding account, another aim for future research could be to design a training experiment whereby familiarity first increases, and then over time, familiar actions revert or fade back to unfamiliar. For example, an action could be observed over the period of the experiment (increasing its familiarity), after which the action is then altered slightly (perhaps the intention changes) so that it becomes novel again. The test of action understanding models would then be to test their adaptability, and coupled with a method of connectivity, such change would further explore how the AON dynamically responds to changes in familiarity.

6.6 Conclusions

In summary, this thesis provides novel insight into the impact of familiarity on effective connectivity within the action observation network. Overall, results suggest that the predictive coding account of action understanding provides a useful and mostly validated framework for how core regions of this network process actions with varying familiarity. Specifically, this dynamic system appears to provide predictions that flow from anterior to posterior regions about how an action should unfold (which become stronger when an observed action is familiar), whereas modulations that originate in posterior regions and project anteriorly update the network's predictions with information about the action that is actually observed or executed (and is stronger for unfamiliar actions). However, the full story of how familiarity modulates the AON is far from concluded, and many interesting questions remain and have arisen from the work reported in this thesis. The use of methodologies such as DCM,

rating scales, training paradigms and ROI regression analyses has provided valuable insights into this field and offer evidence that targeted combinations of methodological and analytical approaches will be required to further advance our understanding of experience-dependent plasticity in the human brain.

References

Abrams, D.A., Ryali, S., Chen, T., Balaban, E., Levitin, D.J. & Menon, V. (2013). Multivariate activation and connectivity patterns discriminate speech intelligibility in Wernicke's, Broca's and Geschwind's areas. *Cerebral Cortex*, 23(7); 1703-1714.

Acs, F. & Greenlee, M.W. (2008). Connectivity modulation of early visual processing areas during covert and overt tracking tasks. *Neuroimage, 41*; 380–388.

Aguirre, G. K., Zarahn, E. & D'Esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *Neuroimage*, 8(4); 360-369.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52*, 317-332.

Attwell, D. & Iadecola, C. (2002). The neural basis of functional brain imaging signals. *Trends Neurosci. 25*; 621–625.

Basler, N., Lorey, B., Pilgramm, S., Naumann, T., Kindermann, S., Stark, F., Zentgraf, R., Williams, A.M. & Munzert, J. (2014). The influence of expertise on brain activation of the action observation network during anticipation of tennis and volleyball serves. *Front Hum Neurosci.,* 8: 586.

Beal, M. & Ghahramani, Z. (2003). *The variational Bayesian EM algorithms for incomplete data: with application to scoring graphical model structures*. In: Bernardo J., Bayarri M., Berger J., Dawid A., editors. Vol. 7. Cambridge University Press (Bayesian Statistics).

Beilock, S.L. & Gonso, S. (2008). Putting in the mind versus putting on the green: expertise, performance time, and the linking of imagery and action. *Q J Exp Psychol.,* 61(6): 920-32.

Belliveau, J. W., Kennedy, D. J., McKinstry, R.C., et al. (1991) Functional mapping of the human visual cortex by magnetic resonance imaging. *Science*, 254; 716-719.

Bischoff, M., Zentgraf, K., Lorey, B., Pilgramm, S., Basler, N., Baumgartner, E., Hohmann, T., Stark, R., Vaitl, D. & Munzert, J. (2012). Motor familiarity: brain activation when watching kinematic displays of one's own movements. *Neuropsychologia*, 50(8); 2085-92.

Bishop, C. (2006). *Pattern recognition for machine learning*. Springer, Berlin.

Blakemore, S.J. & Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychology*, 43; 260–267.

Bloch, F. (1946). Nuclear induction. *Phys Rev*, 70; 460-474.

Bobick, A. (1997). Movement, Activity and Action: The Role of Knowledge in he perception of Motion. *Phil. Trans. Royal Soc. London B.*, 352; 1257-1997.

Bonini, L., Rozzi, S., Serventi, F.U., Simone, L., Ferrari, P.F & Fogassi, L. (2010). Ventral premotor and inferior parietal cortices make distinct contribution to action organization and intention understanding. *Cerebral Cortex*, 20, 1372-1385.

Bosman, C.A., Schoffelen, J. M., Brunet, N., Oostenveld, R., Bastos. A. M., Womelsdorf, T., Rubehn, B., Stieglitz, T., de Weerd, P. & Fries, P. (2012). Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron, 75(*5); 857-88.

Braams, B.R, van Duijvenvoorde, A.C.K, Peper, J.S. & Crone, E.A. (2015). Longitudinal changes in adolescent risk-taking: A comprehensive study of neural response to rewards, pubertal development, and risk-taking behaviour. *J. Neurosci.*, 35(18); 7226-7238.

Bressler, S. L. & Seth, A. K. (2011). Wiener-Granger causality: a well established methodology. *NeuroImage, 58*(2); 323-9.

Brett, M., Anton, J-L., Valabregue, R. & Poline, J-B. (2002). Regions of interest analysis using an SPM toolbox (abstract). Presented at the 8th International conference of Functional Mapping of the Human Brain, June 2- 6, 2002, Sendai, Japan. Available in CD-ROM in *Neuroimage*, 16(2).

Brodersen, K. H., Haiss, F., Soon Ong, C., Jung, F., Tittgemyer, M., Buhmann, J. M., Weber, B. & Stephan, K. E. (2011). Model-based feature construction for multivariate decoding. *NeuroImage, 56*; 601-615.

Brodersen, K. H., Penny, W. D., Harrison, L. M., Daunizeau, J., Ruff, C. C., Duzel, E., Friston, K. J., et al., (2008). Integrated Bayesian models of learning and decision making for saccadic eye movements. *Neural Netw.* 21(9); 1247–1260.

Brodski, A., Paasch, G-F., Helbling, S. & Wibral, M. (2015). The faces of predictive coding. *The J. Neurosci.* 35(24); 8997-9006.

Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R. & Bressler, S. L. (2004). Beta oscillations in a large-scale sensorimotor cortical system: directional influences revealed by Granger causality. *Proc. Natl. Acad. Sci. USA, 101*; 9849-9854.

Buccino, G., Vogt, S., Ritzl, A., Fink, G. R., Zilles, K., Freund, H-J., & Rizzolatti, G. (2004). Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. *Neuron*, 42; 323–334.

Buchel, C. & Friston, K. J. (1997). Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cerebral Cortex, 7(*8); 768-78.

Burnham, K. P. & Anderson, D. R. (2002). *Model selection and multimodal inference: A practical information-theoretical approach*. New York: Springer-Verlag.

Buxton, R. B. & Frank, L. R. (1997). A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation. *J. Cereb. Blood Flow Metab., 17*; 64–72.

Buxton, R., Uludag, K., Dubowitz ,D. & Liu, T. (2004). Modelling the hemodynamic response to brain activation. *NeuroImage, 23*; 220–233.

Buxton, R. B., Wong, E. C. & Frank, L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Mag. Reson. Med*; 855-864.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E. & Haggard, P. (2005). Action observation and acquired motor skills: an fMRI study with expert dancers. *Cereb. Cortex*, 15; 1243–1249.

Calvo-Merino, B., Grezes, J., Glaser, D.E., Passingham, R.E. & Haggard, P. (2005). Seeing or doing? Influence of visual and motor familairty in action observation. *Curr Biol*., 16(19): 1905-1910.

Carpenter, G.A. & Grossberg, S. (2010). *Adaptive resonance theory*. New York: Springer.

Casile, A. & Giese, M.A. (2006). Nonvisual motor training influences biological motion perception. *Curr. Biol.,* 16(1); 69-74.

Caspers, S., Zilles, K., Laird, A. & Eickhoff, S. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage*, 50; 1148–1167.

Cavada, C. & Goldman-Rakic, P.S. (1989). Posterior partial cortex in rhesus monkey: evidence for segregated corticocortial networks linking sensory and limbic areas with the frontal lobe. *J Comp Neurol*., 287: 422-445.

Chen, C. C., Henson, R. N., Stephan, K. E., Kilner, J. M. & Friston, K. J. (2009). Forward and backward connections in the brain: A DCM study of functional asymmetries. *NeuroImage*, 45; 453-462.

Chen, C. C., Kiebel, S. J. & Friston, K. J. (2008). Dynamic causal modelling of induced responses. *NeuroImage, 41;* 1293–1312.

Cochin, S., Barthelemy, C., Roux, S. & Martineau, J. (1999). Observation and execution of movement: similarities demonstrated by quantified electroencephalography. *Eur J Neurosci*., 11: 1189-1842.

