University of Oxford
Department of Experimental Psychology

Ventral Prefrontal Cortex Structure and Function in Behavioural Change

Dr. Franz-Xaver Neubert

A thesis submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

St. Peter's College
Trinity Term, 2015

Table of Contents

Short Abstract

There is a considerable interest in the neural correlates of cognitive flexibility, language, valuation, credit assignment and decision-making. The ventrolateral, orbital and medial prefrontal cortex together with their long-range connections with the rest of the brain are thought to be critically involved in these cognitive processes. My thesis explores human and monkey ventrolateral, orbital and medial prefrontal cortex and their potential role in flexible adaptation and choice. In the introductory chapter I review neuroeconomic and neuro-ecological views of decision making, as well as modular versus connectionist views of brain structure and function. In the second and third chapter I investigate the connectivity of ventrolateral, orbital and medial prefrontal cortex and compare their sub-regions between humans and monkeys. Overall I report a striking degree of similarity of connectivity profiles across species even though these regions are thought to support uniquely human cognitive abilities such as language, social cognition, prospective planning and strategic decision-making. This may be taken to suggest that higher order cognitive functions “re-use” a neural apparatus that is shared with macaque monkeys. In the fourth chapter I present a parcellation of the white matter into the major long-range association fibre systems based on their projection patterns to the cortical surface. These findings may have implications not only for the cross-species comparison of connectivity-profiles but also for understanding some psychiatric and neurological disorders as “disconnection syndromes”. I conclude by evaluating whether a modular view of brain structure and function is consistent with a view that portraits the brain as changeable, highly adaptive and strongly interconnected.

Acknowledgements

I wish to thank my supervisors Matthew Rushworth and Rogier Mars, my close collaborators Vanessa Johnen, Jerome Sallet, MaryAnn Noonan, Urs Schuffelgen and Lennart Verhagen, and my colleagues Georgios Papageorgiou, Bolton Chau, Marco Wittmann, Miriam Klein-Flugge, , Alessandro Bongioanni, Jill O'Reilly, Lev Tankelevitch, Natalie Nelissen and Rhyanna Dale for support, advise and collegiality.

Sincerest thanks go to the Christopher Welch Scholarship and the Gustav Born Graduate Scholarship for their financial support. Finally I wish to thank my partner, parents and family and a great number of good friends for their help and encouragement.

Long Abstract

This research is particularly interested in the neural systems that support behavioural change, valuation and decision-making. These cognitive functions are often thought to be localised in the frontal lobes. The goal of this thesis is to investigate frontal lobe areas and their connectivity profiles. Moreover it attempts to study long-range associative connections between frontal-lobe areas and the rest of the brain in more detail. Connections are structurally relevant since all cortical regions have a distinct set of connections – they can therefore be used to dissociate regions from one another and match areas between individuals or even across species. Connections are also functionally relevant because they determine the information that a brain region has access to and the influence it can exert. Finally connections may be relevant as foci of pathological change – damage that spares cortical regions but disrupts connections between brain areas may lead to distinct "disconnection" syndromes.

The introductory chapter has two parts: I firstly discuss the theory of decision-making and the neuroscience of choice with reference to two alleged recent "turns" in decision neuroscience. I argue that a neuro-economical turn initiated the neuro-scientific study of decision-making not as a reflex-like stimulus-response mechanism, but as relying on complex representations of the structure of the world (including specific aspects such as expected value, risk, ambiguity). The neuro-ecological turn on the other hand proposed the existence of several different decision-making systems most likely working in parallel. Each of these parallel systems evolved during the course of phylogeny and adapted during the course of ontogeny and their specific mode of action may reflect this evolutionary and life-time history. I finish by suggesting that inter-individual variability in choice behaviour and neural representation can partly be understood as arising from adaptation to environmental variables. This may lead to a different understanding of some psychiatric disorders – rather than arising from "broken" brains they

might be the result of (mal)adaptation to potentially extreme individual histories and environments. Secondly I contrast modular and connectionist views of brain structure, function and dysfunction. I suggest that connections and interactions of brain regions may be critically involved in achieving complex cognitive functions and that many intricate computations are distributed in networks of neurons and even brain regions. I specifically stress the importance of long-range connections between "higher order" brain regions for the integration of information.

The second chapter turns to the ventrolateral frontal cortex (vlFC). The vlFC is identified with cognitive processes pre-eminent in humans such as language and cognitive flexibility. Broca's area is a part of the left-hemispheric vlFC and is associated with language, while other vlFC areas, sometimes in the right hemisphere, have been linked to cognitive control – high-level top-down control of behavior – by influencing processing in other brain regions. In this chapter I use diffusion-weighted magnetic resonance imaging (DW-MRI) parcellation techniques to identify sub-regions in human vlFC based on their connectivity with the rest of the brain and attempt to test for the potential existence of new areas supporting "uniquely" human functions such as language or cognitive control. I also use resting-state functional MRI (fMRI) to explore whether there are inter-species differences in the functional networks vlFC participates in. Such inter-regional coupling-patterns often partly reflect anatomical connectivity. I report twelve sub-regions in human vlFC. All but one could be matched to regions in monkey vlFC based on strikingly similar patterns of functional connectivity. However human lateral frontopolar cortex did not match in functional connectivity to monkey frontopolar cortex (area 10), but if anything was more similar to area 46. Moreover human vlFC coupled more strongly with posterior superior temporal cortex compared to monkey vlFC.

In the second chapter I turn to medial prefrontal and orbital regions implicated in reward-guided learning and decision-making. Because of the interest in neural mechanisms of valuation, credit-assignment and choice they have been studied in both humans and monkeys

but whether and how the key areas correspond between the two species has been uncertain. I compare a set of human regions identified in decision-making and learning tasks to regions in the monkey frontal lobe as a function of the brain circuits in which they participate. Similarly I match monkey regions from neurophysiological recording studies in reward and choice tasks to areas in human frontal cortex based on functional connectivity profiles. Altogether I identify fundamental similarities between species in the connectivity profile of these decision-making and learning areas. These findings suggest that everyday human decision-making capitalizes on a neural apparatus similar to the one that supports monkeys when foraging in the wild.

The third chapter attempts to identify the major white-matter association fibre systems. Long-range association fibres are functionally relevant. They constitute long-range connections between higher-order association cortices and may thus be relevant for information integration across different sensory and cognitive domains. Moreover they have been implicated in a number of neurological and psychiatric syndromes called “disconnection” syndromes. I demonstrate a novel data-driven method of dividing the white matter into its major association fibre systems in both the human and the macaque monkey. I show that this can be used in order to link the topography of white matter association fibres with their respective pattern of cortical projection. I also suggest that this can be used to compare the skeleton of connections and the cortical distribution of their projections across a wider range of species in order to test hypotheses about brain evolution. Finally I demonstrate a new way of assessing potential disconnection syndromes or hodological pathologies more generally.

In the general discussion of the thesis I reconsider some of the ideas presented in the introduction. I evaluate whether a modular view of brain structure and function – with clearly discernible sub-regions (as implied by the connectivity-based parcellations) – is consistent with a view that portraits the brain as changeable, highly adaptive and strongly interconnected.

1 General Introduction

Each experimental chapter of this thesis will have its own introduction and discussion. These are meant to introduce and discuss the specific topic or question of the respective experiments. I would like to use this general introduction to present a few motifs, themes and patterns that will underlie most of the thinking that led to this experimental work. To use Schopenhauer's words: "It is just like with the pastry: They are countless curly and colourful figures, but they are all made from the same dough" (freely translated from (1)). The goal of this general introduction is to list the ingredients of this dough.

The first major topic is that of decision-making. Section 1.1 will talk about different views on decision-making. I will mention philosophical and economical views on choice. Then I will explain how neuro-economics tried to reconcile these ideas with knowledge about the brain's structure and function. I will try to argue that knowledge about decision-making can be gained by looking at how the process is implemented in the biological tissue.

The second major topic is that of modular vs. connectionist views of brain structure and function. I will highlight a few prominent historical perspectives on this debate. I will argue that the way we think about the brain and this particular module-network dichotomy largely depends on the methods that we use to investigate brain function. To this end I will therefore briefly introduce some major methodological approaches in cognitive neuroscience.

I would like to stress here that I am not suggesting that these themes and topics are the only or even the most important ones to be considered when interpreting the results presented here. They are just what guided my curiosity when conducting this work and form corner stones for my own thinking about it.

1.1 Decision Theory and Neuroeconomics

"Don't ask the barber if you need a haircut—and don't ask an academic if what he does is relevant," says Nassim Nicholas Taleb in his 2007 book, "The Black Swan" (2). When writing

about one's research, for example to summarise a couple of years of scientific work in form of a D.Phil. thesis, it is particularly elegant if one can link the area of research in one way or another to questions of general relevance to society. In the case of neuroscience and psychology this is easily achieved by invoking medical relevance – for diagnosis or therapy of neurological and psychiatric disorders. Alternatively, one might suggest implications for theory and practice of education: how best to teach children, how to implement cognitive training or how to promote rehabilitation in adults. Finally, one might choose to imply "philosophical" importance by alluding to the question of "what makes us who we are?" However, here I argue that decision neuroscience and so-called "neuro-economics" can even be regarded to be of "political" relevance.

Let us consider the following example: The "war on drugs" has been fiercely fought throughout the world. The term "war on drugs" was coined by United States President Richard Nixon in 1972, who declared drug abuse the "public enemy number one" in a special message to Congress on "Drug Abuse Prevention and Control". The Drug Policy Alliance, a New York City based advocacy group, estimates that the United States still spends $51 billion annually on the war on Drugs. Some of it is spent on military aid to other countries in order to combat guerrilla groups that are involved in illegal drug trade (3).

In most Western countries this "war on drugs" has somewhat eased in the last decade: Several states in the US, including Washington D.C., have legalised cannabis. Much of Europe has decriminalised marihuana. Heroin addicts in Western countries usually have access to clean needles, substitutes such as methadone and, in parts of Europe, heroin prescriptions. Many Western governments are starting to believe that managing drug use causes less harm than trying to battle it.

However, in much of the rest of the world things look rather different. Many countries across Asia and the Middle East routinely execute drug offenders. Indonesia has just executed eight drug traffickers in April 2015, despite lobbying of other governments and the UN. Saudi Arabia

beheads smugglers of cannabis. China's president, Xi Jinping recently promised "forceful measures to wipe out [drugs]". In Iran drug offenders account for a majority of all those put to death (540 in 2011). Russia has adopted a hard-line role against drugs. It is advocating spraying of opium-poppy fields in Afghanistan to annihilate crops, and trying to convince other Eastern European countries to follow it banning methadone. Politicians take up such hard lines on drug offenders because they believe that cracking down on drugs will make addiction go away. However, one might wonder if this is a realistic belief. Can cognitive neuroscience help us find an answer to this question of whether addicts are patients or criminals? In other words could decision neuroscience provide insights into drug-related behaviour in a way that would test some of the assumptions behind drug policy? For example a politician might be worried that the amount of excess speeding on motorways adversely affects their safety. She hopes that increasing the cost for speeding through speed cameras will bias the "decision" whether to speed or not in a way so that people refrain from speeding. This policy contains certain assumptions about the "decision" to speed (e.g. that it can be affected by costs). It is apparent that insights into drug-related behaviours as provided by decision neuroscience should inform assumptions behind drug policies and therefore potentially influence the policies themselves.

A considerable number of recent neuroscientific studies have aimed to look at the neural correlates of drug addiction. A majority of them seems to support the notion that drug addicts need treatment instead of prison. Elucidating some of the neural underpinnings of addiction suggest that addiction is a disease that must be treated, rather than a condition for which addicts can be blamed. Neuroscientists have identified a range of neural mechanisms and neuroadaptations that are correlated with and quite possibly causally involved in addiction. They include long-term depression of the reward circuitry and increased activity in the "anti-reward circuitry" (4, 5), which are argued to explain the persistent changes in motivation that are associated with vulnerability to relapse in addicts. The enhancement of the anti-reward circuitry during the addiction process is attributed to changes in the amygdala and the basal

forebrain and linked to neurotransmitters such as norepinephrine, dynorphin, neuropeptide Y, and nociceptin (5). Depression of the reward-circuitry has been linked to changes in the dopamine system in the substantia nigra and ventral tegmental area and its projection zones in the striatum (particularly ventral striatum) and frontal cortex (particularly orbitofrontal cortex [OFC] and anterior cingulate cortex [ACC]) (6) (see Figure 1-1).

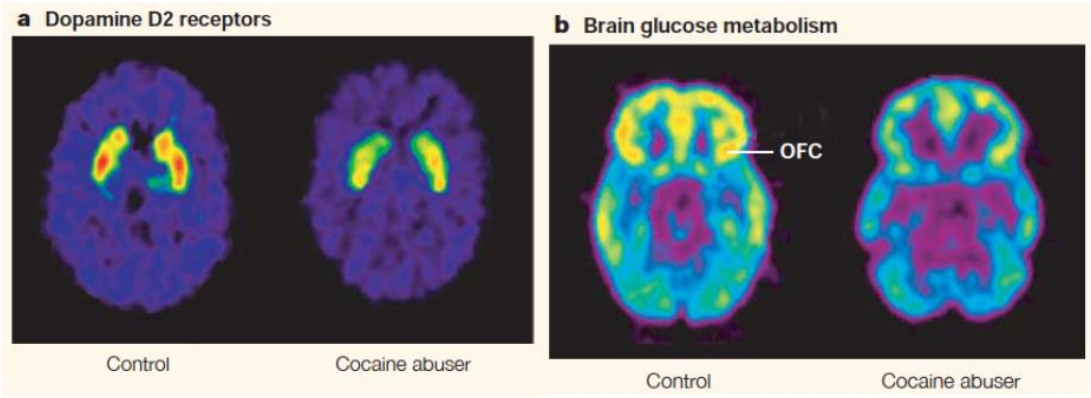


Figure 1-1: Positron emission tomography (PET) shows decreased dopamine D2 receptor density in striatum and reduced brain glucose metabolism in OFC.

Dopamine has long been known to play a key part in our so-called “reward prediction system”. It allows us to learn the value of a reward and the relationship between stimuli and rewards. Dopamine neurons in the ventral tegmental area fire strongly if an outcome is better than expected and are strongly inhibited when outcomes “disappoint” our expectations (7). Crucially, when reward can be expected based on predictive stimuli, dopamine neurons do not fire when the reward is finally obtained.

The enhanced reinforcing effect of drugs is due to their ability to trigger dopaminergic firing, exceeding in magnitude and duration the dopamine increases that occur with natural reinforces such as food and sex. Cocaine, amphetamine, methamphetamine and ecstasy increase dopamine by inhibiting dopamine reuptake or promoting dopamine release through their effects on dopamine transporters. Other drugs, such as nicotine, alcohol, opiates and marijuana, work indirectly by stimulating neurons that modulate dopamine cell firing. The

manner in which drugs “hijack” the mesolimbic dopamine system in order to increase the dopamine signal is widely thought to be central to the explanation of how addiction develops and why it is a chronic, relapsing condition (6).

Notably some hypotheses of addiction do not even claim that it is a “brain disease” in the narrow sense (8). A dopamine system, operating in exactly the manner it should, would attenuate dopamine activation in response to expected reward. Eating a surprisingly delicious chocolate bar for the first time would trigger a bigger dopaminergic reward prediction error compared to eating it for the fourth or fifth time – when its delicious taste is to be expected. The dopamine-reward prediction error hypothesis holds that if the world is exactly as expected, there ought to be no dopamine response. If drugs worked like chocolate bars, we could expect them to trigger an initial dopamine response to consumption, but an attenuation of this response as consumption is repeated. In this case, the dopamine response would be moved to predictors of drug consumption rather than consumption itself. However, what we do find is that there are two dopamine responses in sequence: one response to signals of drug availability and a second response following consumption of the drug whereby the chemical action directly triggers the release of dopamine. “Pathological learning” leads to the drug being of ever increasing value. This account does not need to assume brain pathology in a narrow sense, but would suggest that addiction arises from a brain living in a world to which it had not adapted to (chemical substances that are able to directly release dopamine did not play a role in the evolution of the dopamine system).

This last hypothesis might be promising on a more general level. We might consider understanding some psychiatric disorders, such as addiction or impulsivity, to be emerging from an “ecologically valid“ adaptation of a person (and their brain) to their specific environment. That means rather than being the result of a dysfunctional and mis-wired brain, psychiatric disorders might result from a functional brain which has adapted to a (potentially extreme) environment which then in turn causes dysfunctional behaviour. Such a framework

goes beyond a resilience–vulnerability dichotomy but suggests that maladaptive behaviour emerges as the result of a functional brain living in an environment that is largely different from the average one for which it was adapted.

