The Unconscious Formation of Motor and Abstract Intentions

DISSERTATION

zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt

der Fakultät Mathematik und Naturwissenschaften
der Technischen Universität Dresden

von

B.A.(Hon.) Philo/Psych., B.Sc. Phys. Soon Chun Siong

geboren am 26 Jul 1975 in Singapur

Gutachter:
Prof. Dr. John-Dylan Haynes (Berlin)
Prof. Dr. Thomas Goschke (Dresden)

Tag der Verteidigung: 10.04.2017

Die Dissertation wurde in der Zeit von Februar 2006 bis Juli 2009
im Max-Planck-Institut für Kognitions- und Neurowissenschaften angefertigt.

This thesis is dedicated to

my wife Huang Shuyun and my son Soon Chenxuan

for enduring forty-two long months of my absence

while I pursued my dreams.

Acknowledgement

This thesis has taken many years in the making, since the outline of the first study was chiselled out on the first day I set foot in the cold winter of Leipzig from the perpetual summer of Singapore in 2006. I must express my deepest gratitude for Prof. Dr. John-Dylan Haynes for making this possible despite the many obstacles along the way, and continuing to explore new avenues for me even after I left his lab. I will never forget the many summers and winters spent with my wonderful lab partners in Leipzig, Stefan Bode, Carsten Bogler, Christian Kalberlah, Chen Yi, Marcus Grüshow, Anna He, Annette Horstmann, Anita Tusche, Susann Szukalski and their partners / families in Germany and overseas, and lifelong personal friends, Lin, Raymond, Nora, Thomas, and Nicole, the wonderful people who gave me a home away from home. Max Planck Institute for Human Cognitive and Brain Sciences provided the perfect environment, technical, manpower and financial support to conduct our research. Finally, I am truly grateful to Technische Universität Dresden for accepting me as a PhD candidate, given the somewhat unusual circumstances. Special thanks must be given to Prof. Dr. Thomas Goschke for his support of my candidature, without whom this would not have been possible.

Statement

The following experimental studies included in this dissertation have already been published:

Study 1 (Chapter 2):

Soon, C. S., Brass, M., Heinze, H. J., & Haynes, J. D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, 11(5), 543-545.
doi:10.1038/nn.2112

Study 2 (Chapter 3):

Bode, S., He, A. H., Soon, C. S., Trampel, R., Turner, R., & Haynes, J. D. (2011). Tracking the unconscious generation of free decisions using ultra-high field fMRI. *PLoS One*, *6*(6), e21612.
doi:10.1371/journal.pone.0021612

Study 3 (Chapter 4):

Soon, C. S., He, A. H., Bode, S., & Haynes, J. D. (2013). Predicting free choices for abstract intentions. *Proc Natl Acad Sci USA*, *110*(15), 6217-6222.
doi:10.1073/pnas.1212218110

Table of Contents

Contact: soonchunsiong@gmail.com

List of Figures

List of Tables

Abbreviations

BOLD	Blood-Oxygen-Level Dependent
DMN	Default Mode Network
EEG	Electroencephalography
EMG	Electromyography
EPI	Echo-Planar Imaging
FDR	False-Discovery Rate
FIR	Finite Impulse Response
fMRI	functional Magnetic Resonance Imaging
FOV	Field Of View
FPC	Frontopolar Cortex
FWHM	Full Width at Half Maximum
LRP	Lateralized Readiness Potential
MEG	Magnetoencephalography
MVPA	Multi-Voxel Pattern Analyses
PCC	Posterior Cingulate Cortex
pre-SMA	pre-Supplementary Motor Area
RF	Readiness Field (a.k.a. *Bereitschaftsmagnetfeld*)
RP	Readiness Potential (a.k.a. *Bereitschaftspotential)*
SD	Standard Deviation
SE	Standard Error
SMA	Supplementary Motor Area
SVM	Support Vector Machine
TE	Time of Echo
TR	Time of Repetition

“Volition is surely an act of the mind, with which we are sufficiently acquainted. Reflect upon it. Consider it on all sides. Do you find anything in it like this creative power, by which it raises from nothing a new idea, and with a kind of Fiat, imitates the omnipotence of its Maker, if I may be allowed so to speak, who called forth into existence all the various scenes of nature? So far from being conscious of this energy in the will, it requires as certain experience as that of which we are possessed, to convince us that such extraordinary effects do ever result from a simple act of volition.”

Hume, 1748

An Enquiry Concerning Human Understanding, §VII

Abstract

The capacity to act voluntarily is an integral part of the subjective human experience: my mind makes a conscious decision to act, and this decision is executed by my physical body. Intuitively, I feel that my body is not acting under the command of an external force or agent, but under my conscious volitional control. This assumption, so fundamental to the human psyche, has been challenged in the last few decades by neuroscientific findings showing that even before one consciously decides to make a simple button press, a negative electrical potential that signifies voluntary movement preparation can already be detected over the cortical midline (Haggard and Eimer, 1999; Libet et al., 1983). This temporal precedence implies that unconscious neural activity – rather than conscious intentions – may play a substantial role in voluntary actions. The corollary is that our subjective sense of volition may be illusory: the conscious mind may simply *become aware* of unconsciously formed decisions.

However, to date these findings remain controversial due to various conceptual and technical issues (Baker et al., 2012; Banks & Pockett, 2007; Breitmeyer, 1985; Eccles, 1985; Gomes, 1998; McCallum, 1998; Pockett, 2002; Mele, 2009; Schlegel et al., 2013; Trevena and Miller, 2002). Crucially, the onset of the electrical potential precedes the conscious decision by only a few hundred milliseconds, ~300 ms (Haggard and Eimer, 1999; Libet et al., 1983; Trevena and

Miller, 2002)[1], and could be due to inaccuracies in judging or reporting the exact time of conscious decision with such high precision (Lau et al., 2006, 2007). Thus, it is difficult to conclusively establish the temporal order of the conscious decision and measurable neural changes at such short delays.

This thesis aims to provide more conclusive evidence regarding the timing of neural changes relative to conscious decisions, and to further elucidate the specific functional roles of early predictive neural activity in shaping the timing and content of upcoming decisions. Crucially, in addition to studying regions restricted to motor preparation, we sought to uncover neural networks across the brain involved in the genesis of high-level intentions. Three separate functional magnetic resonance imaging (fMRI) studies were conducted to study the neural dynamics of free decision formation. The higher spatial resolution of fMRI facilitated the use of sensitive novel pattern classification techniques that maximized information extraction from fine-grained local activation patterns in blood-oxygen-level-dependent (BOLD) signals. In each experiment, as participants spontaneously decided between two possible courses of action, we performed a whole-brain search for activity patterns that could predict their eventual choices, using highly sensitive multivariate pattern recognition techniques. If the specific outcomes of their decisions were encoded in neural activity even *before* they reached awareness, it would suggest that, at least in some decision contexts, our conscious choices may be related to preceding unconscious neural activity.

---

[1] But recent depth electrode recordings have shown that the firing rates of single neurons may start changing up to 9000 ms before the conscious decision (Fried et al., 2011).

In Study 1, we first searched across the brain for spatiotemporal patterns that could predict the specific outcome and timing of a free motor decision to make a left or right button press (Soon et al., 2008). In Study 2, we replicated Study 1 using ultra-high field fMRI for improved temporal and spatial resolution to more accurately characterize the evolution of decision-predictive information in prefrontal cortex (Bode et al., 2011). In Study 3, to unequivocally dissociate high-level intentions from motor preparation and execution, we investigated the neural precursors of *abstract* intentions as participants spontaneously decided to perform either of two mental arithmetic tasks: addition or subtraction (Soon et al., 2013).

Across the three studies, we consistently found that upcoming decisions could be predicted with ~60% accuracy from fine-grained spatial activation patterns occurring a few seconds before the decisions reached awareness, with very similar profiles for both motor and abstract intentions. The content and timing of the decisions appeared to be encoded in two functionally dissociable sets of regions: frontopolar and posterior cingulate/ precuneus cortex encoded the content but not the timing of the decisions, while the pre-supplementary motor area encoded the timing but not the content of the decisions. The choice-predictive regions in both motor and abstract decision tasks overlapped partially with the default mode network (DMN). Interestingly, the predictive information in the choice-predictive regions also peaked at around the same time as the activity in the DMN. This spatial and temporal coincidence hinted at a potential involvement of the DMN in unconscious choice preparation. High-resolution imaging in Study 2 further revealed that as the time-point of conscious decision approached, activity patterns in frontopolar cortex became increasingly stable with respect to the final choice.

By providing a more detailed characterization of the information dynamics of neural changes underlying the formation of free decisions, both motoric and abstract, our findings have hopefully shed more light on the question: do some voluntary actions in fact find their causal roots in unconscious – or preconscious – neural activity, rather than conscious decisions.

General Introduction

1.1 Neural precursors of voluntary decisions

A thought arises in my conscious mind, an intention to act. This mental decision is translated into an effect in the physical world via the actions of my musculoskeletal system. Almost every waking moment of our life, our intentions get acted out in this manner, sometimes according to what we envision, occasionally falling short. Subjectively, I feel that my body is not acting under the command of external forces or agents, but under *my* conscious volitional control. Under normal circumstance, while there may be external constraints over the possible courses of action available to me, ultimately my conscious mind appears to have sole dictatorship over my decisions: what I choose to do, and when to do it. This 'ability' to cause changes in the physical world according to my volition underlies my intuitive sense of freedom.

However, as Hume's (1748) astute analyses revealed, causal relations are never directly observed, and are nothing more than our assumptions based on the constant conjunction of two (or more) types of events. Similarly, perhaps my sense of conscious control over my actions is nothing more than a statistical inference based on observed regularities between my intentions to act and the intended bodily actions (Wegner, 2003). What I cannot observe, like neural activity in my brain, would naturally be left out of this inference of 'causation'. But is this merely an esoteric problem for epistemological skeptics, or does it point to a deeper truth underlying the relationship between mind and matter?

Indeed, there are many exceptions to the regularity of intentions and intended actions, for which our conscious intentions are not deemed to be necessary as causal explanations for our body's actions. There is general consensus that under a variety of conditions, the body's actions are not considered truly voluntary: autonomic reflexes; somnambulism (sleepwalking); acting under the influence of psychoactive substances like alcohol or lysergic acid diethylamide (LSD); during psychotic episodes or other mental disorders; etc. In these cases, the mind is either unconscious or in an altered state of consciousness, and the body's actions are not the results of conscious decisions. Thus, the body's actions are not considered truly free. This illustrates how important consciousness is to our intuitive concepts of free will: only actions that can be traced back to decisions made in a (normal waking) state of consciousness are candidates of freedom. In general, for other actions that are not the consequence of conscious volition, we seem to be quite willing to accept causal explanations that are purely physical, e.g., it is accepted as fact that the patellar reflex is due to neural signal transduction in the reflex arc mediated by the spinal cord. In other words, purely physical factors can compete with the mind over the control of our actions, thereby rendering them, at least to some extent, involuntary. It is curious that physical explanations for volitional actions, on the other hand, are often met with considerably more resistance.

Two lines of neuroscientific findings, however, provide evidence that our simple conscious intentions may themselves have physical causes, or at least neural precursors, that are themselves not directly associated with consciousness. One line of evidence comes from electrical stimulation of the brain. Stimulating frontal regions, including the pre-supplementary motor area (pre-SMA) and supplementary motor area (SMA), can elicit a conscious *urge* to perform a movement, sometimes coupled with actual movement (Fried et al., 1991). More recently, similar effects have been found in parietal cortex stimulation (Desmurget et al., 2009). Thus, it appears that one's volition can be directly shaped by stimulating specific regions of the brain. In other words, the 'causal' relationship from the mental to the physical can be reversed. One might argue that this is the exception rather than the rule: direct electrical stimulation of the brain is extraordinary, and its effects may not be representative of how normal intentions arise. The possibility of causing 'urges' artificially does not necessarily mean that every conscious intention is the result of physical causes.

However, another line of evidence from electroencephalography (EEG) studies suggests that even self-initiated intentions have unconscious neural precursors. The *Bereitschaftspotential* (pre-motor or readiness potential, RP), an electrical signature of voluntary movement initiation over the cortical midline (Gilden et al., 1966; Groll-Knapp et al., 1977; Kornhuber and Deecke, 1965; Libet et al., 1982), can be detected a few hundred milliseconds before motor execution, as measured by electromyography (EMG) detecting electrical activity in skeletal muscles. Crucially, the onset of the RP can be detected even *before* one has consciously decided to act (Haggard and Eimer, 1999; Libet et al., 1983). In other words, a voluntary motor action may be initiated by unconscious neural activity that precedes even the associated conscious intention. In the next section, we consider in greater detail what these studies reveal about the initiation of simple voluntary motor actions, and the respective roles of unconscious neural activity and conscious decisions.

1.2 Clocking intentions

To establish whether unconscious factors play a causal role in conscious decisions, one needs to show that information not readily available to conscious awareness can nevertheless shape our choices. For example, information contained in subliminal stimuli (or primes) can affect our perceptual decisions despite the lack of conscious representation (e.g., Dehaene et al., 1998; Greenwald, Klinger, & Schuh, 1995; Hsieh, Colas & Kanwisher, 2011; Lau & Passingham, 2007; Lin & Murray, 2013; Marcel, 1983a, 1983b; Murawski et al., 2012; Vorberg et al., 2003). However, in these studies participants are actively *responding* to external stimuli (but see Schlaghecken & Eimer, 2004), and may thus be more sensitive to the effects of subthreshold stimulation, which experimenters can manipulate. Also, such decisions are not truly representative of human volition since participants are simply indicating what they perceive rather than exercising their free choice. The situation is more complex when trying to study how self-initiated intentions are affected by unconscious neural activity that experimenters have no direct control over. Even if it can be shown that neural activity patterns contain information about specific choices that we make (e.g., Haynes et al., 2007), it remains unclear whether the activity is the cause or effect of one's conscious intentions. Also, such intention-related neural activity could be directly associated with conscious decision processes, or may reflect unconscious intention preparation.

Benjamin Libet's invention of an 'intention clock' effectively addressed these issues, pioneering the main neuroscientific approach for studying the initiation of voluntary actions (Libet et al., 1983) – he simply asked participants to report when they first became conscious of their decision to act. In his seminal study, this clock consisted of a spot of light revolving in a circle at 2.56 s cycles on the screen of a cathode ray oscilloscope. For each trial, participants made a spontaneous flexion of the fingers and/or wrist of his right hand (the act of volition) any time after one full clock cycle had passed. They then reported the clock position of the light when they first subjectively experienced ‘wanting’ or intending to act (W). Thus, Libet was able to determine with high temporal precision when the decision first entered into conscious awareness. Assuming that an effect does not precede its cause, the neural activity before the time-point W could not have been caused by the conscious decision. And yet, neural changes in terms of an early RP over the cortical midline could already be detected a few hundred milliseconds preceding W. As the intention to act did not yet exist in conscious awareness, any intention-related neural information in this time period could be considered ‘unconscious’. The authors “concluded that the cerebral initiation of a spontaneous, freely voluntary act can begin unconsciously, that is, before there is any (at least recallable) subjective awareness that a 'decision' to act has already been initiated cerebrally” (Libet et al., 1983, p. 623). This finding has since been consistently replicated using EEG (Haggard & Eimer, 1999; Lau et al., 2004; Schlegel et al., 2013) and direct cortical recordings (Fried, et al., 2011).

Haggard and Eimer (1999) expanded the intention clock paradigm to investigate not just the timing, but also the content of a spontaneous voluntary action. Instead of making a fixed voluntary action, participants were free to choose between pressing the left or right button with their respective index fingers. In addition to the RP, they also measured the lateralized readiness potential (LRP), the difference in event-related potential between the contra- and ipsi-lateral motor cortex for movement with a specific hand. This allowed the authors to dissociate the neural activity related to a specific motor decision from general state changes such as arousal or expectancy, which might increase before one decides to act (McCallum, 1988). This distinction was crucial in revealing neural changes that reflected the timing of the reported decision awareness time W. The onset of the LRP for early W trials was 193 ms before that of late W trials, i.e., the time of decision awareness was related to when motor preparations for the specific chosen hand began. In contrast, no differences in RP onsets were found between early and late awareness trials. These findings led the authors to conclude that our awareness of initiating an action relates to the preparation of a *specific* movement, rather than a general abstract intention to perform an action of some kind. (Note, however, that a recent study by Schlegel et al. (2013) failed to replicate these findings.)

1.3 Limitations of the EEG intention clock

While these pivotal EEG studies paved the way for the neuroscientific study of unconscious initiation of simple voluntary motor actions, they also invited wide-ranging criticisms (see Open Peer Commentary to Libet, 1985). A major line of criticism highlights potential errors in pinpointing the exact timing of conscious decisions with high precision (Breitmeyer, 1985; Joordens, van Duijn and Spalek, 2002; Latto, 1985; van de Grind, 2002). The perceptual difficulty of judging the precise position of a dot rotating at high angular velocity would only have been exacerbated by the additional dual-task demands of having to simultaneously monitor one’s internal mental state for an event (conscious decision) that may arise over time rather than happen instantaneously (Breitmeyer, 1985; Latto, 1985; Lau et al., 2006, 2007). Thus, a difference of ~300 ms does not seem sufficient to conclusively establish that the RP onset actually preceded the conscious decisions.

This concern might be alleviated if neural precursors of conscious decisions could be traced further back in time. Crucially, these EEG studies essentially focused only on the later stages of motor planning. It has since been confirmed by studying the *Bereitschaftsmagnetfeld* or readiness field (RF), the magnetoencephalographic equivalent of the RP (Deecke et al., 1982), in a lesion study (Deecke et al., 1987) and by using a conjunction of EEG and fMRI (Ball et al., 1999), that the RP originates from cortical regions involved in motor planning and execution, pre-SMA and SMA. It has been proposed that the motor RP might itself be preceded by unconscious neural activity in other regions related to the formation of intentions at a higher, non-motoric level (Haggar & Eimer, 1999; Libet et al., 1983; Groll-Knapp et al., 1977). However, in spite of its high temporal resolution, EEG may not be the optimal medium for providing the neural measurements needed to answer these questions, for the following reasons.

The first major issue is that of temporal range; various experimental constraints conspired to limit the temporal range that the EEG intention clock paradigm could interrogate. In fact, Haggard and Eimer (1999) reported that they could not find any clear baseline within the 2.6 s (one clock cycle) pre-movement epoch, i.e., the RP onset appeared to have occurred even earlier. Unfortunately, neural activity further back than one revolution of the intention clock could not be conclusively attributed to the spontaneous action. Participants triggered each trial with a key press, and had to wait for at least one full clock cycle before making a spontaneous motor decision. Thus, any neural activity earlier than one clock cycle before could potentially be due to the trial-initiating key press rather than the upcoming spontaneous action (Haggard and Eimer, 1999).

This limitation could conceivably be circumvented with some tweaks in the experimental design, but the issue of blinking, which causes strong EEG artefacts, poses greater challenges. Participants in the original Libet study (Libet et al., 1983) were explicitly asked not to blink once the clock revolution started, until they had made the spontaneous button press. If they did blink, they had to wait for another full revolution of the clock before making a ‘spontaneous’ button press, i.e., their ‘spontaneity’ could only occur between blinks. The seemingly innocuous requirement to avoid blinking would most likely have interfered with the spontaneity of the motor decision. The need to blink also effectively curtailed the temporal range of intention formation that EEG signals could interrogate based on this paradigm (Libet, 1985).

The second issue relates to the poor spatial resolution of EEG signals. To link any neural activity to intention formation, it would be more convincing if one could distinguish the signals for different choice options. Haggard and Eimer (1999) were able to use the LRP to differentiate the preparatory neural activity specific to each voluntary action (left and right button presses) because the signals came from spatially segregated motor cortices. However, this strategy might not be effective when probing the initiation of voluntary actions at a non-motoric level, as neural signals specific to comparable intentions would likely involve similar gross brain regions, albeit encoded by distinct fine-grained activation patterns that spatially overlap (Haynes et al., 2007). The poor spatial precision of EEG and the inverse problem of localizing the current source(s) (Niedermeyer & da Silva, 2004) means that neighbouring current sources cannot be effectively dissociated. While more advanced EEG source reconstruction methods are now available (e.g., Ball et al., 1999; Knösche et al., 1996), the spatial specificity and reliability are unlikely to be sufficient for differentiating abstract decisions.

1.4 General Paradigm

The above reasons prompted us to re-design the intention clock for use with pattern classification techniques in fMRI to search for unconscious neural signals related to high-level intention formation. The goal was to probe further back in time for the initiation of spontaneous decisions, rather than observe its manifestation at a late motoric stage. While fMRI has poorer temporal resolution than EEG, a longer time range of pre-decision neural activity could be interrogated, as there was no need to avoid eyeblinks since they were not associated with fMRI artefacts. Its superior spatial resolution also allowed us to capitalize on new multivariate classification techniques to uncover information encoded in fine-grained local activation patterns that would be lost in surface EEG signals (Haxby et al., 2001; Haynes and Rees, 2005; Kamitani and Tong, 2005; Kriegeskorte et al., 2006). If distinctive BOLD activation patterns within a cortical region could be used to decode different choice options, it would imply that the underlying neural activity was related to specific decisions and not generic preparation.

1.4.1 Time for a new intention clock

Given the challenges of reporting the exact time of conscious decisions with high precision, we chose instead to sacrifice temporal resolution for greater certainty in our intention clock. In all three studies, our intention clock was specifically designed to 1) optimize the certainty and accuracy of reporting conscious decision time, albeit at a lower temporal resolution, and 2) interrogate a longer period of time for intention formation (unlike previous EEG studies utilizing the Libet clock paradigm, we did not have to avoid eyeblinks).

To avoid inaccuracies in time judgement that could occur with stimuli rotating at high angular velocity (van de Grind, 2002), our intention clock consisted of a visual stream of consonants, presented serially at 2 Hz in Studies 1 and 2, and 1 Hz in Study 3. These temporal resolutions were slow enough for accurate identification of consonants, but more than sufficient for identifying the functional brain image acquired when the decision was made (time of repetition for each whole-brain image, TR, was 2.0 s in Studies 1 and 3, and 1.5 s in Study 2). Participants were asked to simply remember the consonant displayed on the screen when they became aware of their conscious decision. Consonants were presented in a pseudo-randomized order such that no repetition occurred within 8 consecutive consonants. The unpredictable sequence ensured that participants would not be biased by expectation effects when they reported which consonant was presented as their decision became conscious.

To identify neural signals specific to a decision rather than generic task preparation, in all studies, participants were completely free to choose between two different but comparable options: left or right button press in Studies 1 and 2; simple single-digit addition or subtraction in Study 3. The options were designed to be relatively balanced – in the sense that there was no specific reason to prefer one option over the other – with the aim to increase the likelihood of participants making free choices without worrying about reward or penalty in decision outcome. In other words, unlike economic decisions, there was no need for deliberative evaluation of the pros and cons of each option, thus facilitating the spontaneity of decisions. Nevertheless, in order to ensure sufficient trials for each choice, and to avoid biased sampling during classification, we conducted a behavioural selection test before each fMRI experiment. Only participants who chose both options with roughly similar probability – in spite of explicit instructions that balancing their choices was unnecessary, and that each choice should be made independently without regard to their history of choices – were selected for scanning. Participants were never informed that balanced choices were preferred throughout both behavioural and fMRI experiments.

As in Libet et al. (1983), we asked participants to relax as each trial began. In addition, they were specifically asked not to be too eager to perform a task, or to maintain a heightened state of readiness because they were taking part in an experiment. Rather, they should stay in a relaxed state of mind while monitoring when they first felt the urge to initiate a chosen voluntary action. This allowed us to observe earlier build-up of neural activity related to the formation of the specific intention, rather than just later stages of its execution. The goal was to evaluate the stable fMRI baseline up to a few seconds – rather than just hundreds of milliseconds – before the conscious decision for early signal changes that encoded the decision outcome.

1.4.2 Searching for intentions in dynamic brain patterns

Conventional univariate analyses of fMRI data typically involve running an independent t-test or general linear model (GLM) at every voxel, looking for information at the single-voxel level. Raw data is usually smoothed over a few neighbouring voxels (between 4 to 12 mm) to increase signal to noise ratio, i.e., signal differences between contiguous voxels are essentially treated as noise within a functional region. However, in recent years it has been shown that such local differences in activity may actually constitute fine-grained spatial patterns that encode valuable information, which can be revealed using multi-voxel pattern analyses (MVPA) (Carlson et al., 2003; Cox and Savoy, 2003; Haxby et al., 2001; Haynes and Rees, 2005; Kamitani and Tong, 2005; Kriegeskorte et al., 2006; Mitchell et al., 2004).

To uncover functional regions involved in the formation of free decisions in an unbiased fashion, we used a 'searchlight' approach (Haynes et al., 2007; Kriegeskorte et al., 2006) which examined the information contained in local spatial patterns of brain activity, one searchlight (local cluster of voxels) at a time. This was done independently for consecutive time-points before and after conscious decisions were made to yield a spatiotemporal brain map of decision-related information.

For each brain voxel we investigated whether its local environment contained spatial activation patterns that would allow decoding of the participant's decision. For a given voxel v_i we first defined a cluster of N voxels $c_{1\ldots N}$ found within a small sphere of fixed radius centered on v_i. The signal magnitudes of each brain voxel $c_{1\ldots N}$ in the fixed local cluster were extracted separately for each decision option A and B: left versus right button press in Studies 1 and 2; simple addition versus subtraction in Study 3. This was done separately for a range of time-points t before and after the conscious intention arose, and for each functional imaging run acquired from every participant. (Each experiment consisted of 10 short 5-min runs.) This yielded two N-dimensional pattern vectors $\mathbf{x}_{A,t,r,1\ldots N}$ and $\mathbf{x}_{B,t,r,1\ldots N}$ for each run r and time-point t, representing the time-resolved spatial response patterns in the local cluster in trials where the participant chose option A versus option B.

For a given time-point t and spatial position v_i we used linear support vector machine (SVM) classification (Müller et al., 2001; implemented in LIBSVM) to assess how much decision-related information was encoded in the activity pattern of the searchlight cluster. The pattern vectors $\mathbf{x}_{A,t,r,1\ldots N}$ and $\mathbf{x}_{B,t,r,1\ldots N}$ for one functional run was kept aside as an independent 'test' data set, while the other vectors were assigned to a 'training' data set. The signal magnitude for each voxel $c_{1\ldots N}$ in the searchlight cluster constituted one axis of an N-dimensional space. Each N-dimensional vector is a point in this N-dimensional space. The SVM algorithm essentially attempted to determine the hyperplane in this multi-dimensional space that optimally separated the two classes of training vectors, decision options A and B (with fixed regularisation parameter C=1).

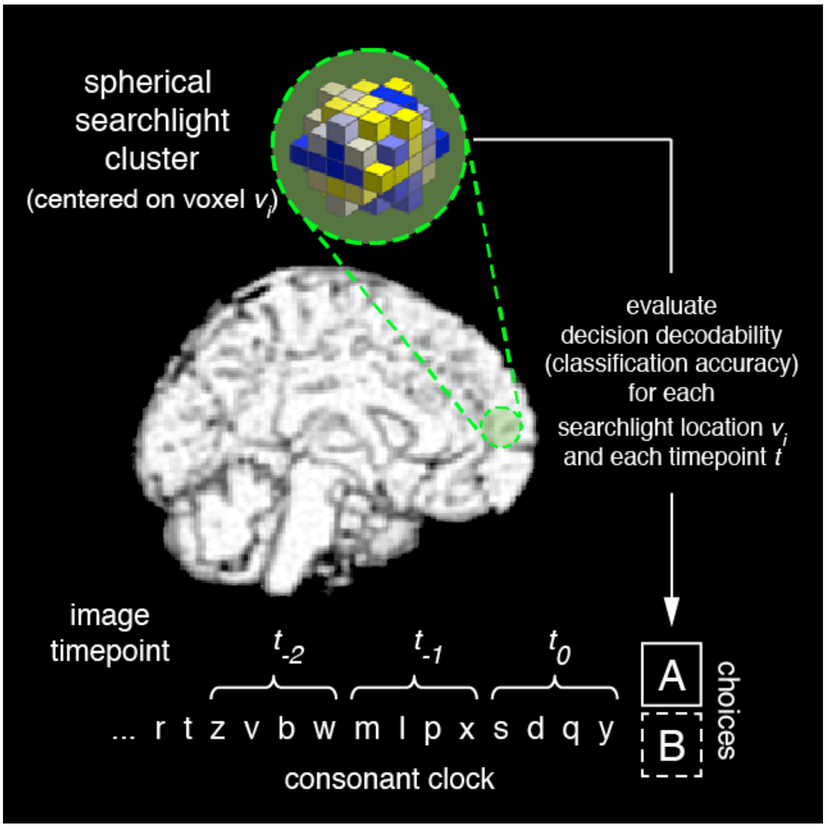


Fig. 1-1. Multi-voxel decoding from local spatial patterns using a moving 'searchlight'. Surrounding each voxel v_i a searchlight consisting of a spherical cluster of voxels is defined. A multivariate decoding algorithm based on support vector classifiers (Müller et al., 2001) is used to assess how accurately this local pattern of brain activity encoded the participant's decision to choose option A or B. For illustrative purposes, voxels more active for A are rendered in blue, and those more active for B in yellow. The decoding accuracy is estimated separately for a range of time-points immediately before and after the moment of conscious decision, which is reported via the consonant clock.

The amount of intention-related information present within this local cluster could then be assessed by examining how accurately the remaining vectors from the independent ‘test’ data set could be classified as being associated with choosing option A, or B. Good classification implied that the local cluster of voxels spatially encoded information about the specific decision of the participant. To ensure that the classification accuracy was not spurious, the training and test procedure was repeated multiple times, with each cross-validation cycle having a different run assigned as the test data set. The result was a decoding accuracy in the local environment of the central voxel v_i, averaged over all cross-validation cycles. This procedure was repeated for each time-point t and each spatial position v_i. This enabled us to create a 3-dimensional spatial map of decoding accuracy for each time-point t before and after the conscious decision. From these accuracy maps we could identify brain regions that encoded the specific decision outcome with above-chance accuracy at various time-points. We were also able to visualize for each region the information timecourse leading up to the conscious decision and following its execution. Comparing the temporal order of information across different decision-related regions provided an insight of the network of activity underlying intention formation and execution.

This searchlight approach was used to search across the whole brain for intention-related information in local brain activity patterns in all three experimental studies. Besides decoding *which* options participants chose, similar searchlight decoding analyses were also performed independently to predict *when* a conscious decision would occur in Studies 1 and 3. This allowed us to determine whether the same brain regions were involved in choosing what to do, and when to do it.

1.5 Overview of studies and research questions

In a series of three fMRI studies, we investigated the neural activity preceding conscious decisions to perform simple motor and abstract actions to look for decision-predictive information. The existence of such pre-conscious information that was predictive of subjects' choices would imply that our conscious decisions may be shaped by unconscious neural factors, via mechanisms which we are not explicitly aware of.

1.5.1 Study 1: Decoding the unconscious formation of motor intentions

We first looked for neural activity patterns across the brain that were predictive of spontaneous, free motor decisions (Soon et al., 2008). Specifically, we were interested in high-level networks involved in intention formation in addition to motor execution.

Research questions:

1. Can the outcome of a free motor decision (left- or right-handed button press) be predicted from neural activity patterns occurring before it enters conscious awareness?

2. Can the timing of a spontaneous motor decision (self-paced) be predicted from neural activity patterns occurring before it enters conscious awareness?

1.5.2 Study 2: Temporal stability of neural patterns involved in intention formation

In the second study, we sought to better characterize the spatiotemporal dynamics of neural information related to intention formation in prefrontal cortex using higher spatial and temporal resolution made possible by a more powerful 7T scanner (Bode et al., 2011[2]).

1. Can the findings from Study 1 be replicated?
2. Can improved spatial and temporal fMRI resolution shed more light on how intention-related information evolves over time and lead to a conscious decision?

1.5.3 Study 3: Decoding the unconscious formation of abstract intentions

In the final study, we used a novel abstract intention task that completely dissociated the abstract choice from motor responses to verify whether the neural networks involved in the formation of motor decisions were also involved in shaping abstract decisions (Soon et al., 2013).

1. Can the outcome and timing of spontaneous free abstract decisions (to perform addition or subtraction) be predicted from neural activity patterns occurring before it enters conscious awareness?
2. Are choice-predictive signals for both motor and abstract decisions related to default mode network activity?

---

[2] I was a co-author on this paper, and contributed to the design of the experiment and manuscript preparation.

Study 1

Decoding the Unconscious Formation of Motor Intentions

Soon, C. S., Brass, M., Heinze, H. J., & Haynes, J. D. (2008) Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*(5), 543-545.

doi:10.1038/nn.2112

Article first published online: 13 April 2008

2.1 Abstract

There has been a long controversy whether subjectively ‘free’ decisions are in fact determined by brain activity ahead of time. Previous reports that neural changes in motor-related cortex precede conscious decisions have been criticized for inaccuracies in subjective reports of decision time. These studies also left unanswered whether intentions to act are initiated in motor-related cortex, or in high-level brain areas. In an fMRI study, we found that the outcome of a decision could be decoded from brain activity of prefrontal and parietal cortex up to 10 s before entering awareness. Such a long delay could not be attributed to inaccuracies in measuring the time of conscious decisions. Instead it presumably reflected the operation of a network of high-level control areas that began to prepare upcoming decisions long before they entered awareness. This suggests that our free choices are determined by brain activity much earlier than commonly appreciated.

Key words: intention; decision; free will; self-paced; Libet; readiness potential

2.2 Introduction

The impression that we are able to freely choose between different possible courses of action is fundamental to our mental life. However, it has been suggested that this subjective experience of freedom is no more than an illusion and that our actions are initiated by unconscious mental processes long before we become aware of our intention to act (Haggard, 2005; Libet et al., 1983; Wegner, 2003). In a previous experiment (Libet et al., 1983), electrical brain activity was recorded while participants made a flexion their fingers and/or wrist as soon as they felt the urge to do so. Notably, their conscious decision to initiate a movement was preceded by a few hundred milliseconds by a negative brain potential, the so-called 'readiness potential' (RP) that originates from the supplementary motor area (SMA), a brain region involved in motor preparation. Since brain activity in the SMA consistently preceded the conscious decision, it has been argued that the brain had already unconsciously made a decision to move even before the participant became aware of it.

However, these intriguing experiments have left a number of controversial questions open (Glynn, 1990; Joordens et al., 2002; van de Grind, 2002). First, the RP is generated by the SMA, and hence only provides information about late stages of motor planning. Thus, it is unclear whether the SMA is indeed the cortical site where the decision for a movement originates or whether high-level planning stages might be involved in unconsciously preparing the decision (Bechara et al., 1997), as was seen in studies on conscious action planning (Burgess et al., 2001; Haynes et al., 2007; Hampton & Doherty, 2007; Koechlin et al., 1999). Second, the time delay between the onset of the RP and the decision is only a few hundred milliseconds (Haggard & Eimer, 1999; Libet et al., 1983). It has been repeatedly argued that potential inaccuracies in the behavioural measurement of the decision time at such short delays could lead one to misjudge the relative timing of brain activity and intention (Glynn, 1990; Haggard, 2005; Joordens et al., 2002; van de Grind, 2002). Third, does any leading brain activity indeed selectively predict the specific outcome of a choice ahead of time? To rule out the idea that such activity merely reflects unspecific preparatory activation (Lau et al., 2004), it is necessary to study free decisions involving more than one behavioural option (Haggard & Eimer, 1999; Haynes et al., 2007).