Cronbach, L. J. & Meehl, P. E. (1995). Construct validity in psychological tests. *Psychological Bulletin*, 52; 281-302.

Cross, E. S., Hamilton, A. F. & Grafton, S. T. (2006). Building a motor simulation de novo: observation of dance by dancers. *NeuroImage*, 31: 1257–1267.

Cross, E. S., Kraemer, D. J., Hamilton, A. F., Kelley, W. M., & Grafton, S. T. (2009). Sensitivity of the action observation network to physical and observational learning. *Cerebral Cortex*, 19(2): 315–326.

Cross, E. S., Liepelt, R., Hamilton, A.F.de.C., Parkinson, J., Ramsey, R., Stadler, W., & Prinz, W. (2012). Robotic movement preferentially engages the action observation network. *Human Brain Mapping*, 33(9): 2238-2254.

Cross, E.S., Kirsch, L., Ticini, L.F. & Schütz-Bosbach, S. (2011). The impact of aesthetic evaluation and physical ability on dance perception. Frontiers in Human Neuroscience, 5(102); 1-10.

Cross, E.S., Stadler, W., Parkinson, J., Schütz-Bosbach, S. & Prinz, W. (2013). The influence of visual training on predicting complex action sequences. *Human Brain Mapping*, 34; 467–486.

Coutanche, M.N. & Thompson-Schill, S.L. (2013). Informational connectivity: identifying synchronised discriminability of multi-voxel patterns across the brain. *Frontiers in Human Neuroscience,* 7(15): doi:10.3389/fnhum.2013.00015

Csibra, G. (1993). *Action mirroring and action understanding: an alternative account,* in Sensorimotor Foundations of Higher Cognition, eds Haggard P., Rossetti Y., Kawato M., editors. Oxford: Oxford University Press: 435–459.

Csibra, G. (2005). Mirror neurons and action observation. Is simulation involved? In *www.interdisciplines.org/mirror*.

Damdian, R. (1971). Tumor detection by nuclear magnetic resonance. *Science*, 171; 1151- 1153.

Daunizeau, J., David, O. & Stephan, K. E. (2011). Dynamic causal modelling: a critical review of the biophysical and statistical foundations. *NeuroImage, 58(*2), 312-22.

Daunizeau, J., Friston, K. J. & Kiebel, S. J. (2009a). Variational Bayesian identification and prediction of stochastic nonlinear dynamic causal models. *Physica., 238;* 2089–2118.

Daunizeau, J., Kiebel, S. J. & Friston, K. J. (2009b). Dynamic causal modelling of distributed electromagnetic responses. *NeuroImage, 47;* 590–601.

Daunizeau, J., Lemieux, L., Vaudano, A. E., Friston, K. J. & Stephan, K. E (2012). An electrophysiological validation of stochastic DCM for fMRI. *Front. Comput. Neurosci., 6(*103).

David, O. (2011). fMRI connectivity, meaning and empiricism Comments on: Roebroeck et al. The identification of interacting systems in the brain using fMRI: model selection, causality and deconvolution. *NeuroImage, 58*(2); 306-9.

David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C. & Depaulis, A. (2008). Identifying neural drivers with functional MRI: an electrophysiological validation. *PLoS Biol, 6*; 2683-2697.

David, O., Kiebel, S. J., Harrison, L. M., Mattout, J., Kilner, J. M. & Friston, K. J. (2006). Dynamic causal modelling of evoked responses in EEG and MEG. *NeuroImage, 30;* 1255–1272.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum likelihood from incomplete data via EM algorithm. *Journal of the Royal Statistical Society Series B-Methodological,* 39; 1-38.

Diersch, N., Mueller, K., Cross, E.S., Stadler, W., Rieger, M. & Schutz-Bosbach, S. (2013). Action prediction in younger versus older adults: neural correlates of motor familiarity. *PLOSone*, 8, e64195 10.1371/journal.pone.0064195.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res.*, 91(1); 176–180.

Dinstein, I., Hasson, U., Rubin, N. & Heeger, D.J. (2007). Brain areas selective in both observed and executed movements. *J Neurophysiol.*, 98(3): 1415-27.

Dirac, P. A. M. (1928). The quantum theory of the electron. *Proc R Soc Lond A*, 117; 610-624.

Eickhoff, S.B., Paus, T., Caspers, S., Grosbars, M.H., Evans, A.C., Ziles, K. & Amunts, K. (2007). Assignment of functional activations of probabilistic cytoarchitectonic areas revisited. *Neuroimage,* 36 (3), 511-21.

Falck-Ytter, T., Gredebäck, G. & von Hofsten, C. (2006). Infants predict other people's action goals. *Nat. Neurosci.*, 9; 878–879.

Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F. & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, 308: 662-667.

Frackowiak, R., Friston , K. J., Frith, C., Dolan, R., Price, C., Zeki, S., et al ... (2003).: *Human Brain Function*. 2nd ed. Amsterdam: Elsevier Academic Press.

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 360; 815–836.

Friston, K. J. (2002). Beyond phrenology: what can neuroimaging tell us about distributed circuitry? *Ann. Rev. Neurosci., 25*; 221–250.

Friston, K. J. (2007). *Statistical Parametric Mapping*. London: Academic Press.

Friston, K. (2009a). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biol. 7*, e33.

Friston, K. J. (2009b). The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci., 13;* 239-301.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E. & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, 6(3); 218-229.

Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G. & Mattout, J. (2008). Multiple sparse priors for the M/EEG inverse problem. *NeuroImage, 39;* 1104–1120.

Friston, K. J., Harrison, L. & Penny, W. D. (2003). Dynamic causal modelling. *Neuroimage*, 19; 1273-1302.

Friston, K. J., Kahan, J., Biswal, B. & Razi, A. (2014a). A DCM for resting state fMRI. *NeuroImage,* 94; 396–407.

Friston, K. J., Kahan, J., Razi, A., Stephan, K. E. & Sporns, O. (2014b). On nodes and modes in resting state fMRI. *NeuroImage*, 99; 533–547.

Friston, K., Moran, R. & Seth, A. K. (2012). Analysing connectivity with Granger causality and dynamic causal modelling. *Curr. Opin. Neurobiol., 23*; 1-7.

Friston, K. J. & Penny, W. (2011). Post hoc Bayesian model selection. *Neuroimage, 56*; 2089-2099.

Friston, K. J. & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, 159; 417-458.

Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain,* 119 (Pt 2); 593–609.

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

Gallese, V., Keysers, C. & Rizzolatti, G. (2004. A unifying view of the basis of' social cognition. *Trends Cogn Sci .,* 8(9): 396-403.

Gardner, T., Goulden, N. & Cross, E.S. (2015). Dynamic modulation of the action observation network by movement familiarity. *J. Neurosci.*, 35(4); 1561-72.

Garrido, M. I., Kilner, J. M., Kiebel, S. J., Stephan, K. E. & Friston K. J. (2007). Dynamic causal modelling of evoked potentials: a reproducibility study. *NeuroImage, 36;* 571– 580.

Gazzola, V. & Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fMRI data. *Cereb. Cortex*, 19; 1239-1255.

Gazzola, V., Rizzolatti, G., Wicker, B. & Keysers, C. (2007). The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage*, 35; 1674–1684.

Grafton, S. T. (2009). Embodied cognition and the simulation of action to understand others. *Ann N Y Acad Sci.,* 1156; 97-117.

Grafton, S.T. & Tipper, C.M. (2012). Decoding intention.: a neuroergonomic perceptive. *Neuroimage*, 59: 14-24.

Gelman, A., Carlin, J. B., Stern. H. S. & Rubin, D. B., (1995). *Bayesian Data Analysis*. Chapman and Hall; Boca Raton.

Gerlach,W. & Stern, O. (1922). Das magnetische moment des silberatoms. *Zeitschrift für Physik,* 9(1); 353-355.

Ghahramani, Z. (2004) Unsupervised Learning. In Bousquet, O., von Luxburg, U. and Raetsch, G. *Advanced Lectures in Machine Learning*. 72-112.

Gitelman, D. R., Penny, W. D., Ashburner, J. & Friston, K. J. (2003). Modelling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. *NeuroImage, 19;* 200–207.

Glover, G. H. (2011). Overview of Functional Magnetic Resonance Imaging. *Neurosurgery Clinics of North America*, *22*(2); 133–139.