It remains to be clarified whether addiction should be viewed as a brain disease in the narrow "neurological sense" or whether it should be viewed as a "functional disorder" and "mis-adaptation". In any case, this example illustrates that we can be hopeful that decision neuroscience will be able to decisively contribute to the understanding of the condition of addiction and, by doing so, affect political decisions and society - it would not be the first time. Psychiatric patients used to be imprisoned together with criminals until, at the end of the 18th century, (neuro-) scientific insight led to the removal of chains from patients with mental illness and the introduction of a "traitement moral".

Figure 1-2: Dr. Philippe Pinel at the Salpêtrière, 1795, by Tony Robert-Fleury. Pinel orders the removal of chains from psychiatric patients at the Asylum for mentally-ill women in Paris.

I will introduce the field of decision theory and mention some major findings from decision neuroscience. I will ask the question whether any of these neuroscientific insights have the potential to inform basic decision theory or further our understanding of microeconomics. In the next section, which describes a "neuroethological" turn in decision neuroscience, I will try to come back to some of the above-mentioned interactions between a specific context and a decision making apparatus shaped by a history of past environments (see particularly Section 1.1.5.4).

1.1.1 Decision Theory

Decision theory is the interdisciplinary study of rational decision-making, drawing on knowledge in economics, psychology, philosophy, mathematics, computer science and statistics (9). For example, in 1492 Columbus decided to sail westwards in order to find a new route to the Far East. His geographical theory predicted that the earth was a much smaller sphere and that the route to East India measured about 2,300 miles. However, he was not sure about his distance calculations (and as it turned out India was actually more than 12,000 miles away, and no ship of the 15th century could have carried provisions for such a long journey). This uncertainty means that other potential states of the world needed consideration – for example, that there was a landmass between Europe and India or that there was just a gigantic ocean.

Decision theory is interested in studying exactly those decisions. A decision maker (here Columbus) chooses from a set of alternatives (stay at home or sail westwards), which – due to the unknown state of the world - lead to different outcomes (see decision matrix in Table 1). If India is 2,300 miles away, sailing westwards will yield wealth and fame. If India is >12,000 miles away, but America is in between India and Europe, his sailing westwards will yield fame but not wealth. Finally, if India is >12,000 miles away and America does not exist, sailing westwards will result in certain death. Staying at home would have the same outcome regardless of the location of India.

Table 1

	Geographical hypothesis is true (India is 2,300 miles away)	Geographical hypothesis is not true but there is landmass between Spain and India	Geographical hypothesis is not true and there is no landmass between Spain and India
Sail westwards	Rich and famous	Famous but not rich	Certain death
Stay at home	Status quo	Status quo	Status quo

The goal of decision theory is to analyse such decisions in terms of their outcome's values and uncertainties and formulate precise and useful theories about rational and optimal choice in these situations. It is often dissociated between normative and descriptive decision theory (9). Descriptive decision theory tries to study how people actually make decisions. It is an empirical discipline, largely stemming from experimental psychology. Normative decision theory on the other hand seeks to explain what a rational decision maker ought to do.

In decision theory this second "normative view" is often emphasized. This might be because of several reasons: It is thought that normative decision theory is of greater philosophical interest since it will tell us what makes a rational decision rational. Actual decision-making is likely to change over time and across cultures, but a sufficiently general normative theory of rational choice is expected to withstand time and cultural differences. Another argument in favour of studying decisions normatively is a pragmatic one: Humans (and other animals) behave rationally, at least most of the time. Without a normative theory of rational and consistent choice it would be hard to describe occasional "deviations" from rationality or advise agents on "better" decision strategies. A common theme of both, descriptive and normative decision theories, is the idea – proposed by David Hume (10) – that decisions are somehow triggered by an agent's beliefs and desires. Descriptive decision theory has tried to devise mathematical models on how exactly beliefs and desires would trigger choices. Normative decision theory

has tried to propose frameworks that describe how beliefs and desires ought to be combined to yield rational decisions.

"Rational" in that context refers to "instrumental rationality", which can be thought of as the most efficient or cost-effective means to achieve a specific end, but not in itself reflecting on the value of that end. An agent chooses "rationally in the instrumental sense" if she does what she has most reason to do in order to achieve her aims. The aims themselves are outside decision theory and cannot be irrational. However, a set of aims can be irrational, for example, if they are mutually inconsistent. Instrumental rationality has been criticised and ways of deriving aims and goals in rational ways have been proposed (11-13). Interestingly psychologists have put forward a similar criticism of instrumentally rational decision theory noting that rational decisions cannot be studied independent of their aims (14, 15).

Different types of decisions are studied in decision theory. Decisions under risk denote decisions with known probabilities for all possible outcomes, whereas in decisions under ignorance these outcome probabilities are not known. Note that the term risk is sometimes taken to mean high outcome variance, e.g. outcomes which manifest in 50% of the cases. Take for example a risky gamble of winning £200 or losing £100 with 50% probability each or a certain £10 win. The first option would be considered more risky than the second. Sometimes the term "risk" is used to refer to high probabilities of adverse events (e.g. high risk of getting cancer). I here refer to "risk" as probabilistic outcomes with known likelihoods, but I will come back to the other aspect of "risk" in Section 1.1.5.2.

Decisions under ignorance. Although decisions under ignorance are based on less information than decisions under risk, they are not necessarily more difficult. Several principles have been proposed for rational decisions under ignorance (9). The *principle of dominance* states that an option should not be chosen if the agent knows that she will be at least as well off if she chooses an alternative regardless of what turns out to be the true state of the world. If

Columbus could have excluded the possibility that there is just a huge ocean westward of the Canary Islands, he might have thought that the decision to sail westwards would actually "dominate" the decision of staying at home. This is because in any of the two remaining states of the world (India or another landmass lying westwards in a distance of less than 3,000 miles) Columbus might have preferred the outcome of sailing westwards (becoming rich and famous in the case of India and becoming famous in the America case) to the status quo at home.

Another principle for rationally choosing under ignorance is the *"maximin" principle*, which states that an agent should maximize the minimal value obtainable with each act. If the worst possible outcome of one act is better than the worst possible outcome of another, then this act is preferable. In Columbus' case this would clearly be "staying at home" because the worst outcome of this decision is clearly better than the worst outcome of sailing away (=death, in the event that only a big ocean was found westwards). The lexical maximin rule ("*leximin*") states that if the worst outcomes are equal for two options, then one should consider the second worst outcome (and if they are equal, the third worst etc.). One strong argument in favour of maximin and leximin is that they transform a decision under ignorance into a decision under some form of certainty. An agent following these rules knows with certainty at least what worst possible outcome she has to expect.

The "*maximax*" rule holds that agents should try to maximize the maximal value obtainable with an act. Following this principle Columbus should have definitely sailed westwards since only that choice contains the possibility of getting rich and famous.

The optimism-pessimism rule requires the decision maker to consider both the best and the worst possible outcome for each choice and to then weigh these according to their "optimism-pessimism bias" (e.g. some value between 0 and 1). This can be demonstrated by assigning specific interval scaled values to all outcomes (see Table 2). If Columbus' had a 0.7-optimism bias, he would for the case of sailing westwards multiply 0.7 with the value of becoming rich and famous (0.7*200 = 140) and 0.3 with the value of dying (0.3*(-200) = -75) and compare the

result (65) to the value of the status quo at home (0). As a result (65>0) he would decide to sail westwards. A different optimism-pessimism bias could yield different choices.

Table 2

	Geographical hypothesis is true (India is 2,300 miles away)	Geographical hypothesis is not true but there is landmass between Spain and India	Geographical hypothesis is not true and there is no landmass between Spain and India
Sail westwards	Rich and famous [200]	Famous but not rich [100]	Certain death [-250]
Stay at home	Status quo [0]	Status quo [0]	Status quo [0]

The "*minimax regret*" principle suggests that the best alternative would be one, which minimizes the maximum amount of regret. Regret for a given decision and a certain state of the world is the difference between the value of the given outcome and the value of the best possible outcome. The minimax regret rule chooses the alternative under which the maximum regret is minimized. Table 3 denotes the regret for each choice and state of world. If Columbus was to minimize maximum regret, he should stay at home because, although he will badly regret having stayed at home – in the event that India lies westwards of Spain (regret = 0 – 200 = -200), he will regret having left even more – in the event that there is no land to the West (regret = -250 – 0 = -250). A prominent objection against the minimax regret rule is that it can be shown that new irrelevant alternatives (alternatives that are themselves not worth being considered seriously) can still change the "regret table" and thereby the choice according to the minimax regret rule. The fact that non-optimal alternatives should affect decision-making seems very counterintuitive to many researchers. However, others have argued that this objection is not as strong as it may seem: In some situations it might be rational to compare specific options to the entire set of alternatives, even if some of these alternatives will certainly not be chosen. By adding a new alternative one may find the relational properties of the initial objects to be changed, and sometimes our preferences are

based on relational properties (e.g. the uniqueness of a given object which might disappear if a similar, though unappealing object, is added to the choice set).

Table 3

	Geographical hypothesis is true (India is 2,300 miles away)	Geographical hypothesis is not true but there is landmass between Spain and India	Geographical hypothesis is not true and there is no landmass between Spain and India
Sail westwards	Regret = 0	Regret = 0	Regret = -250
Stay at home	Regret = -200	Regret = -100	Regret = 0

Finally, the principle of *insufficient reasoning* suggests that if the probability for states of the world is not known (such as in decision making under ignorance) the probability of each of the *n* states should be taken to be 1/n. This rule basically transforms a decision under ignorance into a decision under risk. In Columbus' case the three different states of the world would all have a probability of 1/3. Applying the ideas of expected value (EV [sail westwards] = 1/3 * 200 + 1/3 * 100 + 1/3 * -250 = 16.67) Columbus would decide to sail westwards (because 16.67 > 0). However, this rule is often criticised. One obvious problem is its susceptibility to how states are individuated. For example, if we were to redraw Columbus' decision matrix to now only include two possible states of the world – (1) there is land in reach westwards, (2) there is no land in reach westwards – the probability of death when sailing westwards would suddenly increase (from 1/3 to 1/2) and the expected value of sailing away would drop below that of staying at home. Suddenly, Columbus would decide to stay in Spain.

One might wonder now whether some of the proposed rules for rational choice under ignorance are actually quite similar to some heuristics that people use when making decisions. It is possible that humans actually try to minimize potential losses or regret in some cases, or maximize potential gains in other situations. It is not implausible to think that they would explicitly or implicitly combine extreme gains and losses and weigh them with their optimism bias or that they might think that probabilities of n potential states of the world might each be

somewhat close to 1/n. It would be “desirable” for humans to use such rules as heuristics when making decisions under ignorance since these rules are suggested to yield at least somewhat rational choices under these circumstances. However one might speculate that the use of such and other heuristics might even extend to decision-making under risk since sometimes probabilities are difficult to handle or compute and therefore the expected value approach might be too time-consuming or just not cost-efficient. It certainly is the case that even decisions under risk are affected by irrelevant options, and I will mention this again in one of the following sections (16-18). Moreover, choices are influenced by how they are framed and the way different states of the world are individuated (17, 18). It would be interesting to see if agents in these cases revert to these rules for rational decision making under ignorance and how this “meta-decision” (to make the decision as being basically “under ignorance” rather than “under risk”) is reached itself.

Decision-making under risk. Decision-making under risk is conceptually simpler. Whereas there is significant disagreement among decision theorists about which of the above mentioned rules should be used to achieve rational decisions under ignorance, most theorists agree that under risk agents should try to maximize value. However, it is not necessarily clear what that exactly means.

For example, it could be thought that an agent should basically try to maximize expected monetary value. However, when given the choice between a certain gain of £1,000,000 and a 50/50 gamble between either £3,000,000 and £0 most people would choose the first although the gamble would actually maximize expected monetary value. This can be reconciled with the idea that the marginal value of money decreases and that therefore the difference between £1M and £0 is much bigger than the difference between £3M and £1M. The idea of expected value takes into account the decreasing marginal value of money.

However money is not all there is in the world and with reference to the famous Oscar Wilde quote - 'Nowadays people know the price of everything and the value of nothing' – one might think it helpful to introduce the concept of utility (Wilde should have maybe said "[…] the utility of nothing"). Utility is an abstract entity that cannot directly be observed and refers to how valuable an outcome is from a decision-makers point of view. Crucially this concept can now also be applied to situations involving non-monetary outcomes. For example, the moral value of donating some of your money to a good cause might outweigh the value of the monetary loss and therefore have higher utility.

Why should people try to maximize utility in every single decision? One argument states that "on the long run" one would be better off. On the long run playing roulette loses you money. That might make it irrational or non-utility-maximizing (if you disregard non-monetary gains playing it). The law of large numbers suggests that utility-maximizing decisions are likely to make you better off if repeated an infinite amount of times. However, Keynes famously objected to this argument asserting "in the long run we are all dead". We are not likely to repeat the same decision an infinite amount of times. Particularly life decisions such as whom to marry are hard to reconcile with the law of large numbers. Moreover, it might be doubted whether an exact same decision can even be faced several times. Another set of arguments in favour of utility maximization is based on a set of basic axioms. Following four simple axioms it can be derived that the utility of an act equals the expected utility of an act (9).

The expected utility idea is not uncontroversial and some of the more sophisticated objections are put forward in the form of "paradoxes". The most famous of them are Allais' paradox, the Ellsberg paradox, the St Petersburg paradox and the two-envelope-paradox. As an example I will briefly describe the St Petersburg paradox: It was first described by Daniel Bernoulli and arises from the St Petersburg game. In that game a coin is tossed until it lands heads up, and the player receives a prize worth 2^n units of utility. For example, if the coin lands tails 4 times and only then lands heads (= 5^{th} time), the player will get 32 units of utility (2^5). The puzzle

arises now if one thinks how much a utility maximizing player should be willing to pay in order to play the game. Theoretically it should be any finite amount of utility because

$$\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n * 2^n = \infty$$

This seems absurd and theorists have long tried to address it in several different ways. One popular way of dealing with it, is to suggest we should discard highly improbable outcomes (i.e. very long trains of tails that would yield huge amounts of utility). This and other paradoxes present serious challenges to the idea of expected utility and they still await satisfactory resolution.

Figure 1-3: The Swiss mathematician and physicist described the St Petersburg paradox in his Exposition of a New Theory on the Measurement of Risk.

The decisions discussed so far were decisions of single actors. Some decisions are made by a group of people or at least need to take into account the decisions of other agents. Two other areas of decision theory are *social choice theory* and *game theory*. Social choice theory tries to establish how decisions that involve more than one decision maker ought to be made. It tries to derive a theoretical framework for the combination of individual preferences, interests, or

welfares to reach a collective decision. An example is how political leaders are elected in democratic countries. Social choice theory is able to show that the exact voting procedures in many democratic countries are quite unsatisfactory from a theoretical perspective and fail to fulfil some basic requirements how such social decision ought to be made.

Game theory is the analysis of strategic or interactive decisions, i.e. when decision outcomes depend on other decision maker's choices. Famous examples are the stag hunt or the prisoner's dilemma. Game theory can explain why humans often settle with sub-optimal outcomes when playing non-cooperative games, in which each player knows the equilibrium strategies of the other players, and no player has anything to gain by changing only his own strategy (a phenomenon referred to as *Nash-equilibrium*).

The next section will turn to the study of how these decisions might be carried out by the human brain. Although even descriptive decision theory was largely agnostic about biological implementations of these processes at first, I will try to argue that eventually decision theory could benefit from biological insights into decision making.

1.1.2 The neural correlates of decisions

The study of the human brain is largely influenced by the scientific tools we have at hand (see Section 1.2 for a detailed discussion of this). However, it is no less influenced by ideas, terminology and metaphors in use for describing their problems and findings (in some way these might be viewed as tools in a broader sense). It has been suggested that the way scientists think and talk about the human brain is influenced by engineering, inventions and technology (19). Famously Galen used the metaphor of a roman fountain when talking about the brain. Descartes thought that many processes that we take to be mental are actually physical and can be ascribed to a complex hydraulic system of fluids flowing through the nerves and the nervous system. Jacques de Vaucanson's mechanical duck inspired philosophers to envisage human behaviour as being driven by mechanical processes. Thomas

Hobbes for example proposed that ideas and associations resulted from minute mechanical motions in the head. Galvani's and Volta's pioneering research on electricity in the 1790s made scientists think of electricity as a driving force for mental processes. The first microscopic images of neurons and their axons (see Section 1.2.2 for details) fuelled telegraph and telephone metaphors. Today we often speak about mental processes using the "computer metaphor" and current neuroscientific terminology pays tribute to this idea. We speak of processing of information in different modules or latent variables computed by the brain. Some scientists think of mental processes as the application of logical or mathematical rules to symbols. This influence of technology on how we think about mental processes might not be a one-way street. It is likely that neuroscientific discoveries have constituted inspirations for engineers. They certainly inform computer science and robotics for the solution of problems that humans are very good in solving (e.g. spotting a face in the crowd). They might also shape how we think about the workings of, for example, machines (think of the "memory" metaphor with respect to e.g. computers' RAM and ROM)

The way that neuroscience until recently thought about decision making was still deeply rooted in rather simple "mechanical" concepts of brain function (20). But inspirations from economics allowed for a fundamental turn in decision neuroscience. Neurophysiologists thought about decision processes, in the tradition of Charles Scott Sherrington, as largely reflexive – the very complex but largely predetermined and "mechanical" response to a stimulus (20). As mentioned above this was not at all the way that philosophers and economists talked about decisions. In microeconomics it was thought that a choosing agent would have some internal representation of values and preferences and that these would be revealed by their choice behaviour. The idea of expected utility even extended this idea to include representations of outcome probabilities and stated that choosing agents should always prefer actions that maximize expected utility, i.e. the "best" combination of subjective values and probabilities in a set of lotteries.