In an fMRI study, we investigated how neural activity started shaping a spontaneous, free motor decision before it entered awareness. We directly evaluated which regions of the brain were involved in the formation of an intention to act, and the time at which their influences began. While their brain activity was measured using fMRI, participants carried out a spontaneous motor-decision task, freely choosing to make a button press with either their left or right hand, and reported exactly when the conscious decision was made (Haggard & Eimer, 1999; Libet et al., 1983). Searchlight-based MVPA was used to evaluate whether it was possible to predict the specific outcomes and timings of the free motor decisions from local brain activation patterns before their conscious inception. The amount of decision-related information was independently assessed for each brain region at various time-points before (and after) the participants' conscious decisions. This highly sensitive technique allowed us to uncover a network of frontopolar and precuneus regions that were found to encode the decision outcomes up to 7 s before participants made their conscious decisions. The unexpectedly long temporal precedence of the unconscious neural information directly questioned the role of conscious decisions in simple acts of volition.

2.3 Methods

2.3.1 Participants

Thirty-six healthy right-handed participants who gave informed written consent (18 female, age range 21 – 30 years) were tested in a behavioural selection test, from which 14 participants (seven female, age range 21 – 30 years) were selected to participate in the fMRI experiment. All participants were healthy and had normal or corrected to normal vision, and had no history of neurological disease. Two participants were subsequently excluded from further analysis due to their behavioural performance during the fMRI session (see Results).

The study was approved by the local ethics committee at the Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. The study was carried out in accordance to the Declaration of Helsinki. Informed written consent was obtained from each participant before the study.

2.3.2 Behavioural paradigm

Participants performed a freely paced motor decision task while their brain activity was measured using functional magnetic resonance imaging (Fig. 2-1). At the beginning of each trial period, consonants were presented in the middle of the screen, one at a time for 500 ms without gap, and participants were asked to passively observe this letter stream. This modification to Libet's clock measurements was made to render the sequence unpredictable, and also to avoid inaccuracies in time judgement that can occur with rotating stimuli (van de Grind, 2002). The order of presentation was randomized under the constraint that there were no repetitions within a sequence of 8 consonants. Participants were told to relax and to press either the left or right button with the index finger of the corresponding hand immediately when they became aware of the urge to do so. They were to remember the consonant that was on the screen when they made the conscious decision which button to press (and not when the button was actually pressed). After the button was pressed, the screen went blank for 2,000 ms. Then, a response mapping screen was presented for 2,000 ms. This showed three consonants and a hash symbol (‘#’) arranged in a square configuration. The three consonants were the last three consonants that were presented (‘0-back’, ‘1-back’ and ‘2-back’ relative to the button press). The configuration of choices was randomized so that participants could not plan and prepare which motor action to execute before the response mapping screen appeared. Participants were asked to indicate which consonant was on the screen when they made the decision which button to press. This gave an indication of the time when participants became conscious of the intention which button to press, covering a possible range of 0 – 1,500 ms prior to the actual button press. This time period was chosen based on the pilot experiments. For example, selecting the ‘1-back’ consonant would indicate that the participant was conscious of the intention about 500 – 1,000 ms prior to the button press. If the consonant presented at the onset of the conscious intention was not available, participants were to select the hash (‘#’) symbol. Choice of this symbol indicated a delay of longer than 1,500 ms and suggested that the movement was not executed immediately when the conscious intention was felt, or that the participant was not paying attention to the consonants. The screen then went blank for 2,000 ms again before the next trial period began with the presentation of a new stream of consonants.

Importantly, in order to facilitate spontaneous behaviour, we did not ask participants to balance the left and right button selections. This would have required keeping track of the distribution of button selections (Spence and Frith, 1999) in memory and would also have encouraged preplanning of choices. Instead, we selected participants who *spontaneously* chose a balanced number of left and right button presses without prior instruction based on a behavioural selection test before. This was performed on a previous day and was carried out to select participants who were able to perform the task according to the above instructions. Each participant went through 10 runs, each lasting 5 minutes. A lateralization index was calculated to measure the ratio between total left button presses (L) and total right button presses (R) using the formula: (L–R)/(L+R). To increase the chances of getting approximately balanced distributions of left and right button choices during the fMRI session, only participants who had lateralization indices below 0.30 were selected for the fMRI experiment. Participants were not told of this criterion. In addition, participants were given only minor instructions to encourage the spontaneity of movement choice and execution. As in previous studies (Haggard and Eimer, 1999) we explicitly asked participants not to make button selections based on any kind of pattern. They were specifically asked not to be too eager to initiate a button press when the consonants first appeared, or to maintain a constant state of readiness for the movement. Instead, they should stay as relaxed as possible while looking at the consonants. This served two purposes. The first was to let their mental activity settle down to a stable state, so that any build-up of neural activity prior to the movement could be clearly observed. The second reason was to avoid the overlapping of haemodynamic responses from different trials, so as to facilitate unambiguous interpretation of the blood-oxygen-level dependent (BOLD) signal prior to the button press. It was stressed to participants that the time and choice of movement was completely up to them, but that it should be executed without hesitation once they made the decision which button to press. They were also asked to avoid any form of preplanning for choice of movement or time of execution.

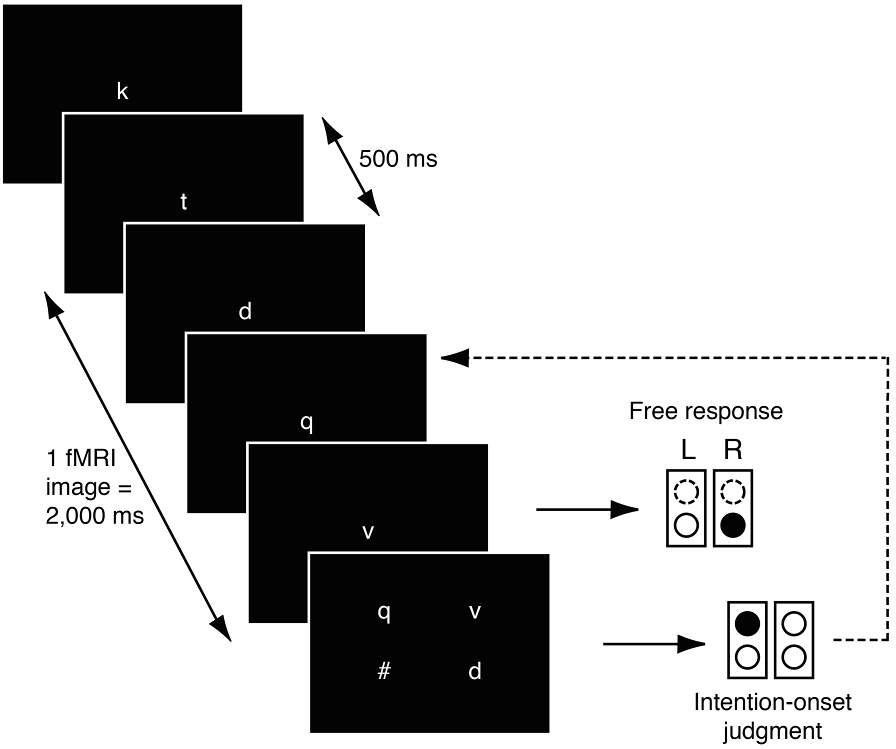


Fig. 2-1. Measuring the onset time of free conscious motor intentions. Participants viewed a letter stream that was updated every 500 ms (shown here only for a few frames). At some point they spontaneously made the decision to press either the left or right button using their corresponding index finger (free response). Subsequently, they were presented with a response-mapping screen that instructed participants as to which second button to press to report the time at which they consciously made the motor decision.

2.3.3 Functional imaging and preprocessing

A Bruker 3T Medspec 30/100 scanner (Ettlingen, Germany) was used to acquire functional MR EPI volumes with 30 slices at an isotropic resolution of 3×3×3 mm resolution covering prefrontal, parietal and most of temporal cortex (TR=2,000 ms; TE=30 ms; tilt 10 degrees axial to coronal; FOV 192×192×90). Ten runs of 150 functional MRI volumes were acquired for each participant. A 46-slice whole brain EPI image was also acquired to facilitate spatial normalization.

Data were preprocessed using SPM2 (http://www.fil.ion.ucl.ac.uk/spm). After discarding the first two images to allow for magnetic saturation effects, the remaining functional images were then realigned, correcting for head motion during the experiment. The functional images were then transformed into standard MNI space by first coregistering with a full-brain EPI image, followed by spatial normalization to the MNI EPI template. Then the functional images were subjected to two different types of analyses: 1) conventional univariate general linear model, and 2) novel searchlight-based multivariate pattern classification.

2.3.4 General linear model

We first performed conventional univariate analyses designed to identify brain regions where overall increases in neural activity occurred prior to the onset of a conscious decision. The preprocessed functional images were smoothed with a 6 mm FWHM Gaussian filter. Then a general linear model (GLM) was estimated for each participant with images concatenated across all runs. Since in our case there was no specific prediction regarding the temporal profile of the shape of the fMRI response timecourse, we used a finite impulse response (FIR) predictor to model fMRI responses (Henson, 2004). The freely chosen button selections were modelled using 26 FIR regressors, 13 for left and 13 for right button presses, covering a time range from 10 s before until 16 s after the button press. The second button presses with which participants indicated the onset time of their conscious intention were modelled as covariates consisting of single events convolved with a standard haemodynamic response function (HRF). The parameter estimates from the 26 FIR regressors of each participant were then entered into a second-level random-effects one-way ANOVA. The localisation of SMA and pre-SMA was based on standard criteria (Picard and Strick, 1996).

2.3.5 Multivariate pattern analyses

Next, we performed a set of multivariate pattern classification analyses based on the searchlight approach described in Chapter 1 (Haynes et al., 2007; Kriegeskorte et al., 2006; Norman et al., 2006) to identify cortical regions that contained decision-related information even before the conscious decision was made.

Decoding the content of motor decisions

The goal of our main multivariate searchlight analysis was to classify whether the participant was about to press the left or right button even prior to their conscious decision to do so. We began by estimating a modified general linear model for each run with regressors as described above but now based on unsmoothed data. This change was made to maximize sensitivity and allow extraction of the full information present in the spatial patterns, which would have been reduced by smoothing.

As described in Chapter 1, we interrogated local brain activity patterns for information related to the decision to press the left or right button using the searchlight approach (Fig. 1-1). As the searchlight picked out each spherical cluster (3-voxel radius) of neighbouring voxels, we extracted from the GLM of each run the 26 unsmoothed parameter estimates, 13 for left button decisions, and 13 for right, covering 10 s before to 16 s after the button press. To estimate the timecourse of intention-related information within this local cluster, we performed 13 independent sets of linear SVM analyses (Müller et al., 2001; LIBSVM implementation, http://www.csie.ntu.edu.tw/~cjlin/libsvm), one at each two-second time-point *t*.

In each set of classification analysis, SVM classifiers were trained on data from nine runs to distinguish the left button decisions from the right, and tested on the last run. This was repeated in ten cross-validation cycles, each time with a different run as the test run (10-fold cross-validation). The average decoding accuracy for each searchlight cluster was then assigned to the center voxel, and used to create 3-dimensional information brain maps of motor intention-related activity. This yielded 13 brain maps of predictive accuracy for each participant, one for each time-point relative to the onset of the conscious intention.

As the participants' EPI images had previously been normalized to a common stereotactic template it was possible to perform a second-level analysis where we computed on a voxel-by-voxel basis how well decoding could be performed on average across all participants from each time-point and each position in the brain. For this purpose the decoding images were smoothed with a 6 mm FWHM Gaussian filter. These smoothed brain maps of local decoding accuracy were entered into a one-way ANOVA with 13 levels, one for each time-point. Regions that predicted the subsequently chosen button were identified using a t-contrast based on all time-points prior to the decision onset (using a family-wise error correction for multiple comparisons, 50-voxel cluster-size threshold).

Control analysis: Decoding decision outcome from motion parameters

In order to exclude that decoding was influenced by potential head motion correlated with the participant's decision we investigated whether it was possible to decode the outcome of the decision from movement parameters obtained from motion correction. For this analysis the 6 motion correction parameters (x-translation, y-translation, z-translation, x-rotation, y-rotation and z-rotation) were extracted separately for left and right choice trials, and separately for 13 time-points *t* from 10 s before until 16 s after the conscious decision. This covered the same temporal range as the main decoding analysis. For each time-point *t*, this yielded two sets of 6-dimensional pattern vectors of each run, representing the estimated motion related to participants choosing left versus right button presses. We then used multivariate pattern recognition with a linear SVM to assess whether the participant's head motion provided any information about left versus right button choices at each time-point *t*. Again, the model was trained using data from nine runs, and tested on the independent tenth run. This training and testing cycle was repeated 10 times, with a different run as the test data set each time (10-fold cross-validation). For each time-point *t*, the classification accuracies across participants were then assessed using Student's t-test.

Decoding the timing of motor decisions

We also conducted a further decoding analysis where we assessed to which degree the timing of the decision, rather than its outcome, could be decoded. A multi-class pattern classification analysis was performed to identify cortical areas which contained early predictive information about when participants would decide to make a button press. As in the main decoding analysis described above, a spherical searchlight approach was used. For each spherical cluster (radius of 3 voxels), a multi-class support vector machine was trained to classify the time bin which a particular data point came from. Correct identification of the time bin prior to the time of conscious intention indicated that it was possible to predict how much later the conscious intention would occur. Six time bins (six possible classes) were interrogated, covering 10 s to 0 s before the button press. Again, 10-fold cross-validation was performed, each cycle involving nine training runs and one independent test run. The searchlight went through the whole brain, yielding 6 brain maps of classification accuracy for each participant, one for each time-point before the conscious intention. To assess statistical significance across participants, a one-way ANOVA with 6 levels, one for each time-point, was performed on smoothed decoding accuracy images (6 mm FWHM Gaussian filter).

2.3.6 Post-scan behavioural control: Unconstrained clock time

A post-scan behavioural control experiment was conducted to further corroborate the timing measurements obtained during the fMRI session. First, we addressed the question of whether the brief delay between intention and motor response reported by the participants might be due to the limited number of response alternatives for reporting the letter that was on the screen when the decision was made. Second, using the same procedure we also asked participants to rate the time when they decided to press *any* button rather than which *specific* button to press to find out if they involved different decision processes. This also explicitly tested for whether participants pre-decided either when to press a button or which button to press.

Eight out of 12 participants from the fMRI experiment who were still available returned for the post-scan behavioural control experiment. Participants performed the same free motor decision task as in the fMRI experiment. However, they were allowed to respond with any consonant on the keyboard in order to report the time of their motor decision (i.e., yielding 21 possible alternative times or up to 10.5 seconds).

In five consecutive runs, they reported when they decided to press a *specific* button, left or right (‘Left/Right decision’). In another five runs, they reported when they decided to press *any* (nonspecific) button (‘Time decision’). The order of runs was counterbalanced across participants.

2.3.7 Control fMRI experiment: Delayed motor intentions

Finally, in order to further investigate the functional roles of frontopolar cortex and precuneus, we conducted a control fMRI experiment to investigate which area first contained information about the decision outcome during the *conscious* free selection of a motor action. In this paradigm, the time-points of motor action selection and execution (after a few seconds’ delay) were both cued and thus under experimental control. This allowed us to identify informative brain regions at the precise time of selection, and to dissociate these from brain regions where information arose during the storage or response period.

Seven participants performed 10 runs of a delayed motor intention task. On each trial participants freely decided whether to make a left or right button press when shown the cue ‘select’. The chosen response was not executed immediately, but the choice had to be maintained over a variable delay period (randomly distributed between 4 to 10 s), and was executed when a second cue, ‘respond’, was presented. Each participant was cued to perform 16 such trials per run, resulting in 160 trials in total. In each run 120 volumes were scanned. Otherwise scanning parameters were the same as in the main experiment (30 slices with an isotropic resolution of 3×3×3 mm resolution covering prefrontal, parietal and most of temporal cortex; TR = 2,000 ms; TE = 30 ms; tilt 10 degrees axial to coronal; field of view 192×192×90).

After preprocessing as in the main experiment, a GLM was estimated for each run. The three phases ‘Select’, ‘Delay’ and ‘Response’ were modelled separately for left and right decision trials. The variable delay allowed for effective deconvolution of the three phases in each trial. We then investigated which cortical regions contained information about the decision (left or right button press) during each of the three phases using the searchlight approach, as described before.

2.4 Results

2.4.1 Behavioural results

Out of the 14 participants who were scanned, two were subsequently excluded from further analysis based on their behavioural performance during the fMRI session. One participant showed disproportionately more frequent selection of right button presses (lateralization index of –0.37), which might lead to unbalanced estimation accuracy of the BOLD response. The other participant selected the ‘#’ symbol for 24.5% of button presses, suggesting that many of her button presses were not spontaneous. The remaining participants had an average lateralization index of 0.01.

The freely paced button presses occurred, on average, 21.6 s after trial onset, thus leaving sufficient time to estimate any potential buildup of a ‘cortical decision’ without contamination by previous trials. Both the left and right response buttons were pressed equally often and most of the decisions (88.6%) were reported to be consciously formed within 1,000 ms before the button press (Fig. 2-2). On a small number of trials (8.5%) participants reported that the decision had preceded the response by between 1,000 and 1,500 ms. The '#' option was rarely chosen (1.4%) and mainly ensured that the participant was performing the task as instructed. Hence, in most trials the reported conscious decision occurred within the fMRI volume preceding the button press. Please note that the temporal resolution with which the judgement was measured was four times higher than the resolution of our fMRI measurements and thus fully sufficient for the present purposes of investigating long-term generation of conscious intentions to act.

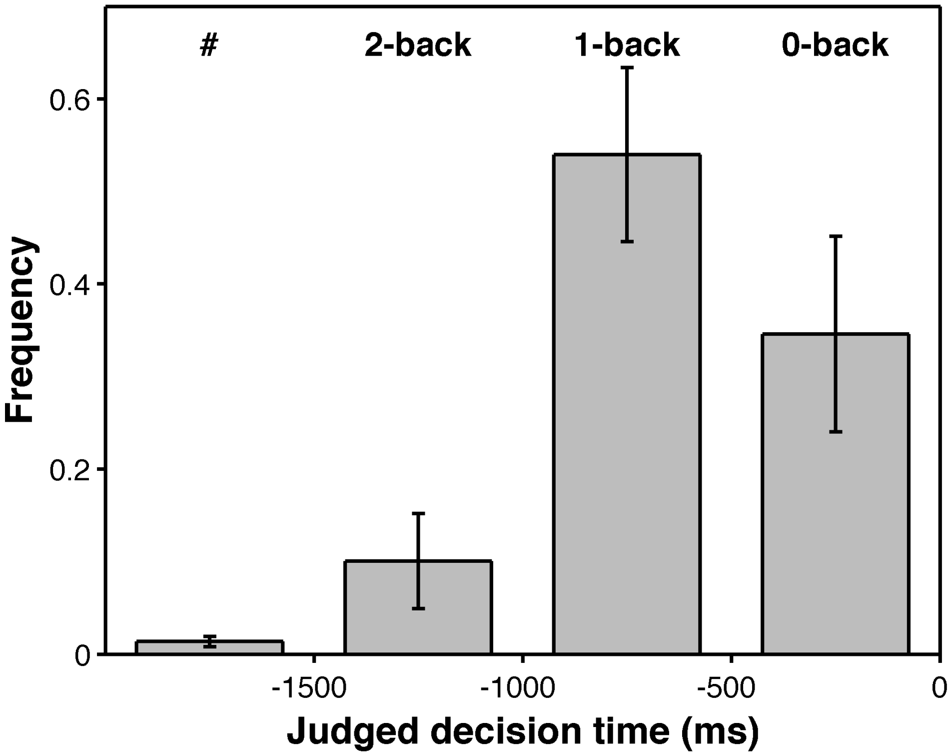


Fig. 2-2. Distribution of judged conscious decision time. In 90.1% of the trials, participants were conscious of their motor decisions within 1,000 ms before the freely selected button press.

2.4.2 Functional imaging results

Decoding the content of motor decisions

We directly assessed how much information each brain region contained about the specific outcome of a motor decision at various time-points before and after it reached awareness. We first investigated which brain regions the spontaneous motor decision could be decoded from *after* it had reached awareness and the chosen motor action was being executed. As expected, two sets of brain regions encoded the outcome of the participant's motor decision during the execution phase: bilateral primary motor cortex and SMA ($p<.05$, family-wise error corrected; Fig. 2-3).

Next, we addressed the key question of this study, whether any brain region encoded the participant's motor decision ahead of time. Indeed, we found that two brain regions encoded with high accuracy whether the participant was about to choose the left or right response prior to the conscious decision ($p<.05$, family-wise error corrected; Fig. 2-3, 2-4 & 2-5). The first region was in frontopolar cortex, BA10. The predictive information in the fMRI signals from this brain region was already present 7 s before the motor decision entered awareness. Taking into account the sluggishness of BOLD responses with respect to neural activity, the predictive neural information would have preceded the conscious motor decision by up to 10 s. There was a second predictive region located in parietal cortex stretching from the precuneus into posterior cingulate cortex.

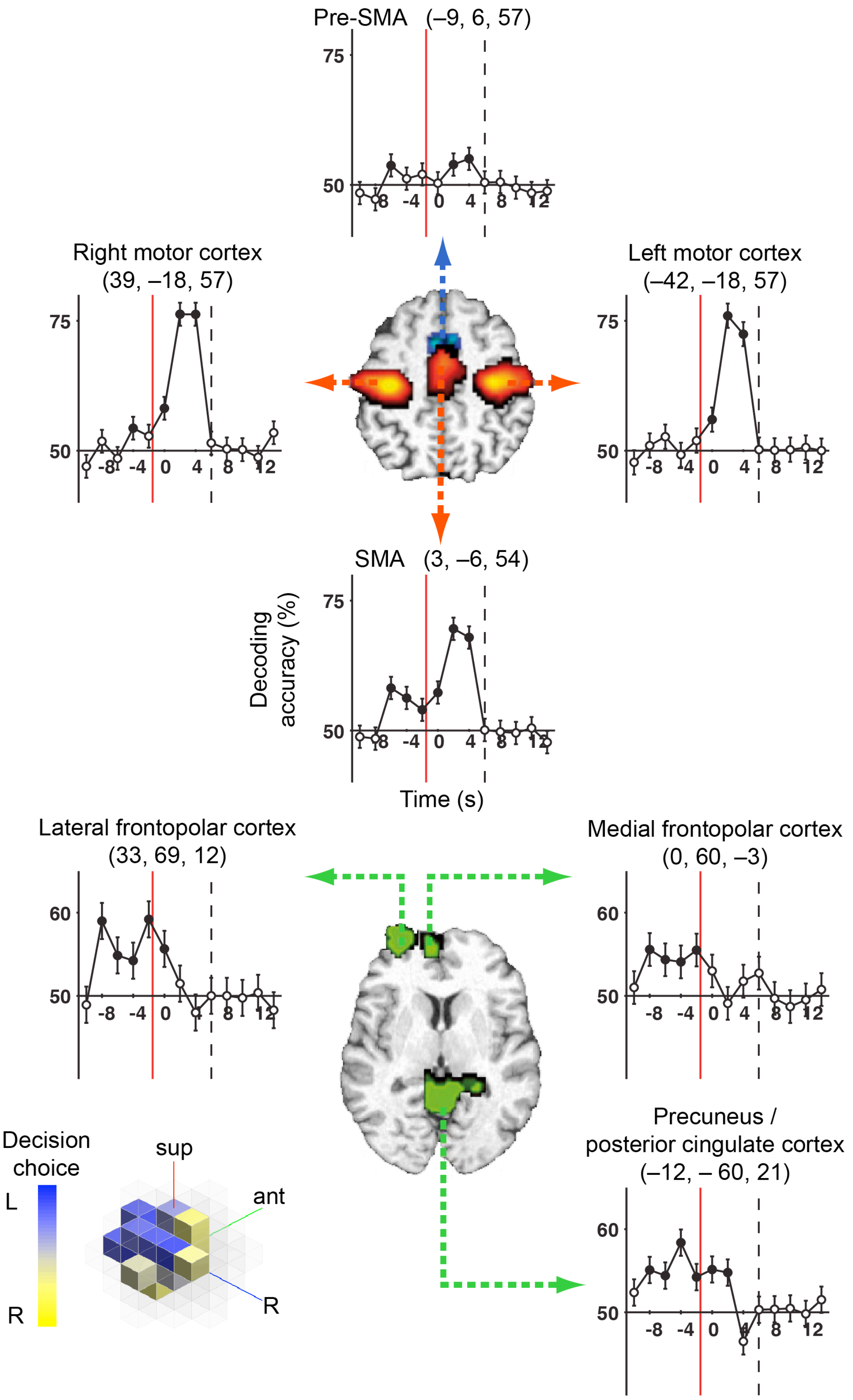

Pre-SMA (–9, 6, 57)
Right motor cortex (39, –18, 57)
Left motor cortex (–42, –18, 57)
SMA (3, –6, 54)
Decoding accuracy (%)
Time (s)
75
50
-8
-4
0
4
8
12
Lateral frontopolar cortex (33, 69, 12)
Medial frontopolar cortex (0, 60, –3)
Precuneus / posterior cingulate cortex (–12, – 60, 21)
60
Decision choice
L
R
sup
ant
R

Fig. 2-3. Decoding the outcome of motor decisions before and after they reach conscious awareness. Color-coded brain areas show regions where the specific outcome of a motor decision could be decoded before (bottom, green) and after (top, red) it had been made. The graphs separately depict for each time-point the accuracy with which the participants' free choices to press the left or right button could be decoded from the spatial pattern of brain activity in that region (solid line, left axis; filled symbols, significant at p<.05; open symbols, not significant; error bars, s.e.m.; chance level, 50%). The vertical red line shows the earliest time at which the participants became aware of their choices. The dashed (right) vertical line in each graph shows the onset of the next trial. The inset in the bottom left shows the representative spatial pattern of preference of the most discriminative searchlight position in frontopolar cortex for one participant (ant: anterior; sup: superior; see Fig. 2-6).

When the statistical threshold was relaxed and the cluster-size threshold removed, several other smaller regions became apparent that had significant predictive information, albeit less pronounced (Table 2-1). These regions were mainly aligned along the medial wall of prefrontal cortex, especially in anterior medial prefrontal cortex, and to a lesser degree near the SMA.

Table 2-1. Brain areas encoding intention prior to conscious decision

MNI coordinates			Z Score	Cortical region
X	Y	Z		
33	69	12	6.58	Frontopolar cortex
0	60	–3	5.12	Anterior medial prefrontal cortex
3	60	–15	4.90	Anterior medial prefrontal cortex
–21	45	9	5.26	Anterior cingulate cortex
3	18	51	4.97	Pre-SMA
–12	–60	21	6.98	Posterior cingulate cortex
3	–57	39	5.20	Precuneus

Note. Relaxed threshold: p<.00001, uncorrected; no cluster-size threshold.

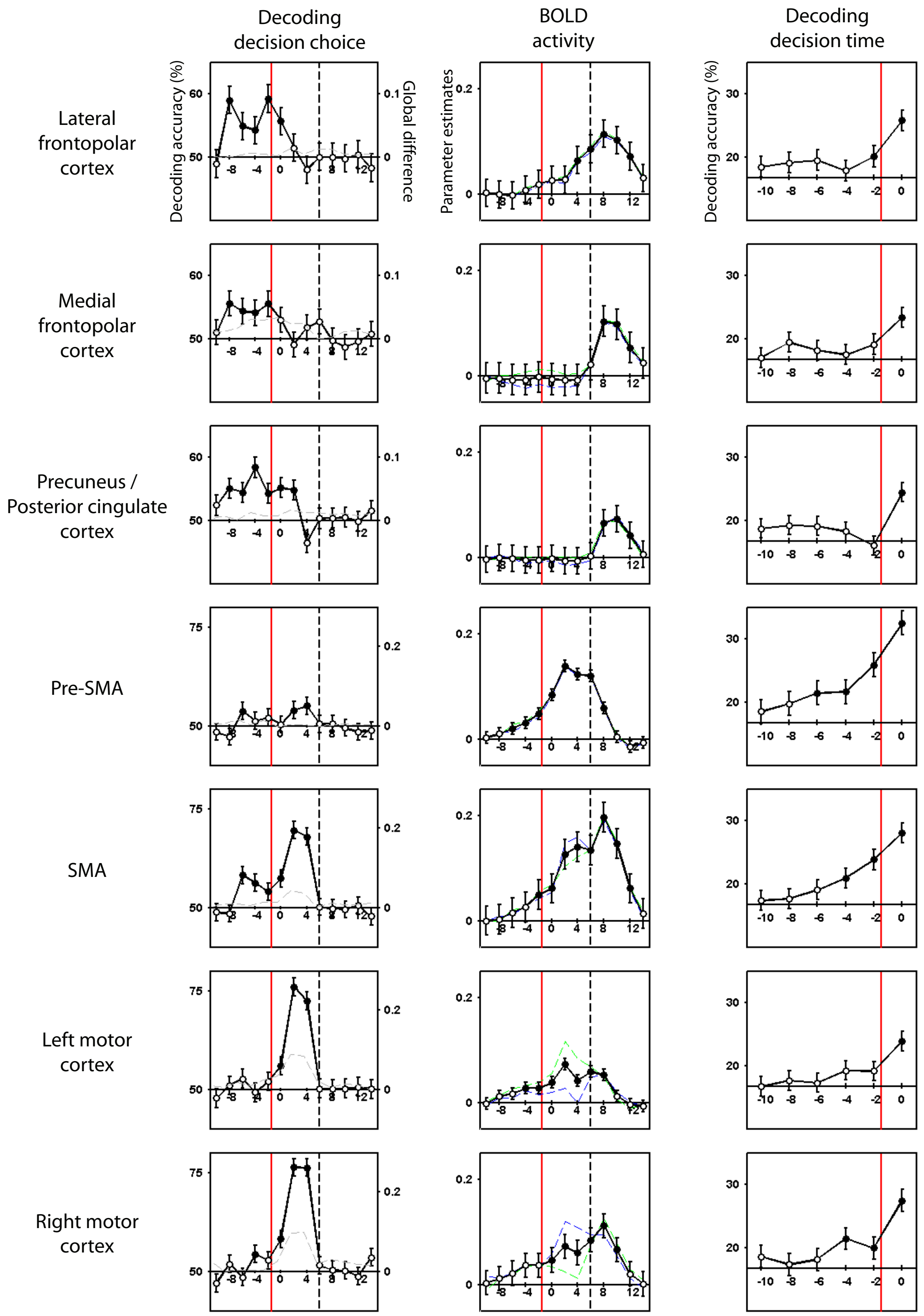


Fig. 2-4. BOLD and information timecourses in all regions containing intention-related information. The *left column* shows brain regions from which the *outcome* of a decision could be decoded either before or after it had reached awareness. The plot depicts separately for each time-point the

accuracy with which the participants' free decisions to press the left or right button could be decoded from the spatial pattern of brain activity in that region (solid line, left axis; filled symbols, significant at $p<.05$; open symbols, not significant; error bars, s.e.m.; chance level, 50%). The vertical red line shows the *earliest* time when the participants made their decision. The dotted (right) vertical line in each graph shows the onset of the *next* trial. The *middle* column shows BOLD activity averaged across searchlight voxels for each region and each time-point. The overall signal at each time-point was also plotted separately for the left and right button presses (green and blue dashed lines). The grey dashed line in the *left* column directly plots the difference between these overall activity levels (right axis). The *right* column shows the results of an additional analysis where a classifier was trained to recognize at which of 6 time windows the decision would be made rather than which outcome it would have (thus, chance level is here 16.7%).

In order to visualise the information dynamics across different brain regions, we plotted the time-resolved classification accuracies and BOLD timecourses of each region of interest (Fig. 2-4). To assess whether any predictive information was caused by overall signal differences between the two conditions, rather than by the local micro-pattern of brain activity, the overall signals for the left and right button presses (middle column blue and green dashed lines respectively), and the difference between them (left column grey dashed line) were also plotted separately at each time-point.

In all brain regions except the primary motor cortex, the overall activity for left and right decisions was virtually identical to the average response, and hence not visible on the graphs (Fig. 2-4). Only the contralateral motor cortex showed clear differences in overall signal for the two conditions, with right motor cortex being more strongly activated by left button presses and vice versa. Correspondingly, the decision could be accurately decoded at the time-points showing such global differences. As expected, this occurred *after* the participant had decided for the left or right button press, and reflected the execution of the chosen motor action. While these motor-related brain regions showed a gradual increase in activity across the pre-decision period, this activity was unspecific for the choices and carried no predictive information for the outcome of the decision (chance-level classification). In contrast, during this preparation period, the frontopolar and precuneus / posterior cingulate regions already encoded the content of the upcoming decision, even though there was no overall signal increase or difference between left and right choices. This suggested that the decoded information was based not on global differences as in the motor cortex, but on differences in local micro-patterns of fMRI signals (see for example, Fig. 2-6), which is presumably why it has not been noticed before in univariate analyses.

As might be expected, the decoding accuracy was generally higher in cortical areas involved in the motor execution of the response than in areas shaping the upcoming decision before it reached awareness (Fig. 2-3 & 2-4; note the difference in scale). However, when the BOLD signals from the early predictive regions, lateral frontopolar and precuneus / posterior cingulate cortex, were combined in a separate classification analysis, the prediction accuracy improved, as would be expected when pooling informative pattern signals (Fig. 2-5).

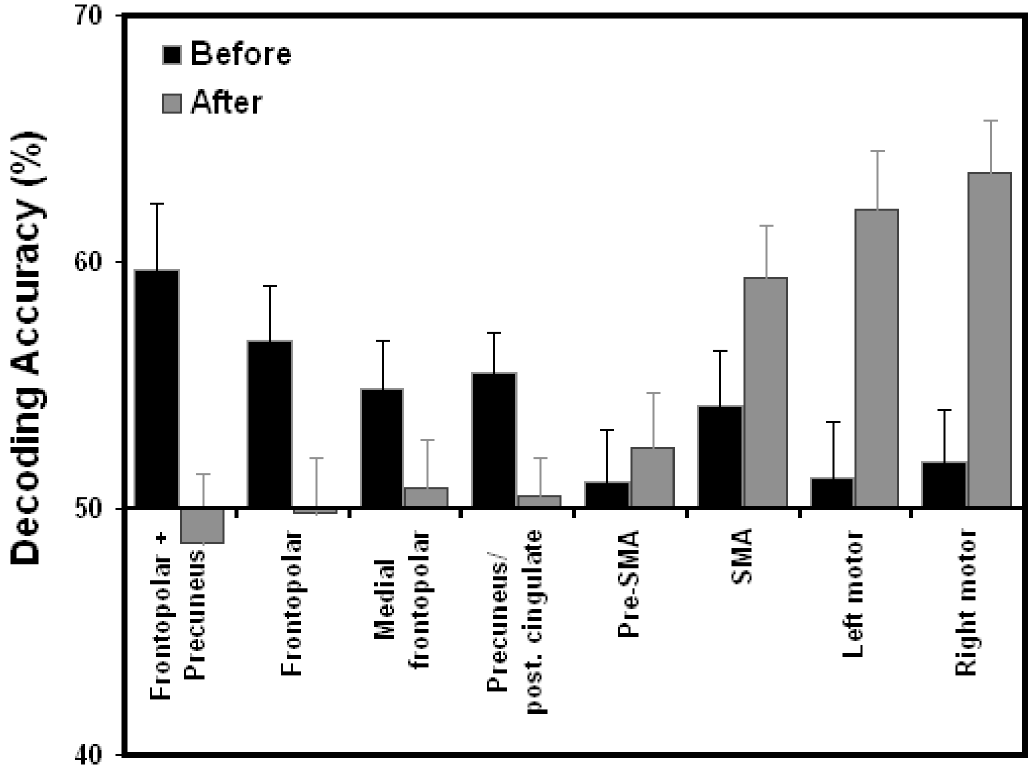


Fig. 2-5. Information profiles across multiple brain regions. Average decoding accuracy across 4 time-points before the decision (black) and 4 time-points after the decision (grey). The two leftmost bar plots show the *combined* decoding accuracy when pooling signals from the two most predictive brain regions: frontopolar and precuneus / posterior cingulate cortex.