Goebel, R., Roebroeck, A., Kim, D. S. & Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modelling and Granger causality mapping. *Magn. Reson. Imaging., 21*; 1251–1261.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*; 424-438

Green, D., Kochukhova, O., & Gredebäck, G. (2014) Extrapolation and Direct Matching Mediate Anticipation in Infancy. *Infant Behaviour and Development,* 37: 111-118.

Gregoriou, G.G., Borra, E., Matelli, M. & Luppino, G. (2006). Architectonic organization of the inferior parietal convexity of the macaque monkey. *J Comp Neurol*., 496: 422–451

Grossberg, S. (2007). Towards a unified theory of neocortex: laminar cortical circuits for vision and cognition. *Prog Brain Res*., 165: 79–104.

Grossberg, S. (2013). Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world. *Neural Netw*., 37: 1–47

Hamilton, A. F. (2013). The mirror neuron system contributes to social responding. *Cortex*, 49(10); 2957-2959.

Hamilton, A. F. & Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *J Neurosci*. 26; 1133–1137.

Hamilton, A.F. & Grafton, S.T. (2008). Action outcomes are represented in human inferior frontoparietal cortex. *Cereb. Cortex*, 8: 1160–1168.

Hari, R., Bourguignon, M., Piitulainen, H., Smeds, E., De Tiège, X. & Jousmäki, V. (2014). Human primary motor correct is both activated and stabilized during observation of other persons phasic motor actions. *Phil. Trans. R. Soc. B*, 369; 1644.

Harrison, L., Penny, W. & Friston, K. J. (2003): Multivariate autoregressive modelling of fMRI time series. *Neuroimage, 19;* 1477–1491.

Hauser, M. & Wood, J. (2010). Evolving the capacity to understand actions, intentions, and goals. *Annu Rev Psychol*., 61:303–324.

Hebb, D. (1949). *The organisation of behaviour*. New York, NY: John Wiley and Sons.

Heeger, D. J., Huk, A. C., Geisler, W.S. & Albrecht, D. G. (2000). Spikes versus BOLD: what does neuroimaging tell us about neuronal activity? *Nat Neurosci*., 3; 631–633.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci*. 21: 1229–1243.

Hickok, G. & Hauser, M. (2010). (Mis)understanding mirror neurons. *Curr Biol*., 20(14): 593-594.

Hoeting, J. A., Madigan, D., Raftery, A. E. & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial with discussion. *Statistical Science, 14*; 382–417.

Hommel, B., Müsseler, J., Aschersleben, G. & Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci*., 24(5); 849-878.

Horwitz, B. & Tagamets, M. A. (1999). Predicting human functional maps with neural net modelling. *Hum. Brain Mapp*., 8; 137–142.

Horwitz, B., Tagamets, M. A. & McIntosh, A. R. (1999). Neural modelling, functional brain imaging, and cognition. *Trends Cogn. Sci,*. 3; 91-98.

Iacoboni, M., Koski, L.M., Brass, M., Bekkering, H., Woods, R.P., Dubeau, M.C., Mazziotta, J.C. & Rizzolatti, G. (2001). Reafferent copies of imitated actions

in the right superior temporal cortex. *Proc Natl Acad Sci U S A.,* 98: 13995-13999.

Kandel, E. R., Schwartz, J. H. & Jessell, T. M. (2000). *Principles of neural science*. McGraw-Hill, New York, 4th Ed.

Kaps, M. & Lamberson, W. R. (2004). *Biostatistics for animal science*. CABI Publishing, Cambridge.

Kass, R. E. & Raftery, A. E., (1995). Bayes factors. *J. Am. Stat. Assoc. 90*; 773–795.

Kasses, C. H., Stephan, K. E., Weissenbacher, A., Pezawas, L., Moser, E. & Windischberger, C. (2010). Multi-subject analyses with dynamic causal modelling. *Neuroimage, 49(4);* 3065-74.

Kawabata, H. & Zeki, S. (2004). Neural Correlates of Beauty. *J. Neurophysiol*., 91; 1699-1705.

Keysers, C. & Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Curr. Opin. Neurobiol*., 19; 666–671.

Keysers, C. & Gazzola, V. (2014). Hebbian learning and predictive mirror neurons for actions, sensations and emotions. *Philos Trans R Soc Lond B Biol Sci*., 369; 20130175.

Keysers, C. & Perrett, D. I. (2004). Demystifying social cognition: a Hebbian perspective. *Trends Cogn. Sci*., 8; 501-507.

Kiebel, S.J., Daunizeau, J. & Friston, K.J. (2008). A hierarchy of time-scales and the brain. *PLoS Comput Biol*., 4(11): e1000209.

Kiebel, S. J., David, O. & Friston, K. J., (2006). Dynamic causal modelling of evoked responses in EEG/MEG with lead field parameterization. *NeuroImage, 30;* 1273–1284.

Kiebel, S. J., Kloppel, S., Weiskopf, N. & Friston, K. J., (2007). Dynamic causal modelling: a generative model of slice timing in fMRI. *NeuroImage, 34;* 1487–1496.

Kilner, J. M., Friston, K. J. & Frith, C. D. (2007a). The mirror-neuron system: a Bayesian perspective. *NeuroReport*, 18; 619–623.

Kilner, J. M., Friston, K. J. & Frith, C. D. (2007b). Predictive coding: an account of the mirror neuron system. *Cogn. Process*, 8; 159-166.

Kilner, J.M., Kraskov, A. & Lemon, R.N. (2014). Do monkey F5 mirror neurons show changes in firing rate during repeated observation of natural actions? *Journal of Neurophysiology*, 111(6): 1214-1226.

Kilner, J.M., Marchant, J. L. & Frith, C. D. (2009b). Relationship between activity in human primary motor cortex during action observation and the mirror neuron system. *PLoS One*, 4(3); e4925.

Kilner. J. M., Neal, A., Weiskopf, N., Friston, K. J. & Frith, C. D. (2009a). Evidence of Mirror Neurons in Human Inferior Frontal Gyrus. *The Journal of Neuroscience*, 29(32); 10153-10159.

Kilner, J.M., Vargas, C., Blakemore, S.J. & Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nat Neuroscience*., 1(2): 143-8.

Kim, T-H. & Cruz, A. (2011). Differences in brain activation during motor imagery and action observation of golf putting. *Scientific Research and Essays*, 6(15): 3132-3138.

Kirsch, L.P. & Cross, E.S. (2015). Additive routes to action learning: Layering experience shapes engagement of the action observation network. *Cerebral Cortex*, 25; 4799-4811.

Kirsch, L.P., Dawson, K. & Cross, E.S. (2015). Dance experience sculpts aesthetic perception and related brain circuits. *Annals of the New York Academy of Sciences*, 1337*;* 130-139.

Knoblich, G. & Flach, R. (2001). Predicting the effects of actions: interactions of perception and actions. *Psychol Sci*., 12(6): 467-72.

Kohler, E., Keysers, C., Umilta, M.A., Fogassi, L., Gallese, V. & Rizzolatti, G. (2002). Hearing sounds understanding actions: action representation in mirror neurons. *Science*, 297: 846-848.

Kumar, A., Welti, D. & Ernst, R. R. (1975). NMR fourier zeugmatography. *J Mag Reson*, 18; 69-83.

Lamm, C., Nusbaum, H. C., Meltzoff, A. N. & Decety, J. (2007). What are you feeling? Using functional magnetic resonance imaging to assess the modulation of sensory and affective responses during empathy for pain. *Brain Res*., 1227; 153–161.

Läppchen, C.H., Ringer, T., Blessin, J., Schulz, K., Seidel, G., Lange, R. & Hamzei, F. (2015). Daily iTBS worsens hand motor training--a combined TMS, fMRI and mirror training study. *Neuroimage*, 107; 257-65.

Laβberg, C., Beykirch, K., Mohler, B.J. & Bülthoff, H.H. (2014). Intersegmental eye-head-body interactions during complex whole body movements. PLos One, 9(4): e95450.

Lauritzen, M. (2004). Reading vascular changes in brain imaging: is dendritic calcium the key? *Nat. Rev. Neurosci., 6*; 77–85.

Lauterbur, P. C. (1973). Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature*, 242; 190-191.

Lauterbur, P. C. (1974). Magnetic resonance zeugmatography. *Pure Appl Chem*, 40; 149-157.

Lauterbur, P. C. (1986). In *NMR in biology and medicine* (ed. Chien, S. & Ho, C.), Raven Press, New York.

Lee, L., Friston, K. & Horwitz, B. (2006). Large-scale neural models and dynamic causal modelling. *NeuroImage, 30;* 1243-1254.