Although it is apparent that economics proposes “representations” of outcome probabilities and subjective values, economists were largely agnostic about how this might actually be implemented in (human) brains. Recently a “neuroeconomic turn” started to combine neurophysiological ideas about decision processes with economical notions of choice, preferences and subjective values and their mathematical frameworks. This combination of neurophysiological models of “representation” and “processing” with economical ideas about choice, preference, subjective value, risk, ambiguity, delay discounting and other things reconciled a multitude of previous theories and findings in neurobiology and inspired a large set of exciting new questions (20).

One of the first major discoveries emerged from studies on the representation of value in the brain. In 1997 Schultz, Montague, and Dayan (7) published a collection of experiments demonstrating that the firing of dopamine neurons in the primate midbrain signals “reward prediction errors”. This was taken as evidence to support theories of reinforcement learning and suggested that the brain’s internal motivational representations might yield to quantitative economic analysis. At about the same time, Knutson and colleagues showed the first signs of robust hemodynamic activation of ventral striatum in human subjects incentivized with money (21), a non-primary reinforcer with no intrinsic relevance for survival. This suggested that the brain evaluated and encoded both primary and secondary rewards in similar ways, perhaps even in a “common currency”. In another early set of studies, Platt and Glimcher recorded neuronal activity from the lateral intraparietal area (LIP), a cortical area involved in connecting visual stimulation to attention and orienting movements of the eyes. In their experiments monkeys were freely choosing between two targets associated with fluid reward (22). Studying an agent’s choices in that way was a novel approach as previous neurophysiological experiments - largely influenced by “Sherringtonian” views of the nervous system, i.e. linking stimuli with responses - had mainly studied experimenter-determined behaviours in response to physical cues. In this experiment, the two options differed in size or

probability of reward across blocks of trials. The authors reported that the firing rates of LIP neurons representing a specific eye movement scaled with the product of reward size and reward probability, i.e. the ‘‘expected value’’ of this eye movement. Consistent with this notion, firing rates predicted monkeys’ choices. This work drew attention to the idea of utility theory being a useful tool for studying the processes by which the brain organizes action.

Crucially all of these early studies suggested that economic variables such as expected reward, subjective value or utility were in fact being represented in the human brain or at least affect how information about stimuli and responses were encoded. For example in the last set of experiments by Platt and Glimcher, monkeys were trained to choose between stimuli associated with different subjective values and probabilities. If the brain was just a mere stimulus-response mapping device, as suggested by earlier neuroscientists and biologists, then the encoding of a specific visual stimulus or a certain eye movement should not be affected by economic variables such as subjective value or probability. However, Platt and Glimcher showed that the representation of eye movements or their respective targets in LIP did exactly that: their neural responses varied largely with their utility and the monkey’s preferences.

These early discoveries inspired decision- neurophysiology and a new field called “neuroeconomics” was born. Soon researchers tried to find neural correlates for other economic variables and phenomena – such as risk, ambiguity, delayed rewards, counterfactual values etc. Moreover, they naturally cared about the exact “mechanistic” implementation of these processes – and this might eventually prove to be of interest for behavioural economics in turn (23, 24).

With this curiosity for “neural implementation” neuroeconomists started to wonder how, for example, value comparison is carried out by the human brain. In the case of choosing between eye movements, they put forward sophisticated models of evidence accumulation in LIP neurons, together with dynamically varying degrees of interaction between the topographically organized representations of the different options and adjustable thresholds

for determining choice (25).In the case of eye movements it is also becoming apparent that signatures for such value comparison processes can be found in several regions – among them not only LIP but also the frontal eye field and the superior colliculus. This was taken by some to suggest that value comparison might not be achieved by a single brain region but might rather be thought of as a network process.

The description of concrete, "mechanistic" ways of how value comparison might be implemented in the brain might allow explaining some puzzling micro-economic phenomena. One example is the "distractor effect on value comparison" (16), which is thought to be a failure of optimal decision-making arising from the presence of irrelevant and non-choosable "distractor options". Surprisingly choosing between two options in the presence of a third, very poor alternative is more difficult than in the presence of a very good alternative. Chau and colleagues (16) linked this phenomenon with brain imaging data which suggested that value difference signals in ventromedial prefrontal cortex (vmPFC) reflected this phenomenon. Furthermore, they used a biophysical model proposed by Wang et al (26) to explain how such a paradoxical effect might emerge as a result of how value comparison is neurally implemented in the brain. Simply put they argue that value comparison is carried out by different pools of neurons, which represent the value of their respective option through self-excitation and mutually inhibit the other neural pools via an additional inhibitory neural pool. A low-value option now will yield very low firing in the respective pool that represents its value. This however will decrease the overall inhibitory activity in the system, which will render comparison of two similarly valuable options more difficult.

In the early years, neuroeconomics focused on studying these relatively simple decisions. The fact that even these simple choices were not "reflex-like" but relied on the representation of "economic variables" was encouraging. However, they hardly lay at the core of human decision-making. This is why recently neuroeconomists have started to apply their approach to study more elaborate decision processes. Some studies focus on decisions with multiple

alternatives (27). Others have started to investigate decisions in competitive interactive settings (28).

During mixed-strategy interactive games (such as work-or-shirk or rock-paper-scissors) humans tend to exhibit “unpredictable” behaviour in equilibrium state, although they self-report that they behave in a volitional manner. Similar behaviour is observed in monkeys. This “deliberately unpredictable” behaviour exhibited by humans and monkeys has been linked with the small random fluctuations in the firing rates of neurons suggesting that these constitute the neural implementation of “unpredictable behaviour” in competitive games predicted by decision theorists (29, 30).

I mentioned two examples (value comparison and “unpredictable” behaviour) in which neuroeconomics – i.e. the study of the neural implementation of economical behaviour – helped to explain behavioural phenomena that may at first surprise axiomatic decision theory. Another example comes from the studies of value representation. Levy and colleagues showed on an individual subject’s level that idiosyncratic preferences relate in a direct way to how values are represented in the brain (31). When subjects were presented with certain goods, the BOLD response of a region in the medial prefrontal cortex predicted how they would choose between them later on, even though they all had different utility functions.

In a different experiment Kable and Glimcher (32) showed that individual behavioural patterns of temporal discounting were reflected in patterns of neural value coding. One subject (a MD/PhD student according to Glimcher) had a very flat discounting function and the BOLD response in medial prefrontal cortex decreased only very slowly with increasing delays. Another subject (a professional skydiver) steeply discounted temporally delayed rewards, and this behavioural pattern was reflected in the way his medial prefrontal cortex encoded rewards. These observations suggest that these preference profiles relate to more stable personality traits which in turn might cause certain career and life decisions (although this might well be a mutual interaction). These more stable personality traits and life histories

might be reflected, potentially even caused, by individual brain structure and function. This could be taken to suggest that studying the implementation of economical processes in the brain might help us to understand choice behaviour even on the individual subject's level.

It is a well-known phenomenon that despite individual differences, almost all people change risk and ambiguity preferences as well as temporal discounting behaviour over the course of the life span. It would be interesting to reveal the neural "implementation" of such changes.

Despite it seeming simplistic and somewhat outdated, the Descartian and Sherringtonian view of choice had one crucial advantage: It was very precise and mechanistic in how decision processes were implemented in the nervous system (as predictable stimulus-response connections). It therefore rendered "choice" a predictable process carried out by the brain. The neuroeconomic turn allowed the reconciliation of neurophysiological and economic views of choice. It took mathematical frameworks of economics to obtain a similarly precise and mechanistic account of how economic decisions were implemented in the brain.

So far this approach has largely proven the significance of economic theory to neuroscience. As discussed above, in some cases the understanding of how economic decision processes are implemented in the brain might ground economic theory more firmly in biology and help to explain some behavioural deviations from axiomatic decision theory. One example of such deviations might be the problem of "internal reference points". It is widely known in psychology that humans often compare offers to internal reference points and choices are largely influenced by these offsets. This behaviour is not easily reconciled with economic predictions of rational decision-making but might be explained by theories of efficient encoding of values in the cortex (25). Similarly some aspects of inter-temporal decision-making seem consistently highly irrational. Humans discount delayed rewards very steeply for short delays but less steeply for very distant time points. Moreover, they are known to often reconsider or overturn their own choices later in time. A better understanding of how reward rates and temporally distant rewards are encoded in the brain as well as insights into how

value-comparison-systems interact with “executive control” systems could resolve this puzzle. Finally, another prominent phenomenon is the “curse of choice”. Humans are known to be less-decisive and unhappier with their eventual choices when offered a bigger range of options. Understanding how value representation and comparison for multi-option decisions is implemented in the neural tissue is likely to help understand this paradoxical effect.

1.1.3 Do we need alternative frameworks in decision neuroscience and does it matter?

Cab drivers have good and bad days. Depending on weather and special events, they can either be very busy or just wait around for most of the time (33). If they were rational, they should finish work early on lazy days and work for longer on busy days. However, they do the exact opposite. As if they had a mental target for their daily income or tracked the “rate of reward” with a time-constant of roughly a day, they decide to work longer on quiet days and stop work early on busy ones.

Humans do not always behave “logical” (in the narrow economic sense). They adapt their risk-taking dynamically to contextual changes. They value things they already own more highly than things they would need to purchase. They respond to questionnaires largely depending on how questions are framed. They “prune” decision trees when making sequential decisions if early steps encounter great losses, even if the sequence would offer much bigger gains in the long run. They become more certain of their own decisions when the number of options is smaller rather than larger. They find surcharges on credit-card payments unfair, but believe that a discount for paying in cash is reasonable. They discount delayed rewards more steeply for shorter compared to longer time periods (which is often referred to as hyperbolic discounting) and therefore later on often overturn their own resolutions.

This may not be a big surprise to many psychologists. Over the last decades even behavioural economists have started to study how individuals actually make decisions rather than how they should make them. But still many macroeconomic and political models do not yet account for these apparent fallacies. At first they dismissed the results of such experiments as being trivial or the consequences of artificial laboratory experiments. Moreover, it is sometimes thought that ordinary people might not always behave optimally, but that professionals making big decisions would be entirely rational. But in his book "Moneyball" (34) Michael Lewis shows how baseball coaches' and scouts' valuation of players can be largely distorted by irrelevant information. Another example is the decision of the British and French governments to continue to fund the joint development of the "Concorde" even after it became apparent that there was no longer an economic case for the aircraft, which led people to call such types of irrational decision-making "Concorde Fallacies".

Even professionals behave irrationally and some policies are now incorporating such seemingly "irrational" human behaviour. It is becoming accepted that the way choices are offered does affect decisions, e.g. asking people to opt out of rather than opt in for organ donations increases the number of effective donors per million inhabitants (34 donors / million inhabitants in Spain and 16 donors / million inhabitants in Germany, with an opt-out and opt-in system, respectively).

All of this seems to suggest that behavioural or "real-world" microeconomic insights are becoming more accepted and appreciated in the economics community. But it is still difficult to have an impact on macroeconomic thinking. This division has been compared to that between quantum physics and classical physics. It is not clear how the two levels of observation should be reconciled. The failure of most macroeconomists to predict the last financial crisis suggests that one should try to incorporate behavioural theories into macroeconomic thinking.

One way to make such behavioural theories more attractive and reconcilable could be by trying to ground them firmly in a bigger scientific framework such as that of psychology or biology. It has been suggested that principles derived from cognitive psychology or behavioural ecology would help to explain seemingly irrational or "false" decisions as being rational within the ecological niche that we have evolved to live in. For example, although most introductory economics textbooks say an individual should ignore sunk costs to make a rational choice (the "Concorde fallacy"), it has been argued that in a broad range of environments, reacting to sunk costs is actually rational (35). For example in a world in which the outcomes and their variance are unknown, a specific return – even if it is a loss – can be informative of the variance of available outcomes and thus be indicative of the expected value of continuing a project. A large loss for example might indicate large outcome variance and consequently suggest high gains with further investments. Moreover in social situations, people might rationally want to conceal bad choices to appear more talented, which may lead them to make further investments, thereby hoping to hide their bad investments. If investments finally turn around their choices might even appear determined and far-sighted rather than irrational and stubborn. Finally, financial or time constraints may lead companies or individuals to stick with projects that no longer appear to be the best choice because they do not have enough financial and/or time resources to embark on alternatives.

Although often treated as rational agents, humans are, to a large part, biological systems manoeuvring in a complex world. In the following section I will try to explore how ethological and evolutionary accounts can explain seemingly irrational human behaviour. One may hope that such broader accounts of individual human decision making might have an easier time penetrating into sociological, macroeconomic and political theories of aggregate human behaviours.

1.1.4 The Neuroethological Turn: What can ecology tell us about human decision making?

A recent review by Pearson, Watson and Platt (36) suggests a turn in decision neuroscience from economically inspired theories of decision making towards frameworks derived from behavioural ecology and ethology. Such “turns” or the introduction of new or external theoretical foundations and approaches can be fruitful for several different reasons. They might help to reconcile existing results and theories in interesting ways. They might even allow for better compatibility of different findings, different views or different levels of investigation (e.g. a macroscopic and microscopic level) for the first time. But they can also allow for the asking of new questions or for new ways of how researchers perceive the subject of their own investigations. It is obvious that science does not take “the world as a whole” in its focus but specifies certain aspects or problems or at least prescribes how questions should be asked. This may change with these “paradigm shifts”. Finally, they may change methodological approaches and how researchers think their questions should be answered. All of these might lead to scientific progress being non-cumulative but more like a qualitative leap.

Nevertheless these shifts also pose their challenges. Decisions between different “schools of thought” are not always made based on scientific data but rather on whether one way seems more attractive and promises to be powerful in generating questions and ideas. Sometimes these comparisons cannot be conducted easily. Different “schools” often use the same words but with slightly different meanings, thereby making direct contrasts tricky. A famous example is the comparison between Ptolemy's and Kopernikus’ astronomical theories. Both use the term “planet” but refer to different things – namely “moving stars” or “entities that orbit the sun”, respectively. One should therefore be wary whether neuroeconomic and neuroethological frameworks speak about the same things when discussing “decisions”, “optimality” or “contextual variables”. Can this so-called “neuroethological turn” be

considered as a major shift in decision neuroscience and does it have to offer some of the above-mentioned things?

Recently decision theory and neuroeconomics have made substantial contributions to our understanding of decision-making. It added sophisticated theories and quantitative tools to the biological study of decisions in cognitive neuroscience. Decision theory is concerned with how agents select beliefs or courses of action among several alternative possibilities. Normative decision theory tries to identify the "optimal" decision – the decision that would be fully rational in considering all relevant information but disregarding irrelevant information, and in this way maximize the probability of achieving outcomes that meet the agent's own preferences. Biologists might extend this rather narrow concept of "decisions" to include other interesting classes of phenomena such as implicit biases or rapid physical responses, as in sports. Biologists might also wonder whether decisions must always occur over multiple options or whether withholding an action or disengaging with a decision process altogether also qualifies as a decision. Biologist might even have a different view of "optimal decision-making" since they are ascribing it to an agent selecting behaviours that maximize their fitness or survival probability in the ecological niche that they occupy.

Concepts from decision theory and economics had been taken up enthusiastically by biologists when it was realized that utility and subjective value, delay discounting, risk, and ambiguity showed correlates in the nervous systems of both humans and nonhuman animals. Frameworks from, for example, reinforcement learning have provided new models for analysing relationships between neural codes and behaviour. This was taken to suggest that economic theories of decision-making such as utility theory did not only prescribe how decisions ought to be made, but also predicted actual human decision making because they related in a direct way to how the human nervous system organizes behaviour and choice.

Earlier neurophysiological recording and imaging studies had already implicated the OFC, the ACC, the inferior frontal cortex (IFC), the posterior parietal cortex (PPC), the amygdala, the

ventral striatum and the dopaminergic midbrain in reward, choice and learning. However, economic utility theory now offered a way to reconcile and unify this multitude of data. Moreover, by arguing that different options are represented in a common currency to allow value comparison according to the economic concept of subjective utility, the theory proposed a scientific agenda – namely to understand the "neural code" of value representation in a common currency and the process of value comparison. Finally, it offered a quantitative framework for pursuing this endeavour.