Interestingly, there was a clear dissociation between brain regions encoding the outcome of the upcoming decision before it was consciously made and brain regions encoding this information following the decision (Fig. 2-4 & 2-5). Only SMA appeared to be involved in both phases. Even before the execution phase, SMA already showed an earlier choice-encoding phase beginning around 5 s prior to the conscious decision during which the upcoming decision could be predicted *before* it entered awareness (Fig. 2-4).

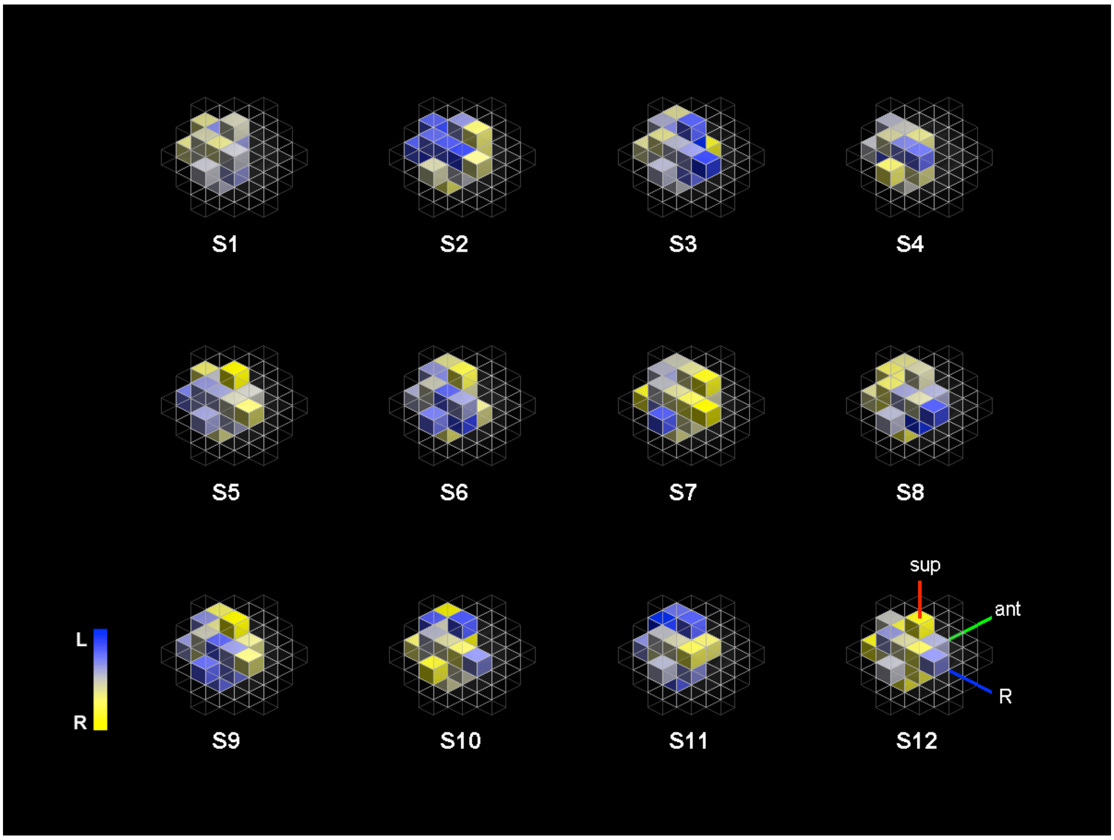


Fig. 2-6. Example of voxel selectivity for a representative searchlight. The spherical clusters at the position with peak decoding accuracy in frontopolar cortex are shown for all 12 participants. The selectivity for each voxel for either a left or right decision is colour coded in blue and yellow respectively. The selectivity profiles clearly indicate that some voxels are activated stronger preceding either left or right decisions, thus pointing towards a distributed encoding of long-term predictive information.

Control analysis: Decoding decision outcome from motion parameters

To verify whether our decoding results were influenced by motion we also attempted to decode the outcome of a decision directly from the estimated movement of the participant. However, the motion parameters contained no information related to the decision (*p*>0.2 for all time-points), hence precluding that our results were driven by head motion artefacts (Fig. 2-7).

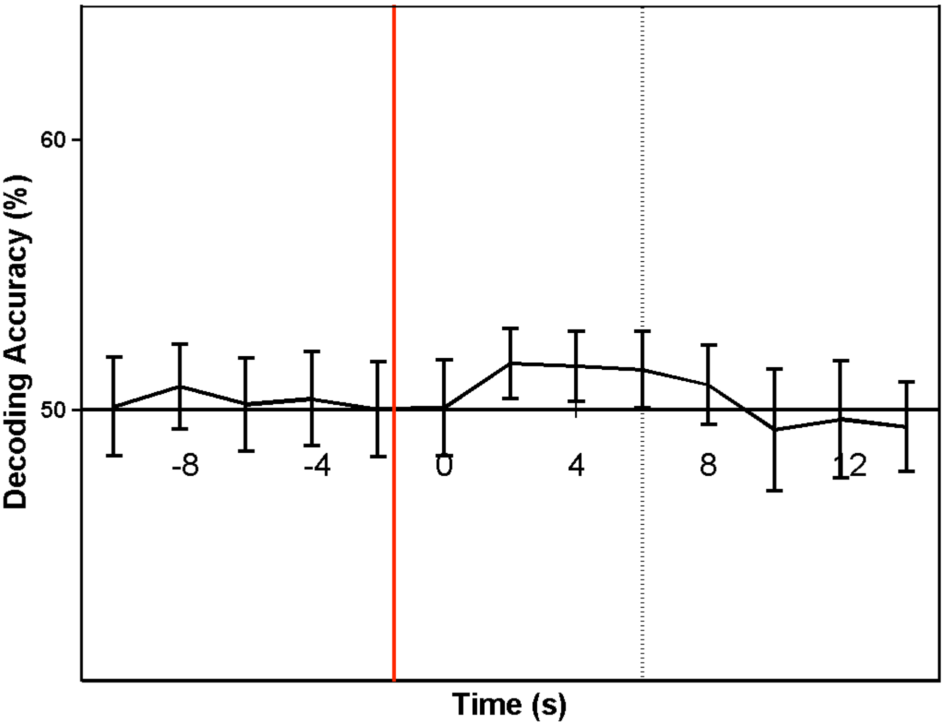


Fig. 2-7. Decoding the outcome of a decision from motion parameters. Head motion did not significantly predict button choice at any time-point (p>0.2 for all time-points).

Decoding the timing of decisions

We also assessed the degree to which the timing of the decision could be predicted ahead of time. Indeed, the neural activity in pre-SMA and SMA leading up to the motor decision was found to be informative about the decision timing (t-contrast based on all time-points, family-wise error correction, 50-voxel cluster threshold). Time-points from as early as 5 s before the motor decision onwards could be correctly identified (Fig. 2-4, right column). There was a linear increase in classification accuracy as the conscious intention approached. In contrast, in the frontopolar cortex timing information was only possible just before the motor decision.

As the time bins chosen for the analysis and the time relative to the decision were not independent, we performed an additional analysis to test whether decoding was indeed possible when focusing only on two early time-points. We conducted a new time decoding analysis using only two of the earlier time-points: 1) the earliest time-point showing above chance classification (6 seconds prior to decision), and 2) the earliest time-point interrogated (10 seconds prior to decision). As in the earlier analyses, a spherical searchlight (radius of 3 voxels) approach was used. For each spherical cluster, an SVM model was trained to classify which of the two time bins a particular data point came from. Again, 10-fold cross-validation was performed, each cycle involving 9 training runs and one independent test run. This yielded 2 images of classification accuracy for each participant, one for each time bin tested. These images were smoothed (6 mm FWHM Gaussian filter) before a second level random effects group analysis was performed. As in the earlier time-decoding analysis, decoding accuracy in the pre-SMA was above chance (56.7%, $p<.00001$, uncorrected). This indicated that, akin to EEG analyses on free decision generation (Haggard & Eimer, 1999; Libet et al., 1983), before a conscious decision was made, neural changes occurred in pre-SMA, resulting in activation patterns that were distinguishable from the preceding baseline activity. This in turn revealed that a decision to act would occur in a few seconds.

2.4.3 Results of post-scan behavioural control: Unconstrained clock time

In the post-scan behavioural control experiment, when allowed to indicate the time of conscious decision with any consonant instead of a limited number of choices, participants still indicated that the delay between decision and button press was below one second (Fig. 2-8). In separate experimental blocks, participants indicated either the time point of decisions for a specific button (as in the fMRI experiment) or the time point for the less specific decision to press *any* button. Both the decisions for the button and for the timing were made approximately at the same time just preceding the actual button press. There was no significant difference between the 'Left / Right' decisions (–962 ms, s.e.m. = 149 ms) and the 'Time' decisions (–866 ms, s.e.m. = 65 ms), $t_7<1$, and both decisions were made within one second before the freely selected button press. This confirmed that participants were making the motor decisions at a much later stage than the long-leading brain activity that was predictive of their response.

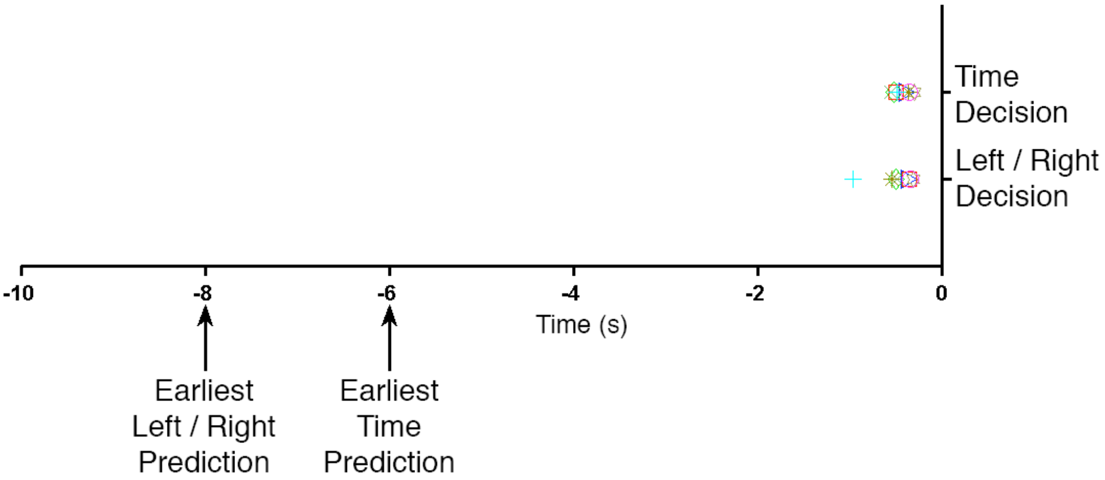


Fig. 2-8. Post-scan behavioural control experiment on decision timing. Participants were required to judge either the timing of their decision to press either the left or right button ('Left / Right Decision') or the timing of their decision to press any button ('Time Decision'). The symbols show the average reported timing judgements separately for each participant. Reports for both of these decision times were within 1,000 ms preceding the motor response (0 ms) and long after the earliest prediction of the Left / Right decision from brain activity.

2.4.4 Results of control fMRI experiment: Delayed motor intention

In a separate control experiment, we investigated whether the content of a delayed motor decision could be decoded from neural activity in frontopolar cortex and precuneus 1) at the point of decision ('Select'), 2) during a variable delay period after decision but before execution ('Delay'), and 3) during the actual execution of the motor action ('Response').

In a small percentage of trials (3.30%), participants failed to make a response within 2 s after the 'respond' cue was presented. These trials were discarded from the analyses because participants may not have been paying sufficient attention to the task during these trials. Overall, participants chose roughly the same number of left and right responses (all individual lateralization indices <0.20; average lateralization index = −0.04).

Through delaying the motor execution of the decision, a temporal order of decision-related information across brain regions became evident. Frontopolar cortex was the first cortical region which contained information about the decision once participants were cued to decide (Fig. 2-9). The decision outcome could already be decoded from frontopolar cortex – but not precuneus – during the 'Select' phase (p<.001, uncorrected). During the variable 'Delay' period both frontopolar cortex and precuneus contained predictive information about the decision (p<.001, uncorrected). During the 'Response' phase only precuneus contained information about the decision. If functional involvement could be inferred from information availability, this would imply different roles for frontopolar cortex (decision formation and maintenance) and precuneus (decision maintenance and execution).

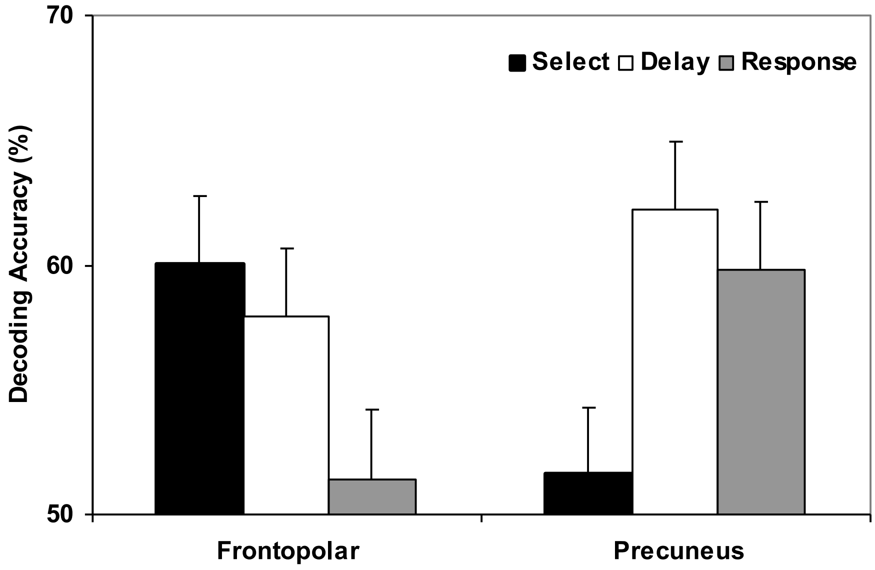


Fig. 2-9. Decoding the outcome of delayed motor intentions. Decoding accuracy in frontopolar cortex and precuneus in a control experiment where participants were cued when to freely select one of two hands for a response. Predictive information regarding the selected hand arose first in frontopolar cortex (MNI coordinates 39, 45, 12), already during the selection, and only later in precuneus (MNI –9, –57, 45). In contrast to frontopolar cortex, the precuneus continued to encode the chosen hand during the response period.

In addition, SMA (MNI 12, –3, 51) also contained information about which hand was selected, especially during the delay period, presumably because participants already started preparing for the upcoming motor action (Select: 59.3%, Delay: 67.3%, Response 60.0 %). Similarly, this information was available in the primary motor cortex once the decision was made (Select: 59.2%, Delay: 58.2%, Response 81.3 %; averaged across left and right). As in the main experiment, pre-SMA (–9, 6, 57) had no predictive information during the selection period, but some information was present during the delay period (Select: 51.8 %, Delay: 57.0 %, Response: 54.2%).

2.5 Discussion

In this fMRI study, we showed using time-resolved searchlight MVPA that a few seconds *before* a spontaneous free decision was consciously made, its outcome could already be predicted from local spatial activation patterns in frontopolar and precuneus / posterior cingulate cortex, even though there was no overall increase in BOLD signal in these regions. Once the decision reached conscious awareness, regions that were involved in executing the voluntary motor action, including the pre-SMA, SMA and bilateral primary motor cortex began to encode the content of the decision. In contrast, the timing of the decision could be decoded from the pre-SMA and SMA a few seconds before reaching awareness, but not from the frontopolar and precuneus / posterior cingulate regions. Thus, there appeared to be a double dissociation in the very early stages of motor intention formation, between brain regions shaping the specific decision outcome and those determining the timing of the decision. In the light of these and other related findings (Haggard & Eimer, 1999; Libet et al., 1983), there might be a need to examine the relationship between conscious will and brain processes.

Timing of intention-related information

Up to ~7 s before a participant consciously decided to press a button, decision-predictive information could already be detected in a network of frontal and parietal regions. Taking into consideration the delay in the haemodynamic response, this would suggest that when a participant's decision reached awareness it had already been influenced by unconscious brain activity for up to 10 s. This preparatory timeframe in high-level control regions was considerably longer than the few hundred milliseconds between RP onset and conscious decisions typically reported in previous

EEG studies for motor-related brain regions (Haggard & Eimer, 1999; Libet et al., 1983). Thus, it is highly doubtful that our findings were due to inaccuracies in the behavioural measurements of the subjective decision time, a common criticism of EEG studies of free decisions (Glynn, 1990; Haggard, 2005; van de Grind, 2002). In our study, the main level of activity in *motor-related* brain regions began to increase as early as 5 s prior to the decision (Fig. 2-4). On the one hand, this was an order of magnitude earlier than previous work on the RP (Haggard & Eimer, 1999; Libet et al., 1983). On the other hand, this was actually not inconsistent with Haggard and Eimer's (1999) observation that the motor-related RP had no clear baseline within 2.6 s before the movement, which implied that the onset of neural changes could have occurred much earlier. One potential explanation is that decision time in our study was unconstrained, and participants did not have to be concerned with eyeblinks, so possibly unspecific preparatory activity was able to build up over a longer period of time. Instructions to relax rather than maintain a state of readiness to act probably also allowed us to observe the neural build-up of the intention arise from a lower baseline.

It was unlikely that participants decided for one of the two response hands long before their button press. If an early conscious decision had been made for one of the two motor outputs, covert motor preparation for the chosen movement would be expected, as seen in the control experiment involving delayed motor intentions, where the response hand decision was made in advance and withheld until cued for execution. In other words, the response should have been decodable from primary motor cortex long before a button was pressed. However, in the main experiment the primary motor cortex only contained information much later than frontopolar cortex and precuneus, at the time of execution. This pointed strongly towards a late decision for one of the two responses, as also indicated by the psychophysical findings.

We also ensured that there was no carry-over of information between trials, so that the high decoding performance preceding the motor decision could not reflect decoding related to the previous trial. First, the classification accuracy increased with distance from the previous trial. Second, decoding of the current intention was at chance level as the subsequent trial began (Fig. 2-3, time-points after the dashed line), suggesting there was no carry-over of information to the next trial.

Please note that due to the temporal delay of the haemodynamic response, the small lead times in SMA/pre-SMA of up to several hundred milliseconds reported in previous studies (Haggard & Eimer, 1999; Kornhuber & Deecke, 1965; Libet et al., 1983) are below the temporal resolution of our method. Hence, we could not exclude that other regions contained predictive information in the short period immediately preceding the conscious intention. It would be difficult to ascertain whether such signals occurred before or after the decision in fMRI signals.

Functional specificity of intention-predictive regions

By studying free decisions with more than one behavioural option, we showed that their *specific* outcomes could be predicted from two regions in the frontal and parietal cortex of the human brain, namely frontopolar and precuneus / posterior cingulate. Our findings went substantially further than those of previous studies (Haggard & Eimer, 1999; Libet et al., 1983) by showing that the earliest predictive information was encoded in high-level control regions, and not in SMA or other motor-related regions. Although the LRP was also predictive of specific motor decisions, its onset was much later, around 1 s before the motor execution (Haggard & Eimer, 1999). Whereas the LRP originating from the primary motor cortex reflected the execution of the decision, predictive neural activity patterns in frontopolar and precuneus / posterior cingulate were likely involved in shaping the upcoming decision. We speculate that these intention predictive regions may be potential candidates as cortical origins for unconscious changes in skin conductance preceding risky decisions (Bechara et al., 1997).

The additional fMRI study of delayed motor intentions shed further light on the respective roles of the early predictive regions. Whereas frontopolar cortex was already predictive during the free selection period when the intention was formed, the predictive information was only available in the precuneus *after* the selection was made, during the delay and execution period. Taken together, the results from the two experiments suggested a functional dissociation between these two regions: the frontopolar cortex was the first cortical stage at which the actual decision was made, whereas precuneus was involved in the storage of the decision until it reached awareness, and eventual execution. This is consistent with previous work suggesting that the precuneus is involved in memory, self-referential processing and awareness (Cavanna & Trimble, 2006). An intriguing possibility is that the precuneus may be involved in binding one's intentions with the actual motor outcome, which is important for establishing a sense of agency over one's bodily actions (Aarts et al., 2004; Haggard, 2008; Sirigu et al., 2004; Wegner, 2003).

Notably, the choice was selected consciously in our control experiment, suggesting that similar networks might be involved in conscious and unconscious preparation of decisions. Interestingly, in our main experiment generic motor preparation appeared to begin in midline motor regions, SMA and pre-SMA, even before the decisions became conscious, in agreement with previous findings on the cortical midline RP (Haggard & Eimer, 1999; Libet et al., 1983). These regions were also most predictive of the time of conscious decision. The double dissociation between brain regions with early prediction of the timing but not choice (pre-SMA, SMA) and brain regions with late prediction of timing but early prediction of choice (frontopolar cortex, precuneus / posterior cingulate) is suggestive. We speculate here that as an action decision is made by the frontopolar and precuneus / posterior cingulate network, neural activity in the pre-SMA and SMA determine when the decided action will be executed, perhaps after a threshold is crossed (Schurger et al., 2012). Further studies are needed to verify if these different networks indeed decide separately what to do and when to do it, and whether such functional segregation generalizes to decision scenarios other than spontaneous motor actions.

Conclusion

Taken together, two specific regions in the frontal and parietal cortex of the human brain had considerable information that predicted the outcome of a motor decision the subject had not yet consciously made. This suggests that when the subject's decision reached awareness it had been influenced by unconscious brain activity for up to 10 s. This preparatory time period in high-level control regions is considerably longer than that reported previously for motor-related brain regions (Haggard & Eimer, 1999; Libet et al., 1983), and is considerably longer than the predictive time shown by the SMA in the current study (Fig. 2-4). Also, in contrast with most previous studies (Lau et al., 2004; Libet et al., 1983), the preparatory time period reveals that this prior activity is not an unspecific preparation of a response. Instead, it specifically encodes how a subject is going to decide. Thus, the SMA is presumably not the ultimate cortical decision stage where the conscious intention is initiated, as has been previously suggested (Eccles, 1982). Notably, the lead times are too long to be explained by any timing inaccuracies in reporting the onset of awareness, which was a major criticism of previous studies (Glynn, 1990; Joordens et al., 2002; van de Grind, 2002). The temporal ordering of information suggests a tentative causal model of information flow, where the earliest unconscious precursors of the motor decision originated in frontopolar cortex, from where they influenced the buildup of decision-related information in the precuneus and later in SMA, where it remained unconscious for up to a few seconds. This substantially extends previous work that has shown that BA10 is involved in storage of conscious action plans (Burgess et al., 2001; Haynes et al., 2007; Koechlin et al., 1999) and shifts in strategy following negative feedback (Hampton & Doherty, 2007). Thus, a network of high-level control areas can begin to shape an upcoming decision long before it enters awareness.

Acknowledgements

The authors would like to thank D. Passingham and H. Lau for valuable comments and T. Mildner and S. Zysset for help with scanning. This work was funded by the Max Planck Society and the German Federal Ministry of Education and Research.

Study 2

Temporal Stability of Neural Patterns Involved in Intention Formation

Bode, S., He, A. H., **Soon, C. S.**, Trampel, R., Turner, R., & Haynes, J. D. (2011). Tracking the unconscious generation of free decisions using ultra-high field fMRI. *PLoS One*, *6*(6), e21612.

doi:10.1371/journal.pone.0021612

Article first published online: 27 June 2011

3.1 Abstract

Recently, we demonstrated using fMRI that the outcome of free decisions can be decoded from brain activity several seconds before reaching conscious awareness (Soon et al., 2008, Study 1). Activity patterns in anterior frontopolar cortex (BA 10) were temporally the first to carry intention-related information and thus a candidate region for the unconscious generation of free decisions. In the present study, the original paradigm was replicated and multivariate pattern classification was applied to functional images of frontopolar cortex, acquired using ultra-high field fMRI at 7 Tesla. Here, we showed that predictive activity patterns became increasingly stable with closer temporal proximity to the conscious decision. Furthermore, detailed questionnaires exploring participants' thoughts before and during the decision confirmed that decisions were made spontaneously and participants were unaware of the evolution of their decision outcomes. These results give further evidence that FPC stands at the top of the prefrontal executive hierarchy in the unconscious generation of free decisions.

Key words: free decisions, Libet, multivariate pattern classification, ultra-high field fMRI

3.2 Introduction

As humans, we experience the ability to consciously choose our actions as well as the time at which we perform them. It has been postulated, however, that this subjective experience of freedom may be no more than an illusion (Haggard, 2005; Wegner, 2003), and even our goals and motivations can operate outside of our consciousness (Custers & Aarts, 2010). In Study 1 (Soon et al., 2008), we found empirical support for this view by showing that our conscious decisions may be shaped by prior *unconscious* neural activity, thus casting doubts on the role of consciousness in spontaneous voluntary motor acts. We adapted Libet et al.'s (1983) intention clock paradigm to fMRI to look for neural precursors of simple motor decisions. We were able to detect predictive neural changes up to a few seconds before participants consciously decided. This was on a much longer time scale than that of a few hundred milliseconds reported in previous EEG studies (Haggard & Eimer, 1999; Libet et al., 1983), and thus afforded a greater level of certainty that the decision-related information really occurred before the decision entered conscious awareness – our results could not be explained away by potential errors in reporting the exact time of conscious decision (Breitmeyer, 1985; Lau et al., 2007; Trevena & Miller, 2002; van de Grind, 2002).

By applying searchlight MVPA techniques to functional brain data from discrete time-bins before and after each decision, we were able to extract decision-related information from fine-grained spatial activation patterns that were not detectable using more conventional univariate analyses (Cox & Savoy, 2003; Haxby et al., 2001; Haynes & Ress, 2006; Kriegeskorte et al., 2006; Norman et al., 2006). We found a double dissociation between areas predictive of the decision outcome, i.e., ‘what’ to do, and areas that contained information about the decision timing, i.e., ‘when’ to do it. In agreement with previous studies, neural changes could be detected in pre-SMA / SMA indicative that a voluntary motor action would be initiated soon (Haggard & Eimer, 1999; Libet et al., 1983). However, these early changes likely reflected generic preparatory activity for decision execution, and were not predictive of which specific option participants would choose out of two possibilities. In contrast, the *earliest* information predictive of decision outcome was detected in high-level brain areas, frontopolar cortex (FPC) and precuneus / posterior cingulate (PCC) rather than motor-related regions. Further analyses suggested that these regions were likely involved in shaping the decision itself, rather than simply reflecting late stages of motor execution. Specifically, Study 1’s results clearly pointed to FPC as a prime candidate for the origin of the free decisions.

Here we replicated the Study 1 using ultra-high field fMRI on a 7-Tesla scanner, which allowed us to acquire images with 1×1×1 mm voxels. Specifically, we were interested in the role of FPC and thus only brain images from anterior FPC were recorded, which allowed a higher spatial resolution of the target region and a better temporal resolution (1.5 s TR) of the early components of the decision making process. These improvements allowed us to explicitly investigate the temporal stability of these early decision-related patterns, which was not addressed in Study 1. Furthermore, after the scanning session we assessed our participants’ behaviour and their thoughts during the experiment to investigate factors that may have biased the decision outcomes. This provided evidence in determining whether early predictive activity patterns already reflected conscious aspects of the decision process or whether these were related to truly unconscious components of evolving intentions.

3.3 Methods

3.3.1 Participants

Twelve right-handed participants (5 female, average age 24 years, age range 22–29 years) participated in the fMRI experiment. All participants were students of the University of Leipzig, enrolled in various fields of study. All were healthy and had normal or corrected-to-normal vision, and had no history of neurological disease. Suitable participants were selected by means of behavioural pre-tests, conducted within the 2 weeks preceding the fMRI session. These pre-tests consisted of 5 blocks of the same task as used for the fMRI experiment and ensured that only those participants who inherently fulfilled important criteria were selected for the fMRI session. First, the frequency with which a participant chose each of the two possible outcomes (left button or right button) needed to be balanced, meaning that one option should not have been chosen more than twice as often as the other. Second, we selected participants that ‘naturally’ performed trials at a moderate pace (i.e., at a speed of 15 to 50 seconds per trial). This pace allowed an optimal separation of fMRI signals for different trials. These first two criteria were not known to the participants such that they had maximal freedom in their decisions, but it was specifically emphasised that their decisions should be unbiased and spontaneous. Third, based on post-experimental questionnaires it was ensured that participants made spontaneous decisions and did not pre-plan them. The behavioural performance from the fMRI session was evaluated using the same criteria. We did not pre-select participants according to their level of intelligence or any other cognitive capacity. Data from one participant (S4) was discarded from all analyses due to relatively unbalanced decisions and exceptionally long trial durations. Data from one run of another participant (S12) had to be discarded due to technical problems with recording buttons.

The study was approved by the local ethics committee at the Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. The study was carried out in accordance to the Declaration of Helsinki. Informed written consent was obtained from each participant before the study.

3.3.2 Behavioural paradigm

The behavioural paradigm was essentially similar to Study 1. A stream of stimulus screens was presented at a rate of 2 Hz (Fig. 3-1). Each stimulus frame displayed a central letter on a dark background. Only consonants were used and were presented in a pseudo-randomised order such that the same letter never occurred twice within 8 consecutive frames. Participants were instructed to passively view the stream of letters, relax, and refrain from thinking about the upcoming task. The index and middle fingers of both hands rested on 4 buttons of two joysticks. Participants were free to decide, at any time, to press the left or the right button with the corresponding index finger. As soon as they were aware of their decision, participants were to note the letter presented on the screen. The time at which participants are first aware of their decision will hereafter be referred to as the ‘decision time’ in short. Participants were instructed to then immediately perform the chosen action without any delay. Once a button was pressed with the index finger, a decision-time response screen was presented after a variable delay of 1, 1.5, 2 or 2.5 s. The delay variability was introduced to facilitate the deconvolution of the BOLD responses for the two button presses: 1) free-decision, and 2) decision-time report. This screen contained three letters and an asterisk arranged in a square. Participants were then to indicate the letter noted at the decision time by pressing the corresponding button on the joystick, the four buttons now corresponding to the position of the letters (or asterisk) on the screen. The three letters always corresponded to the three letters shown immediately prior to and including the button press (0, −0.5 s and −1.0 s relative to the button press). Their positions were also randomized. If participants were unable to recall the letter presented at the time of the button press, or if the relevant letter was not displayed, they were told to select the asterisk. Using a letter stream as a timing device allowed us to detect whether decisions were planned ahead of time (see Study 1, Soon et al., 2008). Upon completion of the trial, the stream of stimuli resumed. Participants again relaxed and passively viewed the stream of stimuli until the next decision was spontaneously made. Note that participants were only instructed to relax and not to pre-plan their decisions at any time; the pacing and the ratio of left and right decisions were deliberately left to spontaneity in order to avoid any artificial restrictions of the free decision process. As described above, pre-tests were used to select only those participants for the fMRI session who inherently showed optimal behavioural performance.

After the scanning session, participants completed a questionnaire about their subjective experiences with the experiment. They were asked to report their thoughts and behaviour during the experiment, even if these had contradicted the task instructions. Participants were asked to rate on a five-point scale: I) how often they made a decision earlier during the trial but waited before executing the button press (0 = *never*, to 4 = *always*); II) their spontaneity throughout the experiment (0 = *not spontaneous at all*, to 4 = *very spontaneous*); III) how often they explicitly thought about the timing of their decisions (0 = *never*, to 4 = *always*). Additionally, they had to describe IV) the content of their thoughts between trials, and V) if they noticed any changes in their performance during the experiment.