Leff A. P., Schofield T. M., Stephan K. E., Crinion J. T., Friston K. J. & Price C. J. (2008). The cortical dynamics of intelligible speech. *J. Neurosci., 28;* 13209–13215.

Lepage, J.-F. & Théoret, H. (2006). EEG evidence for the presence of an action observation– execution matching system in children. *Eur. J. Neurosci*. 23, 2505–2510.

Lewis, A.G., Schoffelen, J-M., Schriefers, H. & Bastiaansen (2016). A predictive coding perspective of beta oscillations during sentence-level language comprehension. Front Hum Neurosci., 10: 85.

Li, B., Daunizeau, J., Stephan, K. E., Penny, W., Hu, D. & Friston, K. J. (2011). Generalised filtering and stochastic DCM for fMRI. *NeuroImage*, 58(2): 442-457.

Liew, S., Han, S. & Aziz-Zadeh, L. (2011). Familiarity modulates mirror neuron and mentalizing regions during intention understanding. *Hum. Brain Mapp*., 32: 1986–1997.

Liew, S-L., Sheng, T., Margetis, J. & Aziz-Zadeh, L. (2013). Both novelty and expertise increase action observation network activity. *Frontiers in Human Neuroscience*, 7: 541.

Lingnau, A., Geserich, B. & Caramazza, A. (2009). Asymmetric fMRI adaptations reveals no evidence for mirror neurons in humans. *Pro Natl ACA Sci U S A.*, 106: 9925-9930.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature, 412;* 510–516.

Logothetis N. K. & Wandell B. A. (2004). Interpreting the BOLD signal. *Annu. Rev. Physiol., 66*; 735–769.

Lohmann, G., Erfurth, K., Muller, K. & Turner, R. (2012). Critical comments on dynamic causal modelling. *NeuroImage,* 59(3); 2322-2329.

Londei, A., D'Ausilio, A., Basso, D. & Belardinelli, M. O. (2006). A new method for detecting causality in fMRI data of cognitive processing. *Cogn. Process*, 7; 42–52.

Lotze, M. (2013). Kinesthetic imagery of musical performance. Frontiers in Human *Neuroscience*, 7, 28.

MacKay, D. J. C. (1992). A practical Bayesian framework for backpropagation systems. *Neural Comput,. 4;* 448–472.

Macuga, K.L. & Frey, S.H. (2011). Selective responses in right interior frontal and supramarginal gyri differentiate between observed movements of ones vs. another. *Neuropsychologia*, 49(5); 1202-1207.

Marreiros, A. C., Kiebel, S. J. & Friston, K. J. (2008). Dynamic causal modelling for fMRI: a two-state model. *Neuroimage*, 39: 269-278.

Mandeville, J. B., Marota, J. J., Ayata, C., Zararchuk, G., Moskowitz, M. A., Rosen, B., et al. (1999). Evidence of a cerebrovascular postarteriole windkessel with delayed compliance. *J. Cereb. Blood Flow Metab.*, 19; 679–689.

Mansfield, P. (1977). Multi-planar image formation using NMR spin echoes. *J. Phys. C.: Sol State Phys.,* 10(3); 55-58.

Mansfield, P. & Maudsley, A. A. (1976). Planar spin imaging by NMR. *J. Phys. C.: Sol State Phys.,* 9(15); 101-119.

Mattavelli, G., Andrews, T.J., Asghar, A.U.R., Towler, J.R. & Young, A.W. (2012). Response of face-selective brain regions to trustworthiness and gender of faces. *Neuropsychologia*, 50; 2205-2211.

McGonigle, D. J., Howseman, A. M., Athwal, B. S., Friston, K. J., Frackowiak, S. J. & Holmes, P. A. (2000). Variability in fMRI: an examination of intersession differences. *Neuroimage*, 11; 708-734.

McIntosh, A. R. & Gozalez-Lima, F. (1991). Structural modelling of functional neural pathways mapped with 2-deoxyglucose: Effects of acoustic startle habituation on the auditory system. *Brain Res*., 547; 295-302.

McIntosh, A. R. & Gonzalez-Lima, F. (1994). Structural equation modelling and its application to system analysis in functional brain imaging. *Hum. Brain Mapp., 2*; 2–22.

Meltzoff, A. N. (2007). Like me: a foundation for social cognition. *Dev. Sci.,* 10(1): 126-34.

Menon, R. S., Ogawa, S., Hu, X., Strupp, J.S., Andersen, P. & Ugurbil, K. (1995). BOLD based functional MRI at 4 tesla includes a capillary bed contribution: Echo-planar imaging mirrors previous optical imaging using intrinsic signals. *Magn Reson Med., 33*; 453–459.

Molenberghs, P., Cunnington, R. & Mattingley, J.B. (2012). Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. *Neurosci. Biobehav. Rev*., 36; 341–349.

Moran, R. J., Kiebel, S. J., Stephan, K. E., Reilly, R. B., Daunizeau, J. & Friston, K. J. (2007). A neural mass model of spectral responses in electrophysiology. *NeuroImage, 37;* 706–720.

Moran, R. J., Stephan, K. E., Kiebel, S. J., Rombach, N., O'Connor, W. T., Murphy, K. J., Reilly, R. B. & Friston, K. J. (2008). Bayesian estimation of synaptic physiology from the spectral responses of neural masses. *NeuroImage, 42*; 272–284.

Moran, R. J., Stephan, K. E., Seidenbecher, T., Pape, H. C., Dolan, R. J. & Friston, K. J. (2009).Dynamic causal models of steady-state responses. *NeuroImage, 44;* 796–811.

Mukamel, R., Ekstrom, A.D., Kaplan, J., Iacoboni, M. & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Curr Biol*., 20: 750-756.

Mumford, D. (1992). On the computational architecture of the neocortex. II. *Biol. Cybern*., 66: 241–251.

Neumann, J. & Lohmann, G. (2003). Bayesian second-level analysis of functional magnetic images. *Neuroimage, 2;* 1346-55.

Obata, T., Liu, T. T., Miller, K. L., Luh, W. M., Wong, E. C., Frank, L. R. & Buxton, R. B. (2004). Discrepancies between BOLD and flow dynamics in primary and supplementary motor areas: application of the Balloon model to the interpretation of BOLD transients. *NeuroImage, 21;* 144–153.

Ogawa, S., Lee, T. M. Kay, R. A. & Tank, D. W. (1990a). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Pro Nat Acad Sci U S A*, 87(24); 9868-9872.

Ogawa, S., Lee, T. M., Nayak, A. S. & Glynn, P. (1990b). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magn Reson Med, 14;* 68–78.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9: 97-113.

Oztop, E., Kawato, M. & Arbib, M.A. (2013). Mirror neurons: functions, mechanism and models. *Neurosci. Letters*, 540: 43-55.

Pauli, W. (1927). Zur quantenmechanik des magnetischen Elektrons. *Z Phy A Hadr Nucl*, 43(9-10); 601-623.

Pauling, L. & Coryell, C. D. (1936). The Magnetic Properties and Structure of Hemoglobin, Oxyhemoglobin and Carbonmonoxyhemoglobin. *Proc Natl Acad Sci U S A.,* 22(4); 210–216.

Pearl, J. (1998). Graphs, Causality, and Structural Equation Models. *Sociological Methods Research, 27*(2); 226-284.

Penny, W. D., Stephan, K. E., Mechelli, A. & Friston, K. J. (2004). Modelling functional integration: a comparison of structural equation and dynamic causal models. *NeuroImage,* 23: 264-274.

Penny, W. D., Ghahramani, Z. & Friston, K. J. (2005). Bilinear dynamical systems. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360; 983–993.

Penny, W. D., Litvak, V., Fuentemilla, L., Duzel, E. & Friston, K. (2009). Dynamic causal models for phase coupling. *J. Neurosci. Methods, 183;* 19–30.

Penny, W. D., Stephan, K. E., Daunizeau, J., Rosa, M. J., Friston, K. J., Schofield, T. M. & Leff, A. P. (2010). Comparing families of Dynamic Causal Models. *PLoS Computational Biology* 6: e1000709.

Pezzulo, G., Rigoli, F. & Friston, K. (2015). Active inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology,* 134; 17-35.

Phillips, C., Rugg, M. D. & Friston, K. J. (2002). Anatomically informed basis functions of EEG source localization: combining functional and anatomical constraints. Neuroimage, 16(3); 678-695.

Pitt, M. A. & Myung, I. J. (2002). When a good fit can be bad. *Trends Cogn. Sci. 6;* 421–425.