Although initial evidence seemed to support the idea of OFC being the module for common currency value representation and comparison (20), it now turns out that its borders are fuzzy and between-species correspondences are not straight-forward (37). Besides, value representation in OFC does not seem to be context-independent in all cases and neural coding in OFC does not always correspond with choice behaviour (38). OFC neurons often signal variables other than reward or punishment, and slowly evidence is emerging that supports the notion of distributed processing of value and value comparison (see below for a general discussion of modular and network views of processing).

Rather than linking it with value comparison, other lines of research have linked OFC with flexible updating of stimulus-reward associations that link environmental stimuli to outcomes (39). It has been argued that the demand for rapid updating of appropriate responses in social contexts may have been a primary evolutionary driver for the enlargement and specialization of OFC and its related circuitry in highly social species (36). In foraging tasks, patch depletion or changes in average reward rates might happen rather slowly but mood or goals or behavioural interactions in a group might change more rapidly.

The other major neuroeconomic theory – the dopamine- reward prediction error theory – has been complicated by recent findings. Alternative theories hold that dopamine rather codes motivational value more broadly as opposed to hedonic value and that dopamine's primary role is in signalling ''incentive salience" (40). They suggest that the dopaminergic midbrain

transmits a repertoire of diverse, anatomically specific signals rather than broadcasting a single prediction error.

The neuroethological turn holds that the brain is an organ that evolved to generate adaptive behaviour in response to its environment, which would eventually support evolutionary fitness in a given environmental niche. Viewed this way neuroethology wants to understand the brain's functioning by looking at the problems it needs to solve, the algorithms that it uses to solve them and the underlying architecture that implements them. This approach has been called "bottom-up" or "engineering- like" because it cares about cost-effective and robust ways of implementing problem-solving tools and hence focuses on constraints, hacks, and specializations for tackling common tasks. Decision theory, on the other hand, has been formulated as a universal approach to mathematizing behaviour. It is axiomatic, top-down, and general, focusing on theorems, optimality, and rationality. It does not necessarily care about implementation, cost-effectiveness or robustness.

A famous example is the discussion about "deviations" from rationality or optimality. Both humans and other animals share some irrational or sub-optimal biases. In neuroeconomics these are often treated as exceptions or special cases or as reflecting the imperfection of biological implementation. The neuroethological turn is arguably able to reconcile these phenomena with reference to constraints under which agents operate, and the modelling of these constraints in e.g. foraging theory, mate selection theory. Biological optimality in that sense might be different from economical optimality. Moreover, prototypical problems used to inspire models of choice behaviour are fundamentally different: whereas neuroeconomics thinks of choices between risky lotteries, intuitions about probabilities, purchasing decisions, neuroethology thinks of "the four Fs: fighting, foraging, fleeing, reproduction" (36). The latter approach thinks that these were at the core of evolutionary fitness and would have therefore shaped the brain and its function. Ethologists might also think that many of the neuroeconomic problems constitute difficult ''edge cases'' for the brain employing different

biological sub-systems to varying degrees. As a result behavioural variability and inconsistency in performance is expected.

The neuroethological turn has suggested grounding the study of decisions more fully in their biological and evolutionary context. Moreover, it borrows ideas and tools from theoretical behavioural ecology - the mathematical study of animal behaviour that encompasses not only foraging, but diet selection, mate choice, and collective dynamics. Finally, it proposes starting the study of choice behaviour not with “edge cases” but with fundamental decisions common to a wide variety of animals: should I chase this prey or not? Should I keep foraging where I am or look for something better? Should I leave the group and forage on my own? Even though we humans nowadays engage in very complicated decisions, it is likely that we still use a rather old biological apparatus that actually evolved to address these basic problems. Therefore understanding them better might also afford insights in how humans manage portfolios and political decisions. Also if it is true that the brain has evolved to solve foraging and mate selection problems under great evolutionary pressure, it is likely that “optimality” and optimal behaviour can be found when studying these core capacities. A more practical benefit of “ecological” tasks might be that they are more realistic, more engaging and easier to teach to experimental animals and humans.

In that the neuroethological turn promises to reconcile a multitude of existing data and potentially harmonise decision paradoxes with axioms of optimality. Moreover, it poses new questions and puzzles for decision-neuroscience and puts forward mathematical ways of solving these. Below I will give a few very brief examples of major phenomena in decision theory and decision neuroscience and discuss how the neuroethological turn has changed the way they are discussed, studied and tentatively answered.

1.1.5 Examples

The previous section tried to argue that although “neuroeconomics” has proven to be a powerful approach for studying decision-making, a “neuroethological turn” promises ways to reconcile “neuroeconomics” with biological views on the process of decision-making. Moreover I suggested that such a neuroethological turn promises to address some unresolved puzzles in neuroeconomics and allows us to pose interesting new questions about decision-making.

I now want to briefly mention some concrete problems or areas of decision-neuroscience where a neuroethological approach has already fulfilled or might be close to fulfilling such promises. One of the core principles of many economic theories concerns the representation and comparison of subjective values of different goods (31, 41-43). One of the major propositions of neuroeconomics is that the biological correlate of such a process lays in specific brain-system for value-comparison in common currency. I will briefly discuss alternative, ethologically-inspired theories of multiple parallel decision-making systems in the brain.

Next I will consider dynamic modulation of “risk-taking” in decisions. Risk-aversion and proneness is often considered an irrational personality-trait of most human decision-makers. However, some models derived from behavioural ecology suggest that risk-taking can be highly adaptive in some situations whereas risk-avoidance should be considered rational in other circumstances. Biological systems should therefore try to dynamically vary their risk-attitudes.

In the next section I briefly return to the problem of inter-temporal choice. I will speculate why temporal discounting and deviations from exponential discount functions can be adaptive and how reconsideration can be viewed as rational. I will consider the changeability of the environment as an important factor to be considered by agents when learning about the world and making decisions. I will finally apply some of these ideas and considerations to the field of

psychiatry and speculate whether neuroethologically inspired theories of decision-making could have some implications for our understanding of psychiatric disorders.

1.1.5.1 Do humans have one or many serial or parallel decision-making systems?

Decision theory has long-argued on theoretical grounds that choices between different options should be made by comparing the different options in terms of their expected utility in a common currency. As mentioned above one of the major insights of neuroeconomics was that expected values and associated probabilities were represented in the brain (22). Moreover, neuroeconomists have argued that a brain system can be identified that implements value-comparison in the vmPFC and ACC (41, 44). In a later section I will contrast modular and network-views of brain function and consider evidence for serial and parallel processing in the brain more generally. I here just want to consider theories of serial and parallel processing for value representation and comparison.

One famous model of value-guided choice draws an analogy to perceptual decision-making: In standard versions of the random dot motion task monkeys have to integrate over the net motion of a cloud of moving dots and then saccade towards the direction of net dot motion (45). A famous theory of this process holds that visual information from the cloud of dots first reaches V1. The motion-sensitive area V5 / MT represents momentary evidence of dot motion. In the previously mentioned area LIP this momentary evidence is integrated over time, and evidence in favour of different “hypotheses” is compared in order to form a decision variable. Finally, a response is selected and executed via the basal ganglia and the frontal eye fields (FEF). Value guided choice might happen in very similar ways (46). In vmPFC values of different options might be estimated and converted into a common currency. ACC could then subsequently compare evidence in favour of any option and form a decision variable.

This model has been challenged by findings that vmPFC activity not only correlates with the value of a preferred option but it also correlates negatively with the value of a non-preferred option suggesting that it may already play a role in value comparison. Moreover, lesions to vmPFC and ACC appear to affect different types of decision-making: Whereas vmPFC lesions seem to affect value-guided choice (47, 48) and disrupt stimulus-reward associations, ACC lesions seem to affect action-value comparisons and action-value learning {Rudebeck, 2008 #2023;Camille, 2011 #2025;Procyk, 2000 #1368;Procyk, 2000 #2589}. Interestingly, axiomatic decision theory is explicitly stating the equivalence of the utility of a good and the utility of the act of picking that good (9). These findings could be taken to suggest that values of stimuli and values of actions are represented by different neural systems and distinct value comparison processes might be taking place.

Similarly it has been suggested that learning the "reward" value of stimuli and the value and reliability of "social" information from other agents and their advice would be carried out by distinct neural systems, although following largely similar computational principles (51). It is plausible to think that social decisions are to some degree also distinct from decisions between stimuli or actions (52). Finally, model-based and model-free decision-making has been proposed to employ partly distinct neural systems (53).

On an even more abstract level, neuroethologically inspired models of decision making have suggested a distinction between two types of decision problems: (1) Binary decisions and (2) "foraging" decisions (54, 55){Quilodran, 2008 #2591}. Although neuroeconomic studies of decision making largely focus on binary decisions, it might be argued that animals evolved to forage and rarely encountered binary decision problems when foraging. Rather they needed to be able to on the one hand weigh the costs and benefits of sticking with an option / staying in a foraging patch where the rates of return slowly decreased (or the patch slowly depleted) or on the other hand taking the time and effort of going somewhere else. Moreover, they need to be able to represent the estimated value of their environment to be able to compare that to

the value of currently encountered options or the momentary reward rate {Amiez, 2006 #1365}. Hence, animals might have evolved several independent value-representation and comparison systems. There seems to be evidence for the representation of “counterfactual values” or the value of the next best option in human frontopolar cortex (27, 56). This further supports the notion that different options are not always evaluated in the same way but that some alternatives are represented in a “special” manner (46) and some decisions are made in a distinctive fashion. One example might be the representation of the “average value” of the environment; this representation might occur in slightly different ways from the representation of the exact value of one of two options. In turn, the decision whether to stay with the current mode of action or search elsewhere (the average richness of the environment is high), might be distinct from binary value-comparison.

Why would the human brain function in such a way? Why would it represent social value, action value, search value or counterfactual value differently and not in a common currency? Why would it make decisions between actions in a different way from decisions between stimuli? Why would it use different value-comparison systems for binary decisions and foraging decisions? Decision theory suggested that rational choice could be achieved by having a precise representation of the subjective value of all options in a choice set and then compare all of them in a common currency. One reason why humans might have several parallel decision- making systems could be an evolutionary one. At different time-points during the course of evolution animals could have faced various different decision-problems. Foraging appears to be a rather old decision problem, although the average value of the environment and the search cost would have been largely different for different species and ecological niches. Inter-temporal choice might also be thought of as a slightly older decision problem. Some interactive, game-theoretic decision-problems might have only arisen later during evolution. Similarly some “economic” (binary) decision problems seem to have occurred mostly later in our phylogenetic history. When faced with a new decision-problem, animals

might have either “tried” to solve it by using an older neural decision making system (note that in some cases lesions to one decision-making system might cause the animal to fall back onto another system (57)). Alternatively, the animal might develop a new decision-making system to address the new decision-problem more quickly and cost-effectively. This might explain why we humans have such a “patchy” decision-making apparatus. Comparative neuro-anatomy and comparative neuroscience might be able to test this hypothesis. In addition, such comparative approach might, ideally, attempt to elucidate whether separate decision-making systems ever cooperate in any one species or even converge again in the course of evolution. Finally, decision-making systems might also disappear if they are no longer called for in a specific environment.

1.1.5.2 Modulation of risk-taking behavior with changing contexts

As mentioned earlier several different notions of “risk” and risk aversion (and in turn proneness) exist in decision theory. One of the most famous ones is the notion of “actuarial” risk. This notion of risk relates to the previously defined “decisions under risk”, and risk aversion in the actuarial sense has been described as follows: “A risk averter is defined as one who, starting from a position of certainty, is unwilling to take a bet which is actuarially fair [e.g. better from an expected monetary value standpoint]” (58). This reminds us of the previously mentioned example of a decision-maker, who prefers a certain £1M win over a 50/50% gamble between £3M and £0 (although the latter has an expected monetary value of £1.5M). We explained this earlier by referring to the concept of diminishing marginal utilities. The decision-maker’s behaviour can be perfectly reconciled with the expected utility principle. This notion of risk can be mathematically defined as a claim about the shape of the decision maker’s utility function. In the case of the above-mentioned risk-averse agent this slope is concave (slopes upwards but with decreasing steepness) and she prefers smaller more likely gains over bigger more unlikely ones. In the below example of dynamic modulation of risk-

taking I will mostly be referring to this notion of risk. In that context it is important to note that the utility function does not need to be concave. Although slightly less intuitive, agents might well prefer bigger more uncertain wins over smaller certain ones, even if they are actuarially fair. Keeping to the previous example, an agent might reject the certain £1M (expected monetary value of £1M) and prefer a 25/75% gamble of winning either £3M or £0 (EV=£750,000). He might do that because he is terminally sick and the medication that could potentially cure him costs £3M. In any case it is assumed that the shape of the utility function is more or less stable for a given agent; although it already becomes obvious from the examples how they may change with context.

A second prominent notion of risk is the idea that large gains and losses are not equivalent. Somebody might be considered risk-averse if the potential losses of a large number of units of value cannot be compensated for by an equally large chance of winning the same number of units of value. This reminds us of the maximin principle mentioned above, which tries to maximize the minimal outcomes when making a choice. However, note that just as before such an overweighing of losses can be transformed into a utility scale and is compatible with the expected utility principle. It is plausible to think that such loss-aversion, although sometimes similarly considered to be a stable personality trait, can be subject to dynamic modulation depending on the exact context.

Another concept of risk is that of epistemic uncertainty. The Ellsberg paradox is somewhat related to this notion of risk. In this paradox, humans are thought to violate the expected utility principle and behave as if they based their choice on something like a “heuristic” (most of them tend to avoid ambiguity). In the paradox people are presented with one urn with 30 red balls and 60 either yellow or black balls and face the following two gambles:

Gamble 1: Receive £10 if a red ball is drawn

Gamble 2: Receive £10 if a black ball is drawn

Most people in this case prefer to play gamble 1. But the paradox arises even when they choose gamble 2. Then they are offered another two gambles:

Gamble 3: Receive £10 if a red or yellow ball is drawn.

Gamble 4: Receive £10 if a yellow or black ball is drawn.

People who have chosen gamble 1 (ambiguity averse) now often carry on and choose gamble 4. Other agents (more ambiguity prone) may choose gamble 2 and 3. In any case this suggests that they do not make their decision based on a clear representation of the expected value of the different options because in both cases this would be an inconsistent combination of choices: if gamble 1 is better than gamble 2 this suggests we think that there are fewer than 30 black balls in the urn. In that case, however, we ought to choose gamble 3, not gamble 4. The behaviour that is often observed (most often choices of 1 and 4, potentially 2 and 3) suggests the use of a different "rule" or heuristic relating to the level of ambiguity associated with both gambles; i.e. either preferring options with less ambiguity (in most cases) or more ambiguity (in some cases).

The general notion of epistemic risk aversion is somewhat linked to the previous two notions of (1) actuarial risk aversion – the unwillingness to give up certain smaller gains for higher but risky (probabilistic) ones – and (2) the utility risk – the aversion against the chance of big losses. Consider the following example:

You are offered a bet on three different tennis matches A, B, C (9). For tennis match A you know the two players, and they are both equally good so the probability for one specific player winning is roughly 50%. In match B you do not know the players but you have heard that one of them is very good and the other one is very bad. So the probability of the good player to win is 90%. Unfortunately you do not know which player is which. Finally, in C you do not know

anything about the players and their probabilities of winning. Now you can bet on any player to win and will gain £10 if you turn out to be right and lose £9 if you turn out to be wrong. For match A you have only one probability distribution to consider and therefore the expected value for this bet is 0.5*10 + 0.5 * (-9) = 0.5. In match B you have to consider two probability distributions (0.1/0.9 and 0.9/0.1). The expected utility for bet B is therefore [0.1 * 10 + 0.9 * (-9)] = -7.1 for the first distribution and [0.9 * 10 + 0.1 * (-9)] = 8.1 for the second distribution. In bet C we do not know anything about the probability distributions and could therefore reasonably propose any expected utility between -9 and 10. When faced with these bets, one might consider applying a "maximin criterion for expected utilities" (9). Note that this parallels the maximin rule for choices under ignorance mentioned earlier. We would try to avoid betting when the smallest expected utility is very low. We would bet in case A, where the only expected utility is 0.5 and bigger than 0. We would, however, avoid a bet in B where the smallest expected utility is -7.1 and considerably smaller than 0. In case C the smallest possible expected utility is -9 and we should definitely reject such a bet. Note that this notion of epistemic risk aversion is somewhat linked with the previous two notions of risk (actuarial risk and "risk of big losses"). Similar to actuarial risk aversion epistemic risk aversion has a "negative attitude" towards uncertainty – but rather than against outcome uncertainty (as for actuarial risk aversion) it is biased against utility uncertainty. Similar to the loss-aversion type of risk aversion it (over-) emphasizes the negative – but rather than being biased against the negative outcomes (or losses), it is biased against negative utilities. As mentioned this "maximin criterion for expected utility" (i.e. focus on the worst possible utility) parallels the previously mentioned maximin rule for rational choice under ignorance (or in the case of bet C is identical to it). Because of that it also has similar problems: There is no clear reason why we should focus on the worst possible utility and not on the best (8.1 in the case of bet B, which is much higher than in bet A). One might even wonder whether humans can dynamically modulate their attitude towards ambiguity.