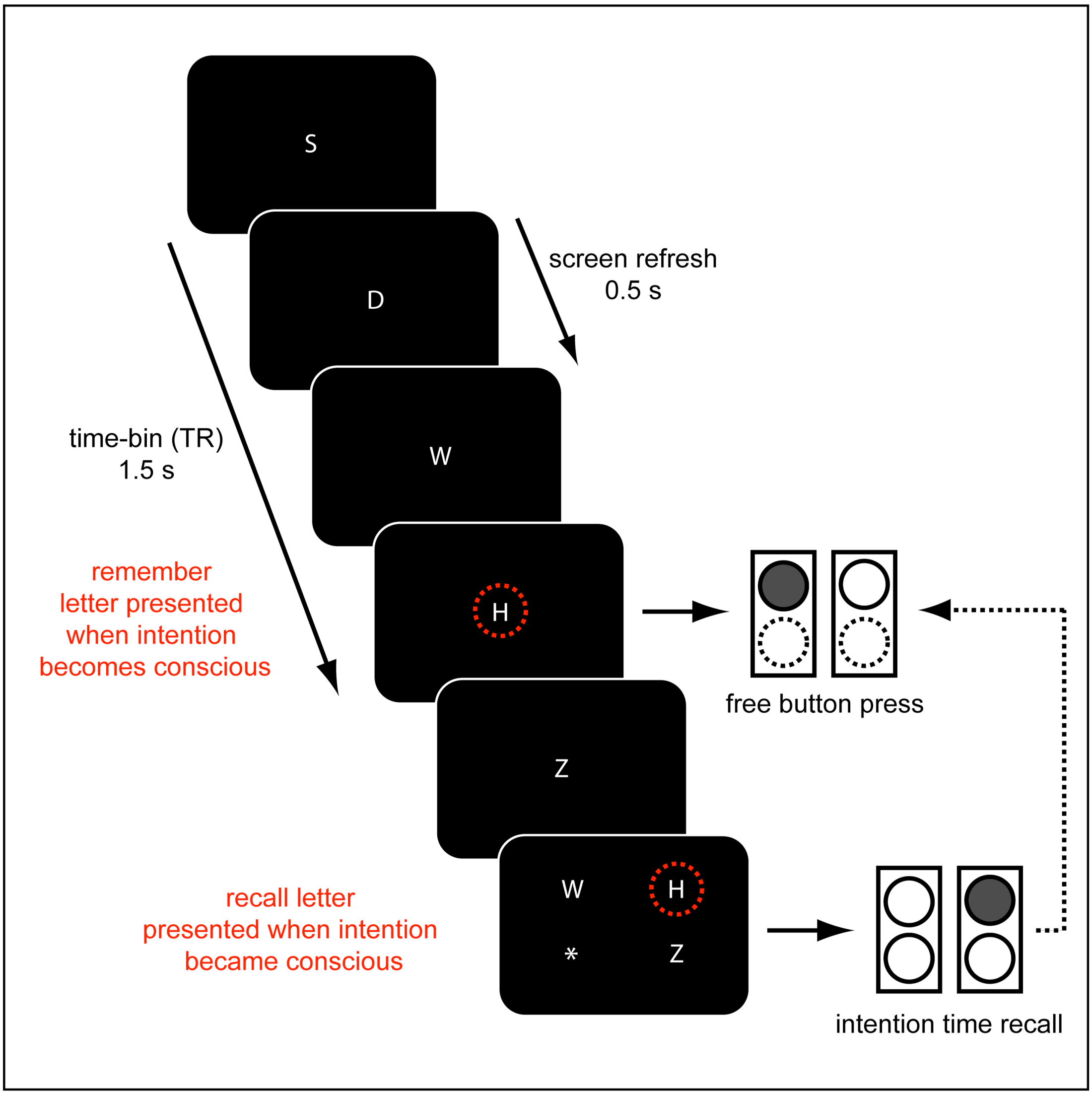


Fig. 3-1. Measuring the onset time of free conscious motor intentions. Participants were presented with a stream of constantly changing white letters on dark background. The screen was refreshed every 500 ms. The task was to freely and spontaneously decide to press a response button with the left or the right index finger (illustrated by upper circles; decision for left button in example illustrated by filled circle) whenever they felt the urge to do so. They were instructed to note the letter displayed on the screen when they became aware of their intention and to immediately perform the button press (the letter H in the example; red circles are for illustration and were not presented). Following the button press, a response screen was presented. Three letters and an asterisk were presented in the four corners of the screen, the letters being those shown during and immediately prior to the button press. Participants indicated the letter that was visible at the time of the decision by pressing the

button corresponding to its position on screen (recalled letter H, indicated by upper right button in example). If they could not remember the letter or if the relevant letter was not present, they indicated this with the asterisk. After the response was given, the next trial started and participants were instructed to return to a relaxed state before making a new decision. The general paradigm was taken from Study 1 (Soon et al., 2008).

3.3.3 Functional imaging and preprocessing

Functional imaging data was acquired using a 7-Tesla whole- body MR scanner (MAGNETOM 7T; Siemens, Germany) with an 8 channel array head coil (RAPID Biomedical, Rimpar, Germany). A gradient echo planar imaging (EPI) sequence was used for functional imaging (TR = 1500 ms, TE = 23 ms, flip angle = 90°, matrix size 64×64, in-plane resolution 1×1 mm). 21–28 slices were acquired (1 mm thickness, no gap), depending on the SAR limit of individual participants, and covered the most anterior part of prefrontal cortex. In order to minimise signal dropout due to the frontal sinuses, the slices were tilted away from the coronal orientation by an angle of 30.2° to the transverse plane (due to the anatomy of individual participants and their position in the scanner, the angle was 37.6° in two cases and 36.0° in one case). Particularly for ultra-high field strength, signal dropout and distortions in frontopolar cortex can be substantial. This setup was found to maximally reduce signal distortions and dropouts for the present study because it allowed us to use a small field of view (FOV), and thus a short echo train length, in order to cover most of anterior prefrontal cortex with maximal exclusion of the air-filled cavities compared to axial slices (Fig. 3-2). However, using this setup no region beyond frontopolar cortex could be covered.

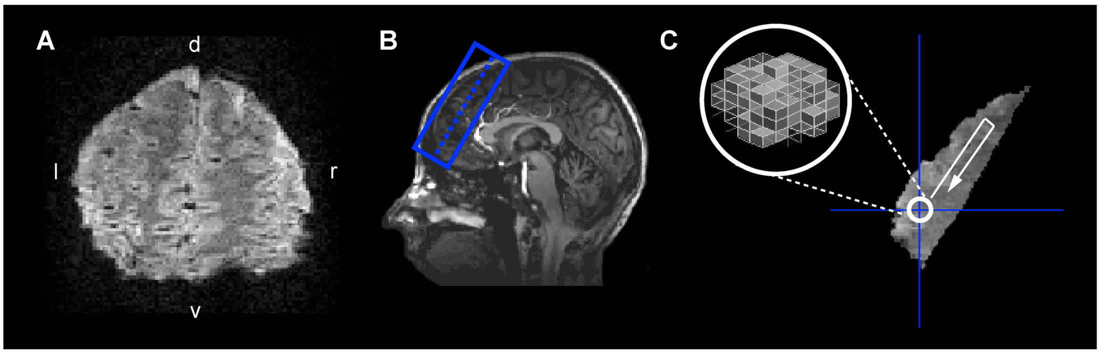


Fig. 3-2. Illustration of EPI image, slice positioning and decoding approach. A) Example of one slice of one participant's EPI image. B) Structural T1 image from the same participant displaying the positioning of the example slice (dotted line) and slice coverage (blue box). For each participant, 21–25 coronal slices (1×1×1 mm, without gap) were positioned such that the most anterior part of frontopolar cortex was covered. Due to the optimized slice positioning, which allowed the use of a small field of view (FOV) and a short echo train length, a relatively small part of the air-filled cavities was included. This improved the quality of the EPIs and reduced signal dropouts and distortions. C) The parameter estimates from the FIR model were used for multivariate pattern classification. A moving 'searchlight' algorithm was implemented using a radius of 3 voxels in order to decode the outcome of the upcoming decision from each position in frontopolar cortex.

Data was acquired for 10 functional runs, each lasting 5 minutes (200 volumes per run). The first two volumes of each run were discarded by default to allow for magnetic saturation effects. Additionally, a structural T1-weighted image was acquired for each participant for co-registration (176 transverse slices; TR = 3840 ms, TE = 268 ms; voxel resolution 1×1×1 mm). During the scanning sessions, stimuli were presented via a projector with a resolution of 1024×768 pixels (refresh rate of 60 Hz) that projected from behind the head-end of the scanner onto a screen. Participants laid supine in the scanner and viewed the projection via a mirror. Responses were recorded using a set of two custom-engineered, deconstructed Nintendo Wii joysticks, each with two buttons operated with the index and middle fingers of either hand.

The first stage of data processing involved motion correction to the first image of the first run using SPM2 (http://www.fil.ion.ucl.ac.uk/spm/). No additional normalization or smoothing was performed at that stage in order to maximize the sensitivity for information encoded in the fine-grained voxel patterns (Haynes & Rees, 2005, 2006; Kamitani & Tong, 2005; Soon et al., 2008).

3.3.4 General linear model

As in Study 1, a finite impulse response (FIR) predictor was used to model fMRI responses, as it was not known whether the profile of the fMRI time course would adhere to the generic haemodynamic stimulus-response function in this situation. This procedure also allowed time-resolved decoding to be implemented (Bode & Haynes, 2009; Soon et al., 2008). Left-button trials and right-button trials were modelled as two separate conditions, each with 20 FIR regressors. Each regressor modelled a time-bin of 1.5 s (1 TR), covering a 30-second time period around each trial. The 10th time-bin was defined as that in which a decision was made. The first 10 regressors therefore modelled the 15 seconds preceding (and including) each decision, the last 10 covered the 15 seconds following the decision. Invalid trials (in which participants were unable to recall a letter) were modelled separately, again using 20 FIR predictors and assigning the 10th predictor as that in which the button was pressed. These trials were excluded from the pattern classification analyses. To minimise unaccounted-for variance in the fMRI data, the second button-presses with which participants indicated the letter present at the decision time were modelled as covariates. Left-handed and right-handed button presses were modelled separately, and convolved with a standard haemodynamic response function (HRF).

3.3.5 Multivariate pattern analyses

The data was analysed with searchlight-based multi-voxel pattern analyses (MVPA), as described in Chapter 1. The analyses sought to identify regions within FPC that allowed participants' decisions for left and right to be decoded from fine-grained patterns of activity as measured by the BOLD signal preceding the participants' conscious awareness of their decisions, identical to Study 1. The choice-predictive information encoded within a spherical cluster of voxels at each position in the brain could be estimated without making any a-priori assumptions as to the exact location (voxel) of the information (Fig. 3-2C).

Decoding the content of motor decisions

As a direct replication of Study 1, the searchlight radius was kept at $r = 3$ voxels. Note that in the original study the voxel size was 3×3×3 mm while the reduction of voxel size in the present study to 1×1×1 mm yielded a searchlight 27 times smaller. Nevertheless, information decoding was based on the same number of dimensions (voxels) in both cases. The total number of voxels in the whole search volume in anterior PFC was comparable to that in a whole-brain data set with the standard 3×3×3 mm voxel resolution acquired on a 3T scanner.

For each spherical volume searchlight of N voxels, the parameter estimates for all 20 time-bins were extracted from each run, separately from left-decision and right-decision trials. These were transformed into two N-dimensional pattern vectors (one corresponding to left-decision trials, the other to right-decision trials) for each of the 20 time-bins, representing the spatial activation patterns for both decisions from 15 s before to 15 s after the conscious decision. An independent classification analysis was performed at each time bin to assess how much discriminative intention-related information was contained in the patterns of that time period. Pattern vectors from nine of the ten runs (eight of the nine runs for the participant with one excluded run) were first assigned to a ‘training data set’. This set was used to train a linear support vector machine (SVM) pattern classifier (LIBSVM implementation, http://www.csie.ntu.edu.tw/~cjlin/libsvm) to discriminate between patterns corresponding to the two different decision outcomes (or intentional states), using a fixed regularization parameter C=1. The classifier estimated a decision boundary separating the two classes of patterns in N-dimensional space. The amount of intention-related information contained in these patterns was then assessed by using the decision boundary to classify the vectors in the independent ‘test data set’ taken from the remaining run. Chance level for correct prediction of whether each vector corresponded to a left-or right-button decision was 50%. This procedure was repeated 10 times, each time using a different run as the independent test data set, resulting in a 10-fold cross-validation. The pattern classification results were averaged across repetitions and assigned to the central voxel of the searchlight cluster as its decoding accuracy.

The entire procedure was repeated by assigning in turn every voxel in the brain volume as the central voxel of the searchlight cluster, yielding a 3D map of decoding accuracies throughout the imaged volume. Each map represented the amount of intention-related information encoded in the local neural neighbourhood at every location in the brain of a participant, at each time-bin before and after the spontaneous decision. Combining accuracy maps across all 20 time-bins revealed the spatio-temporal dynamics of intention-related information flowing across the brain.

In the next step, the participants' individual decoding accuracy maps were normalized to MNI-space for group-level analyses. For this, the functional images were first co-registered to the individual high-resolution T1-weighted structural whole-brain image acquired during the same scanning session. The T1-weighted image was normalized to the MNI T1-template image as implemented in SPM2. The normalization parameters were then applied to the decoding accuracy maps. These were further smoothed with a Gaussian kernel of 3 mm FWHM. Voxels that were not shared by all participants were masked out. For each time-bin, group level analyses were performed across participants. The decoding accuracy maps from the time-bins *preceding* the decision (time-bins 1–10) were analysed, yielding a statistical parametric map of voxel clusters (using a 5-voxel cluster size threshold) that displayed decoding accuracies greater than chance level (50% for two decisions) during the 15 seconds preceding (and up to) the conscious decision time (using a threshold of $p<.05$, FDR-corrected). It was therefore possible to track changes in the amount of information encoded in different regions over time, and in particular, to search for a build-up of intention-related information prior to participants' conscious awareness of their own intentions, as observed by Soon et al. (2008).

Temporal pattern stability analysis

The goal of this analysis was to investigate the spatio-temporal profile (Mourão-Miranda et al., 2007) of the time-bins that allowed the prediction of free decisions before they reach conscious awareness. Individual data from the searchlight yielding the highest decoding accuracy across participants (see Results and Fig. 3-3A) preceding the decision was analysed across time-bins for each participant. This coordinate, which was established from MNI-normalized group-level statistical maps, was transformed back into the original image space of individual participants. A spherical cluster with radius of three voxels was again constructed around this position for each participant, and the spatial activation patterns for the two types of decisions were extracted and transformed into pattern vectors, separately for each run.

The pattern vectors from single time-bins were then combined for each decision by I) simply *averaging* vectors in steps of (i) two, (ii) three or (iii) four time-bins; or by II) *concatenating* vectors in the same steps as I. The multivariate pattern classification analysis was run again on these new vectors, exactly as described before; the difference was that there were a smaller number of time steps per analysis. If averaging across earlier time-bins did not reduce the decoding accuracy, this would mean the spatial activation patterns displayed a consistently high temporal stability in those time-bins. Finally, correlation analyses were also conducted between the pattern vectors of adjacent time-bins (separately for each participant and each condition) in order to assess the temporal stability of these patterns in more detail.

Since the results did not differ for left and right decisions, they were combined. Please note that all these subsequent analyses only aimed to specify the role of the best searchlight cluster and not to select voxels for further dependent statistical analyses, which would have been circular (Kriegeskorte et al., 2009). Also note that the chosen cluster was the best decoding cluster averaged across participants. This cluster therefore did not necessarily represent the optimal decoding cluster in individual participants. Analysing the optimal clusters in individual participants, however, would have carried the risk of arbitrariness and was therefore strictly avoided.

Univariate control analysis

The parameter estimates obtained from a GLM, based on normalised and smoothed (3 mm FWHM) data, were used in a conventional mass-univariate analysis. Again, an FIR predictor was used to model fMRI responses (identical to the GLM analysis described above) and group-level analyses were performed across participants for each time-bin separately. The purpose of this analysis was to investigate whether any voxels at any time-bin showed significant differences in activation between the left-decision and right-decision trials.

3.4 Results

3.4.1 Behavioural results

On average, participants chose the left button on 51% of all trials, and the right button 49%. There were only very few trials in which participants could not recall the letter present when their decision was made (average 1% of all trials). In nearly all trials, participants indicated that the decision reached conscious awareness during the presentation of the same letter or one letter before they pressed the button (see Table 3-1 for details). Any participant who showed highly unbalanced decisions (i.e., one option was chosen more than twice as often as the other), or had exceptionally short or long trial durations (i.e., on average <15 s or >50 s) were excluded from further analyses. One participant (S4) was thus excluded. The average time lapsed between consecutive trials was 29.7 s (SD = 9.29); an average of 11 trials was performed per 5 min run.

Table 3-1. Behavioural results

Participant	Left	Right	Mean trial Duration (SD) s	Letter indicated (mean)	'*' indicated
S1	56%	44%	20.7 (7.8)	-0.19	1%
S2	59%	41%	23.2 (5.5)	-0.17	0%
S3	46%	54%	24.0 (6.3)	-0.13	1%
S4	*39%*	*61%*	*58.4(20.4)*	*-0.11*	*7%*
S5	50%	50%	46.4 (8.2)	-0.17	0%
S6	56%	44%	42.4 (8.0)	-0.09	0%
S7	50%	50%	26.8 (6.2)	-0.13	0%
S8	43%	57%	28.1 (7.1)	-0.12	0%
S9	50%	50%	36.2(14.7)	-0.07	4%
S10	55%	45%	24.4 (8.9)	-0.13	2%
S11	45%	55%	36.6(15.0)	-0.93	5%
S12	52%	48%	18.0 (4.0)	-1.36	2%

Note. 'Letter indicated' refers to the number of letters between the time of conscious decision and button press (500 ms duration per letter). On average, the intention reached conscious awareness during the same time-bin (0) or one time-bin earlier (-1) relative to the button press. '*' was chosen when participants could not remember the letter presented at the point of conscious decision, or when the letter was not displayed (i.e., decision was >1s before button press.

In the post-experimental interviews participants indicated that they were able to relax and make spontaneous decisions (Table 3-2). The button press was generally executed with minimal delay after they became aware of their intentions (M = 0.9; SD = 0.5), as instructed. They indicated having been very spontaneous (M = 3.3; SD = 0.65) and they did not pay much attention to the timing (M = 1.0; SD = 0.9). Most participants reported that they did not have specific thoughts they could remember. Some reported having thought about (or mentally read) the letters, some reported having occasionally thought about daily activities but none reported having thought about the decisions. Most participants reported that they became more relaxed through the experiment and that they either became more spontaneous or that there was no change in spontaneity. This was not surprising given that participants were highly familiar with the task, having completed 10 runs of prior training, and were able to perform the task effortlessly.

Table 3-2. Individual post-experimental interview results

Participant	Qn. I	Qn. II	Qn. III	Qn. IV	Qn. V
S1	2	4	2	the day, girlfriend, relaxing	nothing
S2	1	4	1	letters, nothing	more relaxed
S3	1	3	3	nothing	more relaxed
S4	*excluded*	–	–	–	–
S5	1	3	1	nothing	nothing
S6	1	3	0	nothing	nothing
S7	1	3	0	nothing	decisions slightly faster
S8	1	3	1	letters	more relaxed, spontaneous
S9	0	4	1	letters, uni, holidays	forgot letters few trials
S10	1	2	1	letters	nothing
S11	0	3	0	nothing	decisions slighty slower
S12	1	4	1	nothing	nothing
Mean	0.9	3.3	1.0		
SD	0.5	0.6	0.9		

Note. One participant (S4) was excluded from all analyses because of behavioural criteria. Actual questions:

I) “How often did you make a decision earlier during the trial but waited with the button press?” (0 = *never*; 4 = *always*);

II) “How would you rate your spontaneity throughout the experiment?” (0 = *not spontaneous at all*; 4 = *very spontaneous*);

III) “How often did you explicitly think about the timing of your decisions?” (0 = *never*; 4 = *always*);

IV) “What did you think about during the experiment, as far as you can remember?” (open-ended);

V) “Did you notice any changes in your behaviour during the experiment? If yes, what changed?” (open-ended)

Additional analyses were performed to check whether participants made random choices on each trial or followed specific patterns. We first correlated the sequences of decisions from each run of each participant with the sequence of decisions in the following run, in order to control for the possibility that participants might have simply repeated fixed sequences of decisions over the experiment. None of these correlations were significant for any of the participants (all tests $p>.05$; average correlation $r = -0.11$; range −0.27 to 0.07). For each participant and within each functional run, we further analysed whether the sequences of left and right decisions violated the assumption of a random order (runs test as implemented in MATLAB, Math-Works Inc., corrected for multiple tests). The results showed that out of all 110 functional runs from all 11 participants, violations of the randomness assumption could only be found in one single run ($p=.004$; all others $p>.05$), providing additional evidence that behaviour was spontaneous. (However, note that due to the nature of our task, there were very few successive trials per functional run, thus potentially limiting the informative value of this test.) These results indicated that participants performed correctly and that preplanning or other unaccounted-for conscious cognitive activity could not account for the predictive information detected before the conscious decision.

3.4.2 Functional imaging results

Searchlight-based MVPA was used to search for brain regions encoding participants' decision outcomes. We identified a cluster in FPC from which participants' decisions could be decoded *before* their intentions became conscious (i.e., time-bin 10, Fig. 3-3), with statistically significant decoding accuracies of up to 57% (SE = 1.69; $p<.05$, FDR-corrected) just before the decision was made (time-bin 9; see

Fig. 3-3). This region was located in left frontopolar cortex (MNI coordinates −23 59 −9; see Fig. 3-4 for individual searchlight clusters). The earliest time at which decoding was possible was ~7.5 seconds (time-bin 4) before the decision was reported to be consciously made. Taking into account the temporal delay of the BOLD signal (which is in the order of a few seconds), it was possible that these signals reflected processes up to 10 seconds before the actual decision.

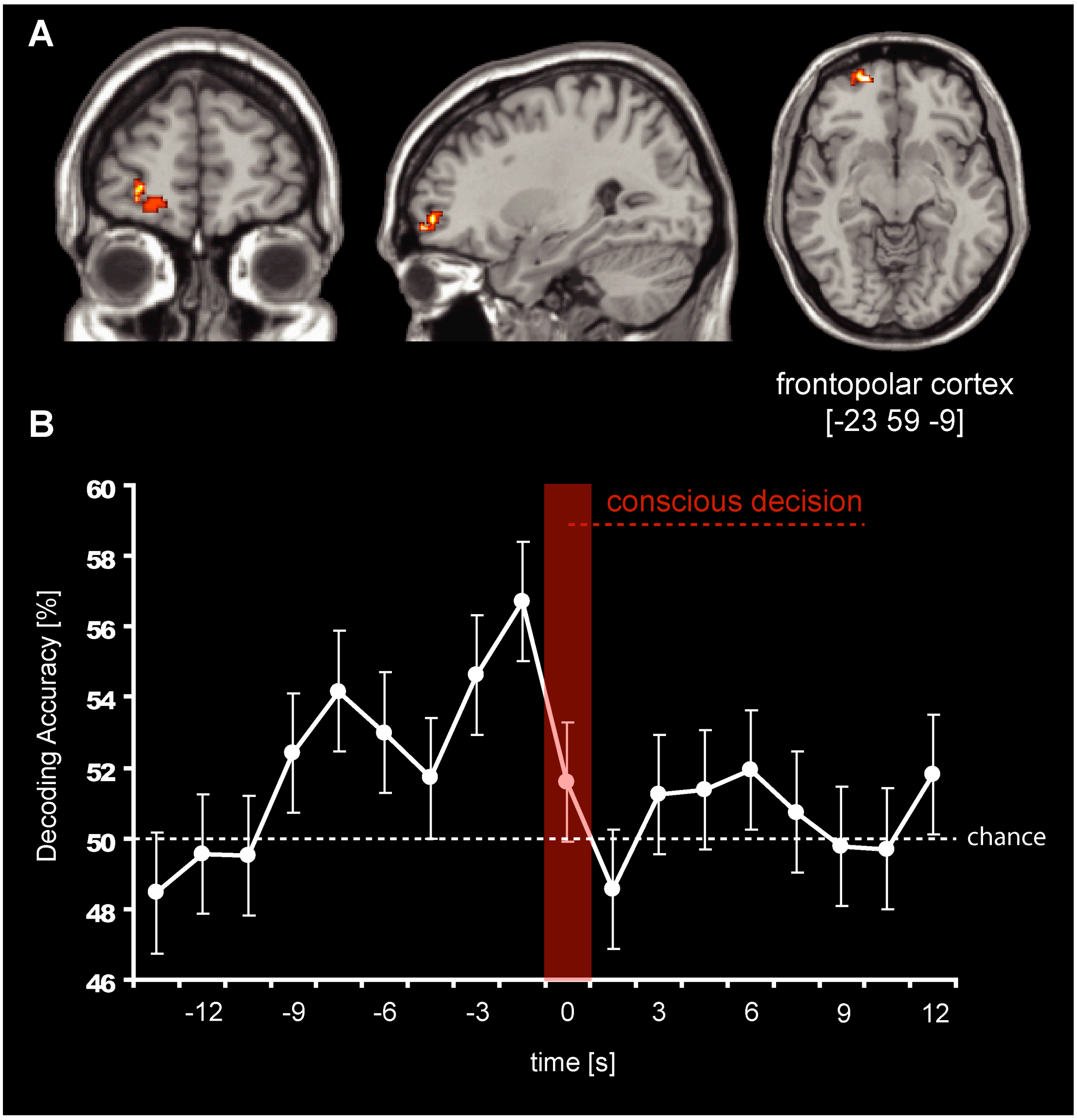


Fig. 3-3. Decoding of upcoming motor decisions from frontopolar cortex. A) The figure displays a region in left frontopolar cortex [-23 59 -9] from which decoding was significantly above chance (50%; *p*<.05, FDR corrected;

5-voxel cluster-size threshold). FPC only showed significant decoding accuracies in the time-bins preceding the decision. B) The graph displays the average time-course of decoding accuracies, taken from the central voxel of the searchlight cluster that showed the highest decoding accuracy. Error bars represent standard errors. The time-bin of the conscious intention is indicated by the red bar and is labelled as time 0. Time-bins preceding the conscious awareness of the intention are labelled as negative numbers (units = seconds, relative to decision); time-bins following the decision are therefore positive. One time- bin corresponds to 1.5 s. Coordinates displayed are MNI coordinates.

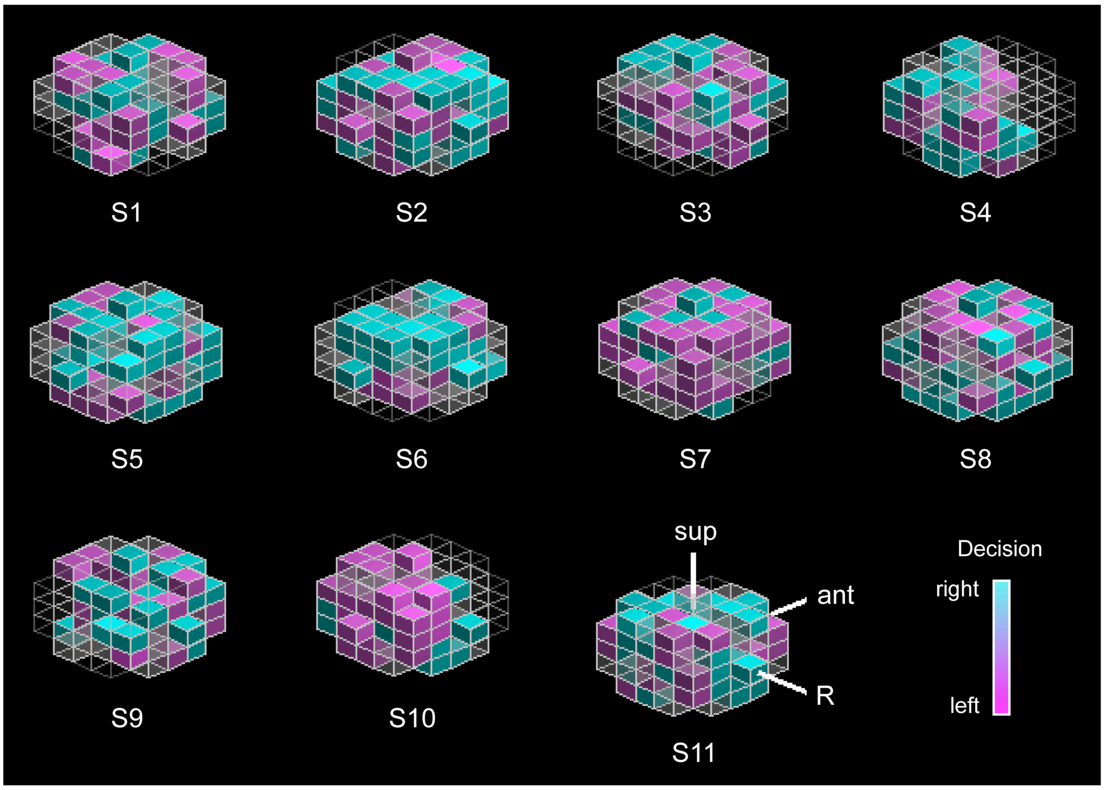


Fig. 3-4. Individual searchlight clusters. Displayed are the spherical voxel clusters (with radius r = 3 voxels) in frontopolar cortex of all participants that yielded the highest decoding accuracy in the time-bin directly preceding the decision (−1.5 s). Voxels responding preferentially to one decision are colour-coded (magenta for left, aqua for right; sup = superior, ant = anterior, R = right). Grey transparent voxels did not show decision preference or were not located in grey matter. Colours are scaled for better visualization. Informative patterns were different for each participant.

Using a searchlight radius of 4 voxels led to decreased decoding accuracies and *p*-values for the same region. When the radius was further increased, no significant results could be achieved, possibly due to the increased dimensionality. In a control analysis the accuracy maps from the time-bins *after* the decision was made (time-bins 11–20) were contrasted against chance level. No clusters could be found in FPC encoding any information above chance level during this period. The same held true if separate time-bins around the time of the motor response were considered, matching the findings from Study 1 that FPC only encoded the intentions *before* participants were aware of making a decision.

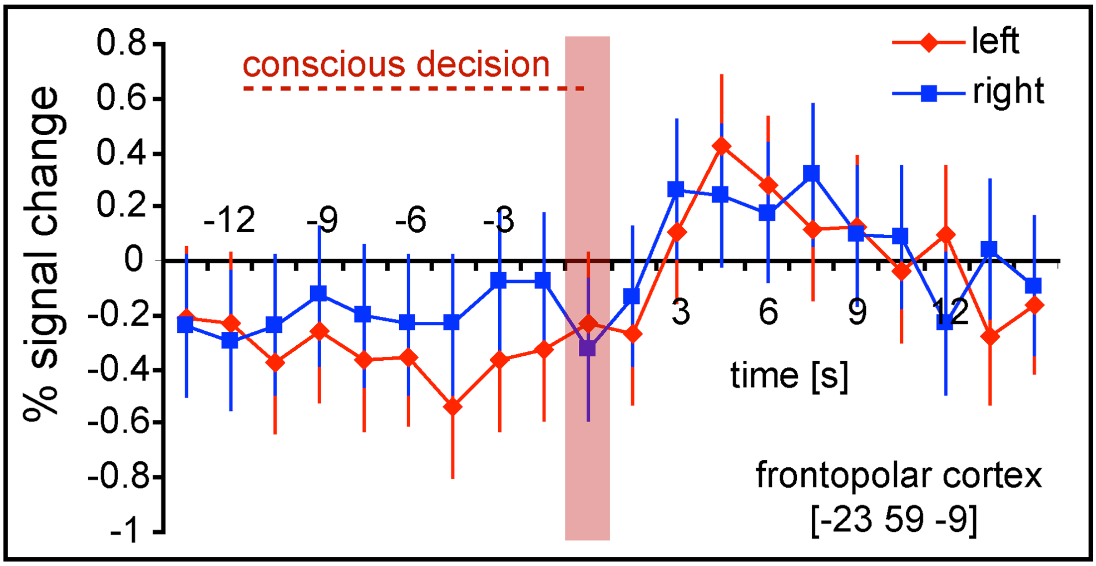


Fig. 3-5. Univariate BOLD signal changes. Average BOLD parameter estimates for left- and right-button decisions in central voxel of searchlight with highest decoding accuracy prior to the conscious decision. For both conditions, the signal increased only *after* the decision (red bar), and returned to baseline within ten seconds. Significant differences between left and right decisions were not found here. Similarly, no region within the imaged volume displayed such a difference, even when a liberal threshold of $p<.001$ (uncorrected) was used. Time-bins preceding the conscious awareness of the intention are labelled as negative (units = seconds, relative to decision). One time-bin corresponds to 1.5 s. Coordinates are given as MNI coordinates.

The information was, as in the original study, only encoded in fine-grained activation patterns rather than in the average signal. Additional univariate analyses confirmed that there was no significant difference between left and right decisions at any time in individual voxels, even when a liberal threshold of $p<.001$ (uncorrected) was applied. This held true for the whole frontopolar region as well as for the region from which decoding was possible. The average BOLD signal did not increase at all until after the decision was made (Fig. 3-5).

Temporal pattern stability analysis results

The subsequent decoding analysis using the best searchlight cluster across all participants showed that *averaging* across adjacent time-bins led to lower decoding accuracies, the more time-bins that were combined (Fig. 3-6). In contrast, concatenating time-bins, which effectively combined spatial and temporal information before the decision, still predicted the decision outcome with high accuracy (Fig. 3-6). Concatenating was superior to averaging by trend, suggesting that, spatial patterns were not uniform throughout the period leading up to a decision, but carried more decision-related information with increasing temporal proximity to the decision. Nevertheless, subsequent analyses showed significant correlation of spatial patterns from consecutive time-bins. Moreover, pattern similarity increased with temporal proximity to the decision (peak $r = .44$, $p<.05$; Fig. 3-7). After the time-bin of the decision, the correlations dropped again to a stable level. This auto-correlation curve closely mimicked the time-course of decoding accuracies (Fig. 3-7). Thus, activity patterns became more similar and more informative the closer the decision was to reaching awareness. After the decision was made, some pattern stability was sustained but the patterns no longer carried information about the decision.

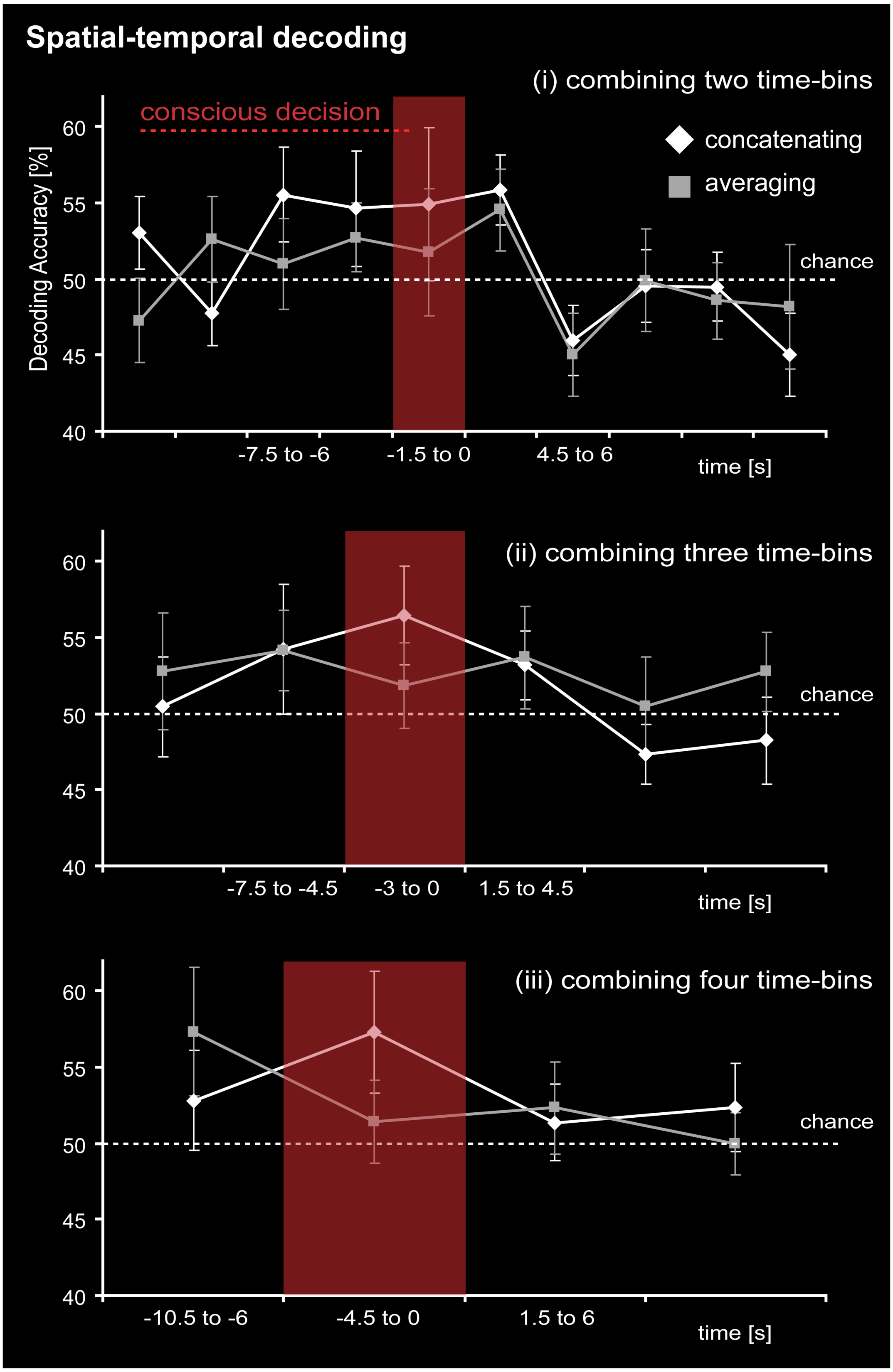


Fig. 3-6. Temporal pattern stability. Temporal-spatial decoding analysis. The spatial activation patterns from the searchlight cluster centered on −23 59 −9 (MNI), which showed the highest predictive accuracy, was extracted from

individual participants' data. The original patterns from the time-bins were combined by concatenating (white) or averaging (grey) the respective pattern vectors in steps of either (i) two, (ii) three, or (iii) four time-bins. The reference time-bin for vector concatenation was the time-point of the decision (time 0 s). The resulting pattern vectors additionally represented temporal information for the best searchlight cluster and were used for multivariate decoding. Spatio-temporal information was found to be highest directly preceding the decision and was still present when four time bins were concatenated. Concatenating was superior to averaging by trend.