Press, C. (2011). Action observation and robotic agents: Learning and anthropomorphism. *Neuroscience and Biobehavioural Reviews*, 35; 1410-1418.

Press, C., Catmur, C., Cook, R., Widmann, H., Heyes, C. & Bird, G. (2012). fMRI evidence of 'mirror' responses to geometric shapes. *PLoS ONE*, 7: 1210.1371/journal.pone.0051934

Press, C., Cook, J., Blakemore, S-J. & Kilner, J. (2011). Dynamic modulation of human motor activity when observing actions. *The Journal of Neuroscience*, 31(8): 2792-2800.

Press, C.M. & Kilner, J.M. (2013). The time course of eye movements during action observation reflects sequence learning. *NeuroReport,* 24(14); 822-26.

Purcell, E. M. (1948). Nuclear magnetism in relation to problems of the liquid and solid states. *Science*, 107(2783); 433-440.

Rabi, I. I., Zacharias, J. R., Millman, S. & Kusch, P. (1938). A new method for measuring nuclear magnetic moment. *Phys Rev*, 53; 318-318.

Raftery, A. (1995). Bayesian model selection in social research. In: Marsden P., *Sociological Methodology*, Cambridge, MA; 111–196.

Rao, R.P. & Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2(1); 79-87.

Ramsey, R. & Hamilton A.F.d.C. (2010). Triangles have goals too: understanding action representation in left aIPS. *Neuropsychologia*, 48(9): 2773-6.

Riera, J. J., Watanabe, J., Kazuki, I., Naoki, M., Aubert, E., Ozaki, T. & Kawashima, R. (2004). A state–space model of the hemodynamic approach: nonlinear filtering of BOLD signals. *NeuroImage 21;* 547–567.

Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M. & Rozzi, S. (2014). Cortical mechanisms underlying the organisation of goal-directed action and mirror neuron-based action understanding. *Physiological Reviews,* 94(2): 655-706.

Rizzolatti G, Craighero L (2004) the mirror-neuron system. Annu. Rev. Neurosci. 27; 169-192.

Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Res Cogn Brain Res*., 3: 131-141.

Rizzolatti, G., Fadiga, L., Gentilucci, M., Luppino, G. & Matelli, M. (1988). Functional organisation of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Exp Brain Res*., 71: 491-507.

Rizzolatti, G., Fogassi, L. & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci*., 2; 661–670.

Rizzolatti, G., Scandolra, C., Matelli, M. & Gentilucci, M. (1981). Afferent properties of periarcuate neurons in macaque monkeys. I. Somatosensory responses. *Behav Brain Res*., 2: 125-146.

Rizzolatti, G. & Sinigaglia, S. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat Rev Neurosci*., 11: 264-274.

Roebroeck, A,. Formisano, E. & Goebel, R. (2005). Mapping directed influences over the brain using Granger causality and fMRI. *Neuroimage, 25;* 230-242.

Roebroeck, A., Formisano, E. & Goebel, R. (2011a). The identification of interacting systems in the brain using fMRI: model selection, causality and deconvolution. *Neuroimage*. 58; 296–302.

Roebroeck, A., Formisano, E. & Goebel, R., (2011b). Reply to Friston and David after comments on: The identification of interacting systems in the brain using fMRI: model selection, causality and deconvolution. *Neuroimage*, 58; 310–311.

Roy, C. S. & Sherrington, C. S. (1890). On the regulation of blood supply to the brain. *J Physiol*, 11(1-2); 85-108.

Rozzi, S., Calzavara, R., Belmalih, A., Borra, E., Gregoriou, G.G., Matelli, M. & Luppino, G. (2006). Cortical connections of the inferior parietal cortical convexity of the macaque monkey. *Cerebral Cortex*, 16: 1389-1417.

Reyt, S., Picq, C., Sinniger, V., Clarençon, D., Bonaz, B. & David, O. (2010). Dynamic Causal Modelling and physiological confound: a functional MRI study of vagus nerve stimulation. *NeuroImage,* 52(4); 1456-64.

Rosa, M. J., Friston, K. & Penny, W. (2012). Post-hoc selection of dynamic causal models. *J Neurosci Methods, 208*; 66–78.

Rowe, J. B., Hughes, L. E., Barker, R. A. & Owen, A. M. (2010). Dynamic causal modelling of effective connectivity from fMRI: Are results reproducible and sensitive to Parkinson's disease and its treatment? *Neuroimage,* 52(3); 1015-1026.

Schanze, T. (1995). Sinc interpolation of discrete periodic signals. *IEEE Trans Signal Process*, 43(6); 1502-1503.

Schippers, M.B. & Keysers, C. (2011). Mapping the flow of information within the putative mirror neuron system during gesture observation. *NeuroImage*, 57(1): 37-44.

Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends Cogn. Sci.*, 11: 211–218.

Schuyler, B., Ollinger, J. M., Oakes, T. R., Johnstone, T. & Davidson, R. J. (2010). Dynamic causal modelling applied to fMRI data shows high reliability. *Neuroimage*, 49; 603–611.

Shimada, S. (2010). Deactivation in the sensorimotor area during observation of a human agent performing robotic actions. *Brain Cognition*, 72: 394–399.

Seth, A. K. (2010). A MATLAB toolbox for Granger causal connectivity analysis. *J. Neurosci. Meth.,* 186; 262-273.

Sinigaglia, C. (2013). What type of action understanding is subserved by mirror neurons? *Neuroscience Letters*, 540; 59-61.

Shmuel, A., Augath, M., Oeltermann, A. & Logothetis, N. K. (2006). Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. *Nature Neuroscience*, 9; 569-577.

Smith, S. M. (2012). The future of FMRI connectivity. *Neuroimage*, 62: 1257– 1266.

Sotero, R. C. & Trujillo-Barreto, N. J. (2007). Modelling the role of excitatory and inhibitory neuronal activity in the generation of the BOLD signal. *NeuroImage*, 35; 149–165.

Southgate, V., Johnson, M.H., Osborne, T., Csibra, G., 2009. Predictive motor activation during action observation in human infants. *Biol. Lett.*, 5: 769–772

Stadler, W., Schubotz, R. I., von Cramon, D. Y., Springer, A., Graf, M. & Prinz, W. (2011). Predicting and memorizing observed action: Differential premotor cortex involvement. *Human Brain Mapping*, 32(5): 677–687.

Steinhorst, A. & Funke, J. (2014). Mirror neuron activity is no proof for action understanding. *Front in Hum Neurosci.*, 8: 333

Stephan, K. E., Fink, G. R. & Marshall, J. C. (2007a). Mechanisms of hemispheric specialization: insights from analyses of connectivity. *Neuropsychologia,* 45; 209-228.

Stephan, K. E., Harrison, L. M., Kiebel, S. J., David, O., Penny, W. D. & Friston, K. J. (2007b). Dynamic causal models of neural system dynamics: current state and future extensions. *Journal of Biosciences,* 32; 129-144.

Stephan, K. E., Kasper, L., Harrison, L. M., Daunizeau, J., den Ouden, H. E., Breakspear, M. & Friston, K. J. (2008). Nonlinear dynamic causal models for fMRI. *Neuroimage*, 42: 649 – 662.

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *NeuroImage*, 46(4): 1004-1017.

Stephan, K. E., Penny, W. D., Moran, R. J., den Ouden, H. E., Daunizeau, J. & Friston, K. J. (2010). Ten simple rules for dynamic causal modelling. *Neuroimage*, 49(4): 3099 3109.

Stephan, K. E., Weiskopf, N., Drysdale, P. M., Robinson, P. A. & Friston, K. J. (2007c). Comparing hemodynamic models with DCM. *NeuroImage,* 38; 387-401.

Summerfield, C. & Koechlin, E. (2008). A neural representation of prior information during perceptual inference. *Neuron, 59;* 336–347.

Tai, Y.F., Scherfler, C., Brooks, D.J., Sawamoto, N. & Castiello, U. (2004). The human premotor cortex is 'mirror' only for biological actions. *Curr Biol.*, 14(2): 117-20.

Thioux, M. & Keyser, C. (2015). Object visibility alters the relative contribution of ventral visual stream and mirror neuron system to goal anticipation during action observation. *NeuroImage*, 105; 380-394.

Umiltà, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al., (2001). I know what you are doing. A neurophysiological study. *Neuron*, 31; 155–165.

Urgesi, C., Maieron, M., Avenanti, A., Tidoni, E., Fabbro, F. & Aglioti, S. M. (2010). Simulating the future of actions in the human corticospinal system. *Cereb. Cortex*, 20(11); 2511-2521.