With respect to the above-discussed notion of "modulation" of attitudes towards risk, we might question whether this rule of avoiding (or potentially approaching) ambiguous options is really a stable trait or whether attitudes towards ambiguity can be modulated with context. It is reasonable to think that animals could face situations that require exploration and others that suggest they should exploit. The neural implementation of such dynamic modulations of attitudes towards ambiguity needs further exploration although elegant studies have been conducted (59).

More generally one might wonder how such dynamic modulations of attitudes towards (1) actuarial risk, (2) utility risk and (3) epistemic risk could be modulated. For example, dynamic modulations of utility risk might be achieved by transformations of the utility scale. However, they might also be achieved by switching between some simpler decision rules or heuristics (e.g. switching from something like the maximin to the maximax rule). Studying the neural implementation of dynamic modulations of different types of risk attitudes could yield important insights into the "causes" of different attitudes towards risk altogether.

One study conducted by Kolling (60) attempted to study the dynamic changes in risk attitudes in humans drawing from risk-sensitive foraging theory. This theory suggests that an animal's attitude to risk should be a function of its momentary resources and its longer term targets. When choosing foraging strategies, animals should not only consider the mean rate of reward from a particular foraging option but also the variability. Whether an animal prefers high or low variability options depends on the animal's needs and the expected reward. If energy needs are less than the expected reward, it will choose "conservatively" (risk-averse) (61). If energy needs exceed average expected reward, it should choose the "risky" strategies. Caraco et al (62) offered yellow-eyed juncos in an aviary the choice between a gamble of obtaining either 0 or 6 seeds with a probability of 0.5 or a certain 3 seeds. When the experiment was carried out at 19 degrees centigrade the birds chose the safe option. When temperature was lowered to 1 degree, they switched to a risky choice behaviour in order to meet their

increased daily energy needs. In his human fMRI-study, Kolling (60) showed a similar behavioural effect. Moreover, he suggests several parallel systems for value-representation either with reference to momentary “risk-pressure” or without. Their findings suggest that such parallel value-representations circumvent the need for a re-coding of the utility scale with a given context. One might speculate that alternative competing utility scales could allow flexible, context-dependent switching between any one of them and therefore generate the dynamic changes in attitudes towards risk that have been observed by behavioural ecologists.

1.1.5.3 Ethological views on intertemporal choice

Animals often have to make decisions about what to do at different points in time, when choices at one point influence the possibilities available at other points in time. Such choices are influenced by the value people assign to different options at different points in time. Many inter-temporal choice-problems involve trade-offs between costs and benefits at different time points.

Humans often have a (differently strong) tendency to discount rewards as they approach a temporal horizon in the future. In other words, they give greater value to rewards available sooner than to rewards available later in time. This is why people forgoing the consumption of a certain good at a specific time point, would normally expect to receive a higher amount as compensation for the same good later in time. This idea relates to the “theory of interest”. People might forgo consumption now if they can expect a return on their savings in the future because rational forward-looking consumers choose consumption for the present and future to maximize their lifetime satisfaction (63). Note that animals often face similar decisions, for example when deciding how much energy to invest in a current brood, potentially at the expense of future fecundity in order to maximize life-time reproductive success (61). Similarly, Modigliani’s life cycle income hypothesis (64) is trying to explain individuals' consumption and saving patterns with reference to insights into their own life-cycles and the desire to maintain

a constant level of consumption or a stable life-style. They hence accumulate savings when they are younger and earn an income and dis-save when they are retired. This life-cycle effect could be explained by steeper temporal discounting in older compared to younger people. Temporal discounting in lab-experiments however does not change in the same manner across the life-span (65), suggesting at least partly independent mechanisms.

However, most economical models proposed- where individuals make intertemporal choices - that value is exponentially discounted with the same interest rate for every period of the future. If we were given the choice whether to take £100 today or receive £110 tomorrow, we should respond to this question in the same manner as when offered £100 in a year and £110 in a year and one day. Yet, this does not match what is often observed in real human behaviour. Some behavioural economists have therefore suggested that individuals are often affected by so-called "temporal myopia". Consumer's response to uncertain offers might be to sharply discount them temporally. They do for some reason not discount exponentially but rather hyperbolically (this effect is called "hyperbolic discounting").

This alternative hyperbolic discounting function has, among others, the following predictions: (1) Compared to economically predicted exponential discount functions humans seem to down-weigh value for shorter delays and inflate values for very long delays. Moreover, such discounting functions can lead to reconsideration and reappraisal, i.e. value differences between two options at different distant points in the future might change as they approach in time and even flip signs. This might lead us to abandon our resolution and decide in favour of the more immediate reward although previously we had decided differently and intended to wait. Imagine you will have a job interview the next day and in the evening you decide you want to get up early to go over your talk once more. When you go to bed at midnight you decide to set the alarm to 6am, because you feel that getting the job in a few months from now is much more valuable than sleeping another hour in six hours from now. However, six hours later when the alarm rings, you evaluate your options differently and because of the

hyperbolic nature of your two discount functions you now feel that sleeping another hour right now is much more valuable than a job in a couple of months' time.

Both the down-weighing of rewards in the immediate future and the inflating of more distant rewards as well as the "crossing" of different hyperbolic discounting functions do not strike us as very rational. Such apparent deviations from rationality have been linked with abovementioned notions of parallel neural systems for choice and valuation. One could imagine that rewards that are available right now (valuation of things right in front of us) are evaluated differently from rewards in the more immediate future (e.g. the next hours or foreseeable future) and even differently from temporally very remote rewards (valuation of things like pension schemes, saving the planet for your grandchildren). For example, it is known that general optimism about the future, but also prospective memory and imagination of future rewards, influences temporal discounting (66). Nevertheless, they might be partly independent mechanisms and therefore influence valuation on distinct timescales differently. It is not implausible to envisage two or more "instrumental" or goal-directed systems for the valuation of future rewards, which would mainly be driven by concrete ideas about what future rewards will be used for or how they will be consumed. Moreover there might be an alternative system, which only has a fuzzy concept of rewards obtainable in the (very remote) future. Another system that has often been proposed to work in tandem with valuation systems in intertemporal-choice problems is the executive control system. The decision against immediate rewards and for temporally distant rewards might be largely driven by this top-down control system which has access to our long-term goals (67); and acts of reconsideration and reappraisal might actually be signs of this self-control system failing. The existence of such a system might relate to the intuition that attention plays a major role in temporal discounting, and the direct encounter of immediately available goods is almost qualitatively different from goods becoming available very soon in the future (but not being at hand yet). Another decision making system that was already mentioned above (54) is the "foraging"

system or the neural system that tracks the average value of the environment. It might be involved in estimating the value of disengaging from a current behavioural strategy (such as waiting) or in evaluating the opportunity cost of the current mode of action (immediate consumption). The foraging system could interact with the executive control system in a specific way: Higher opportunity costs for one of the waiting strategies (i.e. the cost of forgoing a very tempting immediate reward) might require a particular self-control effort. A system tracking the opportunity cost of a waiting strategy could then accordingly recruit top-control systems to avoid reconsideration of that strategy: ACC -> dorsolateral prefrontal cortex (dlPFC) interactions have often been observed (68, 69) and these interactions might be the neural corollary of just such a process. One might in turn speculate that failures in self-control might be failures of ACC to recruit "enough" dlPFC.

One might wonder now why one should have so many different neural systems involved in valuation and value-comparison in intertemporal choice, particularly since their cooperation does not prevent the occurrence of the abovementioned seemingly irrational behaviour (in fact this cooperation might cause it). One very important consideration, however, is that in nature not all these goods will last forever. Discounting functions should be adaptable depending on the goods offered. Food decays very quickly and should be discounted differently from other goods. Fixed discounting functions would not make sense with respect to the variety of goods available in everyday life. Moreover, the interaction of several different neural systems in intertemporal choice might afford the flexibility needed in an ever-changing world. Not only the items themselves but also their availability might change dramatically and the above-described neural systems might allow for the adaptation of temporal discounting to these changes in the structure of the world. But not only the items and their availability are subject to change, the opportunity costs might change as well. Moreover, our own resources and homeostatic state might change (unpredictably) over time. Finally, other agents might influence intertemporal valuation. In such a multi-dimensionally changing world reappraisal

does not seem so irrational after all, and the interaction of multiple neural systems might allow for such flexibility and appears highly adaptive.

In a recent study, Kable and McGuire (70) showed that humans adapt their level of "impulsivity" in uncertain situations to the exact structure of the environment, that is the distribution of delays of rewards. The question arose because in several studies children were shown to be particularly "impulsive" (for example in the famous marshmallow test (71)). One popular theory is that dlPFC is not yet matured in children, and in its absence they are more impulsive. An alternative account holds that they might not have such a well-developed idea of time (and the reasonably abstract idea of delay distributions) when they are for example told: "Don't eat the marshmallow just yet but wait here for a little bit (!) and you will get another one". With a clear grasp of temporal intervals and delays one might expect the reward probabilities to be negatively correlated with elapsed time (e.g. "I have waited for so long, the reward must appear any moment now"). However, that does not have to be the case and with a fuzzier notion of delays one might think that at some point reward probability becomes negatively correlated with elapsed time (e.g. "I have waited for so long, maybe that means it actually takes so so much longer than I had expected and reward is unlikely to arrive any time soon"). In fact, an effect similar to the marshmallow effect regularly happens in adults in relation to slightly more abstract goals. The goal of losing weight in a diet for example might be expected to be achievable within a couple of weeks. However, when after several weeks of dieting the same goal does not materialize, we need to update our temporal expectation (and quite often discontinue the diet). In environments with unknown delay distributions the constant reappraisal of the value of waiting might in fact be highly adaptive.

In the study by Kable and McGuire (70) participants are faced with two different environments with different delay distributions. In one environment with uniform delay distributions (same probability for reward to arrive at any time in the first 40 seconds) it would be optimal to wait until the reward really arrives. In the other environment, which has bimodal delay distributions

(50% probability of reward arriving in first 20 sec. and 50% probability of reward arriving after 90 sec.) it is more rational to abort the trial if the reward has not arrived after 20 seconds and start the next after a 2 seconds inter-trial interval. The authors report that subjects in this task behave in exactly that way. They discontinue waiting early in the bimodal delay distribution environment, but wait much longer for rewards to arrive in the uniform delay distribution trials. Moreover, they propose a model in which subjects constantly update the value of waiting and compare it to the value of discontinuing waiting or progressing to the next trial. A perigenual ACC (pg-ACC) region's BOLD signal correlates with the value of "persisting" and changes dynamically over time as the value for persisting changes. Interestingly, the same region has been implicated in cost-benefit decision-making (72), representing the costs of overcoming a default-bias (54) or dynamically tracking and projecting the rate of reward in a given environment (Wittmann et al, submitted). More generally it may seem that humans are able to evaluate the expected rates for costs and benefits in a given environment dynamically in order to make and revise decisions to stick with a mode of action. It appears that the ability to evaluate and reappraise cost-benefit trade-offs over time is crucial for intertemporal choice with unknown delay-distributions.

1.1.5.4 Neuroethologically-inspired views of psychiatric disorder

The above-discussed neuroeconomic turn had a major impact on how we thought about psychiatric disorders (73). Clinicians thought neuroeconomic paradigms could help characterize some aspects of psychopathology as deviations from theoretical predictions about optimal behaviour. These deviations could then be linked with clearly discernible brain-pathology and potentially genetic aberrations. In other words neuroeconomics was thought to help with discerning endophenotypes of psychiatric symptoms and disorders (74). This promising view is partly founded in a more general movement away from a solely neurochemical view of psychopathology towards an integrated neurochemical and

neuroanatomical view. For example, impulsivity became linked to ventral frontal cortex aberrations (75). Schizophrenia is associated with frontal and temporal lobe change (76, 77). Depression can be attributed to anterior cingulate abnormality (78). This view, largely inspired by neurology, proposes that psychiatric disorders are fundamentally "brain diseases". "Lesions" of brain areas or their disconnection in turn cause deviations from "normal" behaviour (e.g. optimal decision making). This frame of mind is supported by an impressive number of studies reporting differences in brain structure between psychiatric patients and controls.

I will spend a few words here speculating whether a neuroethological turn could have any relevance for how we think about psychiatric disorders. I propose that at least some psychiatric disorders can be understood as an "ecologically valid" adaptation of a person to the specific environment they have encountered. Rather than being the result of a "broken brain", psychiatric symptoms emerge from behavioural and neural adaption to extreme environments, which in turn causes dysfunctional behaviour. This framework expands on views about resilience to and vulnerability for psychiatric disorders. It suggests that a well-functioning brain living in an unusual environment, perhaps one that is different from the environments for which the brain adapted, especially during development, produces aberrant behaviour. This may be especially relevant in today's society due to the rapid pace of environmental and technological change.

Take for example the concept of "impulsivity", which is thought to be a major component of various psychiatric disorders, including attention deficit hyperactivity disorder (ADHD), substance abuse or bipolar disorder. The broader concept of impulsivity is likely to comprise several sub-components, such as (A) the tendency towards actions, which are poorly conceived, risky, premature and undeliberate or (B) propensity to choose short-term gains over long-term ones. Classically these concepts are often tested in lab-settings using executive control tasks (such as the stop-signal task) and intertemporal choice tasks. Interestingly when

“impulsivity” leads to positive outcomes, it tends not to be seen as an indication of malfunction or disorder, but as a sign of courage, spontaneity, boldness or unconventionality. This might be taken to suggest that impulsivity and the “loss“ of self-control (as observed for example in drug-addicts) should not be thought of as the result of a brain's breakdown (as suggested by the phrase “*loss* of self-control”) but as an adaptation to a particular environment.

In the previous section I have already discussed how “deviations” from economically rational choice, such as temporal discounting functions and reappraisal of waiting to delay gratification can be adaptive in certain situations. Quite often addicts deliberate over whether to stay clean or to take a given drug and then make a seemingly “rational“ choice to take the drug because this option appears to be more rewarding. This could suggest that in these situations their brains (i.e. their decision-making apparatus) functions “normally“ within the constraints of their ontogenetic and phylogenetic history. Rather than viewing the impulsivity of drug-addicts as a deviation (patients) from optimal behaviour (“normal” people) one might speculate about a continuum of personal choice patterns with relation to delayed gratification (as is indeed seen in studies like Kable’s and Glimcher’s (32)). Divergent delayed gratification patterns might reflect variation in the structure and properties of the environments that we grew up and lived in (as is anecdotally seen in Kable’s and Glimcher’s study with steeply discounting sky-divers and patient MD/PhD-students). As already mentioned in the previous section one major environmental variable could be its volatility. It is known to influence how we learn from feedback. It may also influence how we make decisions between immediate and more distant rewards. In a highly volatile environment, it might prove “ethologically” sound to let very recent and immediately available rewards guide decisions (as opposed to temporally distant past and future events), given the likelihood that later rewards may no longer exist in the future. In these environments the choice of immediate rewards might be rational and not a reflection of brain dysfunction or disinhibition.

It is easy to conceive how the other aspect of impulsivity mentioned above– poorly conceived and prematurely expressed actions – can be an adaptation to the structure of a given environment. “Response-inhibition skills” can be improved upon by training, which is thought to have an effect on brain-structure and carry over to behaviour in everyday life situations (79). It is plausible to think that such an effect could also go the other way: Consistent patterns of interaction with a given environment might shape brain-systems for response inhibition which can be measured in laboratory experiments. Future research will show whether such “adaptive” accounts of response-inhibition are supported by evidence.

Another example are so-called “attribution styles” or explanatory styles, which are thought to be at the core of psychiatric disorders such as depression (80). Attribution styles are thought to be patterns that indicate how people explain the causal structure of certain events. The notion of attribution is linked to that of credit assignment and causal inference. The hypothesis of person-specific styles of attribution suggests that there are individual biases for how people infer causes for events and how they assign credit. Depression is thought to be consistently associated with (and potentially caused by) attributions to internal, stable and global causes, particularly for negative events. This attribution style biases people to think that negative events (e.g. failing an exam) are caused by themselves (internal) rather than either other people or circumstances (e.g. bad performance during exam rather than noisy environment). Moreover, they tend to infer that the causes are stable rather than unstable (e.g. always being badly prepared for exams) and global rather than local or specific (e.g. being generally stupid rather than not being interested in this particular subject). In experiments on attribution style “normal” people (i.e. non-depressed control subjects) are often shown to exhibit a “self-serving bias”, i.e. they tend to attribute positive events in an internal, stable and global way, but suggest that negative events are being caused by external, unstable and specific incidents. One might suggest now that a dysfunctional attribution style in depressed patients is due to an impairment of the “credit-assignment system” or a lesion to the “self-serving module” might

be the cause for depression. However, at least some data points towards an effect which is sometimes referred to as “depressive realism” (81), i.e. that depressed people are at least in some cases more in touch with reality and their causal inferences are more sound than those of “healthy” controls. This precludes any simple “lesion” or “deviation” account of “depressive” attribution styles. I would rather speculate that there might be a continuum of credit-assignment patterns. They are likely to shape how we learn in any given environment and would largely be the product of a given environment. Learning is known to be shaped by environmental variables (82), and future research will have to show whether long-term environmental patterns influence the way agents assign credit. At this point one might be tempted to speculate that in our current, very stable and protected part of the world infants’ self-driven actions (above I have mentioned action-value learning) are largely shaped by positive reinforcement (act in certain ways to obtain a reward) whereas learning about stimuli and contexts (stimulus-value learning) might still entail negative reinforcement (learn to avoid certain stimuli or contexts). It is possible that our individual learning-history shapes how we deal with positive and negative outcomes – i.e. interpret them either as the result of something that we have done or something that was situational. Finally, it is not implausible to think that social hierarchy would influence attribution style – subordinate individuals might be forced to at least overtly attribute negative events to themselves rather than to others in their group. Rank in social hierarchy is known to leave its mark on brain structure (83), particularly in regions that are implicated in credit-assignment such as the amygdala.