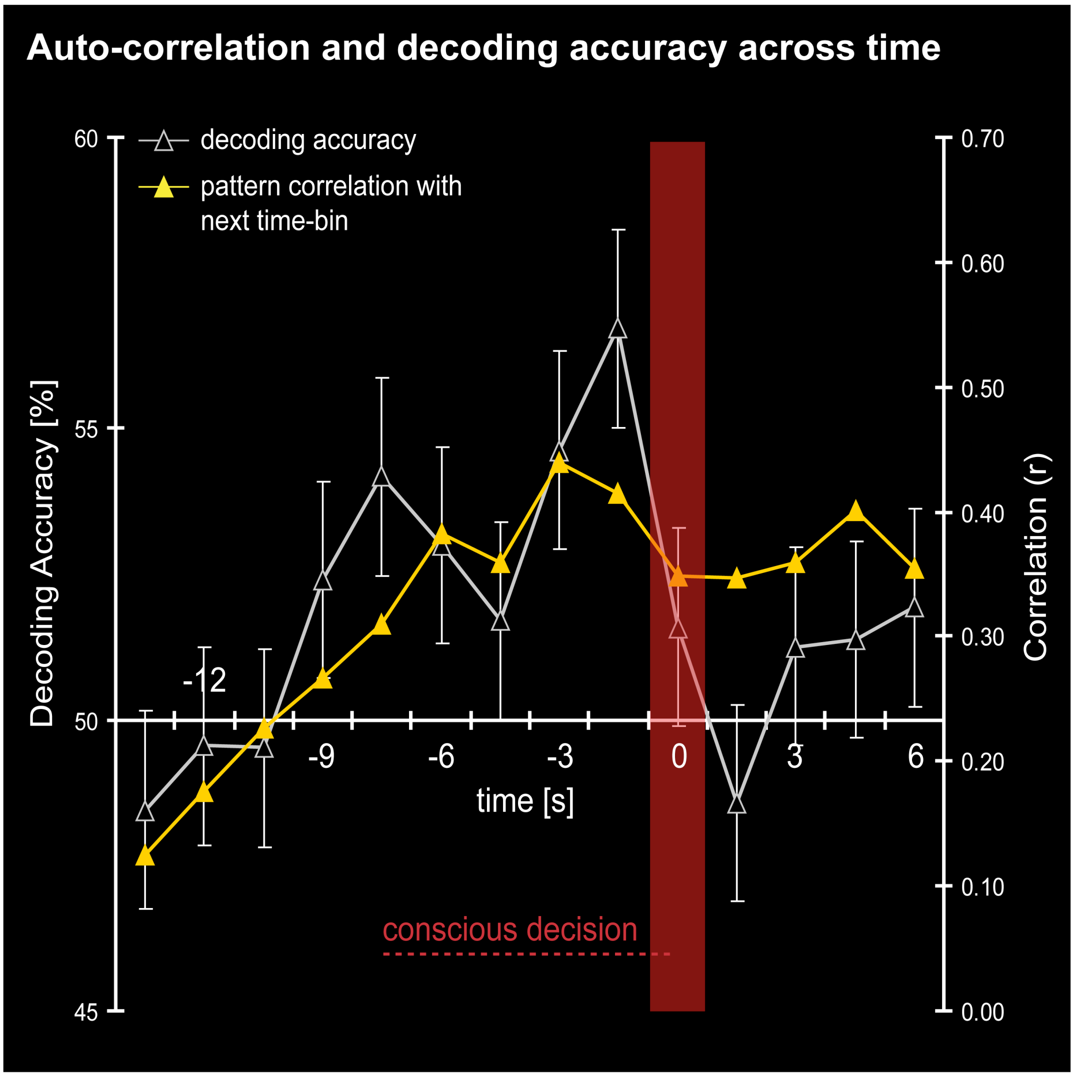


Fig. 3-7. Correlation analysis for spatial activation patterns. Displayed is the decoding accuracy across time from the best cluster (empty gray triangles) as well as the correlation of each time-bin with its preceding time-bin (filled

yellow triangles) as a measure of pattern similarity (averaged across patterns for left and right decisions). Up to the time of the decision (time 0 s) the decoding accuracy and pattern similarity increased in a similar fashion. After the decision, the pattern similarity dropped slightly and patterns did not predict the decision outcome anymore.

3.5 Discussion

This study aimed to assess whether local spatial activity patterns in FPC, which was previously found to encode unconscious intentions in Study 1, would display temporal stability over time. Using ultra-high field fMRI at 7T, allowing a voxel resolution of 1×1×1 mm, we could replicate the findings of the first study. We also demonstrated that activity patterns preceding the time-point of the conscious decision became increasingly similar with closer temporal proximity to the decision. Our behavioural data and questionnaire results provided further support that the decisions were unlikely to have been biased by conscious processes. Thus, early predictive activity patterns were attributable to unconscious components of evolving intentions.

Comparable to Study 1, the participants' decisions could be read out approximately seven seconds before they became conscious. Given the haemodynamic delay, it was likely that this reflected neural processes that occurred even earlier by a few seconds. The site of information encoding was found to be left frontopolar cortex, also referred to as the rostral lateral prefrontal cortex or the anterior prefrontal cortex, and approximately the most anterior part of Brodmann area 10 (Brodmann, 1909; Ongur et al., 2003; Ramnani & Owen, 2004). The same region was identified in Study 1 but in the opposite hemisphere. In this second study, we optimized the slice positioning to minimize distortion effects and signal dropouts, which are a common problem due to the proximity to frontal sinuses, especially at higher field strength. Since the analysis only included voxels that were present in all participants, residual dropout in individual participants could have led to the exclusion of more informative voxels. Hence, our results might underestimate the extent of the decision-related region.

No information about the participants' decisions was found after the decision was made, which is also in line with our original findings in Study 1 that after the time of the decision, information was only encoded in primary motor cortex and pre-motor cortex. These areas were not covered in the present study. Due to the optimized slice positioning in this second study, the precuneus/posterior cingulate cortex, which additionally encoded early decision-related signals in Study 1, was also not covered. As demonstrated before, the procedure used in both studies ensured that decoding could not be explained by activity related to the previous trial. Again, a FIR model was used, designed to separate effects of the current trial from the previous and the following trial. This method is highly efficient as long as both types of responses are roughly equally frequent, as here. Importantly, participants self-paced their decisions, ensuring that the intervals between trials were variable, which makes the estimation of the FIR model even more robust to carry-over effects. Second, the time delay between the onset of predictive information in frontopolar cortex and the end of the previous trial was on average ~15 seconds, and thus far beyond the relaxation time of the haemodynamic response. The average trial duration (29.7 s) in the present study was even longer than in Study 1 (21.6 s), thus making it even more unlikely that spill-over effects from the previous trials might have occurred. For the earliest time-points in a trial we found no predictive information – contrary to what would have been expected if successful decoding was due to carry-over effects from the previous trial. However, as the trial progressed, approaching the decision time predictive information emerged. Third, the temporal resolution was also improved (1.5 s per time-bin compared to 2 s per time-bin in the original study), further validating the original findings. Although we do not believe and do not claim that our participants produced perfectly random sequences, our behavioural results suggest that participants made spontaneous decisions. This was probably because we did not ask participants to balance their decisions, and there were no external constraints for preferring one option over another in any given trial.

Interestingly, we observed an increase in similarity between patterns with increasing temporal proximity to the conscious decision. This increase in correlation was mirrored by the increase in information content about the decision outcome. Thus, one possible explanation for this finding is that during the unconscious phase of intention-formation, the patterns slowly ‘evolved’ towards the final conscious decision, comparable to an ‘evidence accumulation’ process postulated for fast, stimulus-driven decisions (Smith & Ratcliff, 2004). This hypothesis states that once a threshold is crossed (a certain pattern is stable enough), a conscious decision is made and activation patterns lose their predictive power afterwards. The remaining (but reduced) pattern stability might be explained by the dependence of sequentially acquired brain scans. Although there was some tendency for patterns to remain stable for a few seconds after the decision, there was no decodable information at these post-decision time periods. Similarly, patterns during the initial phase of the following trial were not informative. It was only later, closer to the next decision in the next trial, that we again observed a slow increase of pattern similarity and information encoding.

This again spoke against carry-over effects from the previous trial. Our detailed behavioural analysis confirmed that participants did not systematically and consciously prepare their decisions ahead of time. They acted as instructed and were spontaneous. Similarly, in Study 1 we did not observe any encoding of the chosen movement in motor cortex before the decision; however, this could be expected when participants pre-plan a motor response (Bode & Haynes, 2009). Here this analysis was not possible due to the restriction of coverage to PFC which was necessary to achieve a higher spatial resolution. We thus conclude that the early informative spatial activation patterns in frontopolar cortex were related to unconscious components of the intention. This again supported the hypothesis of a slow unconscious drift towards a 'prototypical' pattern in FPC, which was related to the conscious decision.

It might be surprising that decision-related information is encoded in the brain several seconds before the decision becomes conscious, given that the task was rather simple. One possibility is that random activity directly preceding the decision might have biased the decision outcome, as suggested for short time periods (Smith & Ratcliff, 2004). This, however, was less likely for such long periods as observed here. Our study might have facilitated the detection of very early information by encouraging participants to relax and refrain from decision-related thoughts as well as by instructing participants to be as spontaneous as possible in making and executing their decisions. By doing so, unlike most other studies, our experiment was uniquely suited to investigate the early evolution of intentions. Even though there was credible evidence that our participants' behaviour was spontaneous, the possibility remains that there may be some hidden regularities in their responses, which might only become detectable in longer behavioural sequences than produced here. Such biases, even though outside participants' awareness, could potentially contribute to the build-up of early brain activation patterns. It is important to note that any temporal autocorrelation in the signals could cause a correlation between choices in successive trials, even without a conscious, deliberative link. Such autocorrelation might be considered a very basic form of memory, but our conclusion that choices could be predicted before awareness would remain unchanged, as the participants themselves were not aware of following such response regularities.

Please note that our results could not be taken as direct evidence for a causal relationship between the activation in frontopolar cortex and the decision because fMRI measures neural decision-related processes only indirectly and prediction accuracy was far from perfect.

The present study supported the hypothesis that prefrontal cortex is a core region for free decisions. Presently, it is believed that the anterior prefrontal cortex lies at the top of a hierarchically organized prefrontal functional architecture. Prefrontal cortex represents sensory input information in its most abstract form and guides cognitive control (Miller & Cohen, 2001). It maintains the abstract representation of a desired act, together with context-relevant information such as environmental contexts, task-rules, motivation and potential outcomes (Bode & Haynes, 2009; Bunge, 2004; Bunge & Wallis, 2007; Sakai, 2008). The motor plan for the execution of this act is prepared in premotor areas, and broken down into coordinated recruitment of single motor units in primary motor cortex (Passingham, 1995). Medial prefrontal cortex might additionally contribute to action planning by processing self-related information (Amodio & Frith, 2006), in this case, one’s intentions. It has also been found to encode freely chosen decisions during a delay before execution (Haynes et al., 2007).

Of the different regions in prefrontal cortex, however, evidence from cytoarchitectural studies suggests that frontopolar cortex has the necessary architecture to support the highest level of processing within prefrontal cortex. First, it has the greatest number of dentritic spines per cell, and overall spine density is higher than for all other areas of prefrontal cortex. Furthermore, it is the only supramodal area that is connected solely with other supramodal areas, has less laminar differentiation compared to other prefrontal areas, and its connections within PFC point towards a hierarchically high level of processing (Badre & D’Esposito, 2009; Jacobs et al., 2001; Semendeferi et al., 2001). Given these properties, frontopolar cortex is an optimal candidate for the representation of the most abstract contents (Badre & D’Esposito, 2009). Current hypotheses about the function of this region are based mainly on functional imaging studies, as this region is markedly smaller and difficult to access in primate electrophysiology. Presently, the cognitive processes in which frontopolar cortex has been implicated include: processing of internal states (Christoff & Gabrieli, 2000), modulation of episodic memory retrieval (Herron et al., 2004; LePage et al., 2000), prospective memory (Burgess et al., 2001), relational reasoning (Christoff et al., 2001; Kroger et al., 2002), the integration of cognitive processes (Ramnani & Owen, 2004) and cognitive branching (Koechlin & Hyafil, 2007). Frontopolar cortex has also been suggested to control long-term plans and to generate new cognitive sequences (Koechlin & Hyafil, 2007). This is supported by recent findings showing that frontopolar cortex also tracks the advantage of alternative action plans and might initiate switching (Boorman et al., 2009). Burgess et al. (2007) proposed that the type of processing in frontopolar cortex is determined by the context, allowing either stimulus-oriented (i.e., pertaining to the external environment) or stimulus-independent (i.e., pertaining to the internally generated representations) processing to occur. These theories are in line with a role of frontopolar cortex in the generation of free decision as demonstrated by both Studies 1 and 2.

One possibility is that neurons in frontopolar cortex could be tuned to different decision outcomes, while having the capacity to be flexibly re-coded depending on task demands as previously suggested for prefrontal cortex (Duncan, 2001; Sigala et al., 2008). This mechanism would also allow different types of intentions to be encoded without the need for hardwiring of single neurons to any single intention. Most abstract intentions are closely linked to some motor action anyway and might therefore be represented in a similar manner. Additionally, it has been proposed that evolutionarily newer functions, such as cultural inventions, could make use of already existing neural structures evolved for more basic but similar functions (Dehaene & Cohen, 2007). Study 3 will address this question by investigating the encoding of more abstract intentions, such as performing mathematical calculations, and the functional organization of the architecture that gives rise to them.

Conclusion

In summary, we could replicate the findings of Study 1 that motor intentions were encoded in frontopolar cortex up to seven seconds before participants were aware of their decisions. Using ultra-high field fMRI on a 7 Tesla scanner, we showed that these patterns became more stable with increasing temporal proximity to the conscious decision. These findings support the conclusion that frontopolar cortex is part of a network of brain regions that shape conscious decisions long before they reach conscious awareness. This once again questions the specific role of the conscious mind in voluntary decisions.

Acknowledgements

The authors want to thank Domenica Wilfling and Elisabeth Wladimirow for their assistance with scanning, Thomas Siegert for the construction of the joysticks and Carsten Bogler and David K. Sewell for valuable support.

Study 3

Decoding the Unconscious Formation of Abstract Intentions

Soon, C. S., He, A. H., Bode, S., & Haynes, J. D. (2013) Predicting free choices for abstract intentions. *Proc Natl Acad Sci USA*, *110*(15), 6217-6222.

doi:10.1073/pnas.1212218110

Article first published online: 18 March 2013

4.1 Abstract

Unconscious neural activity has been repeatedly shown to precede and potentially even influence subsequent free decisions. However, to date such findings have been mostly restricted to simple motor choices, and despite considerable debate there is no evidence that the outcome of more complex free decisions can be predicted from prior brain signals. Here, we showed that the outcome of a free *abstract* decision to either add or subtract numbers could already be decoded from neural activity in medial prefrontal and parietal cortex four seconds before the participants reported to be consciously making their choice. These choice-predictive signals co-occurred with the so-called default mode brain activity pattern that was still dominant at the time when the choice-predictive signals occurred. Our results suggest that unconscious preparation of free choices is not restricted to motor preparation. Instead, decisions at multiple scales of abstraction evolve from the dynamics of preceding brain activity.

Key words: free decisions, Libet, multivariate pattern classification, self-paced

4.2 Introduction

The subjective experience that our voluntary actions are initiated in the conscious mind has been challenged by the finding that the human brain may already start shaping spontaneous decisions even before they enter into conscious awareness (Libet et al., 1983; Soon et al., 2008, Study 1). Specifically, the human brain can start preparing spontaneous movements up to several seconds before a person believes themselves to be consciously making a decision to move (Bode et al., 2011; Haggard & Eimer, 1999; Libet et al., 1983; Soon et al., 2008).

To date, such early choice-predictive signals have only been investigated for simple movement decisions (Bode et al., 2011; Fried et al., 2011; Haggard & Eimer, 1999; Libet et al., 1983; Schurger et al., 2012; Soon et al., 2008). However, there are several reasons to assess whether preparatory processes also occur for higher-level, more abstract types of decisions. First, the relevance of motor decisions for understanding the neural formation and preparation of intentions has been heavily debated (Roskies, 2010; Smith, 2011), mainly because of their reduced complexity (Breitmeyer, 1985; Jung, 1985) and the limited levels of awareness in motor control (Danto, 1985; Doty, 1985). Second, previous studies on predictive signals for motor choices have revealed early information in prefrontal and parietal brain regions. These regions are not generally considered ‘motor’, but they have been sporadically observed in motor preparation (Boyd et al., 2009; Groll-Knapp et al., 1977). This invites the question of whether these regions provide only unconscious preparation of motor intentions or a common, task-independent network for preparing multiple types of decisions before awareness.

As such, we adapted Libet's intention clock paradigm to a context involving endogenous selection between two options that required more complex cognitive operations. Another fMRI experiment was conducted to investigate the neural precursors of abstract intentions by asking participants to perform an abstract decision task in which they spontaneously and freely chose to perform either of two mental arithmetic tasks: adding or subtracting. Given the fundamentally different neural processes involved in performing motor acts and arithmetic, identifying any overlap between the early choice-predictive signals would be of high relevance because it would point towards a common cerebral starting point for different types of choices. The task was specifically designed to encourage spontaneity of choice, but still requiring a conscious decision. As such, all problems and their answers involved only single digits, to ensure that both addition and subtraction could be performed with minimal effort and preparation. Essentially, these simple problems required only retrieval of over-learned arithmetic facts rather than actual mental calculations (Dehaene et al., 2003). To further ensure that participants' choices were not affected by considerations of difficulty, the problems were only made available *after* participants had freely decided whether to perform addition or subtraction.

We also aimed to address another question regarding the prediction of free choices. Studies 1 (Soon et al., 2008) and 2 (Bode et al., 2011) found early choice-predictive information in areas that overlapped with the so-called default mode network (DMN) (Buckner et al., 2008; Fox & Raichle, 2007; Raichle et al., 2001), thus raising the question of whether they actually reflect similar underlying mechanisms. High levels of DMN activity are typically observed during off-task periods (Greicius et al, 2003; Raichle et al., 2001; Shulman et al., 1997), that is, during prolonged rest periods between episodes of active task performance. This raises the question of whether DMN activity is also high in the period preceding the free decision. For these reasons, we directly investigated the link between choice-predictive signals for both motor (Soon et al., 2008) and abstract decisions (current study) and these 'off-task' brain signals. Interestingly, we identified a partial spatial and temporal overlap of choice-predictive signals with activity in the DMN.

4.3 Methods

4.3.1 Participants

Thirty-four healthy, right-handed participants (15 men; age range, 19–31 y) participated in a behavioural selection test in which they performed a voluntary abstract intention task. Twenty-two participants (12 women; age range, 22–31 y) who spontaneously showed balanced choices were selected to perform the same task while fMRI was conducted. Of these individuals, 4 participants were subsequently excluded from group analyses because of excessive movement, and another was excluded for biased intention selection (did not select subtraction at all in one run).

4.3.2 Behavioural paradigm

Participants spontaneously and freely decided to perform either of two abstract intentions: adding or subtracting. A continuous stimulus stream was presented from the start of each trial period, 1 frame/s without gap. Each stimulus frame consisted of a consonant below a central fixation point, a single-digit number above it, and four single-digit answer options, one in each corner (Fig. 4-1).

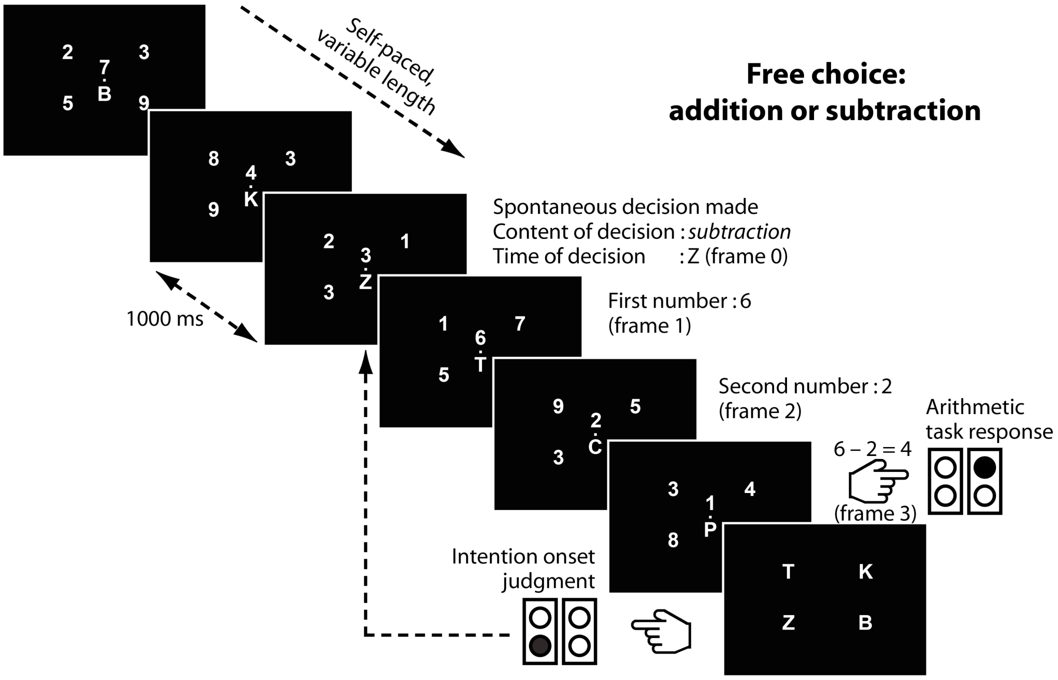


Fig. 4-1. Measuring the onset time of free conscious abstract intentions. Measuring the onset and content of spontaneous abstract intentions. A trial began with a continuous series of stimulus frames refreshed every second, each consisting of a central fixation point, a consonant below it, a single-digit number above it, and four single-digit response options – one in each corner. Immediately when participants felt the spontaneous urge to perform either adding or subtracting, they first noted the letter on the screen (frame 0 relative to time of decision). The chosen arithmetic task was then performed on the numbers presented above the central fixation in the next two stimulus frames (frames 1 and 2). Next, the response options for the numbers in frames 1 and 2 were presented in random order in the four corners of the subsequent stimulus frame (frame 3): the correct addition answer, the correct subtraction answer (absolute difference between the two numbers, i.e., no negative answers were involved), and two incorrect response options. Participants selected the correct answer for the chosen task by pressing one of four corresponding buttons, thereby revealing the content of their abstract decision. After the response was given, four letter options were presented from which participants selected the letter presented at frame 0, thereby revealing the time of conscious decision.

Participants were asked to relax as they looked passively at the letter stream. The order of presentation of letters was randomized under the constraint that there were no repetitions within a sequence of eight consonants. Immediately when they first became conscious of the spontaneous urge to perform either addition or subtraction, participants noted the letter presented on the screen (frame 0). It was stressed to participants that the time and choice of task were completely up to them but that the task should be executed without hesitation once the decision was made. The arithmetic task selected was then performed on the number presented above the central fixation from the subsequent two stimulus frames (frames 1 and 2). Since the problem itself appeared only *after* the decision was made, and participants had to remember the letter presented at the point of conscious decision, they could not choose when to perform a task by selecting particular numbers. The response options for the numbers in the earlier two frames were randomly presented in the four corners of the third stimulus frame: the correct addition answer, the correct subtraction answer (absolute difference between the two numbers, i.e., no negative answers were involved), and two inappropriate response options (Haynes et al., 2007).

Participants selected their answer by pressing one of four corresponding buttons, using either their left or right index or middle fingers. Having two inappropriate response options ensured that participants performed the chosen intention properly and did not simply press a random button. As both problems and answers involved only single-digit numbers, both adding and subtracting operations were essentially over-learned and could be easily performed within the short span of time given without needing much effort or preparation. There was a 3 s gap between the point of conscious decision and the point of response during which the intention was conscious but entirely covert; that is, there was no motor response. This long gap thus alleviated concerns that the ability to judge the time of decision accurately may be biased by motor preparation, execution, and resultant effects. This was a problem in the original Libet paradigm (Libet et al., 1983) that many subsequent studies have tried to address (Haggard & Eimer, 1999; Haggard, 2005; Haggard et al., 2002; Lau et al., 2004; Lau et al., 2006; Trevena & Miller, 2002; van de Grind, 2002). After the first response was given, the screen went blank for 1,000 ms. Then four letter options were presented: the letter presented at the point of conscious decision (frame 0), the two letters immediately before it, and the letter immediately after it. They were shown for 1,000 ms in a randomized configuration in the four corners of the stimulus frame, from which participants selected the letter by pressing the corresponding button (Soon et al., 2008). On the basis of the timing of the first response, the stimulus frame during which the conscious decision was made could already be inferred, but having an explicit behavioural response provided additional confirmation and also ensured that the participants performed the task exactly as instructed. The screen then went blank again for 1,000 ms before the next trial period began with the presentation of a new stimulus stream.

Before the fMRI experiment, a behavioural selection test was conducted to select participants who met a set of criteria and to familiarize them with the task. It was important that the number of instances of both abstract intentions were roughly balanced to avoid biased sampling. To measure how balanced participants were in their choices, a bias index comparing the total number of ‘addition’ trials (A) with the total number of ‘subtraction’ trials (S) was calculated using the formula $(A - S)/(A + S)$. However, asking participants to balance their choices both would have required that they kept track of the distribution of intentions and would have affected their choice spontaneity, perhaps even encouraging preplanning of choices. Instead, participants were explicitly told that they did not have to balance their choices. To increase the chances of getting approximately balanced choices during the fMRI session, only participants who spontaneously showed balanced choices in the behavioural selection test were selected. Participants were also asked to stay as relaxed as possible when the stimulus stream began and to refrain from initiating actions too quickly just because they were eager to fulfill (perceived) experimental demands. Rather, they should only initiate a volitional action when they felt the spontaneous ‘urge’ to do so. This served two purposes: the first was to let their mental activity settle down to a stable state, so that any build-up of neural activity before the conscious decision could be clearly observed against this baseline; and the second was to minimize the overlapping of haemodynamic responses from different trials, so as to facilitate unambiguous interpretation of the BOLD signal before the conscious decision. Participants who were too hasty and made decisions in less than 10 s on average during the behavioural selection test were also excluded from the fMRI experiment. Participants were never told of these selection criteria throughout both behavioural and fMRI experiments.

4.3.3 Functional imaging and preprocessing

A Siemens 3T Magnetom scanner was used to acquire echo-planar image (EPI) volumes with 30 slices at a resolution of 3×3 mm, with a slice thickness of 2 mm with a 1-mm gap, covering prefrontal, parietal, and most of temporal cortex (repetition time, 2,000 ms; echo time, 30 ms; field of view, 192×192×90). Ten runs of 152 fMRI volumes were acquired for each participant. A 46-slice whole-brain EPI image was also acquired to facilitate spatial normalization.

fMRI data were preprocessed using SPM2 (www.fil.ion.ucl.ac.uk/spm). After discarding the first two images to allow for magnetic saturation effects, the remaining functional images were then realigned to correct for head motion. The whole-brain EPI image was spatially normalized to the Montreal Neurological Institute (MNI) EPI template. The resultant transformation matrix was subsequently used to spatially normalize the decoding accuracy maps.

4.3.4 General linear model

A general linear model was estimated for each functional run. As this paradigm involved self-initiated voluntary action, we did not expect the shape of the BOLD time course to fit the standard stimulus-response haemodynamic profile. Thus, we used finite impulse response (FIR) predictors to model the BOLD signal (Henson, 2004). A set of 14 FIR regressors were used to model each trial type (adding or subtracting), covering a time range from 8 s before until 18 s after the conscious decision (a functional scan was acquired every 2 s). In addition, trials with inappropriate responses for either the arithmetic problem or the decision time were modeled separately as error trials. Each of the two button presses were modeled separately as involving either the left or right hand, and were entered into the GLM as covariates after convolving with a standard Gaussian haemodynamic response function.

4.3.5 Multivariate pattern analyses

Decoding the content of abstract decisions

The parameter estimates from the FIR regressors for addition and subtraction trials were then subjected to multi-voxel pattern analyses (MVPA) using a searchlight approach to search for predictive neural activity in different brain regions in an unbiased fashion (Haynes et al., 2007; Kriegeskorte et al., 2006; Soon et al., 2008) (see Chapter 1 for details). This method has previously been shown to be highly sensitive in decoding simple motor intentions and in uncovering the flow of information in decision formation (Bode & Haynes, 2009; Soon et al., 2008). Independent linear support vector machine (SVM) classifications (Müller et al., 2001) were performed for each searchlight (3-voxel radius) and at each time-point to assess the amount of intention-related information present in local neural activation patterns using LIBSVM (www.csie.ntu.edu.tw/~cjlin/libsvm). In each of 10 cross-validation cycles, data vectors from 9 runs were used to train a SVM model to maximally distinguish the spatial activation patterns associated with the two decision classes: addition and subtraction, and the resultant model was then evaluated with the independent test run. Good classification accuracy implied that the local cluster of voxels within the searchlight spatially encoded information about the participant's specific current intention at the tested time-point. Combining the accuracy maps across all 14 time-points (from 8 s before to 18 s after the decision) revealed the temporal flow of intention-related information across the cortex.

A second-level analysis was then performed on a voxel-by-voxel basis to determine how well classification could be performed on average across all participants from each time and each position in the brain. For this purpose, individual classification accuracy maps were spatially normalized to the MNI EPI template. These spatial images of local decoding accuracy were entered into a one-way ANOVA with 14 levels, one for each time-point. Regions that predicted the abstract intention were identified using a t-contrast of time-bins before the decision onset (p<.00005 uncorrected, 5-voxel cluster threshold).

Decoding the timing of abstract decisions

As in Study 1 (Soon et al., 2008), we also assessed whether the timing of the decision – as opposed to the specific choice for adding or subtracting – could also be predicted. This distinction is referred to as when-versus-what decisions (Brass & Haggard, 2007; Haggard, 2008). First, multi-class pattern classification was performed to identify cortical areas that contained early predictive information about when participants would decide to perform adding or subtracting. For each searchlight, a multi-class SVM was trained to classify which of five time bins −8 s to 0 s before the decision a particular vector of parameter estimates came from. Correct identification of time bins before the decision indicated the possibility of predicting how much later it would occur. (Given the haemodynamic response delay, the BOLD signal at 0 s would have come from neural activity occurring a few seconds earlier.) This yielded five classification accuracy maps for each participant: one for each time-point before the conscious intention. A one-way ANOVA with five levels, one per time-point, revealed early predictive information about the timing of the upcoming decision in pre-SMA (t-contrast based on all five time-points, family-wise error correction for multiple comparisons, and 50-voxel cluster threshold).

This result was then used to define the region of interest for the next stage of the analysis, in which we tried to predict on a trial-by-trial basis when each decision was going to be made, using brain activity occurring before the decision became conscious. To maximally exploit available information relevant to decision timing, we combined information from spatial and temporal dimensions to perform a spatiotemporal classification analysis (Mourão-Miranda et al., 2007). The aim was to identify the spatiotemporal activation patterns immediately preceding the conscious decision against spatiotemporal vectors that were created from other time windows.

The previous analysis indicated that the three time-points immediately preceding the conscious decision (−4 to 0 s) were most informative about decision timing. For each trial, preprocessed EPI signals, rather than parameter estimates, from every three consecutive time-points and all of the voxels within the pre-SMA region of interest were concatenated into a spatiotemporal vector. If we could accurately identify the last (target) vector within the trial, made up of time-points −4 to 0 s, we would be able to predict that the participant was making an abstract decision that would result in a behavioural response 3 s later. All trials were randomly divided into 4 sets for each participant. For each of 4 cross-validation cycles, 3 sets were used to train a support vector classifier (radial basis function) to distinguish the target vector from vectors of other concatenated time-points (non-targets). To avoid biased sampling during training, the same number of target and non-target vectors were randomly selected. Every spatiotemporal vector from the remaining data set was then classified as a target or non-target. This allowed us to assess how accurately the exact time a conscious decision was going to be made could be predicted beforehand.

4.3.6 Independent components analyses

Finally, we investigated how the predictive signals were related to activity in the task-negative DMN. To identify the DMN across participants, group-level independent components analysis (ICA) was performed using the Group ICA of fMRI Toolbox (GIFT; http://mialab.mrn.org/software/gift/index.html) with the Infomax algorithm (Bell & Sejnowski, 1995). The DMN was visually selected from 15 resulting components, and its time course was extracted for individual participants. The same analysis was also performed for our earlier experiment involving free motor decisions (Soon et al., 2008, Study 1). For each experiment, a conjunction analysis was used to determine the spatial overlap between the DMN and choice-predictive regions (slightly relaxed threshold of $p<.0001$, 5-voxel cluster threshold).

4.4 Results

4.4.1 Behavioural results

Only participants who were highly proficient in both tasks during the behavioural training session were selected for scanning. In the fMRI experiment, the inappropriate numbers (two out of four response options) were only selected on 1.2% (SE = 0.3%) of the trials, indicating that participants performed the chosen task correctly instead of making random responses. The participants' decisions could thus be reliably inferred, based on the choice of one of the two correct answers. For the appropriate responses, participants randomly selected between adding (51.9% of trials, SE = 2.0%) and subtracting (48.1%; SE = 2.0%) with equal probability ($t_{16}<1$). Given that the four response options were randomly arranged, the specific motor response was uncorrelated to the choice of task. Therefore, any information regarding the content of the decision that was decodable from brain activity could not possibly be a result of covert motor preparation. In 97.4% (SE = 0.5%) of the trials, participants indicated that the decision was made three frames before the response, i.e., in frame 0, showing that they followed the task execution sequence exactly as instructed. On average, participants made a spontaneous decision 17.8 s (SE = 1.8 s) after trial onset, resulting in 12.2 (SE = 0.8) decisions per experimental run. This long delay facilitated the search for unconscious neural precursors of the decision while avoiding haemodynamic signal contamination from the preceding trial.