Valdés-Sosa, P. A. (2004). Spatio-temporal autoregressive models defined over brain manifolds. *Neuroinformatics,* 2; 239–250.

Valdes-Sosa, P. A., Roebroeck, A., Daunizeau, J., & Friston, K. (2011). Effective connectivity: influence, causality and biophysical modelling. *Neuroimage,* 58; 339-361.

Vogt, S., Buccino, G., Wohlschläger, A. M., Canessa, N., Shah, N. J., Zilles, K., Eickhoff, S. B., Freund, H., Rizzolatti, G. & Fink, G. R. (2007). Prefrontal Involvement in imitation learning of hand actions: Effects of practice and expertise. *NeuroImage*, 37(4); 1371-1383.

Wolpert, D. M., Doya, K. & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 358; 593–602.

Wolpert, D.M., Ghahramani, Z. & Jordan, M.I. (1995). An internal model for sensorimotor integration. *Science,* 269; 1880-1882.

7 Appendices

7.1 Appendix 1- Glossary of terms

Activity dependent plasticity – continuously modified activity-dependent efficacy of synaptic transmission.

Akaike Information Criteria (AIC) – A method to compare model evidence which is maximised when the approximating likelihood is closest to the true likelihood (Akaike, 1973). It approximates the model evidence and is asymptotically equivalent to those based on Bayesian factors.

Anatomical/structural connectivity – examines the presence of axonal connections. e.g. Diffusion Tensor Imaging (DTI) and Diffusion Spectrum Imaging (DSI).

Autoregulation – the perfusion of blood, thereby oxygen, to regions of the brain that require it – due to neuronal activity (a term that applies to all major organs; in this case, the brain).

Balloon Model – a series of differential equations that describe the coupling among hemodynamic state variables; relates to increasing blood flow, blood volume and deoxyhemoglobin content.

Balloon Model (Extended) – as above but extended by Friston (2000) and Stephan et al (2007c) to include non-linear functions and free parameters to provide a more plausible hemodynamic model.

Bayes factor – the result of comparing the likelihood of two models. It can be thought of as a likelihood-ratio test of the models to the given data, however, it integrates over-all parameters. A Bayes factor below 3 is not worth mentioning (Kass & Raftery, 1995).

Bayes Theorem (or Bayes rule) – Bayesian probability is a mathematical technique that allows us to reason about beliefs under conditions of uncertainty, by calculating conditional probabilities, e.g. probability of A given B.

Bayesian hypothesis testing – By using Bayes Theorem, can test the likelihood of models which have uncertainty to explain competing system hypotheses.

Bayesian Information Criterion (BIC) – a special case of the Laplace approximation (to the log-evidence) which drops all terms that do not scale with the number of data points (Schwarz, 1978).

Bayesian inversion – Combining measured data and a priori information (assumptions) to gain the posterior distribution (Prior distribution times the likelihood distribution or measurement).

Bayesian Model Averaging (BMA) - Averages across models, which in turn, incorporates the uncertainty about models into conclusions about parameters and predictions. It uses the entire model space chosen and computes weighted averages of each model parameter, where the weighting is given by the posterior probability of each model.

Bayesian Model Selection (BMS) – the probability of data given a model is integrated over unknown parameter values in that model. The resulting model evidence is then compared to identify the best model fit.

Bayesian Parameter Averaging (BPA) – Used in FFX, this computes a joint posterior density by combining individual posterior densities, treating the posterior from one participant as the prior for the next.

Bilinear (DCM) – the most common form of DCM; an input-state-output model of deterministic neuronal systems which has a known input and modulatory activity is restricted to direct modulations (as opposed to node modulations; see Nonlinear DCM).

Conditional Probability – the probability that one proposition is true provided that another proposition is true.

Conduction delays – the time required for an action potential to travel from the soma to the axon terminals (from origin to where synapses are formed).

Construct Validity – the question of whether the measurement used is truly addressing the hypothesis.

Control theory - a branch of mathematics and engineering whereby a reference is manipulated by a controller, entered into a system and an system output is measured. If the output is not desired effect then it can be reentered into the system to control the dynamic behaviour of the system.

Coupling parameter – the effective connectivity between two nodes.

Dendric Backpropagation – the Backpropagation of action potentials from the axon hillock back to the end of the dendric tree.

Effective connectivity – causal (directed) influences between neurons and neuronal populations; regional effects in terms of interregional connectivity, e.g. DCM, Granger causality, PPI.

Eigenvariate – the first eigenvariate is the weighted mean of the region of interest data that result in the time series with maximum possible variance (the mean of the voxels in the ROI).

Empirical priors – these are made of the hemodynamic parameters.

Endogenous activity – the activity between nodes where there is no modulation and just a stimulus input.

Exceedance probability – Our belief that a particular model is more likely than another. These probabilities all sum to one and can be used to rank models at the group level.

Exogenous activity – activity within the system that is affected by the modulatory input of a contextual variable.

Expectation Maximization (EM) Algorithm – an iterative method to find the maximum likelihood or maximum a posteriori (MAP) estimates. The expectation step creates a function for the log-likelihood for the parameters, and the maximisation step computes parameters maximising the expected log-likelihood found in the *E* step until convergence.

Face Validity – whether the measurement looks like it would attend to the hypothesis in question.

Field DCMs – could account for macro-scale propagation effects by incorporating elements of neural field theory.

Fixed effects analysis (FFX) – a method of group level inference where the optimal model is assumed to be the same for each subject in the population.

Forward model – a model of how neuronal activity causes the signals you observe (BOLD signal): a combination of neural and hemodynamic states.

Free Energy (also known as variational, negative or Gibbs) – places a lower bound on the model log-evidence and can be estimated using the Laplace approximation.

Free parameters – If a term is non-linear, as with DCM, a Bayesian framework can be used to estimate parameters based on constraints; making them more biologically plausible.

Frequentist statistics – inference is based on the frequency or proportion of data; it assumes that parameters are not random (e.g. ANOVA and *t*-tests).

Functional Connectivity – investigates statistical dependencies (temporal correlations) between regions, e.g. ICA.

Functional Integration – asks how brain regions influence each other, e.g. connectivity analysis.

Functional Segregation – asks which brain regions respond to a particular experimental input, e.g. conventional GLM analysis.

Gauss-Newton algorithm – when non-linear expressions are used (e.g. with DCM), this method can solve the 'least squares problem' as to fit a model to the data by minimising the sum of squares errors between the data and model's prediction.

Gaussian approximation – using standard units to estimate the area under the bell curve.

Gaussian distribution – an alternative term for normal distribution (a bell curve).

Generalized Filtering (GF) – a Bayesian filtering scheme that dispenses of the mean-field approximation and treat all unknown parameters and precisions as very small. This constraint is implemented by representing all unknown variables in generalised coordinates of motion - allowing one to optimise the moments of the joint posterior as data arrive. The resulting scheme enables an efficient assimilation of data and the possibility of online and real-time deconvolution.

Generative models – a model with hidden parameters that is usually then compared to its fit to real data (in relation to other models) by giving a conditional distribution (based on Bayes' Rule).

Granger Causality Modelling – tests for the signature of experimental influences by looking for correlations in the activity of two or more regions.

Group Bayes Factor (GBF) - the result of comparing the likelihood of two models at a group level.

Grubbs exponent – the relationship between cerebral blood volume and flow.

Hemodynamic Forward Model – a deterministic model of hemodynamic state equations which describe the coupling between these variables (see the Balloon Model).

Hidden States Variables - give rise to noisy observations through forward mapping (e.g., neurovascular coupling in fMRI) and this cannot be observed.

Hyperparameters – the parameters of the prior distribution.

Independent Component Analysis (ICA) – assumes that the components are non-Gaussian and independent, allowing for identification of their additive contribution to the signal.

Jacobian Matrix – a matrix of first order partial derivatives.

Knoecker product - an operation performed on to matrices that results in a block matrix (not to be confused with multiplication) – giving the matrix of the tensor product.

Kullback – Liebler Divergence – a measure in difference between two probability distributions. In DCM, it measures the information lost when x (the data) is used to approximate q (the model).

Laplace approximation - a method for using a Gaussian approximation to represent a given posterior density function.

Levenberg–Marquardt algorithm – a method for minimising a function over a space of parameters of the function. It can be thought of as Gauss-Newton but uses an approach where a subset of objective function is found, it is then expanded around it.

Log Likelihood ratio – the ratio of support for one model over another at a group level.