Depressive attribution styles might thus just be the result of an adaptation to a specific learning history. How such styles can be adaptive is easy to conceive. Self-value and self-esteem are intimately associated with attribution. The level of self-esteem as well as the changeability of this level influences attribution (84). People with lower self-esteem are more likely to attribute negative events to themselves. On the other hand “negative” attribution styles can decrease self-esteem and more extreme attributions of everyday events can cause

more unstable self-value. Low self-values or very changeable self-values, if they occur only for a limited amount of time, might be strong drivers for behavioural change or rapid behavioural adaptation. This ties in with some views of depression as an adaptation to major life changes, which are solved by internally focused attention and analytical rumination (85).

Delusions are one of the core symptoms of schizophrenia. They are often described as strongly held convictions despite superior evidence to the contrary and are therefore prime examples of irrational and sub-optimal behaviour or thought. Among the most common delusions are persecutory delusions and delusions of reference. But perhaps even in the case of delusions it is possible to understand them not as the result of a brain's breakdown but as adaptive interaction with a certain environment. Several lines of evidence support the notion of delusions resulting from normal cognitive processes and can be thought of as to some degree "rational" theories about the world. Firstly, just as normal beliefs, delusions are known to fluctuate over time (86, 87). Secondly, their content is often observed to reflect earlier or childhood experiences (88). Thirdly, paranoid patients have been reported to potentially be more sensitive to non-verbal cues and facial expressions of other people (89) suggesting that some of their beliefs about other people might actually be correct (an argument somewhat paralleling the depressive realism idea). Finally, just as patients with depression, they are known to exhibit different attribution styles that are markedly different from those of "normal" people (90, 91). Yet in the case of schizophrenia patients, the attribution style is changed in different ways: their self-serving bias is substantially increased owing to internal attribution for positive events and external attribution to negative events. Rather than blaming negative events on the situation (external-situational attribution) they tend to attribute them to other people (external-personal attribution) (92). Together with a tendency towards higher epistemological impulsivity (93) – the tendency to jump to conclusions more quickly – this might constitute a specific pattern of causal inference in schizophrenia patients. I will not have the space here to speculate in which circumstances such a pattern of credit assignment can be

considered adaptive. However, I will hint at a few aspects briefly: First of all, it is known that when we account for the behaviour of other people, we initially make external-personal attributions and only then reconsider and search for situational explanations (94). Patients with persecutory delusions might stop at the level of external-personal attributions because of their epistemic impulsivity. Alternatively they might attribute events to other people because they are in fact better at reading them (95). It is a known phenomenon that people tend to attribute cognitions that are particularly effortless to the external world. A link with effort seems to provide a major cue for discerning when a thought or event is ours i.e. made-up rather than experienced and deriving from the outside world (in that sense it is cognitively harder to "imagine" a holiday than to experience it). An example might be a genius-mathematician who claims not to have come up with a solution herself but as having just seen it emerge from the numbers spontaneously. Another example might be Paul McCartney, who thought for several days, that he had heard the tune of "Yesterday" somewhere else (maybe because he just came up with it without much cognitive effort). It is easy to imagine how such "projective processes" would be adaptive (already on the level of perception, as Gestalt-psychologists have pointed out) since they make our interpretation of the world more reliable, robust and efficient allowing for rapid behavioural adaptation in it. If we were aware of the fact that our interpretation of big green things as trees is just an interpretation we would struggle to manage our lives. Similarly deluded people might be particularly gifted at discovering patterns in their (social) environment. They therefore do not consider themselves the authors of such hypotheses and tend to not question them as a consequence. This prevents them from considering alternative situational explanations.

Moreover epistemological impulsivity, which is sometimes related to delusions, could also be seen as an environmental adaptation. Just as with impulsivity in intertemporal choice, epistemological impulsivity might be an adaptation to volatile or ephemeral environments. If the world changes more quickly and if responses are required more readily, it might prove

adaptive to develop more impulsive reasoning styles. Epistemological impulsivity might also be linked with an increased “need for cognitive closure”, which indeed has been reported in schizophrenic patients (96). This “need for cognitive closure” might be thought of as an epistemological variant of ambiguity aversion. As discussed previously, ambiguity aversion in value-guided choice can be an adaptive strategy in certain situations (e.g. when exploration is unsafe). Similar ideas might hold for epistemological ambiguity aversion. Finally, I would like to point out that epistemological impulsivity and misattribution could be linked to the way schizophrenic patients evaluate hypotheses. A popular hypothesis about delusions holds that they arise from flawed and illogical thinking and hypothesis-testing. However, it has been suggested that even normal human hypothesis-testing strategies have formal-logical flaws. For example ordinary people often seek confirmatory data when testing a hypothesis although disconfirmatory evidence would be more informative. This phenomenon is called confirmation-bias. Although formally illogical this approach can be highly adaptive in real-life situations particularly if negative outcomes (disconfirmatory evidence) is highly costly (97). One might now conclude that change of the confirmation bias in response to changes in the costs for disconfirmatory evidence would be highly adaptive, although formally illogical.

Interestingly, several of the above mentioned psychiatric disorders and symptoms were linked to the problem of attribution, causal inference and credit assignment. I tried to argue that patterns of causal inference should change with environmental variables. This idea still needs to be tested empirically but if it is true (mal)-adaptations of attribution might prove to be a major “breaking point” for several psychiatric disorders. This might be for functional and structural reasons. Attribution might be a core cognitive process that is called upon by a wide variety of other cognitive processes and therefore employed constantly (almost like a “cognitive hub”). Moreover, it might be supported by neural structures that are more flexible and adaptable than others.

Another core cognitive component at least in schizophrenia might be that of meta-cognition and source-monitoring. Source-monitoring allows us to determine the source of cognition. For example it allows us to distinguish between cognitions (thoughts, emotions, memories, images) and perceptions/actions (things heard/seen/smelled etc., things done). There is reason to believe that it is not always easy to determine whether things have actually happened or whether we have just made them up (and philosophical disciplines like solipsism have argued it is impossible). Aberrant source-monitoring might play a role in hallucinations. Schizophrenic patients sometimes report perceptions (most often voices) in the absence of appropriate stimuli. Aberrant source-monitoring is also argued to play a role in thought disorder. Thought disorder is a misleading term and describes disordered language and communication: Patients are observed to exhibit an increased amount of speech ("pressure of speech"), replying to questions in tangential ways ("tangentiality"), letting ideas slip of the track on to other obliquely related topics ("derailment"), excessive long-windedness ("circumstantiality"), perseveration and interruptions of trains of speech without completion (blocking), in short: exactly the way that some academics speak. The inability to dissociate things one has thought and things one has said (particularly in highly emotionally charged circumstances) could underlie what psychiatrists call "thought disorder".

How do we know whether something has really happened or was just imagined? Humans often use simple cues and heuristics to determine whether things have actually happened (I mentioned cognitive effort as a cue for internally generated thoughts above). Bentall and Slade suggested signal-detection theory (98) as a framework to think about source-monitoring. Whether something is judged as externally or internally sourced depends on the perceptual bias. This in turn will determine the number of *hits* (a cognition is correctly identified as real and "outside the brain"), *misses* (a cognition is thought to not be "real" but it really exists), *false alarms* (hallucinations) and *correct rejections* (a cognition is correctly identified as imagined). Experiments using signal-detection theory measures showed that hallucinating

patients show no difference in sensitivity (the ability to detect e.g. voices in white noise) but exhibit a different perceptual bias (willing to accept more false alarms in order to avoid misses) (98, 99). Humans are known to adapt perceptual biases to the contingencies (rewards/punishments) in their environments (100). One might predict that source-monitoring styles observed in schizophrenia patients are caused by skewed contingencies with unusually high punishments for misses and unusually high rewards for false alarms. It is pure speculation whether extremely volatile, highly unpredictable, emotionally stressful and haphazardly hostile environments would predominantly pose an emphasis on punishing misses (the missing of true external signs of danger and hostility).

I have only given some very cautious hints for how a neuroethological turn in decision-making and reasoning might change our view of psychiatric disorders. I have argued that several psychiatric symptoms should not be viewed as emerging from brain-breakdown, as might be expected when looking at neurological accounts of behavioural aberration. Therefore, I suggest that they can be studied as short-term adaptations of behaviour to specific environmental architectures, which on the long run however prove maladaptive and dysfunctional. I would certainly expect such adaptations to leave their mark on the brain (I would be surprised if they did not). Ultimately, I would interpret the results of group-comparisons between psychiatric patients and "normal" controls as reflecting a continuum of neural adaptations and behavioural solutions to environmental problems.

1.2 Modules, Networks and Connectivity

The history of neurology and neuroscience has been shaped by a debate between modular and connectionist accounts of the brain. Some researchers and clinicians viewed structure, function and dysfunction of the human central nervous system mainly with reference to highly differentiated and specialised modules and brain areas. Others emphasized the significance of their connections and interactions in order to explain cognitive abilities and their impairment.

Although Franz Joseph Gall is often maligned for trying to use external cranial features as an indirect measure of cortical organ size and therefore cognitive ability or personality traits, he needs to be credited for recognizing the specific functional importance of the cortex (101). He showed that the brain consisted of grey and white matter, proposing that white matter was made up of ascending and descending conducting fibres originating from or projecting to the cortex (102). He developed a system of organology in which functional variations were correlated with the size of specific cortical regions (cortical organs) both across and within species. These ideas share similarities with approaches used today in, for example, voxel-based morphometry studies. Although his ideas soon developed into the pseudo-science of phrenology, they nevertheless had an important influence on later clinicopathological correlation studies.

In his seminal paper of 1861 Paul Broca linked the progressive loss of speech in his patient Leborgne to damage in his left inferior frontal cortex (now known as Broca's area). Karl Wenicke noticed that not all language deficits were the result of damage to Broca's area. He found that damage to the left posterior, superior temporal gyrus resulted in deficits in language comprehension.

Later discoveries, such as that degeneration of cells in the substantia nigra leads to Parkinson's disease or that lesions to an area in the fusiform cortex result in a specific impairment in recognising faces, seemed to support a modular view of brain function.

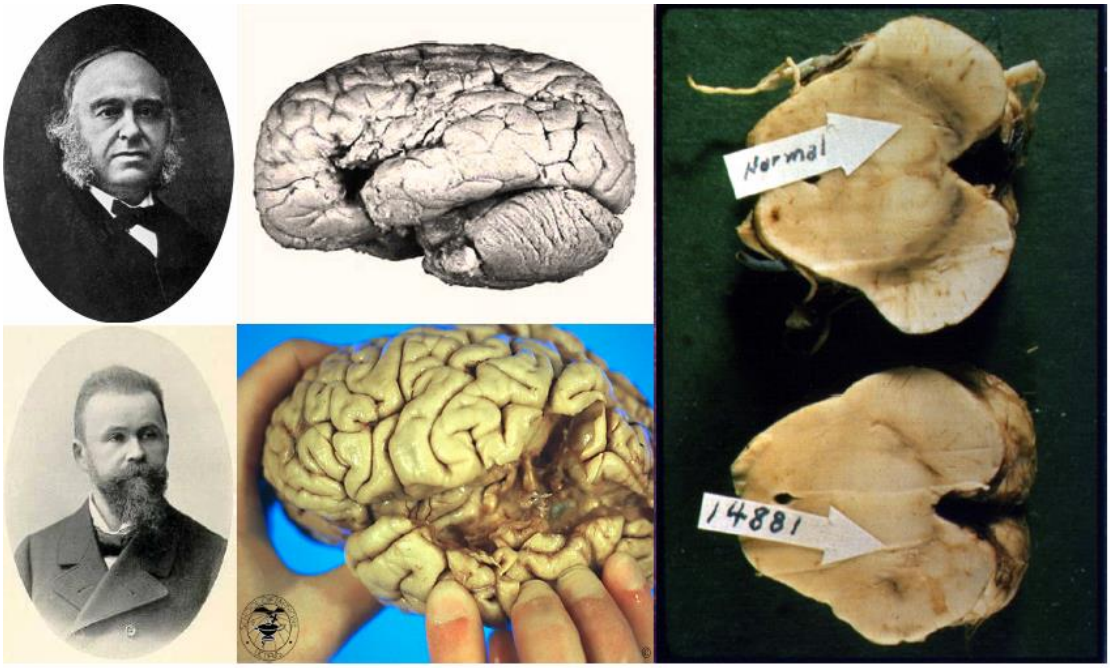


Figure 1-4: The upper left shows Paul Broca and the brain lesion of his patient Leborgne. The lower left shows Carl Wernicke and a classical depiction of a lesion to Wernicke's area. On the right, axial slices through the midbrain depict degeneration of cells in the substantia nigra in Parkinson's disease.

Wernicke thought of the brain as a mosaic-like arrangement of areas containing 'fundamental psychic elements' or 'memory images' related to motor acts and sensory experiences. These memory modules were localised in primary sensory and motor areas according to the following principle:

"the acoustic images find their abode within the cortical terminals of the acoustic nerve; the visual images, within the cortical endings of the optic nerve; and the olfactory images in that of the olfactory nerve; and so on. Likewise the motor images or movement representation could be located in the cortical sites of the motor nerve origins." (102)

Later the discovery that distinct sensory and motor features are indeed represented in circumscribed cortical areas (such as M1, A1, V1 and S1) further strengthened the modular view. Together with the description of secondary sensory areas (such as V2, V3, V4 etc. in the case of visual information) they implicate modular and serial models of brain function. They may later have inspired cognitive neuroscientists to "localise" higher order cognitive functions

such as working memory (to the dorsolateral prefrontal cortex) or conflict resolution (to dorsal ACC).

However, sceptics have asserted that even though distinct sensory and motor features may be represented in specific modules in the brain (and this already can be doubted with reference to, for example, the number of motor areas representing the arm in primary motor, pre-motor, supplementary motor and cingulate cortex (103)), it still needs to be explained how this sensory channel-specific information is integrated to achieve our evidently coherent "percept" of the world. Moreover clinicians have pointed out further limitations of the modular account of brain dysfunction by discovering that, for example, lesions to other areas besides „Broca's area" or even only to the white matter underlying cortical regions can lead to Broca's aphasia. Similarly, degeneration of other structures besides the substantia nigra can lead to Parkinson's disease-like syndromes.

Researchers have therefore turned to connections and proposed that perceptual integration or higher cognitive functions may be the result of processing within distributed networks. With respect to these ideas it is interesting to note that Wernicke, although arguing for his "memory images" modules, also (perhaps in an attempt to distance himself from phrenology) insisted that higher cognitive functions were not localised in specific regions but were the result of associative connections between motor and sensory memory image areas. Thus, Wernicke asserted that

"any higher psychic process, exceeding […] mere primary assumptions, could not, I reasoned, be localised, but rested on the mutual interaction of these fundamental psychic elements mediated by means of their manifold connections via the association fibres." (102) (See Figure 1-5)

In this framework there was no place for cortical specialisations beyond those of primary sensory and motor functions. With these basic ideas Wernicke provided the foundation of his

school of “associationists”, who studied higher cognitive functions as arising from associative connections and described disorders of higher functions as resulting from their “disconnection”. This powerful approach helped explain distinctive patterns of language, praxis and vision deficits that are referred to as *disconnection syndromes*. Examples are conduction aphasia, visual agnosia, different forms of apraxia, and pure alexia.