4.4.2 Functional imaging results

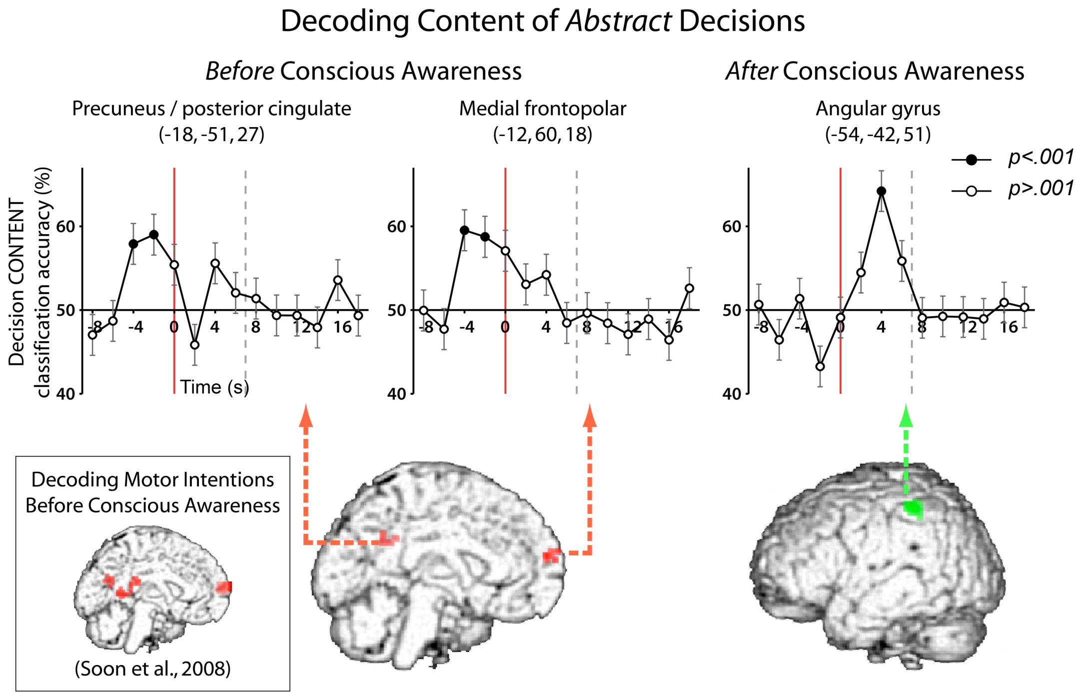


Fig. 4-2. Decoding the outcome of abstract decisions before and after they reach conscious awareness. Projected onto the medial cortical surface are brain regions that predicted the outcome (red) of the abstract decision before it was consciously made (MNI coordinates). Inset shows similar results for the decoding of free motor decisions before conscious awareness in Study 1 (Soon et al., 2008). The lateral surface shows the region that encoded the outcome of the decision after it became conscious. Line graphs depict for each cortical region the accuracy with which the abstract decision to perform addition or subtraction could be decoded at each time (error bars, SE; chance level, 50%). The vertical red line indicates the point of conscious decision, and the vertical gray dashed line indicates the onset of the next trial. Given the haemodynamic delay, information available at 0 s would have been a result of neural activity occurring a few seconds earlier. Please note that none of the points below chance level was statistically significant and should thus be attributed to random fluctuation.

Decoding the content of abstract decisions

We first investigated which cortical regions contained predictive information about the outcome of the abstract decision to perform addition or subtraction, and whether such information was available before or after the decision reached conscious awareness. We found that up to 4 s before the conscious decision, a medial frontopolar region (p<.00005 uncorrected; 5-voxel cluster threshold; 59.5% accuracy) and a region straddling the precuneus and posterior cingulate (p<.00005 uncorrected; 5-voxel cluster threshold; 59.0% accuracy) began to encode the outcome of the upcoming decision (Fig. 4-2). During this early phase, the overall signal in both regions did not show any significant change from baseline ($t_{16}<1$), nor was there any significant difference between addition trials and subtraction trials ($t_{16}<1$), suggesting that the information was encoded in the fine-grained spatial pattern of activation, rather than any global increase or decrease in neural activity (Fig. 4-3).

We also looked for brain areas that encoded the decision *after* it was made, during the task preparation and execution phase. The task choice could be decoded from the angular gyrus 4 s after the time of the conscious decision (Fig. 4-2, right; p<.00005 uncorrected; 5-voxel cluster threshold; 64.2% accuracy). Increase in information began as early as 2 s post-decision and reached statistical significance at 4 s post-decision. Taking into account the haemodynamic delay, this means that the angular gyrus probably began to encode the task choice around the time of conscious decision. This likely reflected the preparation followed by actual performance of the arithmetic task, as the angular gyrus has been found to be involved in the retrieval of overlearned arithmetic facts from memory (Dehaene et al., 2003; Grabner et al., 2009). As expected, there was also a global increase in BOLD signal in the angular gyrus, peaking around 4–6 s after the decision was made (Fig. 4-3). However, no significant difference between the two arithmetic tasks was found, even when thresholds were relaxed (*p*<.001, uncorrected), indicating again that the choice of task was encoded in the local detailed spatial patterns of activation, rather than the global activation magnitude.

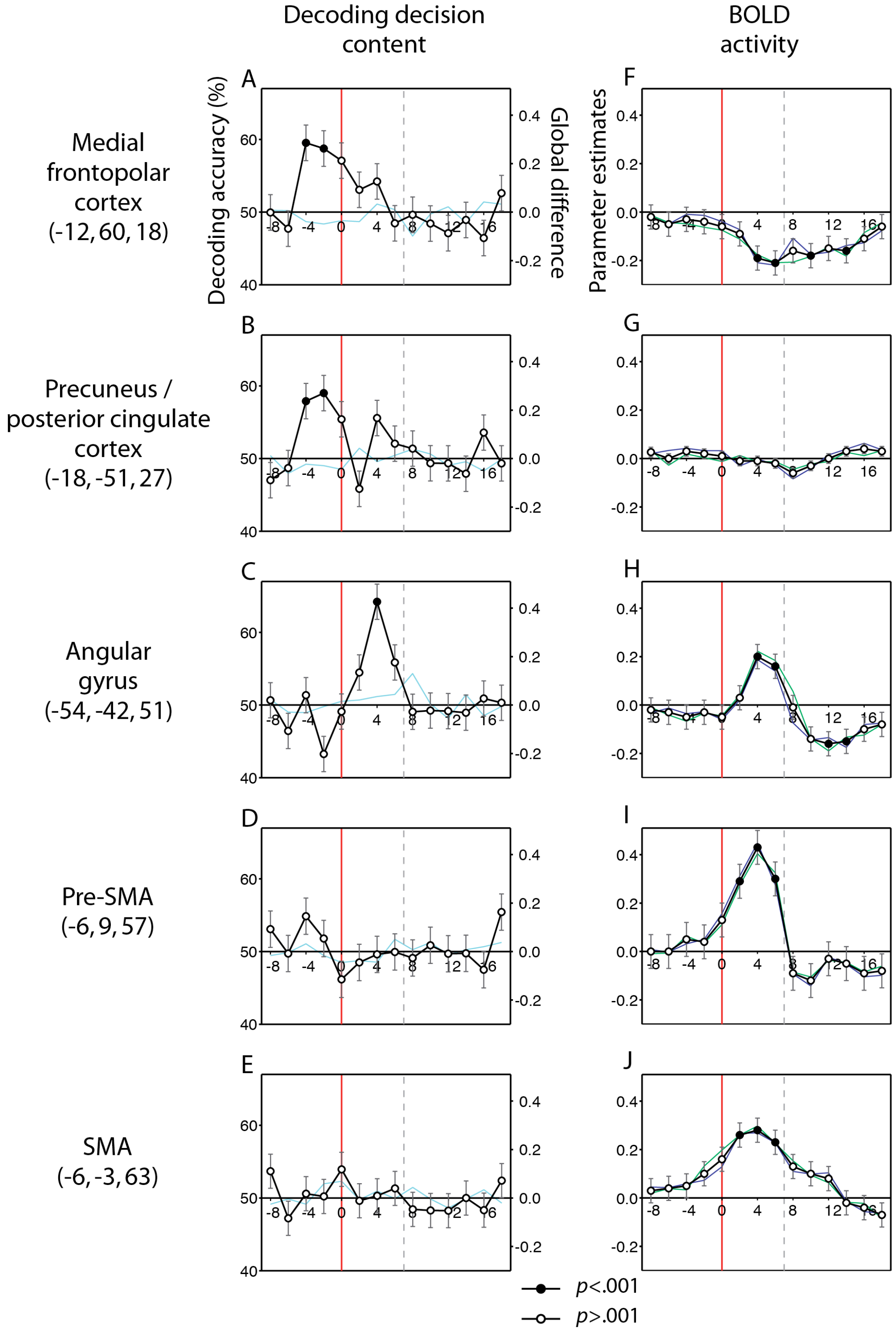


Fig. 4-3. Decoding accuracies and BOLD activation from regions that encoded either the content or timing of abstract decisions. (A–E) Decoding accuracies for the outcome of an abstract decision before and after it reached awareness. Each plot depicts separately for each time the peak accuracy with which the participant's free decision to perform addition or subtraction could be decoded from the spatial pattern of brain activity in that region (black line, left axis; filled symbols, significant at $p<.001$; open symbols, not significant; error bars = standard error; chance level, 50%). The vertical red line shows the point when the decision was first consciously made. Up to 4 s before the decision reached conscious awareness, or ~7 s if the haemodynamic delay is taken into account, predictive information was already found in high-level brain regions (medial frontopolar cortex and precuneus/posterior cingulate). After the conscious decision was made, when the arithmetic task was being performed, the participant's choice was encoded in the angular gyrus, which has been implicated in the retrieval of overlearned arithmetic facts from memory. The vertical gray dashed line in each graph shows the onset of the next trial, with new stimuli appearing. Note that in all regions of interest, decoding of the current intention was at chance level after the next trial began, suggesting there was no 'carry-over' of information to the next trial. (F–J) BOLD activity averaged across voxels in the peak searchlight for each region and each time (black line; filled symbols, significant at $p<.001$; open symbols, not significant; error bars = standard error). As reported previously for the free formation of motor intentions (Soon et al., 2008), there was gradual increase in overall BOLD activation in pre-SMA and SMA in the pre-decision phase, which did not carry information about the decision outcome. In all regions of interest, addition and subtraction trials had similar overall BOLD activation profiles (F–J: green and blue lines) and their overall differences (A–E: right axis, cyan line) were small, suggesting that the specific outcome of the intention was mainly encoded in fine-grained local patterns of activation rather than overall signal differences.

Control analysis: Decoding motor responses

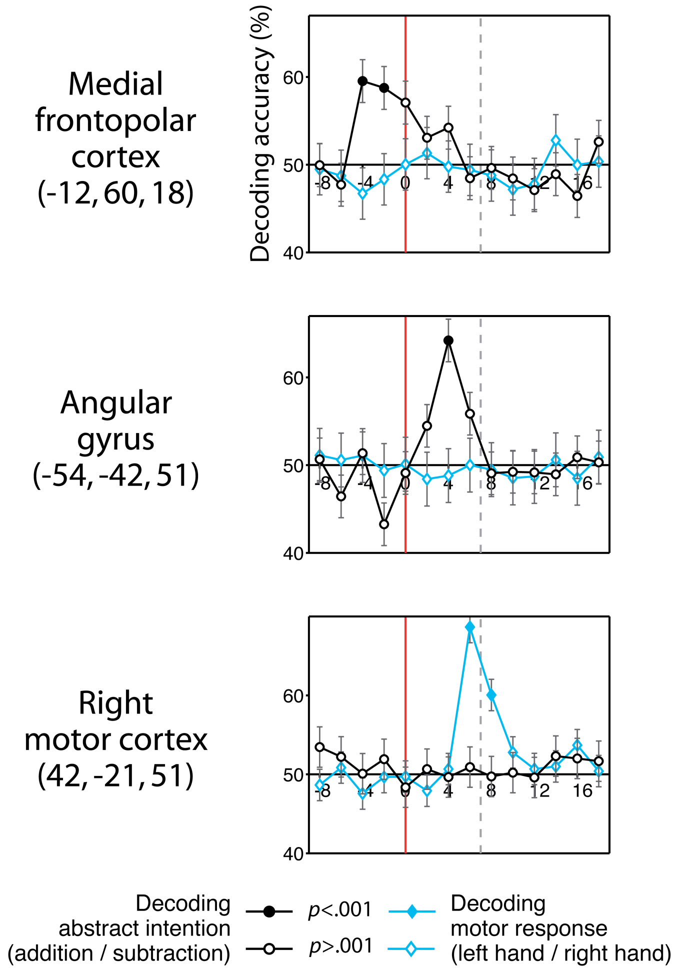


Fig. 4-4. Decoding the response hand for the arithmetic task. The response hand could only be significantly decoded from bilateral primary motor cortex and SMA, 6–10 s after the decision was made. Classification accuracies are shown for three regions of interest that encoded the abstract intention before the conscious decision (Top: medial frontopolar cortex), the performance of the chosen arithmetic task after the decision was made (Middle: angular gyrus), and the motor response (Bottom: right motor cortex). In each plot, the classification accuracy for the abstract intention is depicted in black and the classification accuracy for the motor response is depicted in cyan (left axis; filled symbols, significant at $p<.001$; open symbols, not significant; error bars = SEM; chance level, 50%). The vertical red line shows the time when the decision was first consciously made. Classification accuracies for the motor response were at chance level for medial frontopolar, angular gyrus, and all

other regions of interest that encoded the abstract intention (not shown). In contrast, the motor cortex, which did not encode the abstract intention, began to encode the motor response about 2 s after the angular gyrus encoded the chosen arithmetic operation. Together, these results suggested a temporal flow of abstract intention formation in high-level frontoparietal regions, followed by simple arithmetic task processing involving the angular gyrus and culminating in the motor response by the motor cortex. The vertical gray dashed line in each graph shows the onset of the next trial, with new stimuli appearing.

We also used additional searchlight analyses to investigate the neural encoding of motor responses. As expected, the movement could be decoded from bilateral motor cortex and SMA, 6–10 s after the abstract decision was made (Fig. 4-4). Please note that this delay included the time to complete the arithmetic task plus the haemodynamic response delay. Importantly, these motor regions did not encode the abstract intention (Fig. 4-3 and 4-4). In our previous study (Soon et al., 2008), SMA also had choice-predictive information for motor intentions, albeit later than the frontopolar and parietal cortex. However, it was not possible here to decode the abstract intention from activity patterns in SMA at any time before or after the decision (Fig. 4-3). In return, the cortical regions that encoded the abstract decisions here did not encode the motor response at any time (Fig. 4-4), providing further evidence for a dissociation between the preparation of an abstract intention and motor processing.

Decoding the timing of abstract decisions

A two-stage multivariate classification analysis showed that the timing of the decision, rather than its outcome, could also be predicted. First, we identified brain regions where we could correctly identify which time bin before the decision a pattern vector came from. This information was found to be encoded in the pre-SMA, extending into the SMA and rostral cingulate zone (p<.00005 uncorrected; 5-voxel cluster threshold; 31.4% accuracy; chance level, 20%; Fig. 4-5), as reported previously for simple motor intentions (Soon et al., 2008). In contrast, the medial frontopolar and posterior cingulate/precuneus regions that encoded the content of the decision did not contain predictive information about the timing of the decision (t_{16}<1).

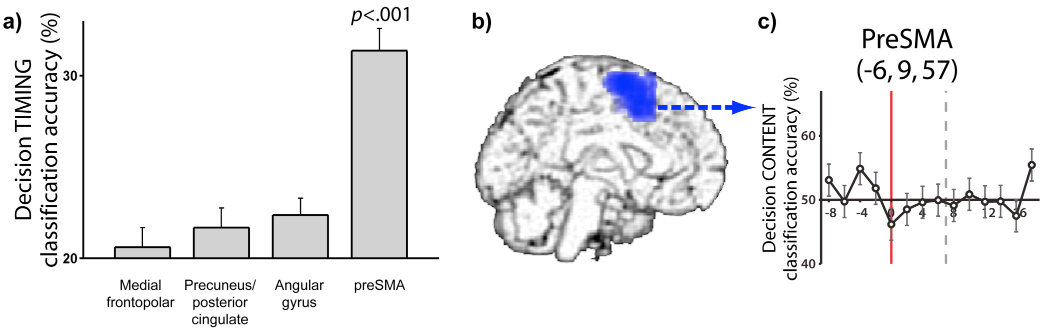
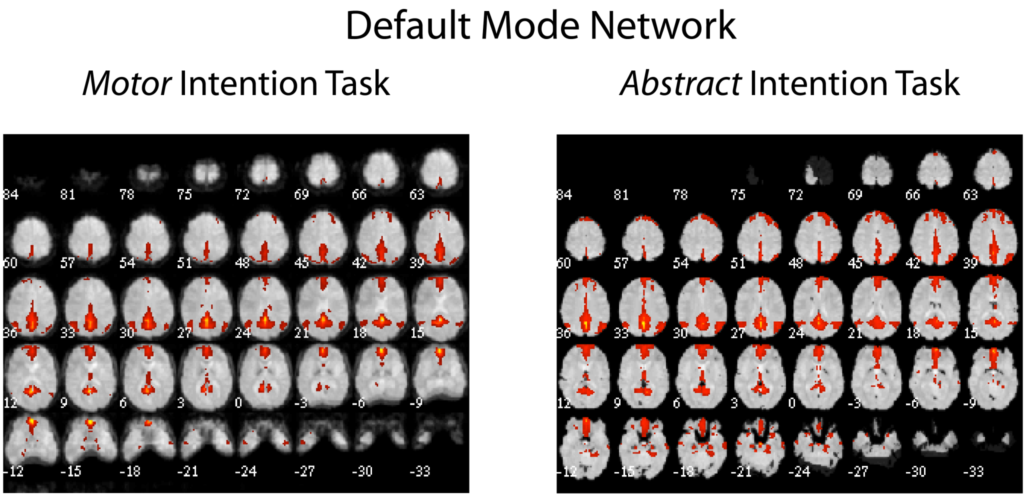


Fig. 4-5. Classification results for timing of upcoming decisions. The searchlight approach was also used to determine whether different brain regions contained early predictive information about the timing of the upcoming decision rather than its outcome. (A) The bar chart shows how accurately the five time bins from −8 to 0 s before the upcoming decision could be classified (chance level, 20%) in various regions. This information was not available in regions that encoded the outcome of the decision but was found in pre-SMA stretching into SMA. (B) The pre-SMA region predictive of decision time projected onto the medial cortical surface is rendered in blue. (C) Line graph showing that pre-SMA was not predictive of the outcome of the upcoming decision (see Fig. 4-2 for plot legend). Thus, there was a double dissociation such that regions that predicted the outcome of the decision did not predict its timing, whereas regions that predicted the timing of the decision did not predict its outcome.

Next, we explored on a trial-by-trial basis how accurately the timing of each decision could be predicted before it was consciously made from the pre-SMA cluster. Classifiers were trained to identify spatiotemporal activity patterns in the pre-SMA region from −4 to 0 s (three time-points) before participants made a conscious decision. By detecting when this pattern occurred during each trial, we were able to accurately predict the exact time that participants were going to make a decision before they had made any behavioural response (71.8%, SE = 1.6%).

4.4.3 DMN and intention formation


Default Mode Network

Motor Intention Task

Abstract Intention Task

Fig. 4-6. Brain images showing the default mode network. The DMN was identified using group-level independent components analyses in both motor intention (Left) and abstract intention (Right) experiments.

The DMN was identified with standard techniques using group-level ICA (Calhoun et al., 2001), and its average time course was extracted (Fig. 4-6 and Fig. 4-7). The same analysis was also applied to data from Study 1 involving free *motor* decisions for comparison (Soon et al., 2008). In both experiments, a typical fronto-parietal DMN profile was observed (Fig. 4-6). To formally assess the spatial overlap between choice-predictive regions and the DMN, we performed a conjunction analysis (Fig. 4-7A). In the abstract intention task, there was a partial overlap with DMN in anterior medial prefrontal cortex. For Study 1, there was a partial overlap with DMN in medial parietal cortex. Please note that it has been repeatedly shown that successful spatial pattern classification does not depend on whether a brain region shows a net positive (or negative) change in magnitude (Bode & Haynes, 2009; Bode et al., 2011; Haynes and Rees, 2006; Norman et al., 2006; Soon et al., 2008).

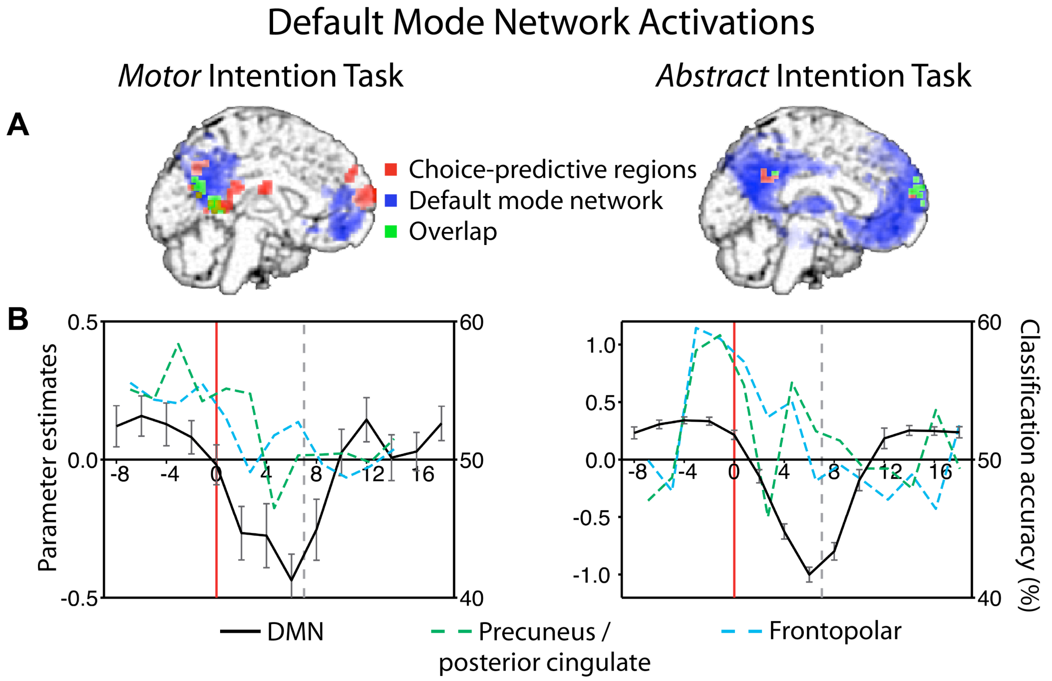


Fig. 4-7. Changes in overall default mode network activity for spontaneous motor and abstract intentions. (A) For both motor (Left) and abstract (Right) tasks, the DMN (blue; $p<.0001$) and choice-predictive (red; $p<.0001$) regions were projected onto the medial cortical surface, with overlapping voxels rendered in green. Please note that volume rendering projected regions from different depths onto the medial plane. (B) Activation time courses of the DMN (black line, left axis) were plotted together with the classification accuracies for precuneus/posterior cingulate (green dotted line, right axis) and frontopolar cortex (cyan dotted line, right axis). For both motor (Left) and abstract (Right) tasks, the default mode activity and the predictive information peaked roughly

> around the same time before conscious decision. The DMN during this early preparatory phase still resembled typical off-task or ‘resting’ period activity, whereas parietal and prefrontal signals already encoded the upcoming choice.

We also directly compared the time courses of DMN activity and early predictive information and found that both peaked roughly around the same time before the spontaneous decision (Fig. 4-7B). The main difference was that the default mode activity began to rise earlier and lasted longer. The comparable peak time of DMN and choice-predictive information was also interesting because it meant that even while parietal and prefrontal signals already had information about the outcome of the next choice, the DMN continued to exhibit off-task-like behavior (high BOLD signal).

4.5 Discussion

Our results showed that regions of medial frontopolar cortex and posterior cingulate/precuneus encode freely chosen abstract intentions before the decisions have been consciously made. Previously, a similar network of regions was reported to be involved in the unconscious preparation of motor intentions (Soon et al., 2008, Study 1). In contrast, in the current experiment, the intended ‘action’ was a nonmotor, abstract mental operation. Importantly, the randomized response mapping ensured that the motor output used to indicate the arithmetic answer was independent of, and did not contain any information about, the abstract decision made. In addition, these regions did not encode the motor response at any time before or after the conscious decision, even when the behavioural response was made. Thus, by fully dissociating high-level intentions from motor preparation and execution, our current findings provided direct evidence that the medial frontopolar and precuneus/posterior cingulate regions were involved in the formation of high-level intentions for voluntary actions in general. This also confirmed that the frontoparietal network previously reported was indeed involved in the formation of free intentions *per se* (Soon et al., 2008), rather than motor preparation. In future studies, it should be possible to enhance the sensitivity for smaller differences between informative brain regions by directly comparing motor and calculation tasks in the same participants.

Closer to and after the point of conscious decision, other regions involved in the actual execution of the specific voluntary action began to encode the intention. In the case of a simple motor action, this was the motor system, including SMA and bilateral primary motor cortex (Fried et al., 2011; Soon et al., 2008). For the abstract mental arithmetic task reported here, this was the angular gyrus, involved in the retrieval of overlearned arithmetic facts from memory (Dehaene et al., 2003; Grabner et al., 2009). These areas were only involved in the execution of specific tasks and did not encode intentions in general.

Previous studies have also shown that abstract decisions can be decoded from the medial frontopolar cortex, but only when these decisions were consciously maintained during a delay period before execution (Hampton & O'Doherty, 2007; Haynes et al., 2007). Thus, it has remained unclear whether the same regions were also involved in the unconscious formation of voluntary intentions or merely store the decisions after they have been consciously formed. In the current study, participants were not cued to make decisions at specific points in time but were allowed to make decisions spontaneously. By asking participants to report when they first consciously decided, we could investigate what happened in the brain before the decisions were consciously made. We found that both medial frontopolar cortex and posterior cingulate / precuneus started to encode the specific outcome of the abstract decisions even before they entered conscious awareness. Our results suggested that, in addition to the representation of conscious abstract decisions (Haynes et al., 2007), the medial frontopolar cortex was also involved in the *unconscious* preparation of abstract decisions.

Both regions that encoded the content of the decision ahead of time have also been implicated in tasks involving prospective memory (Burgess et al., 2001; Burgess et al., 2003; den Ouden et al., 2005; Okuda et al., 2003; Simons et al., 2006) and imagining the past or the future (Addis et al., 2007; Badre & D'Esposito, 2009; Boorman et al., 2009; Okuda et al., 2003; Szpunar et al., 2007). A possible interpretation of the current results is that these regions were involved in unconscious preparation for actions in the near future (Boorman et al., 2009). It remains unclear whether both conscious and unconscious representations are subserved by the exact same neural substrates within these regions or whether they are separable at a finer scale. Future studies could address this question by verifying whether mutual classification between unconscious and conscious decision representation is possible. If they do share common neural substrates, then it should be possible to use classifiers trained on unconscious decision representation to decode consciously maintained decisions and vice versa.

At present, the specific functional roles of the frontopolar and precuneus/posterior cingulate regions in the formation of free decisions are not fully understood. Interestingly, medial frontopolar and posterior cingulate cortex have also been implicated in other types of decision making involving rewards, rather than ‘random’ and ‘free’ choices (Kahnt et al., 2010, 2011; McClure et al., 2007; Tusche et al., 2010; Weber & Huettel, 2008), and in tracking the subjective values of the rewards (Kable & Glimcher, 2007; McCoy & Platt, 2005). In most studies of decision making, participants are typically asked to decide between different alternatives that are seldom as balanced as in the current experiment. The choices are often biased by prior feedback (Hampton & O’Doherty, 2007) or reward characteristics such as reward type, reward level, or reward timing (Kable & Glimcher, 2007; McClure et al., 2007; McCoy & Platt, 2005).

An important finding was that regions that predicted the content of the decision (adding or subtracting) did not predict the exact timing of the decision. Rather, this information was found in the pre-SMA, which has also been shown to be involved in the generation of self-initiated and self-paced motor actions in experiments in which the timing of decisions was not explicitly measured (Cunnington et al., 2003; Cunnington et al., 2005; Debaere et al., 2003; Deiber et al., 1999; Forstmann et al., 2006; Fried et al., 2011; Libet et al., 1983; Mueller et al., 2007; Soon et al. 2008; Wiese et al., 2004). Taken together, when decision time is unconstrained and self-paced, unspecific activity can build up in this region up to a few seconds before the decision for a voluntary action, whether motoric or abstract, is triggered. When different alternatives are available, the specific choice is shaped by activity in other regions such as the frontopolar and precuneus/posterior cingulate. Combining both sources of information about intention content and timing may provide a promising approach to more completely understanding the neural mechanisms underlying voluntary actions.

Consistent with our findings, there is increasing evidence that prior fMRI signals can be used to predict certain upcoming mental states: Spontaneous generation of motor intentions (Study 1) and abstract intentions (current study) can be predicted a few seconds before conscious decisions are made, and behavioural errors can be predicted up to 30 s before they occur (Eichele et al., 2008). Although predictive fMRI signals have to be interpreted carefully (Sirotin & Das, 2009), there is evidence suggesting that they could reflect low-frequency local field potentials (Handwerker & Bandettini, 2011; Logothetis et al., 2001). Moreover, other studies using electrical recordings have also found early predictive signals. In their original EEG study, Libet et al. (1983) reported the onset of the readiness potential to occur a few hundred milliseconds before the conscious motor intention. In a more recent study using single and multiunit depth electrodes, Fried et al. (2011) found that in some cases, the firing rates of certain medial frontal neurons could start changing up to a few seconds before the conscious decision is made (see figure S3e in Fried et al., 2011), lending support to our earlier (Soon et al., 2008) and current fMRI findings.

It is interesting that mental calculation, the more complex task, had less predictive lead time than a simple binary motor choice in a previous study (Soon et al., 2008). This could tentatively reflect a general limitation of unconscious processing in the sense that unconscious processes might be restricted in their ability to develop or stabilize complex representations such as abstract intentions. On the other hand, it could reflect a shorter time of intention formation, as participants also took less time to make spontaneous decisions (17.8 s vs 21.6 s), perhaps due to diffusion processes settling on a winning option more rapidly (Schurger et al., 2012). It is also worth noting that both studies showed the same dissociation between cortical regions that were predictive of the content versus the timing of the decision. This implies that the formation of an intention to act depends on interactions between the choice-predictive and time-predictive regions. The temporal profile of this interaction, which is still poorly understood, is likely to determine when the earliest choice-predictive information is available, and might differ between tasks.

There was a partial spatial overlap between the choice-predictive brain regions and the DMN, a functionally connected network of regions that typically show lower BOLD signal during active, externally oriented tasks (Buckner et al., 2008; Fox & Raichle, 2007; Greicius et al., 2003; Raichle et al., 2001). Interestingly, the state of the DMN during the early preparatory phase still resembled that during off-task or 'resting' periods. This lends further credit to the notion that the preparatory signals were not a result of conscious engagement with the task (Fox et al., 2005; He & Raichle, 2009; Raichle et al., 2001). Furthermore, the spatial and temporal overlaps hint at the potential involvement of the DMN in unconscious choice preparation, consistent with its proposed role in self-related processing (Buckner and Carroll, 2007; Gusnard et al., 2001; Wicker et al., 2003).

Conclusion

To summarize, we directly investigated the formation of spontaneous abstract intentions and showed that the brain may already start preparing for a voluntary action up to a few seconds before the decision enters into conscious awareness. Importantly, these results cannot be explained by motor preparation or general attentional mechanisms. We found that regions in frontopolar and precuneus / posterior cingulate, partially overlapping with the DMN, encoded the content of the upcoming decision, but not the timing. In contrast, the pre-SMA predicted the timing of the decision, but not the content.

Taking together the consistent findings across our three studies, our sense of absolute conscious control over our voluntary actions seems to falter, even for simple free-choice tasks with no real consequences.

Acknowledgements

This work was funded by the Bernstein Computational Neuroscience Program of the German Federal Ministry of Education and Research (BMBF Grant 01GQ0411) and the Max Planck Society.

General Discussion

In this thesis, we investigated the temporal order of conscious decisions and predictive neural activity that occur before simple, spontaneous acts of volition. In a series of three fMRI experiments, sensitive multi-voxel pattern analyses (MVPA) techniques were used to evaluate whether information about the content and timing of self-initiated motor or abstract actions was already available in brain activation patterns even *before* the decisions had been consciously made. Frontopolar and precuneus / posterior cingulate cortex were reliably found to contain predictive information about the content of upcoming decisions but not the timing. In contrast, pre-SMA and SMA encoded the timing but not the content of as yet unconscious decisions. Our findings are consistent with a rich body of neuroscientific evidence showing indisputably that we are not always consciously cognizant of the factors that shape our decisions (e.g., Dehaene et al., 1998; Greenwald, Klinger, & Schuh, 1995; Hsieh, Colas & Kanwisher, 2011; Lau & Passingham, 2007; Lin & Murray, 2013; Murawski et al., 2011; Vorberg et al., 2003), even when we subjectively feel to have freely and consciously made these decisions, and are fully aware of the choices made (Bechara et al., 1997; Colas & Hsieh, 2014; Fried et al., 2011; Haggard & Eimer, 1999; Huang et al., 2014a, 2014b; Libet et al, 1983; Murawski et al., 2011; Schlaghecken & Eimer, 2004). The temporal precedence of such predictive information relative to the conscious decision – despite participants reporting to be highly spontaneous, not premeditating on their decisions – suggested that at least some volitional actions are initiated and shaped by unconscious neural activity. The implications of these findings on the functional role of consciousness in human volition will be discussed.

5.1 Summary of findings

Together, the findings reported in these three studies have uncovered neural networks across the brain involved in the genesis of intentions at an abstract, non-motoric level, going beyond previous studies that investigated only the initiation of spontaneous motor intentions and regions restricted to motor preparation (Haggard & Eimer, 1999; Libet et al., 1983). Our results pointed to high-level planning stages that were likely involved in unconsciously preparing the decision, as previously seen in studies on conscious action planning (Bechara et al., 1997; Burgess et al., 2001; Hampton & O'Doherty, 2007; Haynes et al., 2007; Koechlin et al., 1999). In Studies 1 and 2, we revealed the neural precursors of a two-choice spontaneous motor decision. The long temporal precedence of predictive information, bolstered by control experiments and analyses, suggested that the findings reflected the genesis of high-level intentions rather than preparatory motor action plans. Nevertheless, critics continued to point to the motor nature of the decisions in our and other experiments utilizing the Libet clock paradigm, and questioned their relevance for abstract human intentions (Mele, 2009; Roskies, 2010; Schlegel et al., 2013). As such, there was a need to go beyond indirect inferences to study the formation of high-level *abstract* intentions directly. Thus, in Study 3 we adapted our intention clock task to an abstract decision which allowed us to completely dissociate abstract intention formation from motor preparation. Similar choice-predictive and time-predictive networks were found to shape both motor and abstract intentions. Interestingly, we found across the studies that, in the pre-decision period, there was an information double dissociation: the choice-predictive regions did not contain time-predictive information, while the time-predictive regions did not encode the decision outcome. One possible interpretation is that different networks shape our decision of 'what' to do, and 'when' to do it.