Log Model Evidence – model evidence of models from a group of participants.

Long-Term Synaptic Plasticity (LTP) - the ability to change strength in connections between two neurons in response to either use or disuse transmission over pathways; long term as last minutes or hours and only occurs at excitatory synapses.

Lyapunov exponent - determines a notion of predictability for a dynamical system.

Marginal likelihood ratio – an alternative term for the Bayes factor (20 to 1) is considered strong evidence.

Markov-chain – a random memoryless mathematical system where the next state depends only on the current state and not on the preceding.

Maximum a Posteriori Probability (MAP) – used to obtain the best guess of the unobserved (e.g. posterior). It is the most frequent of the posterior distribution and is obtained by EM.

Mean Field Model – the neuronal state equation contains random fluctuations on the neuronal states that can be modeled in terms of dynamics of the ensuing probability distribution over states if a population (Marrerios et al, 2009). This based on the mean-field theory, which is based on the assumption that the fluctuations around the average value of the order parameter are so small that they can be neglected.

Model evidence - the probability of the data given the model.

Model space – the arbitrary and subjective space that contains all models used for comparison.

Model Space Partitioning – by systematically separating model space into families, one can compare uncertainties of model structure systematically.

Monte Carlo Method of Sampling – a method of optimisation where repeated random sampling is conducted to obtain the same probabilities heuristically.

Nested Models – each node is visited twice, in DCM, this would be a connection to and from the node.

Neuronal dynamics – a model of the neuronal activity (indirectly inferred for DCM) that encumbers neurophysiological processes, e.g. synaptic transmission

Neuronal efficacy – the increase in perfusion signal elicited by neuronal activity.

Neurovascular Coupling – the relationship between neural activity and the changes in rCBF.

Neyman-Pearson Lemma - if no unknown parameters, N-PL is a justification of the likelihood ratio test which the test has the highest power.

Non-Nested Models - each node is only visited once, in DCM, this would be a connection to or from the node.

Nonlinear (DCM) – comparable to bilinear DCM but with the additional matrix *d* which states a nodes modulatory influence on a connection.

Occam's razor – the simplest explanation used, usually the correct one.

Optimal Model - This is the model which given alternative in model space, best explains the data.

Overfitting - where the model favours complexity and therefore begins to fit observational noise.

Partial Least Squares Regression (PLS) – a regression methods that aims to find a linear regression model that projects predicted and observed data**.**

Plastic DCM's – aim to represent different neurobiological mechanisms of synaptic plasticity more explicitly (could help with neuropsychiatric markers).

Positive Evidence Ratio (PER) – a complementary index to the GBF: the ratio of subjects with support for the model against no support.

Posterior family probabilities – as below but when families of models are compared

Posterior Probabilities (or revised probability) – the conditional probability assigned to a parameter (which is unknown) after the data is taken into account

Posterior Variance weighted averaging - a FFX method of group analysis whereby each MAP estimate is weighted by its posterior variance, while covariance between parameters are ignored.

Predictive Validity – the extent to which a measurement predicts results. In DCM, this can be done via simulations and then compare this to real world data to obtain face validity.

Principle Components Analysis (PCA) - a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences.

Principled priors – consists of temporal scaling (important to consider slice timing).

Prior (probability) distribution or prior – the uncertainty about *X* (e.g. a parameter) before the data is considered. This is uncertainty and not randomness, a fact that is important when using DCM.

Prior constraints – the restriction of where inputs can elicit extrinsic responses .

Probability density function (which Posterior and prior densities are calculated from) - a continuous distribution that describes the likelihood of events. One can compute the area under the curve (which is equal to one) to determine the probability of your event by taking the difference between two events.

Random Effects Analysis (RFX) – treats the model as a random variable and estimating the parameters of a Dirichlet distribution describing the probabilities of all models considered;an then compute how likely it is that a specific model generated the data of a randomly chosen subject (and the exceedance probability of one model being more likely than any other).

Regional Cerebral Blood Flow (rCBF) – the blood supply to a specific region in the brain at a given point in time.

Resting oxygen extraction – needs to be considered for DCM as it may be sensitive to the nature of the baseline measure.

Restricted maximum likelihood (REML) - a form of maximum likelihood estimation (estimates of the model's parameters) which is not based on estimates on a maximum likelihood fir of all the data but uses a probability function from a transformed data set (restricted) so nuisance parameters have no effect.

Savage-Dickey Density Ratio – provides a representation of the Bayes factor for testing an embodied model (which contains a nuisance factor) under the assumption that the conditional prior density under the alternative is equal to the prior under the null hypothesis.

Short-Term Synaptic Plasticity (STP) – the ability to change strength in connections between two neurons in response to either use or disuse transmission over pathways; short term as acts within tens of milliseconds to minutes.

Shrinkage priors – enable stability of the system by stating that the coupling parameters are zero unless otherwise stated.

Signal decay – a dampening of the rCBF (mediated by Nitric Oxide (NO)) signal which also suppresses the undershoot.

Single state per region (DCM) – each region has one state variable; a simple summary of the neuronal (i.e. synaptic) activity in a region.

Stiffness parameter (Grubbs vessel) – Volume and flow are always in a state of flux when an evoked event occurs. DCM tries to account for the nonlinearities in the flow-volume of the venous balloon that underpins the nonlinear behaviours trying to be explained

Stochastic (DCM) – alternatively to deterministic, stochastic DCM accounts for stochastic fluctuations in neuronal activity and their interaction with task-specific processes, e.g. can be applied to resting state data with no direct external input into a region.

Structural Equation Modelling (SEM) – Given a set of regions, directed connections are stated (causal relationship assumed a priori) and can set connections strengths so that so to minimise the discrepancy between the observed and implied correlations and thereby fit the model to the data.

Taylor series - a mathematical function that provides the exact value of a function for all values of x where that series converges.

Temporal averaging – averaging across time series. This is dependent on the events and trials being temporal aligned exactly.

Transit time – determines the dynamics of the signal and can be thought of as the resting venous volume divided by resting volume. If the mean transit time increases then the dynamics of the BOLD signal slow down with respect to the flow changes.

True Model – the true model cannot be obtained by DCM alone. The true model is only acquired via invasive techniques, e.g. local field potentials (LFP) – see optimal model.

True Posterior – as with the true model, the true posterior values can only be obtained via methods such as DTI.

Two State per region (DCM) – incorporates two state variables within a region to model the activity of inhibitory and excitatory population.

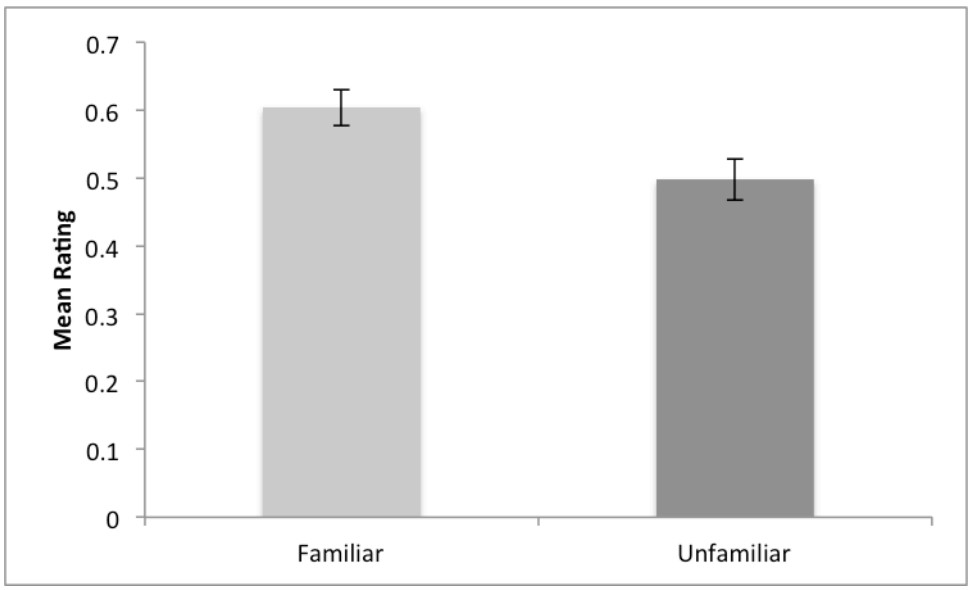
Variational Bayes – an extension of the EM algorithm which computes an approximation to the entire posterior distribution of the parameters.

Variational Laplace - used for Bayesian estimation of any non-linear model and assumes that the prior means and covariances are known (they are not estimated from data).