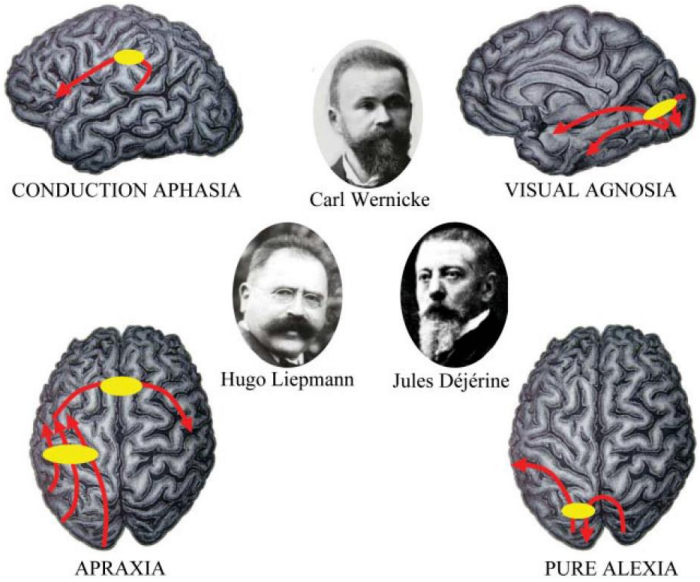


Figure 1-5: Wernicke thought that a disconnection between Broca’s and Wernicke’s area would lead to “conduction aphasia”, an isolated problem of speech repetition. Wernicke and Lissauer thought that a lesion, which spared primary visual cortex but involved its white matter outputs would result in visual agnosia, an impairment of visual recognition. Liepmann argued that the left hemisphere was dominant for complex movement control and that a lesion to the left parietal lobe disconnected the left-hand area from visual, somatosensory and auditory input, leading to bilateral apraxia. In contrast, a lesion of the anterior portion of the corpus callosum disconnected the right hemisphere from the left leading to unilateral left-hand apraxia. Déjérine argued that pure alexia resulted from the disconnection of the visual verbal centre in the angular gyrus from earlier visual areas in both hemispheres (102).

One might wonder why after the great successes of modular accounts of brain function (such as the description of Broca’s and Wernicke’s aphasia) neurologists like Wernicke turned to connections and models of distributed processing in the brain. This “associationist” and connectionist turn is particularly surprising given that our modern neuroanatomical concept of the brain as consisting (among other things) of a multitude of cells (i.e. neurons) connected to

each other via their extensions (i.e. dendrites and axons) was put forward by Santiago Ramon y Cajal only some decades later.

The emergence of the connectivity and network theme as a fundamental component of neurological and neuroscientific thinking might be linked with two developments in the latter half of the 19th century (104). One was the extension of neuroanatomical research from a description of surface morphology to the tracing of projection pathways coursing through the white matter of the cerebral hemispheres. The other was the success of 'associationistic' models of cognitive function.

Korbinian Brodmann for example attempted to map the cerebral cortex based on its cytoarchitecture. Using Nissl-stained sections of a single human hemisphere he proposed 52 distinct brain regions. Theodor Meynert described the major white matter tracts and classified them into three groups. The first group consisted of projection fibres, the ascending or descending pathways arising and terminating in the cortex, the second of commissural fibres, which connected cortex in both hemispheres and the third of association fibres which connected cortical regions within a hemisphere. Each area seemed to receive a different combination of fibres. He emphasised the important functional role played by association fibres.

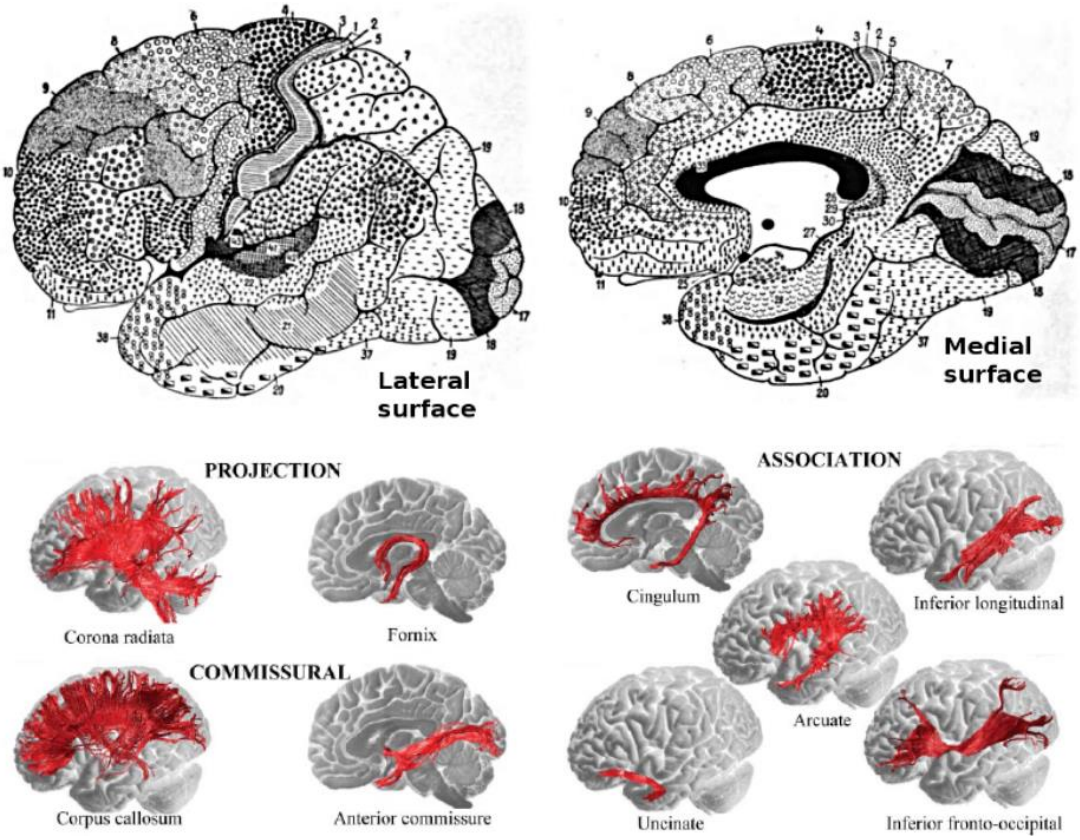


Figure 1-6: (top) Brodmann's diagram of the cerebral cortex with the cytoarchitectonic areas he proposed. (bottom) Meynert's classification of white matter tracts (104).

These investigations led to the exciting notion that the brain was neither uniform nor composed of identical modules. The liver, which is of similar size as the brain and also consists of four lobes, is uniform: (a) microscopic images of its different parts look indistinguishable, (b) "lesions" to its different parts (e.g., partial removal for transplant or because of tumour) lead to exactly the same symptoms and (c) portal venous "projections", which supply blood and nutrients to be "processed" by the liver, reach the different macroscopic parts of it homogenously. The brain on the other hand increasingly emerged as a heterogeneous mosaic of interconnected regions, each presumably serving a distinct purpose, because they all have (a) a different microscopic structure as shown by Brodmann, (b) lesions to different areas led to different symptoms as shown by Broca and Wernicke, (c) each area turned out to have a unique set of inputs and outputs as suggested by the studies of Meynert.

The second development, which might have fuelled the search for connectionist accounts of brain function, was the spread of ‘‘associationistic’’ models of cognitive functioning from the realm of psychology to that of neurology (105). According to such models, the formation of concepts, the recall of memories or the naming of objects all required the associative convergence (or integration) of information from multiple sources. Obvious conduits for this convergence were the white matter axonal pathways being identified by the new neuroanatomy of the late 19th century. Brain damage could thus disrupt higher mental functions either through the destruction of convergence centres where critical associations were formed or through the destruction of conduction pathways that supplied the inputs and outputs of these centres.

Although this thesis will not present any microscopic data, it still is interesting to point out that the module-versus-connection question is also debated on a microscopic and genetic level. A multitude of different cell types reside in the human cerebral cortex. The biggest group comprises (more than 80 percent) glial cells. However, neurologists and psychiatrists trying to explain major disorders such as autism or schizophrenia tend to focus on neurons and differences in their expression of genes and in their development. Interneurons connect other neurons within the same area; whereas projection neurons connect to neurons outside a given region (either to regions in the same hemisphere, in the other hemisphere, the thalamus, the subcortex, the brain stem or the spinal cord, see Figure 1-7). Recent research has implied that major psychiatric and neurological disorders, such as schizophrenia, autism and intellectual disabilities are best understood as spectrums of diseases that have overlapping phenotypes and genetics. Some have proposed that the disruption of local inhibitory circuits might be responsible for many of the clinical features of these disorders (106). Interneurons are responsible for establishing local inhibitory circuits and animal studies have therefore sought to show that these disorders are being caused by specific defects in the development and function of interneurons. Genetic alterations might modify the number or location of cortical

interneurons and affect the excitatory-inhibitory balance in localised cortical circuits. Other researchers however have argued that changes in the expression or translation of genes in projection neurons and the resulting alterations in projection neuron development and migration may be the major cause for intellectual disability and autism (107). Interestingly the latter findings have been linked with ideas of autism, schizophrenia and other major brain disorders as being disconnection syndromes.

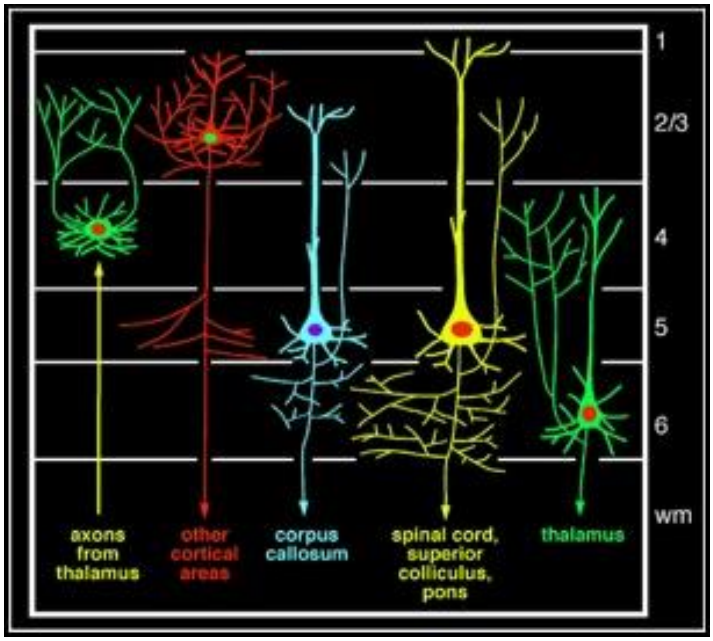


Figure 1-7 depicts different types of cortical projection neurons, which project beyond their area or origin, as a function of the cortical layer in which they are located.

Fundamentally, the view of any scientific subject and even the way many research questions are posed is largely influenced by the methods at hand. Studying brain lesions in neurological patients or non-human primates or inducing “virtual lesions” using transcranial magnetic stimulation (TMS) might bias a researcher’s view more towards a modular and serial view of brain function. Moreover, studying the cytoarchitecture or receptor density of different brain areas emphasizes a brain that is a collection of different modules. Linking neural activity in neurophysiological recordings or the blood oxygenation level dependent signal in functional magnetic resonance imaging (fMRI) studies with specific cognitive contexts might also favour

views of functional specialisation in distinct areas. Other methodological approaches however, such as tracer injections, diffusions-weighted MRI or graph-theoretical descriptions of brain structure and connectivity seem to focus on connectionist views of brain structure, function and dysfunction and envisage processing in distributed networks. Some of these methods shall be described in more detail now.

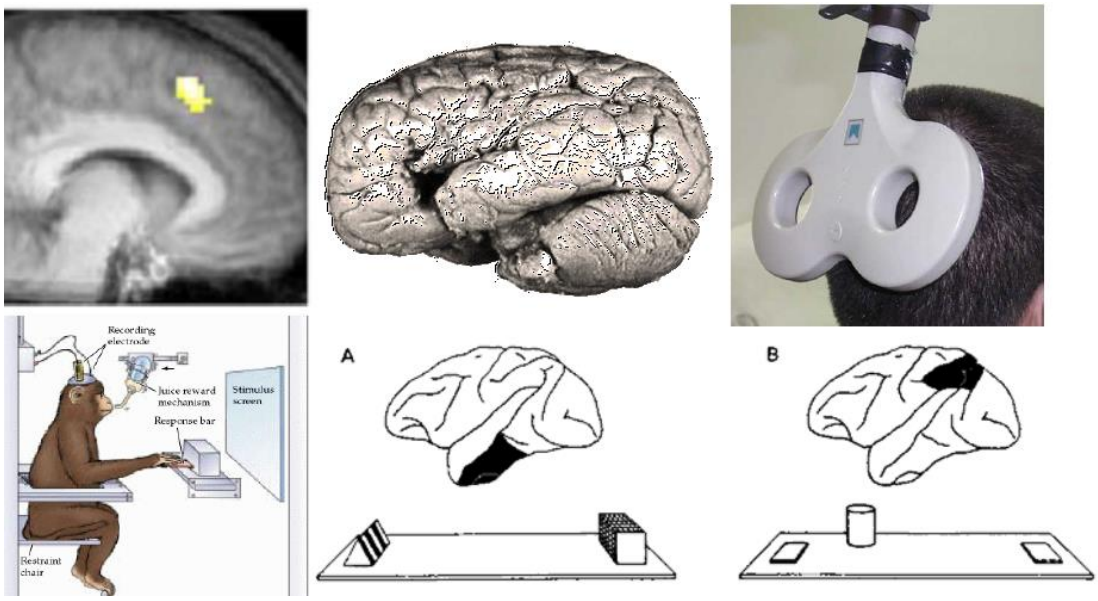


Figure 1-8: A modular view of brain function might be largely influenced by methodological approaches. Right hand site and centre: lesion studies (top centre) in patients, non-human primates (lower centre and lower right) and healthy human volunteers using TMS (upper right) emphasize modular accounts of brain function. Left hand site: Even neurophysiogical recordings (bottom left) and fMRI (top left) might favour such modular accounts to some degree.

1.2.1 Methods for studying brain modules

Some neuroscientific methods have proven to be particularly suitable for studying the brain and its structure and function as a collection of discernible, highly specialized modules. I will briefly introduce some of the methods, which are particularly important for this thesis in the sections below, namely TMS (1) and task-fMRI (1.2.1.2).

TMS is often used to disrupt processing in a given region of the cortex and to relate these disruptions to behavioural changes, such as slowing in reaction times or increased error rates.

Neurophysiological recordings in non-human primates or rodents are used in somewhat similar ways: Researchers attempt to record the activity of single neurons or local field potentials generated by a population of neurons in a given cytoarchitectonically defined area of the brain and relate the activity to the cognitive context in a specific behavioural task performed by the animal. They are sometimes combined with micro-stimulation or pharmacological manipulations of the same neurons/areas in order to further strengthen claims about the functional role of neural activity in a given region.

Cytoarchitectonic studies investigate the overall cellular composition in a given piece of tissue under the microscope and have been used particularly to parse the cerebral cortex into different areas with different architectures. Meynert was one of the first to notice regional differences in the histological structure of different parts of the grey matter in the cerebral cortex. Alfred Walter Campbell used cytoarchitecture in order to obtain a parcellation of the human brain into 14 areas (108). One of the most influential cytoarchitectonic parcellations was obtained by Korbinian Brodmann. He worked on the brains of different mammalian species. His division of the cerebral cortex suggested 52 discrete areas (nowadays often referred to as Brodmann areas). He used numbers to categorize the different architectural areas (109), and believed that each of these regions served a unique functional purpose. More recently these cytoarchitectonic subdivision have been refined and revised, and semi-automated and observer-independent mapping approaches have been developed to discern areal boundaries using multivariate statistical analysis of different cytoarchitectonic features (110, 111). Analysis of neurotransmitter receptor distribution patterns in the cerebral cortex promise even further refinement of current parcellation schemes (112, 113).

However it should be noted that although many of the neuroscientific methods mentioned here as “methods to study the modular organization of the brain” are not solely tuned towards a modular view of brain structure and function. TMS has been used to study “network effects” of focal “virtual lesions” (114). Multi-focal electrophysiological recordings are increasingly

being used to understand interactions of brain regions at the neuron level (115). Moreover, neurophysiological recordings are now also combined with lesions in distant brain areas or pharmacological manipulations in order to understand neural processing in distributed networks. Finally cytoarchitectonic information has also been linked with connectivity. Cytoarchitectonic features of a given region might be related to its "role" in a bigger network of areas (116). Areas with similar cyto-architectonics might also be strongly connected to one another or share graph-theoretical network properties. For example, macroscale degree of connectedness of cortical regions (i.e. the number of traceable long-range connections to other cortical regions) have been linked with microscale layer III pyramidal complexity (measured as dendritic tree size, total spine count, and spine density) (116). This implies that a set of "hubs", i.e. regions with a high degree of connectedness, share some fundamental similarities in their cytoarchitectonic organization. More of these cross-scale studies need to be conducted in order to test the claim that a region's microscale neuronal architecture is related to its role in the global brain network. It would be exciting to see whether, instead of distinguishing brain areas based on cytoarchitecture, it would also be possible to delineate macroscopic networks based on specific microscopic features. It is not implausible to think that not only localised brain areas but whole networks might have cytoarchitectonic fingerprints. Moreover it needs to be established how connectivity and microscopic architecture co-develop and co-evolve in phylogeny and ontogeny. Finally a graph-theoretical network-approach, although very powerful in delineating whole-brain connectional properties, might nevertheless be in danger of telling only half of the story: It may group together several different brain regions because of their similarities in graph-theoretical properties (for example high degree of connectedness). Microscopic studies and other "modular" methods need to help here in order to elucidate how these brain regions differ from one another in terms of anatomical structure, biochemical properties and computational function (117).