5.1.1 Decoding decision outcome

Even before participants consciously decided to act, choosing between two possible options, neural precursors of simple motor (Studies 1 and 2) and abstract (Study 3) intentions could be reliably detected from frontopolar (all three studies) and precuneus / posterior cingulate (Studies 1 and 3) cortex. Multivariate classification analyses could consistently decode from neural activation patterns in these regions whether participants were going to perform a left or right button press (Studies 1 and 2), or to perform addition or subtraction (Study 3) with around 60% accuracy. The choice-predictive information emerged a few seconds before the conscious decision: ~7 s for the motor decisions (Studies 1 and 2), and ~4 s for the abstract decisions (Study 3). The delay between the onset of predictive information – or more accurately, the underlying neural changes – and the conscious decision itself is almost definitely highly dependent on the decision context, constrained by one's higher level goals. More importantly, given that in all three studies predictive information was available much earlier than previous reports of a few hundred milliseconds in EEG studies (Haggard & Eimer, 1999; Libet et al., 1983; Trevena & Miller, 2002), our findings could not be due to inaccuracies in the judgment or reporting of the time of conscious decision.

Higher spatial and temporal resolution of imaging in Study 2 additionally revealed that the choice-predictive activation patterns became increasingly stable with closer temporal proximity to the upcoming decision. The choice-predictive regions (Studies 1 and 3) overlapped partially with the default mode network (DMN), and similarities were observed between the temporal profile of the information timecourse and the mean DMN BOLD signal. The predictive information disappeared from these regions soon after, when the voluntary action was being executed. In contrast, the mean BOLD signal in these regions remained around the baseline level before the conscious decisions, showing no difference between the two choices. The mismatching temporal profiles of the information and BOLD timecourses implied that the information was encoded in fine-grained spatial activation patterns rather than regional BOLD signal changes.

Evidence for a functional dissociation between the frontopolar cortex and the precuneus / posterior cingulate cluster was seen in a control experiment in Study 1 where participants were cued when to freely select one of two hands to respond with, and when to execute the choice after a few seconds' delay. The first region-of-interest to contain information regarding the selected hand was frontopolar cortex in the free selection phase, followed by the precuneus during the delay. In contrast to frontopolar cortex, which only encoded the chosen hand up to the delay but not during the execution phase, the precuneus continued to encode this information during the motor execution phase. This suggested that the decision was shaped and maintained by the frontopolar cortex, while the precuneus was involved in the maintenance and execution of the chosen intentional act.

For the motor decision task, information related to the decision outcome was also seen at a more lenient threshold in SMA in the pre-decision phase, albeit to a lesser degree (Study 1). This was not seen in the abstract decision task, in which the resultant motor response was independent of the decision to perform addition or subtraction, and could not be pre-planned (Study 3). Thus, the choice-specific patterns in SMA were likely involved in preparing for the upcoming execution of the motor decision and did not represent the shaping of the high-level intention *per se*.

Once the choice reached conscious awareness, various task-specific regions involved in its execution began to encode the decision outcome: pre-SMA, SMA and bilateral primary motor cortex for a motor button press (Study 1), and angular gyrus for simple arithmetic (Study 3). Mean regional BOLD signal differences between choices were only seen in bilateral primary motor cortex *post-decision* in Study 1 – although the source neural activity could have occurred around the point of conscious decision. Presumably, the same neural activity also gives rise to the LRP signal previously reported in a similar free-choice motor task with two options (Haggard & Eimer, 1999). Based on the finding that the LRP onset covaried with the conscious decision time, Haggard and Eimer (1999) suggested that the processes underlying the LRP may lead to our awareness of movement initiation[1]. Unfortunately, this could not be verified in our studies, as neither our intention clock nor the haemodynamic response provided sufficient temporal resolution to differentiate 'late' from 'early' decisions relative to movement onset.

5.1.2 Decoding decision time

Despite the limitations in temporal resolution, we were still able to extract information about the *timing* of the upcoming decision from neural activity patterns occurring *before* the conscious decision. This was found in a set of regions that was generally uninformative about decision outcome: pre-SMA, SMA. This time-predictive cluster was first identified for motor decisions (Study 1), and could potentially be the source of the EEG and MEG readiness potential for spontaneous motor actions reported in previous studies (Ball et al., 1999; Deecke et al., 1982;

---

[1] However, in a recent study by Schlegel et al. (2013), these findings could not be replicated.

Haggard & Eimer, 1999; Libet et al., 1983). Mirroring the increase in predictive information for decision time, there was also a trend of increasing mean BOLD signal in pre-SMA and SMA as the conscious decision approached. Thus, it remains unclear whether fine-grained spatial patterns further contributed to the classification accuracy. In contrast to work on the readiness potential (Haggard & Eimer, 1999; Libet et al., 1983), the main level of activity in motor-related brain regions began to increase as early as 5 s prior to the decision. One potential explanation is that decision times in our studies were unconstrained, so unspecific preparatory activity was possibly able to build up over a longer period of time.

Interestingly, this cluster also contained early predictive information about when an *abstract* decision to act was going to occur, and showed a similar trend in rising mean BOLD signal (albeit not significant), even though the decision involved a mental action that did not require any immediate motor response (Study 3). Participants only needed to respond 3 s after deciding to add or subtract, having to first perform the chosen mental operation. By targeting the activation patterns within a time window of −4 to 0 s immediately preceding the decision, the exact time of the abstract decision could be classified with about 70% accuracy on a trial-by-trial basis. Taken together, our results suggested that this time-predictive cluster was involved in generic preparation for upcoming voluntary actions, including not just motor, but also abstract mental operations. On the other hand, the possibility remains that this early preparatory activity was indeed specific to motor outputs, but may arise whenever motor execution of one’s intention is anticipated, whether immediate or delayed.

5.2 Predicting the upcoming decision, not the previous one

An important question is whether the early predictive information could be a result of the carry-over of information from the previous trial(s) rather than due to processes involved in the shaping of the upcoming choice (Lages & Jaworska, 2012; Lages et al., 2014). The basic argument is that participants were not perfectly random when generating a sequence of decisions, and such sequential dependency meant that a participant's history of choices contained information about the upcoming decision. As such, the possibility exists that accurately decoding the previous choice from brain activity could potentially yield information about the upcoming decision. There are many arguments against this possibility, and we have dealt with these in detail in separate publications (Allefeld et al., 2014; Soon et al., 2014), in addition to the original publications of Studies 1 to 3 (Bode et al., 2011; Soon et al., 2008; Soon et al., 2013).

First, the onsets of significant classification accuracy began long after the free decision in the previous trial had been executed (>17 s), and *increased* with distance from the previous trial (Fig. 2-3, 2-4, 3-3, 3-6, 4-2 & 4-3), inconsistent with the idea that the information was decoded from remnant signals of the previous trial. Second, this absence of cross-trial information also held looking forward in time: the classification accuracies in the early predictive regions returned to chance level even before the next trial began, and after that it would be another 15 s or more before the next decision was made. If there was indeed sequential dependency and carry-over of information between trials, then time points that overlapped with the subsequent trial should also encode the choice made in the current trial. However, they did not. This again suggested that the predictive signals could not simply be a result of cross-trial correlation of choices. Third, the trials were modeled using finite impulse response

(FIR) predictors, which can efficiently separate temporally overlapping haemodynamic responses of consecutive trials as long as the intervals between trials are jittered (Miezin et al., 2000). As the decisions were made spontaneously at random self-paced intervals, there was inherent variability in the trial-to-trial intervals, which made the estimation of the FIR GLM even more robust against carry-over effects. If we were really decoding the previous decision rather than the upcoming one, it would mean that the information about the previous choice disappeared for a variable period of time, and then somehow re-emerged again a few seconds before the next decision was made. Fourth, each parameter estimate used for MVPA was derived from trials within the same run that were grouped based on the decision made at the end of each trial, and would have included a mix of trials with different trial histories, i.e., the prior trials of any parameter estimate would have included both ‘left’ and ‘right’ button presses in Studies 1 and 2, and both ‘addition’ and ‘subtraction’ in Study 3. This would have minimized – if not eliminated – any effect of trial-specific history. Fifth, the current choice could be predicted behaviorally from the preceding choice with about 62-64% accuracy, comparable to other studies (Lages & Jaworska, 2012; Lopes, 1982). If this was the *singular* source of the brain-based prediction accuracies of ~60%, i.e., our classification analyses were simply decoding the remnant signals from the preceding trial, it would imply that we could decode the previous trial with near perfect accuracy (Soon et al., 2014). This was highly improbable, given that in Study 1, even though primary motor cortices showed clear post-decision differences in mean activation magnitude between left and right button presses, decoding accuracy remained below 80% (Fig. 2-3). Finally, we *directly* evaluated whether brain activation patterns in the predictive frontal and parietal regions actually contained information about the previous trial (Fig. 5-1; Soon et al., 2014). An independent

MVPA, similar to the original analyses except that the trials were grouped based on the decision in the *previous* trial rather than the current trial, showed that activation patterns did not encode the previous decision.

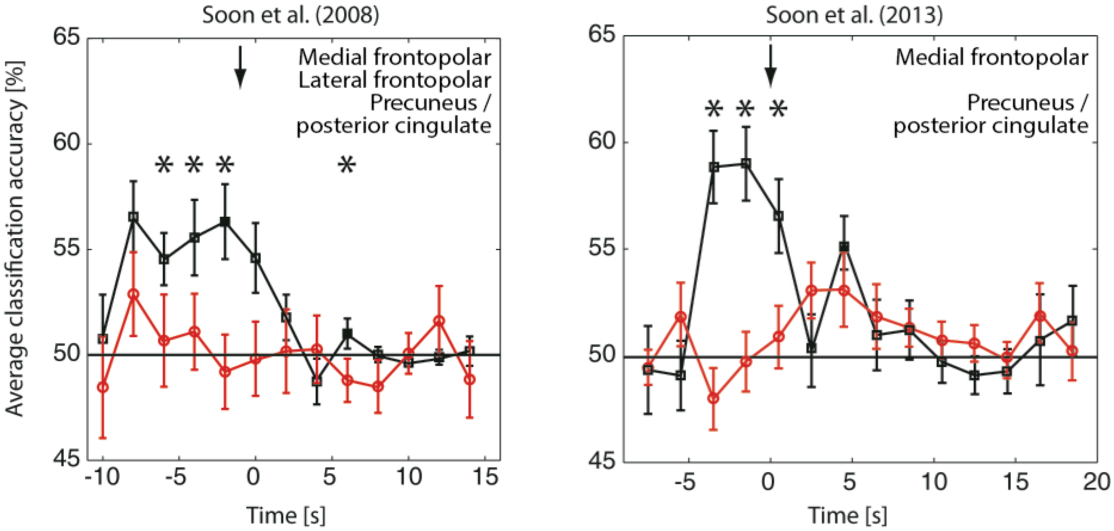


Fig. 5-1. Predictive brain signals do not encode decision in previous trial. For this reanalysis of the original data (left: Soon et al., 2008; right: Soon et al., 2013) we shifted the trial labels by one trial, i.e., trials were grouped based on the decision made in the *previous* trial, rather than the current one. Classification accuracies for the shifted models were not significant at any ROI, implying that spillover from the previous trial did not provide a better account of the predictive brain signals reported in Studies 1 and 3 (red: label-shifted reanalysis; black: original analysis). Here, data was collapsed across ROIs which individually contained choice-predictive signals in the original analyses to increase the statistical power for additionally testing for a *difference* between the original and the shifted analyses (*p<.05). This was necessary as the original analysis was tested against a fixed (i.e., 'noise-free') parameter (theoretical chance-level), whereas the statistical power for testing for a difference between the original and shifted analyses is affected by the noise in the shifted classification. Please also note that the baseline accuracies apparent here (and in the original studies) show that the default accuracy is 50%, as expected for two alternative choices. For this reason, we did not perform additional permutation tests.

However, please note that the above arguments in no way amount to the postulation that the current decision is completely unaffected by the preceding history of choices, which would imply a break in the causal flow of events in the brain. Clearly, a temporally continuous evolution of dynamic brain states must link the sequence of choices. Nevertheless, this does not necessarily mean that the choice-predictive signals we detected across all three studies were direct and explicit representations of the previous trial. Our arguments and analyses directly refute such an interpretation, and support the view that the predictive information reflected neural preparation for the upcoming decision.

5.3 Networks of intention

Having addressed some potential concerns about our paradigms, we now turn to the theoretical implications of our findings. Integrating the results across our three studies and other publications, a rudimentary picture of intention generation emerges (Haggard & Eimer, 1999; Haynes et al., 2007; Libet et al., 1983; Trevena & Miller, 2002). We have shown that spontaneous decisions for simple voluntary acts may be shaped by preceding neural activity across two groups of regions: frontopolar cortex (BA10) and precuneus / posterior cingulate cortex (what to do); and pre-SMA and SMA (when to act). Next we shall consider the specific functional role of each of these regions in preparing the upcoming voluntary action, and how they interact to bring about the final decision to act.

5.3.1 Frontopolar cortex

For both motor and abstract decisions, choice-predictive information first arose in the frontopolar cortex and the posterior cingulate / precuneus. The frontopolar cortex (BA 10) has previously been shown to be involved in free selection (Haynes et al., 2007), self-reflection about intentions (Burgess et al., 2001; Gusnard et al., 2001), and elaboration of past and future events (Addis, et al., 2007; Burgess et al., 2001; Sakai & Passingham, 2003; Schacter et al., 2007). In addition, it has often been implicated in tasks requiring high-level executive control, especially those that involve storing conscious intentions across a delay (Burgess et al., 2001; Haynes et al., 2007; Koechlin et al., 1999; Koechlin & Hyafil, 2007; Ramnani & Owen, 2004; Sakai & Passingham, 2003). Activity in this region can build up even prior to the *execution* of simple movements (Groll-Knapp et al., 1977), as in Studies 1 and 2. However, its relevance for unconscious preparation of intentions has not been demonstrated previously until now.

Of particular interest is the involvement of frontopolar cortex in prospective memory (Burgess et al., 2001; Burgess et al., 2003; den Ouden et al., 2005; Gilbert, 2011; McFarland and Glisky, 2009; Mommenejad & Haynes, 2012, 2013; Okuda et al., 2003; Sakai, 2008; Simons et al., 2006), which requires the retention of an intention over a delay until the appropriate moment for execution. Depending on whether the delayed intention requires proximal or distal execution, they may not always be maintained in conscious awareness, but can get triggered by external cues or spontaneous retrieval (Einstein & McDaniel, 2005; Mommenejad & Haynes, 2012, 2013). In fact, it may be important to keep some delayed intentions out of working memory so that cognitive resources can be freed up for some other task that is of higher priority or requires immediate attention (Burgess et al., 2007; Einstein et al., 2005; Mommenejad & Haynes, 2012, 2013). Prioritizing amongst multiple goals is required for appropriate actions to be executed at the right time. For example, remembering at this moment that I have yet to complete the peer review for a journal manuscript due in two days is not helpful to my current intention: finishing up my thesis. That being said, spontaneous early retrieval or rehearsal could facilitate long-term retention of distal intentions, thereby supporting prospective memory (Einstein et al., 2005). It would not be inappropriate to see prospective memory as part of the executive system for the generic management of high-level current and latent goals and intentions, which includes selecting among different latent intentions for current execution or conscious deliberation (Burgess et al., 2000; Burgess et al., 2007). At any point in time, many latent intentions in our mental to-do list remain outside of conscious awareness. How these latent intentions are represented neurally, and how their neural representations compete to emerge in conscious awareness is still poorly understood. The intention clock paradigm may have allowed us to catch glimpses of this unconscious competition. One possibility is that the activation of frontopolar (BA 10) cortex in prospective memory tasks actually reflects its high-level executive role in the generic management of goals and intentions (Koechlin et al., 1999; Koechlin & Hyafil, 2007; Mommenejad & Haynes, 2012, 2013; Ramnani & Owen, 2004). Perhaps the two options in each decision scenario in each of our studies were represented as separate latent intentions. During the experiment, they would be the two most active and mutually exclusive candidates competing for proximal execution, possibly as described by multi-alternative drift-diffusion accumulator models (Bode et al., 2014; Bogacz, 2007; Krajbich & Rangel, 2011; Smith & Ratcliff, 2004). The winner emerges into conscious awareness and is executed, i.e., a voluntary decision is made to act.

Previous studies suggest a functional dissociation in prospective processing between lateral and medial aspects of frontopolar cortex, based on their respective positive and negative mean BOLD activation profiles (Burgess et al., 2001; Simons et al., 2006). However, this could be because medial frontopolar cortex (also referred to as anterior medial prefrontal cortex) is an important hub of the DMN, which usually shows lower BOLD signal during most active tasks compared to resting baseline (Greicius et al, 2003; Raichle et al., 2001; Shulman et al., 1997). In contrast, our multivariate analyses were indifferent to the mean activation level of a cortical region, and instead focused on the information content encoded in spatial activation patterns. Here we found predictive information for both motor (Study 1) and abstract (Study 3) intentions in medial frontopolar cortex, and for motor intentions in lateral frontopolar cortex (Studies 1 and 2). Our experimental paradigms did not reveal a consistent functional dissociation between medial and lateral BA 10 in information content related to preparing simple upcoming decisions. However, decision-making studies consistently show that an adjacent region of the anterior medial prefrontal cortex represents subjective or economic value (Kable & Glimcher, 2007; Kahnt et al., 2010, 2011; Montague & Berns, 2002; Padoa-Schioppa & Assad, 2006). Perhaps similar subjective 'valuations' of latent intentions were also used to guide the non-economic free decisions in our studies. Expanding the information search to a wider variety of decision contexts may help elicit their functional specificity (Mommenejad & Haynes, 2012, 2013).

It is intriguing that similar cortical regions were involved in the preparation of both motor and abstract decisions. However, it remains unclear whether these different decisions actually involved the same neural populations within the same individual. One possibility is that different types of decisions were coded by independent neural populations coexisting within the same cortical region. Alternatively, the same neural networks may subserve different types of decisions, perhaps with the same sets of competing neural coalitions being flexibly assigned to opposing options across different decision contexts. If true, then it might be possible to predict abstract decisions based on a classification model trained using motor decision activation patterns, and vice versa.

Another important issue relates to the neural origin of the information contained in the predictive spatial patterns. One possible explanation could be that there is a fine-grained clustering of cells with similar preferences for one of the two decision outcomes and that this clustering is smaller than the size of conventional functional areas. In the visual cortex, information encoded in similar fine-grained patterns of visual cortex can be read out using pattern recognition techniques. Simulations based on realistic neural topographies (Haynes & Rees, 2005; Kamitani & Tong, 2005) suggest that this may be due to a 'biased sampling' or 'aliasing' of fine-grained feature columns by the individual fMRI voxels (Haynes & Rees, 2006). This raises the question whether the informative spatial patterns we found might point to the existence of a similar columnar architecture in prefrontal cortex, where cells might be clustered according to similar roles in selective cognitive control. Such a columnar architecture has been highly debated as a general principle of cortical organization (Horton & Adams, 2005; Mountcastle, 1997). It remains unclear whether there is such a topographic organization in prefrontal cortex and other association areas (Averbeck et al., 2006; Constantinidis et al., 2001). Alternatively, our classification patterns might reflect the sampling of a distributed population code for different tasks, as has been proposed from the findings of similar studies on object recognition (Haxby et al., 2001).

5.3.3 Precuneus / Posterior cingulate

It has long been known that parietal cortex plays an important role in processing of motor intentions (Quian-Quiroga et al., 2006; Sirigu et al., 2004). The more inferior medial regions found in our studies, stretching from precuneus to posterior cingulate cortex, have been involved in several tasks closely related to intentions, including prospective processing (Burgess et al., 2001; Mommenejad & Haynes, 2012, 2013), self-referential processing (Cavanna & Trimble, 2006; Kircher et al., 2000; Kjaer et al., 2002; Lou et al., 2004; Vogt & Laureys, 2005), and even as here in free-choice tasks (Larsson et al., 1996). Like the medial frontopolar cortex, the precuneus / posterior cingulate cortex is also a central hub in the DMN (Buckner et al., 2008; Fox & Raichle, 2007; Raichle et al., 2001), and is thought to be heavily involved in monitoring internal and external changes, facilitating novel behaviour or thought in response, or even to regulate the focusing of attention internally or externally (Leech & Sharp, 2013; Pearson et al., 2011). While this region is thought to be a potential neural correlate of consciousness (Vogt & Laureys, 2005), our findings suggest that it also encodes mental representations regarding future decisions that may not be directly or immediately available to conscious awareness.

In Study 1's control experiment, the precuneus / posterior cingulate only began to encode a motor decision after it was available in frontopolar cortex, and stored this information across a delay until the execution phase. This suggested that whereas frontopolar cortex might be involved in task selection, the precuneus' role could be to store the decision across a delay – perhaps also detailing the actions required for task fulfilment (Haggard, 2008) – until the decision reached awareness and eventual execution. We speculate that this region may be instrumental in binding one's intentions with the actual outcome, perhaps encoding the forward model of the expected action and outcome, an important mechanism for ensuring that one's actions lead to intended results, and for establishing a sense of 'willing' and of agency over one's actions (Aarts et al., 2004; Haggard, 2008; Miall & Wolpert, 1996; Sirigu et al., 2004; Wegner, 2003). Our theory is also consistent with its involvement in visuospatial coordination, directing of spatial attention when imagining, preparing or making movements (Cavanna & Trimble, 2006; Kawashima et al., 1995; Wenderoth et al., 2005). Also required would be the monitoring of both internal and external states for changes, and controlling of internal or external focus of attention, in agreement with the stipulated functional roles ascribed to the DMN (Buckner et al., 2008; Fox & Raichle, 2007; Raichle et al., 2001). The centrality of the precuneus / posterior cingulate cluster in the cortical network as a well-connected small-world network hub between parietal and prefrontal regions would be well suited for serving these functions (Bullmore & Sporns, 2009). The finding that this region also encoded the outcome of abstract decisions – before a motor response could be specified – would imply that its proposed role in binding intention to output is not restricted to the motor domain, but also includes the fulfilment of higher level action goals.

Whether the precuneus / posterior cingulate merely stores the intended action and coordinates its execution, or actively shapes the intention itself is less clear. One possibility is that this region may also evaluate and provide feedback on the viability of proximal intentions based on the current state of the body and the external environment (Bullmore & Sporns, 2009; Cavanna & Trimble, 2006; Kawashima et al., 1995; Wenderoth et al., 2005). Supporting evidence is seen in its encoding of participants' decisions when they are forced to 'identify' objects in a noisy scene without any object, suggesting that the precuneus shapes the *guesses*, perhaps by amplifying perceptual noise, or by generating random choices (Bode et al., 2012).

5.3.2 Pre-SMA / SMA

Previous studies on spontaneous free decisions have usually targeted simple motor decisions and cortical regions involved in motor execution, which were also assumed to play a key role in determining the decision (Deecke et al., 1982; Haggard & Eimer, 1999; Lau et al., 2004; Libet et al., 1983; Trevena & Miller, 2002). Our findings that activity patterns and mean BOLD signal changes in pre-SMA and SMA were predictive of the timing of the upcoming decision / spontaneous act corroborated their findings that the onset of the *Bereitschaftspotential* can come before the conscious decision (Deecke et al., 1982; Fried et al., 2011; Haggard & Eimer, 1999; Libet et al., 1983; Trevena & Miller, 2002). In addition, we showed that these signals can actually arise a few seconds before the conscious decision, much earlier than the few hundred milliseconds previously reported. Various reasons facilitated our detection of such early signals. Crucially, our participants did not have to concern themselves with blinking, which had to be avoided in EEG studies due to the large signal artefacts (Libet et al., 1983). As such, their spontaneity was not temporally constrained, and was therefore more ‘free’. This likely allowed participants to settle into a more ‘relaxed’ state, as seen in the longer delay between trial onset and decision time, ~20 s in our studies compared to typical delays of around 5 s in previous studies (Haggard & Eimer, 1999; Lau et al., 2004). Also, how early the RP onset can be detected in electrophysiological signals is dependent on the baseline period selected for comparison. Haggard & Eimer (1999) did not find any clear baseline up to 2600 ms pre-movement, the maximum range they could interrogate due to technical reasons, and thus could not reject the possibility that the RP onset might have occurred earlier. This was recently corroborated in a depth recording study using Libet’s intention clock paradigm (Fried et al., 2011). Fried et al. reported that in trials which patients took longer to make simple motor decisions, thereby allowing for the use of an earlier reference baseline, neural changes in medial frontal regions could be detected up to several seconds before the conscious decision was made.

While information about the decision outcome was also detected in SMA, as implicated previously in free-selection tasks (Cunnington et al, 2005; Deiber et al., 1991; Forstmann et al., 2006; Frith et al., 1991; Haggard & Eimer, 1999; Larsson et al., 1996; Lau et al., 2004; Libet et al., 1983; Pedersen et al., 1998; Tanji, 2001), the onset of this information was much later than in higher-level control regions. Thus, SMA was unlikely to be the cortical site where the specific decision for a movement originated (Eccles, 1982), but more likely to be involved in generic motor preparation in the early phase, and subsequently during motor execution. Please note that due to the temporal delay of the haemodynamic response, the small lead times in SMA / pre-SMA of up to several hundred milliseconds reported in previous studies (Haggard & Eimer, 1999; Kornhuber & Deecke, 1965; Libet et al., 1983) are below the temporal resolution of our method. Hence, we cannot exclude that other regions contain predictive information in the short period immediately preceding the conscious intention. It would be difficult to ascertain whether such signals occurred before or after the decision in fMRI signals.

An intriguing question that remains unanswered is how the choice-predictive regions and the time-predictive regions interact to shape the ‘what’ and ‘when’ of a spontaneous voluntary act (Haggard & Brass, 2008; Mommenejad & Haynes, 2012). In the pre-decision phase, there was an information double-dissociation between these two sets of regions. Even though information about decision time could be gleaned from the pre-SMA / SMA cluster, it remained unclear whether it was actually involved in deciding when to act, or only in preparing for execution of the decision formed in choice-predictive regions. One possibility is that once a winning option began to emerge in choice-predictive regions, this triggered neural preparations for action execution in pre-SMA / SMA. Alternatively, perhaps when activity within this time-predictive cluster crossed a certain threshold, e.g., via diffusion processes (Schurger et al., 2012), an ‘urge’ to act was triggered (Libet et al., 1983), and the ‘preferred’ option in choice-predictive regions at that moment was selected for execution, resulting in a specific voluntary action. This could explain why choice-predictive regions were not predictive of decision time. On the other hand, choice-predictive patterns were found to become increasingly stable over time with higher spatial and temporal imaging resolution (Study 2). In addition, given that both choice-predictive and time-predictive information appeared to build up in tandem over a relatively long duration, it would seem more likely that there was ongoing communication between both sets of regions, such that both ‘when’ and ‘what’ to do reached a conscious decision concurrently in our experimental contexts to culminate in a spontaneous voluntary action. When asked to report the decision times for ‘what’ to do, and ‘when’ to do separately, no significant difference was found. However, it should be pointed out that participants were specifically instructed to be spontaneous, not to pre-decide on a particular choice, and to execute their decision once it was made. This could have inadvertently contributed to the temporal coincidence of ‘what’ and ‘when’ decisions, and perhaps reduced further evaluations of ‘whether’ to execute a chosen action (Brass & Haggard, 2007; Haggard, 2008).

Detailing the functional and information connectivity between choice-predictive and time-predictive networks in various decision contexts using higher resolution imaging could help elucidate how they interact to shape a voluntary action.

5.4 Functional role of consciousness in volitional acts

Our empirical results consistently support the idea that free decisions for spontaneous voluntary acts, both motoric and abstract, result from a cascade of activity within different brain networks, which may occur before the point of conscious decision. It appears unlikely that participants consistently misreported their subjective decision time by a delay of a few seconds. While further evidence is needed before a causal relationship can be concluded, the availability of choice-predictive neural information before the subjective decision time suggest that prior neural activity may have affected the conscious decision. In other words, one may be conscious of making a choice without conscious knowledge of *how* or *why* the decision was actually made (e.g., Bechara et al., 1997). If unconscious neural activity indeed plays a significant role in shaping my decisions before they enter into conscious awareness, and I am not cognizant of such influences as I consciously 'decide', can I still claim that the resultant 'voluntary actions were *caused* by my conscious mind? Or did I simply become aware of unconsciously formed intentions without really knowing how they really came about, akin to how I perceive sensory inputs (Haggard, 2008; Wegner, 2003)? If so, what is the specific functional role of my consciousness in my intentions and actions?

Before we further explore the significance of our results for causal role of consciousness, it should be noted that doubts have been raised about the relevance of studies utilizing variants of Libet's intention clock paradigm, including ours, for understanding human intentions (Mele, 2009; Roskies, 2010; van de Grind, 2002). The general intention to press a button (or in the case of Study 3, to perform addition or subtraction) at some point in the experiment was arguably formed during the briefing at the beginning of the experiment. The subsequent 'choices' made in each trial were just the perpetuation of this high-level distal intention. Moreover, the participants were not completely free in each decision, given that only two options were provided, completely specified by the experimenter. Also, the participants were asked to monitor and execute any spontaneous *urges*, which leads to the question of whether these spontaneous ‘decisions’ should be considered bona fide *intentions* at all. Crucially, the participants were fully conscious when they voluntarily agreed to perform the task according to instructions, way before any of the reported ‘unconscious’ decision neural precursors occurred. Thus, according to this view, these studies pose no threat to the concept of mental causation or free will as they only revealed the neural precursors of the *urges* and not the *intention* itself.

Indeed, any volitional act can and should be seen within the wider context of the underlying hierarchy of goals and intentions (Haggard, 2008). Nevertheless, there is little doubt that the resultant action in every trial of our experiments would itself be considered a voluntary act performed according to one’s volition. Neither the button presses in Studies 1 and 2, nor the mental arithmetic operations in Study 3 would be considered automatized behaviour that are thought to bypass the conscious mind (Bargh & Chartrand, 1999). The causal roots of each action, including ‘what’, ‘when’ and ‘whether’ aspects of the decision (Brass & Haggard, 2008), should presumably be traced back to the conscious mind at the point when one consciously decides to act. Even though the intentional tasks used in our studies indeed involved highly constrained decision-making, what remains clear is that: 1) the participant’s decision to choose one of the two options in any particular trial was neither *completely* determined by the experimenter nor by the distal intention at the beginning of the experiment (Haggard, 2008); and 2) the frequency or exact time to make a voluntary action was up to the individual participant (and in our fMRI studies, not even constrained by eyeblinks). As such, there were still degrees of freedom to each decision in both the exact timing and specific choice of actions, and it would be legitimate to ask whether they were in fact fully determined by the conscious mind.

The fact that our prediction accuracy was only around 60% implied that there was still room for the conscious mind to potentially exert a causal influence. Prima facie, this level of prediction accuracy would indeed suggest that the neural origins of the informative signals contributed to shaping the decisions, but did not determine them completely. However, it should be noted that even though the searchlight MVPA analyses can be highly sensitive, there are still technical limitations. For instance, in Study 1, even in the primary motor cortex, where mean regional BOLD signal differences between the left and right button presses could be easily detected in univariate contrasts during the actual execution, the decoding accuracy remained below 80%. Furthermore, the general goal of our experiments was to demonstrate the existence of significant unconscious predictive information, rather than to push the limits of prediction accuracy. As such, we adopted a relatively conservative approach to classification, and did not, for example, use non-linear classifiers, or optimize classification parameters to increase prediction accuracy. Even so, the choices of some individuals could be predicted with accuracies as high as 80%. It remains to be seen how much more predictive information can be extracted from the pre-decision neural activity patterns with further improvements in data acquisition technology and image analyses techniques.

That being said, our findings cannot and should not be taken as categorically precluding the *possibility* of mental causation, since our experiments did not – and were not designed to – show that the conscious mind is *never* the direct cause of volitional acts (Roskies, 2010; Tse, 2013). However, the very existence of these above-chance predictive signals, which the decision-maker does not have conscious access to, already begs explanation, if one believes that the conscious mind is the sole cause of all voluntary actions.

A conscious decision to act (or not to act) is simply a decision made while one is in a conscious state, and does not entail that the decision process is necessarily conscious, only that the decision outcome is (Rosenthal, 2008). For example, information contained in subliminal stimuli can affect our conscious decisions despite the lack of conscious representation (e.g., Dehaene et al., 1998; Greenwald, Klinger, & Schuh, 1995; Hsieh, Colas & Kanwisher, 2011; Lau & Passingham, 2007; Lin & Murray, 2013; Marcel, 1983a, 1983b; Murawski et al., 2012; Vorberg et al., 2003). Similarly, we may even be unaware of how *supraliminal* stimuli affect our conscious decisions (Bechara et al., 1997). This brings to question the specific causal role of consciousness in the flow of processes from the genesis of an intention to its physical (or mental) execution. On the one hand, our subjective experiences would lead us to believe that our conscious minds are the sole authors of our decisions. On the other extreme, consciousness has been proposed to be a passive witness with no causal role, a view central to some Hindu views of consciousness (Gupta, 1998). More recently, Rosenthal (2008) argued that though the cognitive states that consciousness is associated with have functional significance, their being conscious does not. Others have proposed a more limited causal role for consciousness (Bargh & Chartrand, 1999; Dijksterhuis & Aarts, 2010; Libet et al., 1983). Libet et al. (1983) suggested that consciousness could still exert causal influence via a ‘power of veto’, a notion that has earned the moniker of *free won’t* (Brass and Haggard, 2007). Even while unconscious neural activity shape upcoming decisions, consciousness acts as a gatekeeper that can decide whether to execute the resultant choice or to abort it. Such intentional inhibition has been shown to involve dorsal fronto-median cortex and bilateral anterior insula (Brass & Haggard, 2007). However, the initiation of inhibition may also find its roots in unconscious neural activity (Filevich et al., 2013).

More generally, consciousness is thought to be important for achieving a balance between attentional focus and cognitive flexibility, so as to promote performance and adaptive behaviour, especially when there are obstacles in the pursuit of one's goals (Bargh & Chartrand, 1999; Dijksterhuis & Aarts, 2010). Consciousness may serve as a platform for integrating diverse brain functions that could otherwise operate independently, promoting information exchange for the exercise of global coordination and control, according to the Global Workspace Theory (Baars, 1997; Dehaene & Naccache, 2001). While not every volitional act thus mediated would require or benefit from this purported role of consciousness, having this functionality as a central component of the sophisticated system underlying volitional actions can potentially expand our repertoire of possible actions and increase our adaptability in general (Dijksterhuis & Aarts, 2010). Based on this perspective, different sources of predictive information detected in our studies may reflect the independent operation of various decision-related functions, which remained outside conscious awareness until the point of entry into the 'global workspace'. Rather than being the direct unitary cause of our actions, consciousness may act as a generic enabling platform where 'decisions' (or 'proposals') could potentially be further evaluated before, during and after execution. As our tasks involved simple free decisions of no real consequence, and participants were encouraged to be spontaneous, the final decision very likely played out according to preceding preparatory activity, with little conscious intervention (Bode et al., 2014).