Volterra series – a model of non-linear behaviour; similar to the Taylor series but differs as it is able to capture memory effects (not dependent on a specific time of an input).

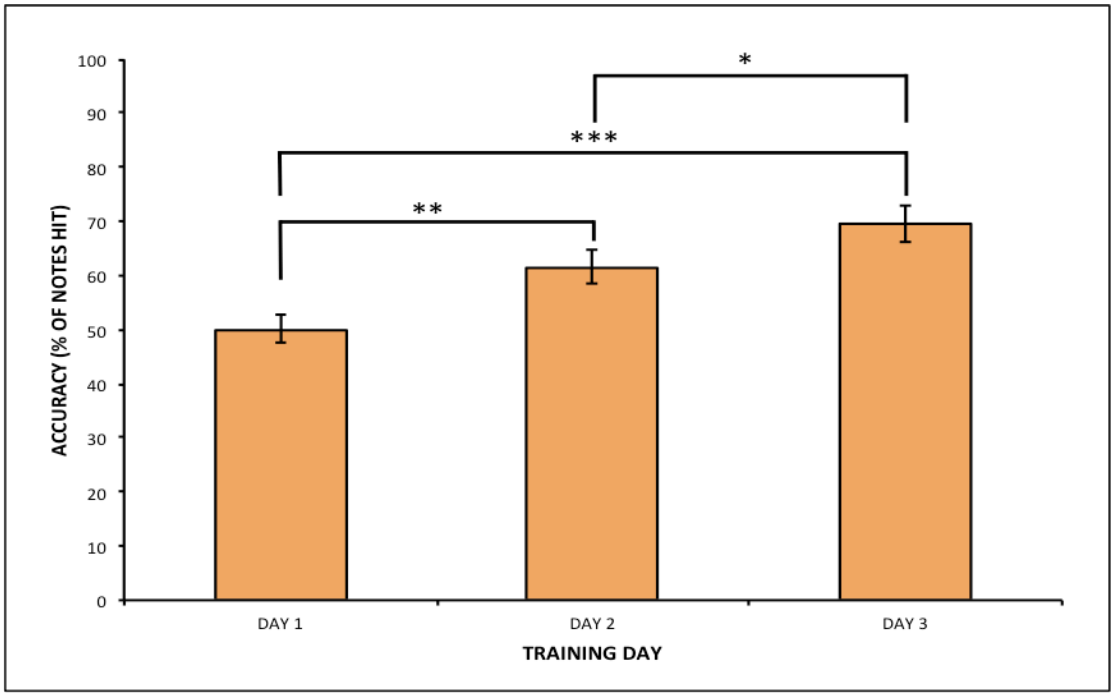
7.2 Appendix 2- Pilot data for Chapter III

Of the 157 videos piloted, mean ratings of these were taken from each participant and the 30 most familiar and 30 most unfamiliar were extracted. This allowed for differences of familiarity within the stimuli, confirmed by a t test, $t(20) = 2.802$, $p<0.01$ (two tailed).



Appendix 2. The mean rating of the reduced stimuli set.

7.3 Appendix 3- Pilot data for Chapters IV & V



Appendix 3. In order to determine whether there was a difference between average accuracy (% of notes hit) across the three days of training, a repeated measures ANOVA was used. The results of this analysis showed that was an effect of training day, $F(2,36)=27.95$, $p<0.001$, $\eta_p^2=0.61$. Post-hoc analyses using Bonferroni correction indicated that accuracy on training Day 2 ($M=63.10$, $SD=13.24$) was higher than Day 1 ($M=50.14$, $SD=13.72$), ($p<0.01$). Accuracy on Day 3 ($M=71.24$, $SD=12.87$) was greater than Day 2 ($p<0.05$) and Day 1 ($p<0.001$). This relationship is characterised by a linear within-subjects contrast, $F(1,18)=49.93$, $p<0.001$, $\eta_p^2=0.74$.

7.4 Appendix 4 - Endogenous connections for observed DCMs (Chapters V)

Appendix 4A. Results of a one-sample *t* test for coupling parameters of endogenous activity of movement: Observation Day 1

	From					
	L STG	**R STG**	**L IPL**	**R IPL**	**L IFG**	**R IFG**
To						
L STG	-	-	**0.3962 (0.0535) (*p*<0.0001)***	-	-	-
R STG	-	-	-	**0.5118 (0.0860) (*p*<0.0001)***	-	-
L IPL	**0.2953 (0.0416) (*p*<0.0001)***	-	-	-	**0.3919 (0.0510) (*p*<0.0001)***	-
R IPL	-	**0.3339 (0.0712) (*p*<0.0001)***	-	-	-	**0.4164 (0.0625) (*p*<0.0001)***
L IFG	-	-	**0.4687 (0.0713) (*p*<0.0001)***	-	-	-
R IFG	-	-	-	**0.5889 (0.0922) (*p*<0.0001)***	-	-

Data are mean (SEM) for each connection. -, Not investigated. The threshold was set at $p<0.002$ (corresponding to a FDR-corrected threshold of $p<0.05$ for multiple comparisons. ***** Significant connection.

Appendix 4B. Results of a one-sample *t* test for coupling parameters of endogenous activity of movement: Observation Day 2

	From					
	L STG	**R STG**	**L IPL**	**R IPL**	**L IFG**	**R IFG**
To						
L STG	-	-	**0.5003 (0.0688) (*p*<0.0001)***	-	-	-
R STG	-	-	-	**0.4955 (0.0653) (*p*<0.0001)***	-	-

L IPL	**0.3211 (0.0410) (*p*<0.0001)***	**-**	-	**-**	**0.4114 (0.0941) (*p*=0.001)***	**-**
R IPL	**-**	**0.3633 (0.0535) (*p*<0.0001)***	**-**	-	**-**	**-**
L IFG	-	-	**0.5863 (0.1173) (*p*<0.0001)***	**-**	-	**-**
R IFG	-	-	**-**	**0.4604 (0.0808) (*p*<0.0001)***	**-**	-

Data are mean (SEM) for each connection. -, Not investigated. The threshold was set at $p<0.002$ (corresponding to a FDR-corrected threshold of $p<0.05$ for multiple comparisons. ***** Significant connection.

7.5 Appendix 5 - Endogenous connections for executed DCMs (Chapters V)

Appendix 5A. Results of a one-sample *t* test for coupling parameters of endogenous activity of movement: Execution Day 1

	From					
	L STG	**R STG**	**L IPL**	**R IPL**	**L IFG**	**R IFG**
To						
L STG	-	-	**0.1222 (0.0145) (*p*<0.0001)***	-	-	-
R STG	-	-	-	**0.1241 (0.0114) (*p*<0.0001)***	-	-
L IPL	**0.0936 (0.0106) (*p*<0.0001)***	-	-	-	**0.1376 (0.0159) (*p*<0.0001)***	-
R IPL	-	**0.0982 (0.0048) (*p*<0.0001)***	-	-	-	**0.1217 (0.0111) (*p*<0.0001)***
L IFG	-	-	**0.2044 (0.0203) (*p*<0.0001)***	-	-	-
R IFG	-	-	-	**0.1979 (0.0195) (*p*<0.0001)***	-	-

Data are mean (SEM) for each connection. -, Not investigated. The threshold was set at $p<0.002$ (corresponding to a FDR-corrected threshold of $p<0.05$ for multiple comparisons. ***** Significant connection.

Appendix 5B. Results of a one-sample *t* test for coupling parameters of endogenous activity of movement: Execution Day 2

	From					
	L STG	**R STG**	**L IPL**	**R IPL**	**L IFG**	**R IFG**
To						
L STG	-	-	**0.1516 (0.0322) (*p*<0.0001)***	-	-	-
R STG	-	-	-	**0.1495 (0.0344)**	-	-

				(*p*=0.001)*		
L IPL	**0.1098 (0.0207) (*p*<0.0001)***	-	-	-	**0.1538 (0.0254) (*p*<0.0001)***	-
R IPL	-	**0.1234 (0.0254) (*p*<0.0001)***	-	-	-	**0.1475 (0.0257) (*p*<0.0001)***
L IFG	-	-	**0.2272 (0.0363) (*p*<0.0001)***	-	-	-
R IFG	-	-	-	**0.2314 (0.0359) (*p*<0.0001)***	-	-

Data are mean (SEM) for each connection. -, Not investigated. The threshold was set at $p<0.002$ (corresponding to a FDR-corrected threshold of $p<0.05$ for multiple comparisons. * Significant connection.