The extent to which prior neural activity shape our conscious decisions is likely to be dependent on the decision context. Perhaps decisions of greater personal consequence would involve more conscious deliberation, and therefore be less predictable from unconscious neural activity. However, it has been shown that even goal selection (Custers & Aarts, 2005), the strength of motivation (Pessiglione et al., 2007) and cognitive resource recruitment (Bijleveld et al., 2009) can be modulated by subliminal reward primes. More recently, endogenous brain activity patterns occurring before participants were exposed to gamble options could predict whether they would be more prone to choosing the risky option or not (Huang et al., 2014). Compared to unconscious thought processes, whether conscious deliberation really leads to better decisions when making complex choices is also a matter of much debate (Dijksterhuis et al., 2006; Rey et al., 2009). In the same vein, while it remains to be shown that our findings based on simple free decisions of no personal consequence can be generalized to consequential decisions that typically engage more conscious deliberation, the possibility remains that such conscious deliberation may be shaped by prior unconscious neural activity or brain states. Thus, the specific causal influence of consciousness in decision making and its purported benefits remains an open question.

5.6 Conclusion and outlook

As volitional agents who are not just concerned with daily survival or reacting to environmental changes reflexively, human beings have to constantly balance complex hierarchies of goals and intentions. This requires prioritizing amongst multiple possible intentions so that appropriate actions can be executed at the right time, while other goals are kept out of working memory, and conscious awareness, while being maintained unconsciously. From this perspective, finding unconscious neural precursors of voluntary decisions should come as no surprise. After all, on what basis do we presuppose that there is a unitary *cause* – the conscious mind – for the diverse repertoire of human voluntary actions that can be found within the most complicated connected dynamic structure in human knowledge – the human brain?

A better approach might be to define the functional parameters for an action to be considered fully volitional. While a conscious brain state would be a necessary qualifying parameter, it would not be the singular cause, but an essential part of the sophisticated neural machinery underlying human volition. Here, we have presented empirical evidence showing that, at the very least, our intuitions that our decisions are always and purely products of the conscious mind may not be reliable. Rather, various aspects of our decisions are shaped by neural activity to which we do not have direct conscious assess, for both motor and abstract decisions. Much still needs to be done to elucidate the complex structure of neural networks involved in intention formation, and the information dynamics that lead to the formation of a decision. Further studies capitalizing on improvements in brain imaging technologies and multivariate pattern analyses are needed to pinpoint the functional specialization of each region and the effective connectivity between them.

While we have uncovered evidence for the unconscious neural initiation of free decisions in the brain, how and why this translates into a conscious decision in the mind remains elusive. On the other hand, embedded in the information patterns that distinguish conscious decisions and their precursors might be a unique chance of identifying the neural correlates of consciousness.

References

Aarts, H., Gollwitzer, P. M., & Hassin, R. (2004). Goal contagion: Perceiving is for pursuing. *J Pers Soc Psychol, 87*(1), 23-37.

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia, 45*(7), 1363-1377.

Allefeld, C., Soon, C. S., Bogler, C., Heinzle, J., & Haynes, J.-D. (2013). Sequential dependencies between trials in free choice tasks. arXiv Preprint arXiv:1311.0753.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial prefrontal cortex and social cognition. *Nat Rev Neurosci, 7*(4), 268-277.

Averbeck, B. B., Sohn, J. W., & Lee, D. (2006). Prefrontal neural correlates of memory for sequences. *Nat Neurosci, 9*(2), 276-282.

Baars, B. J. (1997). *In the Theater of Consciousness: The Workspace of the Mind.* New York: Oxford Univ Press.

Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nat Rev Neurosci, 10*(9), 659-669.

Baker, K. S., Piriyapunyaporn, T., & Cunnington, R. (2012). Neural activity in readiness for incidental and explicitly timed actions. *Neuropsychologia, 50*(5), 715-722.

Ball, T., Schreiber, A., Feige, B., Wagner, M., Lücking, C. H., & Kristeva-Feige, R. (1999). The role of higher-order motor areas in voluntary movement as revealed by high-resolution EEG and fMRI. *Neuroimage, 10*(6), 682-694.

Banks, W. P., & Pockett, S. (2007). Benjamin Libet’s work on the neuroscience of free will. In M. Velmans, S. Schinder (Eds.), *Blackwell Companion to Consciousness* (pp. 657-70). Malden, MA: Blackwell.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *Am Psychol, 54*(7),462-479.

Bargh, J. A., & Morsella, E. (2008). The Unconscious Mind. *Perspect Psychol Sci, 3*(1), 73-79.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*(5304), 1293-1295.

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Comput, 7*(6), 1129-1159.

Bijleveld, E., Custers, R., & Aarts, H. (2009). The unconscious eye opener: Pupil dilation reveals strategic recruitment of mental resources upon subliminal reward cues. *Psychol Sci, 20*(5), 1313-1315.

Bode, S., Bogler, C., Soon, C. S., & Haynes, J.-D. (2012). The neural encoding of guesses in the human brain. *NeuroImage, 59*(2), 1924-1931.

Bode, S., & Haynes, J.-D. (2009). Decoding sequential stages of task preparation in the human brain. *Neuroimage, 45*(2), 606-613.

Bode, S., He, A. H., Soon, C. S., Trampel, R., Turner, R., & Haynes, J.-D. (2011). Tracking the unconscious generation of free decisions using ultra-high field fMRI. *PLoS One, 6*(6), e21612.

Bode, S., Murawski, C., Soon, C. S., Bode, P., Stahl, J., & Smith, P. L. (2014). Demystifying “free will”: The role of contextual information and evidence accumulation for predictive brain activity. *Neurosci Biobehav Rev, 47*, 636-645.

Boorman, E. D., Behrens, T. E., Woolrich, M. W., & Rushworth, M. F. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron, 62*(5), 733-743.

Boyd, L. A., Vidoni, E. D., Siengsukon, C. F., & Wessel, B. D. (2009). Manipulating time-to-plan alters patterns of brain activation during the Fitts’ task. *Exp Brain*

Res, 194(4), 527-539.

Brass, M., & Haggard, P. (2007). To do or not to do: the neural signature of self-control. *J Neurosci, 27*(34), 9141-9145.

Brass, M., & Haggard, P. (2008). The What, When, Whether model of intentional action. *Neuroscientist, 14*(4), 319-325.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues.* Leipzig, Germany: Johann Ambrosius Barth.

Breitmeyer, B. G. (1985). Problems with the psychophysics of intention. *Behav Brain Sci, 8*, 539-540.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann N Y Acad Sci, 1124*, 1-38.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends Cogn Sci, 11*(2), 49-57.

Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci, 10* (3), 186-198.

Bunge, S. A. (2004). How we use rules to select actions: a review of evidence from cognitive neuroscience. *Cogn Affect Behav Neurosci, 4*(4), 564-679.

Bunge, S. A., & Wallis, J. D. (2007). *Neuroscience of rule-guided behavior.* New York, USA: Oxford University Press.

Burgess, P. W., Dumontheil, I., & Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends Cogn Sci, 11*(7), 290-298.

Burgess, P. W., Quayle, A., & Frith, C. D. (2001). Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia, 39*(6), 545-555.

Burgess, P. W., Scott, S. K., & Frith, C. D. (2003). The role of the rostral frontal cortex (area 10) in prospective memory: a lateral versus medial dissociation. *Neuropsychologia, 41*(8), 906-918.

Burgess, P. W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia, 38*(6), 848-863.

Calhoun, V. D., Adali, T., Pearlson, G. D., & Pekar, J. J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Hum Brain Mapp, 14*(3), 140-151.

Carlson, T. A., Schrater, P., & He, S. (2003). Patterns of activity in the categorical representations of objects. *J Cogn Neurosci, 15*(5), 704-717.

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain, 129*(3), 564-583.

Christoff, K., & Gabrieli, J. D. (2000). The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organisation within the human prefrontal cortex. *Psychobiology, 28*(2),168-186.

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K.J., & Gabrieli, J. D. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *Neuroimage, 14*(5), 1136-1149.

Constantinidis, C., Franowicz, M. N., & Goldman-Rakic, P. S. (2001). Coding specificity in cortical microcircuits: a multiple-electrode analysis of primate prefrontal cortex. *J Neurosci, 21*(10), 3646-3655.

Colas, J. T., & Hsieh, P. J. (2013). Pre-existing brain states predict aesthetic judgments. *Hum Brain Mapp*, *35*(7), 2924-2934.

Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) “brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage*, *19*(2 Pt 1), 261-270.

Cunnington, R., Windischberger, C., Deecke, L., & Moser, E. (2002). The preparation and execution of self-initiated and externally-triggered movement: a study of event-related fMRI. *Neuroimage, 15*(2), 373-385.

Cunnington, R., Windischberger, C., Deecke, L., & Moser, E. (2003). The preparation and readiness for voluntary movement: a high-field event-related fMRI study of the Bereitschafts-BOLD response. *Neuroimage, 20*(1), 404-412.

Cunnington, R., Windischberger, C., & Moser, E. (2005). Premovement activity of the pre-supplementary motor area and the readiness for action: studies of time-resolved event-related functional MRI. *Hum Mov Sci, 24*(5-6), 644-656.

Custers, R., & Aarts, H. (2005). Positive affect as implicit motivator: on the nonconscious operation of behavioral goals. *J Personal Soc Psychol, 89*(2), 129-142.

Custers, R., & Aarts, H. (2010). The unconscious will: how the pursuit of goals operates outside of conscious awareness. *Science, 329*(5987), 47-50.

Danto, A. C. (1985). Consciousness and motor control. *Behav Brain Sci, 8*, 540-541.

Debaere, F., Wenderoth, N., Sunaert, S., Van Hecke, P., & Swinnen, S. P. (2003). Internal vs external generation of movements: differential neural pathways involved in bimanual coordination performed in the presence or absence of augmented visual feedback. *Neuroimage, 19*(3), 764-776.

Deecke, L., Lang, W., Heller, H. J., Hufnagl, M., & Kornhuber, H. H. (1987). Bereitschaftspotential in patients with unilateral lesions in supplementary motor area. *J Neurol Neurosurg Psychiatry, 50*(11), 1430-1434.

Deecke, L., Weinberg, H., & Brickett, P. (1982). Magnetic fields of the human brain accompanying voluntary movement: Bereitschaftsmagnetfeld. *Exp Brain Res, 48*(1), 144-148.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*(2), 384-398.

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-

Lambertz, G., van de Moortele, P. F., & Le Bihan, D. (1998). Imaging unconscious priming. *Nature, 395*(6702), 597-600.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cogn Neuropsychol, 20*(3), 487-506.

Deiber, M. P., Honda, M., Ibanez, V., Sadato, N., & Hallett, M. (1999). Mesial motor areas in self-initiated versus externally triggered movements examined with fMRI: effect of movement type and rate. *J Neurophysiol, 81*(6), 3065-3077.

Deiber, M. P., Passingham, R. E., Colebatch, J. G., Friston, K. J., Nixon, P. D. & Frackowiak, R. S. (1991). Cortical areas and the selection of movement: a study withpositron emission tomography. *Exp Brain Res, 84*(2), 393-402.

den Ouden, H. E., Frith, U., Frith, C., & Blakemore, S. J. (2005). Thinking about intentions. *Neuroimage, 28*(4), 787-796.

Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., & Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science, 324*(5928), 811-813.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & van Baaren, R. B. (2006). On making the right choice: The deliberation-without-attention effect. *Science, 311*(5763), 1005-1007.

Dijksterhuis, A., & Aarts, H. (2010). Goals, attention, and (un)consciousness. *Annu Rev Psychol, 61*, 467-490.

Doty, R. W. (1985). The time course of conscious processing: Vetoes by the uninformed? *Behav Brain Sci, 8*, 541-542.

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nat Rev Neurosci, 2*(11), 820-829.

Eccles, J. C. (1982). The initiation of voluntary movements by the supplementary motor area. *Arch Psychiatr Nervenkr 231*(5), 423-441.

Eichele, T., Debener, S., Calhoun, V. D., Specht, K., Engel, A. K., Hugdahl, K., . . .

Ullsperger, M. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *Proc Natl Acad Sci U S A, 105*(16), 6173-6178.

Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science*, *14*(6), 286-290.

Einstein, G. O., McDaniel, M. A., Thomas, R., Mayfield, S., Shank, H., Morrisette, N., & Breneiser, J. (2005). Multiple processes in prospective memory retrieval: factors determining monitoring versus spontaneous retrieval. *J Exp Psychol Gen*, *134*(3), 327-342.

Filevich, E., Kühn, S., & Haggard, P. (2013). There is no free won't: Antecedent brain activity predicts decisions to inhibit. *PLoS One 8*(2), e53053.

Forstmann, B. U., Brass, M., Koch, I., & von Cramon, D. Y. (2006). Voluntary selection of task sets revealed by functional magnetic resonance imaging. *J Cogn Neurosci, 18*(3), 388-398.

Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci, 8*(9), 700-711.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci U S A, 102*(27), 9673-9678.

Fried, I., Katz, A., McCarthy, G., Sass, K. J., Williamson, P., Spencer, S. S., & Spencer, D. D. (1991). Functional organization of human supplementary motor cortex studied by electrical stimulation. *J Neurosci, 11*(11), 3656-3666.

Fried, I., Mukamel, R., & Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron, 69*(3), 548-562.

Frith, C. D., Friston, K., Liddle, P. F., & Frackowiak, R. S. (1991). Willed action andthe prefrontal cortex in man: a study with PET. *Proc Biol Sci, 244*(1311), 241-246.

Gilden, L., Vaughan, H. G. Jr., & Costa, L. D. (1966). Summated human EEG potentials with voluntarymovement. *Electroencephalogr Clin Neurophysiol, 20*(5), 433-438.

Gilbert, S. J. (2011). Decoding the content of delayed intentions. *J Neurosci, 31*(8), 2888-2894.

Glynn, I. M. (1990). Consciousness and time. *Nature, 348*(6301), 477-479.

Gomes, G. (2010). Preparing to move and deciding not to move. *Conscious Cogn, 19*(1), 457-459.

Grabner, R. H., Ansari, D., Koschutnig, K., Reishofer, G., Ebner, F., & Neuper, C. (2009). To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia, 47*(2), 604-608.

Greenwald, A. G., Klinger, M. R., & Schuh, E. S. (1995). Activation by marginally perceptible ('subliminal') stimuli: dissociation of unconscious from conscious cognition. *J Exp Psychol Gen, 124*(1), 22-42.

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc Natl Acad Sci U S A, 100*(1), 253-258.

Groll-Knapp, E., Ganglberger, J. A., & Haider, M. (1977). Voluntary movement-related slow potentials in cortex and thalamus in man. In J. E. Desmedt (Ed.), *Progress in Clinical Neurophysiology, Vol 1* (pp 164-173). Basel: Karger.

Gupta, B. (1998). *The disinterested witness. A fragment of Advaita Vedanta phenomenology.* Evanston, IL: Northwestern University Press.

Gusnard, D. A., Akbudak, E., Shulman, G. L., Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc Natl Acad Sci U S A, 98*(7), 4259-4264.

Haggard, P. (2005). Conscious intention and motor cognition. *Trends Cogn Sci, 9*(6), 290-295.

Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nat Rev Neurosci, 9*(12), 934-946.

Haggard, P., & Brass, M. (2008). The what, when, whether model of intentional action. *Neuroscientist, 14*(4), 319-325.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nat Neurosci, 5*(4), 382-385.

Haggard, P., & Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Exp Brain Res, 126*(1), 128-133.

Hampton, A. N., & O'Doherty J, P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. *Proc Natl Acad Sci U S A, 104*(4), 1377-1382.

Handwerker, D. A., & Bandettini, P. A. (2011). Hemodynamic signals not predicted? Not so: a comment on Sirotin and Das (2009). *Neuroimage, 55*(4), 1409-1412.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293*(5539), 2425-2430.

Haynes, J.-D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat Neurosci, 8*(5), 686-691.

Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nat Rev Neurosci, 7*(7), 523-534.

Haynes, J.-D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, R. E. (2007). Reading hidden intentions in the human brain. *Curr Biol, 17*(4), 323-328.

He, B. J., & Raichle, M. E. (2009). The fMRI signal, slow cortical potential and consciousness. *Trends Cogn Sci, 13*(7), 302-309.

Henson, R. (2004). Analysis of fMRI time series. In Frackowiak, R. S. J. et al. (Eds.), *Human Brain Function, Second Edition* (pp. 793-822). London: Academic.

Herron, J. E., Henson, R. N., & Rugg, M. D. (2004). Probability effects on the neural correlates of retrieval success: an fMRI study. *Neuroimage, 21*(1), 302-310.

Horton, J. C., & Adams, D. L. (2005). The cortical column: a structure without a function. *Philos Trans R Soc Lond B Biol Sci, 360*(1456), 837-862.

Hsieh, P. J., Colas, J. T., & Kanwisher, N. (2011). Pop-out without awareness unseen feature singletons capture attention only when top-down attention is available. *Psychol Sci, 22*(9), 1220-1226.

Huang, Y. F., Soon, C. S., Mullette-Gillman, O. A., & Hsieh, P. J. (2014). Pre-existing brain states predict risky choices. *NeuroImage*, *101*, 466-472.

Huang, Y. F., Tan, E. G. F., Soon, C. S., & Hsieh, P. J. (2014). Unconscious cues bias first saccades in a free-saccade task. *Conscious and Cogn*, *29*, 48-55.

Hume, D. (1748). *An Enquiry Concerning Human Understanding*. Oxford: Clarendon Press, 1978.

Jacobs, B., Schall, M., Prather, M., Kapler, E., Driscoll, L., Baca, S., Jacobs, J., Ford, K., Wainwright, M., & Treml, M. (2001). Regional dendritic and spine variation in human cerebral cortex: a quantitative golgi study. *Cereb Cortex, 11*(6), 558-571.

Joordens, S., van Duijn, M., & Spalek, T. M. (2002). When timing the mind one should also mind the timing: biases in the measurement of voluntary actions. *Conscious and Cogn, 11*(2), 231-40.

Jung, R. (1985). Voluntary intention and conscious selection in complex learned action. *Behav Brain Sci, 8*, 544-545.

Jenkins, I. H., Jahanshahi, M., Jueptner, M., Passingham, R. E., & Brooks, D. J. (2000). Self-initiated versus externally triggered movements: II. The effect of movement predictability on regional cerebral blood flow. *Brain, 123*(6),1216-1228.

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat Neurosci, 10*(12), 1625-1633.

Kahnt, T., Heinzle, J., Park, S. Q., & Haynes, J.-D. (2010). The neural code of reward anticipation in human orbitofrontal cortex. *Proc Natl Acad Sci U S A, 107*(13), 6010-6015.

Kahnt, T., Heinzle, J., Park, S. Q., & Haynes, J.-D. (2011). Decoding different roles for vmPFC and dlPFC in multi-attribute decision making. *Neuroimage, 56*(2), 709-715.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nat Neurosci, 8*(5), 679-685.

Kawashima, R., Roland, P. E., & O'Sullivan, B. T. (1995). Functional anatomy of reaching and visuomotor learning: a positron emission tomography study. *Cereb Cortex, 5*(2), 111-122.

Kircher, T. T. J., Senior, C., Phillips, M. L., Benson, P. J., Bullmore, E. T., Brammer, M., et al. (2000). Towards a functional neuroanatomy of self-processing: effects of faces and words. *Cogn Brain Res, 10*(1-2) 133-144.

Kjaer, T. W., Nowak, M., & Lou, H. C. (2002). Reflective self-awareness and conscious states: PET evidence for a common midline parietofrontal core. *Neuroimage, 17*(2), 1080-1086.

Knösche, T., Praamstra, P., Stegeman, D., and Peters, M. (1996) Linear estimation discriminates midline sources and a motor cortex contribution to the readiness potential. *Electroencephalogr Clin Neurophysiol, 99*(2), 183-190.

Koechlin, E., Basso, G., Pietrini, P., Panzer, S., & Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature, 399*(6732), 148-151.

Koechlin, E., Corrado, G., Pietrini, P., & Grafman, J. (2000). Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. *Proc Natl Acad Sci U S A, 97*(13), 7651-7656.

Koechlin, E., & Hyafil, A. (2007a). Anterior prefrontal function and the limits of human decision-making. *Science*, *318*(5850), 594-598.

Kornhuber, H. H., & Deecke, L. (1965). Hirnpotentialänderungen bei

Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflüglers Arch für die gesamte Phys des Menschen u der Tiere, 284*, 1-17.

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proc Natl Acad Sci U S A, 108*(33), 13852-13857.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proc Natl Acad Sci U S A, 103*(10), 3863-3868.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the danger of double dipping. *Nat Neurosci, 12*(5), 535-540.

Kroger, J. K., Sabb, F. W., Fales, C. L., Bookheimer, S. Y., Cohen, M. S., & Holyoak, K. J. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: a parametric study of relational complexity. *Cereb Cortex, 12*(5), 477-485.

Lages, M., Boyle, S. C., & Jaworska, K. (2013). Flipping a coin in your head without monitoring outcomes? Comments on predicting free choices and a demo program. *Front Psychol*, *4*, 925. doi:10.3389/fpsyg.2013.00925.

Lages, M., & Jaworska, K. (2012). How predictable are “spontaneous decisions” and “hidden intentions”? Comparing classification results based on previous responses with multivariate pattern analysis of fMRI BOLD signals. *Front Psychol, 3*, 56, doi: 10.3389/fpsyg.2012.00056.

Larsson, J., Gulyas, B., & Roland, P. E. (1996). Cortical representation of self-paced finger movement. *Neuroreport, 7*, 463-468.

Latto, R. (1985). Consciousness as an experimental variable: Problems of definition, practice, and interpretation. *Behav Brain Sci, 8*, 545-546.

Lau, H. C., & Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *J Neurosci, 27*(21), 5805-5811.

Lau, H. C., Rogers, R. D., Haggard, P., & Passingham, R. E. (2004). Attention to intention. *Science, 303*(5661), 1208-1210.

Lau, H. C., Rogers, R. D., & Passingham, R. E. (2006). On measuring the perceived onsets of spontaneous actions. *J Neurosci, 26*(27), 7265-7271.

Lau, H. C., Rogers, R. D., & Passingham, R. E. (2007). Manipulating the experienced onset of intention after action execution*. J Cogn Neurosci, 19*(1), 81-90.

Leech, R., & Sharp, D. J. (2013). The role of the posterior cingulate cortex in cognition and disease. *Brain, 137*(1), 12-32.

LePage, M., Ghaffar, O., Nyberg, L., & Tulving, E. (2000). Prefrontal cortex and episodic memory retrieval mode. *Proc Natl Acad Sci U S A, 97*(1), 506-511.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain, 106(3)*, 623-642.

Libet, B., et al. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behav Brain Sci, 8*, 529-566.

Libet, B., Wright, E. W., Jr., & Gleason, C. A. (1982). Readiness-potentials preceding unrestricted ‘spontaneous’ vs. pre-planned voluntary acts. *Electroencephalogr and Clin Neurophysiol, 54*(3), 322-335.

Lin, Z., & Murray, S. O. (2014). Unconscious Processing of an Abstract Concept. *Psychol Sci, 25*(1), 296-298.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature, 412*(6843), 150-157.

Lou, H. C., Luber, B., Crupain, M., Keenan, J. P., Nowak, M., Kjaer, T. W., Sackeim H. A., & Lisanby, S. H. (2004). Parietal cortex and representation of the mental Self. *Proc Natl Acad Sci U S A, 101*(17), 6827-6832.

Marcel, A. J. (1983a). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cogn Psychol*, *15*(2), 197-237.

Marcel, A. J. (1983b). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cogn Psychol*, *15*(2), 238-300.

Marsh, R. L., & Hicks, J. L. (1998) Event-based prospective memory and executive control of working memory. *J Exp Psychol: Learn Mem Cogn, 24*(2), 336-349.

McCallum, W. C. (1988). Potentials related to expectancy, preparation and motor activity. In T. W. Picton (Ed.), *EEG Handbook: Vol. 3. Human event-related potentials* (pp. 427-534). Amsterdam: Elsevier.

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2007). Time discounting for primary rewards. *J Neurosci, 27*(21), 5796-5804.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306*(5695), 503-507.

McCoy, A. N., & Platt, M. L. (2005). Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat Neurosci, 8*(9), 1220-1227.

McFarland, C. P., & Glisky, E. L. (2009). Frontal lobe involvement in a task of time-based prospective memory. *Neuropsychologia 47*(7), 1660-1669.

Mele, A. R. (2009). *Effective Intentions: The Power of Conscious Will.* New York: Oxford Univ Press.

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Netw, 9*(8), 1265-1279.

Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000). Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage, 11*(6), 735-759.

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

Mitchell, T., Hutchinson, R., Niculescu, R. S., Pereira, F., Wang, X., Just, M., & Newman, S. (2004). Learning to Decode Cognitive States from Brain Images. *Machine Learning, 57*, 145-175.

Mommenejad, I., & Haynes, J.-D. (2012). Human anterior prefrontal cortex encodes the 'what' and 'when' of future intentions. *Neuroimage, 61*(1), 139-148.

Mommenejad, I., & Haynes, J.-D. (2013). Encoding of prospective tasks in the human prefrontal cortex under varying task loads. *J Neurosci, 33*(44), 17342-17349.

Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron, 36*(2), 265-284.

Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain, 120*(4), 701-722.

Mourao-Miranda, J., Friston, K. J., & Brammer, M. (2007). Dynamic discrimination analysis: a spatial-temporal SVM. *Neuroimage, 36*(1), 88-99.

Mueller, V. A., Brass, M., Waszak, F., & Prinz, W. (2007). The role of the preSMA and the rostral cingulate zone in internally selected actions. *Neuroimage, 37*(4), 1354-1361.

Müller, K. R., Mika, S., Rätsch, G., Tsuda, K., & Schölkopf, B. (2001). An introduction to kernel-based learning algorithms. *IEEE Trans Neural Netw, 12*(2), 181-201.

Mur, M., Bandettini, P. A., & Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRI – an introductory guide. *Soc Cogn Affect Neurosci, 4*(1), 101-109.

Murawski, C., Harris, P. G., Bode, S., Domínguez, D. J. F., & Egan, G. F. (2012). Led into temptation? Rewarding brand logos bias the neural encoding of economic decisions. *PLoS One 7*(3), e34155.

Nachev, P., Rees, G., Parton, A., Kennard, C., & Husain, M. (2005). Volition and conflict in human medial frontal cortex. *Curr Biol, 15*(2), 122-128.

Nachev, P., Wydell, H., O'Neill, K., Husain, M., & Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *Neuroimage, 36*(Supp 2), 55-63.

Niedermeyer, E., & Lopes da Silva, F. H. (Eds.). (1999). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields.* 4th ed. Baltimore: Williams & Wilkins.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci, 10*(9), 424-430.

Okuda, J., Fujii, T., Ohtake, H., Tsukiura, T., Tanji, K., Suzuki, K., . . . Yamadori, A. (2003). Thinking of the future and past: the roles of the frontal pole and the medial temporal lobes. *Neuroimage, 19*(4), 1369-1380.

Ongur, D., Ferry, A. T., & Price, J. L. (2003). Architectonic subdivision of the human orbital and medial prefrontal cortex. *J Comp Neurol, 460*(3), 425-449.

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature, 441*(7090), 223-226.

Passingham, R. E. (1995). *The Frontal Lobes and Voluntary Action*. Oxford: Oxford University Press.

Pearson, J. M., Heilbronner, S., R., Barack, D. L., Hayden, B. Y., & Platt, M. L. (2011). Posterior cingulate cortex: adapting behavior to a changing world. *Trends Cogn Sci, 15* (4), 143-151.

Pedersen, J. R., Johannsen, P., Bak, C. K., Kofoed, B., Saermark, K., & Gjedde, A. (1998). Origin of human motor readiness field linked to left middle frontal gyrus by MEG and PET. *Neuroimage, 8*(2), 214-220.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage, 45*(1 Suppl), S199-209.

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., & Frith, C. D. (2007). How the brain translates money into force: a neuroimaging study of subliminal motivation. *Science, 316*(5826), 904-906.

Picard, N., & Strick, P. L. (1996). Motor areas of the medial wall: a review of their location and functional activation. *Cereb Cort, 6*(3), 342-353.

Pinker, S. (1997). *How the mind works.* New York, NY: W.W. Norton & Co.

Pockett, S. (2002). On subjective back-referral and how long it takes to become conscious of a stimulus: A reinterpretation of Libet's data. *Conscious Cogn, 11*(2), 144-161.

Quian-Quiroga, R., Cui, H., & Andersen, R. A. J. (2006). Movement intention is better predicted than attention in the posterior parietal cortex. *Neurosci, 26*(13), 3615-3620.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proc Natl Acad Sci U S A, 98*(2), 676-682.

Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nat Rev Neurosci, 5*(3), 184-194.

Rey, A., Goldstein, R. M., & Perruchet, P. (2009). Does unconscious thought improve complex decision making? *Psychol Res, 73*(3), 372-379.

Rosenthal, D. M. (2008). Consciousness and its function. *Neuropsychologia, 46*(3), 829-840.

Roskies, A. L. (2010). How does neuroscience affect our conception of volition? *Annu Rev Neurosci, 33*, 109-130.

Sakai, K. (2008). Task set and prefrontal cortex. *Annu Rev Neurosci, 31*, 219-245.

Sakai, K. & Passingham, R. E. (2003). Prefrontal interactions reflect future task operations. *Nat Neurosci, 6*(1), 75-81.

Schacter, D. L., Addis, D. R. & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Rev. Neurosci. 8*(9), 657-661.

Schlaghecken, F., & Eimer, M. (2004). Masked prime stimuli can bias 'free' choices between response alternatives. *Psychon Bull & Rev, 11*(3), 463-468.

Schlegel, A., Alexander, P., Sinnott-Armstrong, W., Roskies, A., Tse, P. U., & Wheatley, T. (2013). Barking up the wrong free: readiness potentials reflect processes independent of conscious will. *Exp Brain Res, 229*(3), 329-335.

Schurger, A., Sitt, J. D., & Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proc Natl Acad Sci U S A, 109*(42), E2904-2913.

Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K., & van Hoesen, G. (2001). Prefrontal cortex in humans and apes: a comparative study of area 10. *Am J Phys Anthropol, 114*(3), 224-241.

Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., & Petersen, S. E. (1997). Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *J Cogn Neurosci, 9*, 648-663.

Sigala, N., Kusunoki, M., Nimmo-Smith, I., Gaffan, D., & Duncan, J. (2008). Hierarchical coding for sequential task events in the monkey prefrontal cortex. *Proc Natl Acad Sci U S A, 105*(33), 11969-11974.

Simons, J. S., Scholvinck, M. L., Gilbert, S. J., Frith, C. D., & Burgess, P. W. (2006). Differential components of prospective memory? Evidence from fMRI. *Neuropsychologia, 44*(8), 1388-1397.

Sirigu, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., & Posada, A. (2004). Altered awareness of voluntary action after damage to the parietal cortex. *Nat Neurosci 7*(1), 80-84.

Sirotin, Y. B., & Das, A. (2009). Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature, 457*(7228), 475-479.

Smith, K. (2011). Neuroscience vs philosophy: Taking aim at free will. *Nature, 477*(7362), 23-25.

Smith, P. L., & Ratcliff, R. (2004). Psychology and Neurobiology of simple decisions. *Trends Neurosci, 27*(3), 161-168.

Soon, C. S., Allefeld, C., Bogler, C., Heinzle, J., & Haynes, J.-D. (2014). Predictive brain signals best predict upcoming and not previous choices. *Front Psychol, 5*, 406. doi:10.3389/fpsyg.2014.00406.

Soon, C. S., Brass, M., Heinze, H. J., & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nat Neurosci, 11*(5), 543-545.

Soon, C. S., He, A. H., Bode, S., & Haynes, J.-D. (2013). Predicting free choices for abstract intentions. *Proc Natl Acad Sci U S A, 110*(15), 6217-6222.

Spence, S., & Frith, C. D. (1999). Towards a functional anatomy of volition. *J Consc Stud, 6*, 11-29.

Szpunar, K. K., Watson, J. M., & McDermott, K. B. (2007). Neural substrates of envisioning the future. *Proc Natl Acad Sci U S A, 104*(2), 642-647.

Tanji, J. (2001). Sequential organization of multiple movements: involvement of cortical motor areas. *Ann Rev Neurosci, 24*, 631-651.

Trevena, J. A., & Miller, J. (2002). Cortical movement preparation before and after a conscious decision to move. *Conscious Cogn, 11*(2), 162-190.

Tse, P. U. (2013). *The neural basis of free will: Criterial causation.* Mit Press.

Tusche, A., Bode, S., & Haynes, J.-D. (2010). Neural responses to unattended products predict later consumer choices. *J Neurosci, 30*(23), 8024-8031.

van de Grind, W. (2002). Physical, neural, and mental timing. *Conscious Cogn, 11*(2), 241-264.

Vorberg, D., Mattler, U., Heinecke, A., Schmidt, T., & Schwarzbach, J. (2003). Different time courses for visual perception and action priming. *Proc Natl Acad Sci USA, 100*(10), 6275-6280.

Vogt, B. A. & Laureys, S. (2005). Posterior cingulate, precuneal and retrosplenial cortices: cytology and components of the neural network correlates of consciousness. *Prog Brain Res, 150*, 205-217.

Wallis, J. D., & Miller, E. K. (2003). From rule to response: neuronal processes in the premotor and prefrontal cortex. *J Neurophysiol, 90*(3), 1790-1806.

Weber, B. J., & Huettel, S. A. (2008). The neural substrates of probabilistic and intertemporal decision making. *Brain Res, 1234*, 104-115.

Wegner, D.M. (2003). The mind's best trick: how we experience conscious will. *Trends Cogn Sci, 7*(2), 65-69.

Wenderoth, N., Debaere, F., Sunaert, S., & Swinnen, S. P. (2005). The role of anterior cingulate cortex and precuneus in the coordination of motor behaviour. *Eur J Neurosci, 22*(1), 235-246.

Wicker, B., Ruby, P., Royet, J. P., & Fonlupt, P. (2003). A relation between rest and the self in the brain? *Brain Res Rev, 43*(2), 224-230

Wiese, H., Stude, P., Nebel, K., de Greiff, A., Forsting, M., Diener, H. C., & Keidel, M. (2004). Movement preparation in self-initiated versus externally triggered movements: an event-related fMRI-study. *Neurosci Lett, 371*(2-3), 220-225.

Wolpert, D. M., & Miall, R. C. (1996). Forward models for physiological motor control. *Neural Netw, 9*(8), 1265-1279.

Yablo, S. (1992). Mental causation. *Philos Rev, 101*(2), 245-280.

Confirmation

I herewith declare that I have produced this paper without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This paper has not previously been presented in identical or similar form to any other German or foreign examination board.

This thesis was produced at the Max-Planck-Institut für Kognitions- und Neurowissenschaften in Leipzig under the scientific supervision of Prof. Dr. John-Dylan Haynes.

There has been no previous unsuccessful doctorate proceeding.

I recognise the doctoral regulations of the Fakultät Mathematik und Naturwissenschaften of Technischen Universität Dresden dated 23rd February 2011.

March 2016

Soon Chun Siong