1.2.1.1 TMS

D'Arsonval (1896), Beer (1902), Thompson (1910) and Dunlap (1911) had already reported that rapidly changing magnetic fields applied to the retina, the optic nerve or the occipital cortex induced phosphenes (118). In 1965, Bickford and Fremming applied magnetic stimulation to peripheral nerves and muscles and described the observation of twitches in healthy human subjects. A few years later, Merton and Morton excited the human primary motor cortex using transcranial electrical stimulation (TES); TES is a relatively painful procedure of non-invasive brain stimulation. In 1985, Anthony Barker and his colleagues stimulated the human brain over the primary motor cortex using magnetic pulses for the first time (118, 119). The experiment was successful with clearly visible muscle contractions and without much pain and discomfort for the subject. They published a very short report about the experiment a few months later in "The Lancet" (119) (see Figure 1-9).

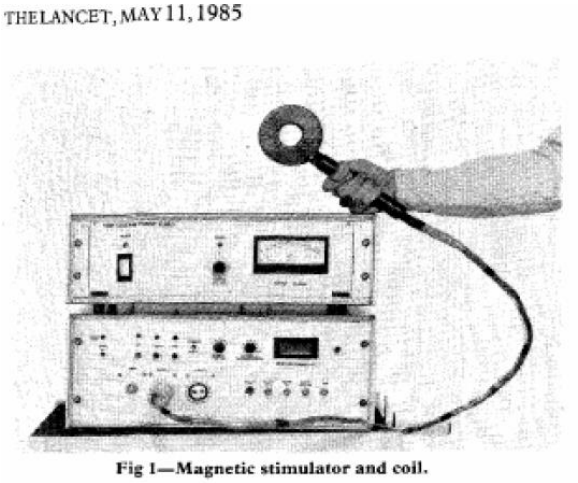

THE LANCET, MAY 11, 1985

Fig 1—Magnetic stimulator and coil.

Figure 1-9: Picture from the original publication by Barker and (119) reporting the first application of TMS to a human brain.

Since these early days of TMS, it has become an increasingly popular tool in neurophysiology and cognitive neuroscience to study motor control, visual perception, perceptual decision-making, language, memory and others; it has also found multiple diagnostic and therapeutic applications in clinical neurology and psychiatry (120). TMS has been combined with other

neuroscientific techniques such as positron emission tomography (PET), fMRI, electroencephalography (EEG), diffusion weighted MRI and single unit recordings in animals (121-124).

In order to generate a single magnetic pulse with TMS, one has to generate an electric current of up to 8 kA using a capacitor and discharge it into a circular or figure-of-eight shaped coil (122, 125). This electric current subsequently produces a magnetic field of up to 3 Tesla (comparable to that of many magnets used for MRI). Such a pulse has a rise-time of about 100 µs and lasts for less than 1ms. The magnetic field induces an electric field in a nearby "coil" (Faraday's electromagnetic induction), which is, in the case of TMS, the brain tissue (see Figure 1-10). The induced electric field is proportional to the change of the magnetic field. It is still debated how neuronal elements are activated by the induced electric field, but at least two mechanisms are proposed. If the field is parallel to the neuronal element, then the field will be most effective where the intensity changes as a function of distance. If the field is not completely parallel to the neuronal element, activation will occur at the bends of the neuron or its axon (121).

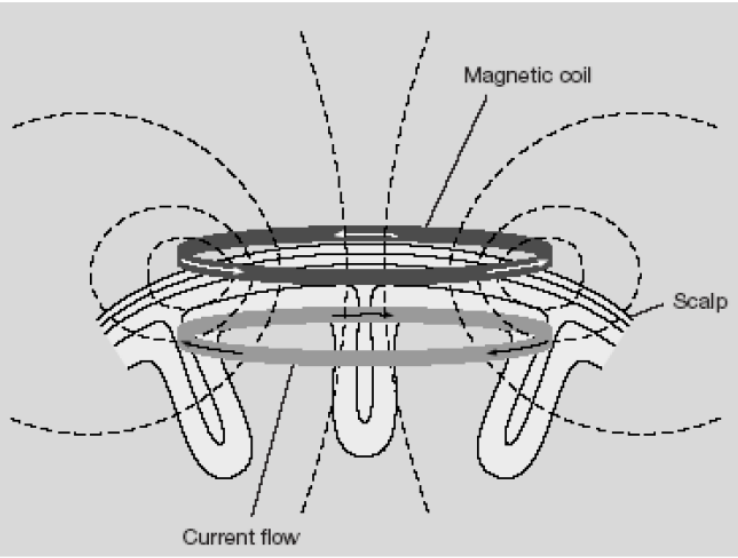


Figure 1-10: A high-current pulse is produced in a coil of wire (dark grey circle) which is placed above the scalp. A magnetic field is induced perpendicular to the plane of the coil (dotted

lines). An electric field is induced perpendicular to the magnetic field and proportional to the change of the magnetic field (light grey circle). Figure taken from (121).

Therefore the effects of the magnetic pulse on neural populations crucially depend on (1) the waveform of the coil's current (most often researchers use biphasic and monophasic waves); (2) the direction of the current with respect to the orientation of the neurons and their axons (electric current in the brain is always directed opposite to the current in the coil); (3) intensity of the stimulation (often expressed as a "percentage of maximum stimulator output" or percentage of the resting motor threshold and adjusted with respect to skull – cortex distance); (4) form and type of the TMS coil (e.g. round coil, figure-of-eight shaped coil, iron-core coils, H-coils).

Among the many techniques used in cognitive neuroscience brain stimulation is one of the few that allows us to interfere actively with human brain function in a spatially and temporally precise manner. Although the experimenter's input to the brain during a TMS experiment is always very similar – a magnetic pulse – the output s/he is actually interested in can be rather diverse, and there are many measures and quantifications of TMS effects on the brain (such as motor or phosphene thresholds, motor evoked potentials, recruitment curves, cortical silent periods or behavioural read-outs after disruption of processing in given regions).

The temporal resolution of TMS when used as "brain mapping" technique can be relatively high compared to other brain imaging tools (such as fMRI or PET). The functional contribution of a certain brain area to a cognitive operation can be elucidated by applying single TMS pulses or short trains of pulses during a cognitive task. It is thought that TMS introduces noise into the neural system and can therefore be used as a "virtual lesion" interrupting information processing in a certain area of the brain at a certain time point during task performance (126). This virtual lesion approach has several advantages over lesion studies in neuropsychological patients and non-human animals. Human brain lesions are often less focal and when the lesion arises as a result of a disease process it is possible that the damaged brain might have

undergone months of plastic reorganisation and compensation of deficits before the time of testing. By contrast the short and relatively focal TMS pulse reversibly interrupts processing of an area in healthy subjects' brains without allowing long-term compensatory mechanisms. Measurements of behavioural impairments can be drawn from reaction times, error rates etc. (just as in classic lesion studies with patients or animals). The virtual lesion (TMS) condition is contrasted with a no-TMS condition, a sham-TMS condition or a virtual lesion condition of another brain area.

Brain stimulation can also be used to induce plastic changes in a brain region. Frequency-dependent plasticity is often used as a model to assess and understand brain plasticity and its underlying mechanism. In some cases it is hoped that these approaches can be used therapeutically. In animal experiments it has been possible to change the efficacy of synaptic transmission by repetitive electrical stimulation of central nervous pathways through mechanisms of long-term potentiation (LTP) and depression (LTD). Researchers have tried to produce related effects in human subjects using repetitive TMS (rTMS) (125, 127). rTMS at slow rates of 0.2 to 1 Hz has been shown to decrease brain excitability, whereas rTMS at faster rates (more than 5 Hz) is supposed to increase excitability. Repetitive stimulation protocols delivering very short, very high frequency trains of pulses (e.g. three pulses of 50 Hz) at theta frequency of about 5 Hz have been shown to yield even stronger excitatory or inhibitory effects (127), depending on whether they are applied continuously (decreased excitability) or intermittently (increased excitability). This protocol is referred to as theta burst stimulation (TBS). Reasonably short TBS applications (40 seconds) have been shown to yield relatively long lasting effects (up to 60 minutes) (127). These longer lasting effects of rTMS applications were hoped to be useful in clinical contexts.

Another interesting field of research has combined TMS with imaging techniques such as fMRI and PET as well as with EEG (124). This allows combining the advantage of TMS being able to

causally interfere with the human brain noninvasively with the ability of fMRI, PET or EEG to disclose whole-brain activity and to discern compensatory mechanisms (114, 128).

Historically, clinicians and researchers had an initial concern about the safety of magnetic brain stimulation. There was a fear of producing short-term and long-term damage to the brain and inducing seizures (120). Now after 30 years of scientific and clinical use, the technique appears to be safe. Single-pulse and paired-pulse TMS protocols are extremely unlikely to cause seizures. Rapid-rate rTMS has a small risk to induce seizures, especially in patients with a history of epilepsy. However, the risk depends on the intensity and the frequency of the stimulus, and careful planning of the study will help to prevent complications.

Earlier I have argued that TMS might favour a modular view of brain function as it is often used in order to briefly disrupt processing or induce longer lasting plastic effects in a given brain region and link these with behavioural changes. However, it needs to be pointed out here that attempts have been made to use TMS as a tool to study connections as well as network architecture and dynamics. The effects of virtual lesions on wider networks of connected brain regions can be discerned using elegant dual-coil approaches (129) or combined TMS–fMRI (114, 130) or TMS-EEG (131). Spike timing-dependent plasticity in cortico-cortical connections has been induced using repetitive paired-pulse TMS (132), and its effects on a wider network of connected and functionally related brain regions has been described using task- as well as resting-state fMRI (133). Promising new lines of research increasingly employ TMS as a tool for studying the brain as an assembly of dynamic networks.

1.2.1.2 Functional Magnetic Resonance Imaging

Functional magnetic resonance imaging (fMRI) is a functional neuroimaging technique using MRI in order to map brain activity. This technique relies on three links, which incidentally span three levels of resolution.

(1) The first is the link between neural activity in a given brain region (which is thought to reflect "processing" of information) with blood flow. In the 19th century the Italian physiologist Angelo Mosso tried to demonstrate the redistribution of blood during emotional and intellectual activity (134). In 1890, Charles Roy and Charles Sherrington first experimentally linked brain function to blood flow. Glucose is the primary source of the brain's energy supply. However nerve cells do not store much glucose. Action potentials are particularly energy consuming, because repolarization after depolarization is achieved via ATP fueled Na+/K+ exchangers. ATP is regenerated from substrates like glucose or via oxidative phosphorylation, which requires oxygen. An increase in neural firing in a specific region is therefore met with an increase in blood flow to this region to transport more glucose and oxygen in the form of oxygenated hemoglobin molecules in erythrocytes. The neural system provides feedback of its need for more glucose to the vascular system partly by the releasing glutamate (as part of neural firing). Glutamate affects nearby astrocytes, in turn causing a change in calcium (Ca2+) concentration. This releases nitric oxide (NO) at contact points to arterioles causing vasodilation (34). With a delay of 1-2 seconds after increase in neural firing this now starts to increase the rate of blood flow and the diameter of blood vessels. The blood-flow change is thought to be localized within 2 to 3 mm of where the change in neural activity takes place. Usually the increase in oxygen supply outweighs the increased oxygen consumption and therefore a net increase in oxy-hemoglobin (Hb) and a decrease in deoxy-hemoglobin (dHB) is observed. However note that this ratio of change in oxygen-consumption on the one hand and change in oxygen inflow on the other hand might differ between regions. It has been argued that inflow by far outweighs consumption in regions such as the amygdala, basal ganglia, thalamus and cingulate cortex (regions that are thought to play a role in rapid responding). On the other hand lateral frontal and lateral parietal lobes seem to exhibit different ratios (and are also thought to be somewhat involved in deliberate responding and planning) (34). This would have major implications for measuring activity in these regions (and I will touch upon a

similar point – the question of different time-courses of representation and processing – below). The hemodynamic response peaks at about 5 seconds after the neural firing rate increase. If increased neural firing is maintained blood flow plateaus. If neural firing decreases again, the blood flow will fall below the original baseline level, a phenomenon sometimes called “undershoot”.

(2) The second link that fMRI crucially relies on was discovered by Linus Pauling and Charles Coryell. The two chemists reported in 1936 that oxygen-rich blood with Hb was weakly repelled by magnetic fields (diamagnetic), while oxygen-depleted blood with dHb was attracted to a magnetic field (paramagnetic).

(3) The third link underlying fMRI was discovered by Seiji Ogawa in the early 1990s. He suggested that one might be able to use the dHb’s and Hb’s different responses to magnetic fields in imaging. As dHb is paramagnetic it slightly distorts the surrounding magnetic field induced by an MRI scanner. This causes the surrounding protons to lose transverse magnetization faster and therefore leads to T2* signal loss (similar effects play a role in diffusion-weighted MRI and will be discussed again below). Hence T2* weighted brain images show more signal in areas where blood is highly oxygenated and less where it is not. The blood-oxygen-level-dependent (BOLD) signal derives from this idea.

Simply put the links between (1) neural firing and blood flow, (2) blood-flow/oxygenation and magnetic properties and (3) magnetic properties and T2 signal loss lie at the heart of BOLD fMRI.

The temporal resolution of fMRI – i.e. the smallest time period of neural activity reliably separated – is thought to be on the order of seconds. Nowadays it is not so much constrained by sampling time (i.e. the TR- the time the brain is “imaged” once; now often on the order of 1-3 seconds), but rather on the physiology of the blood-flow response to neural firing. This is an important aspect of fMRI (it is not only limiting but also has crucial advantages relating to the “smoothing out of cognitive processes”) as many cognitive processes have largely differing

time-courses: Consider for example visual processing. Within milliseconds images will have been registered by the retinal cones and rods. The lateral geniculate nucleus relays the information to the primary visual cortex within tens of milliseconds. Processing of visual information lasts for more than 100ms. A fast response can be made within 200 ms or less. Some emotional or cognitive aspects of the image may last minutes or hours. Learned changes may last days, months, or years. We aim to capture all of these neural processes measuring the BOLD response, which lasts over 10 seconds, peaking at 4 to 6 seconds.

The spatial resolution of fMRI on the other hand is thought to be in the order of millimetres. This largely depend on the voxel-size of the MRI-acquisition protocol. Spatial resolution increases with smaller voxels. But these may take longer to acquire if data are to be collected from across the whole brain and each voxel may have less signal. Moreover, as is the case for temporal resolution, the physiology of the hemodynamic response limits the spatial resolution of BOLD fMRI.

Researchers have tried to elucidate the link between neural signals and BOLD using neurophysiological recordings, EEG and more recently optogenetics in combination with fMRI. Experiments in rats, humans and monkeys suggest the BOLD contrast reflects mainly the inputs to a neuron and the neuron's integrative processing within its body, and less the output firing of neurons (135-137). Moreover studies have tried to verify the spatial resolution of fMRI (arguably one of its major strengths) showing that for example different thalamic nuclei can be reliably dissociated from one another with the appropriate tasks (e.g. the lateral geniculate nucleus from the nearby pulvinar by visual stimulation). Finally studies have tried to show the linear summation of the BOLD signal over closely spaced bursts of neuronal firing, which is a core assumption of, for example. event-related fMRI designs (138).

The design of fMRI-experiments and the way data is pre-processed and analysed owes much to the considerable amount of noise in MR data, i.e. changes in the signal that are caused by effects not of interest to the study. The five main sources of noise in fMRI are thermal noise

(e.g. heat can distort currents in the receiver coil), system noise (magnetic field inhomogeneities, scanner drift), physiological noise (head and brain movements from e.g. breathing, heart beats), random neural activity (neural activity independent of the experimental manipulation) and differences in mental and behavioural strategies to solve a given task, both within and between subjects.

For human fMRI-studies pre-processing and analysis is largely standardised starting with k-space reconstruction and slice-timing correction. After that the data is usually corrected for head motions. Distortion correction (e.g. using fieldmaps) can account for field nonuniformities. Coregistration with individual's structural MR scans and a group-template (such as the Talairach or the MNI template) are a prerequisite for group-level analyses. Temporal filtering removes frequencies of no interest from the signal. Smoothing (spatial filtering) averages the intensities of nearby voxels to produce a smooth spatial map of intensity change across the brain or region of interest. If the true spatial extent of activation matches the width of the smoothing filter this process improves the signal-to-noise ratio.

When analysing fMRI data it is common to consider each voxel separately within the framework of the general linear model (GLM). This model assumes, that the BOLD response is equal to the scaled and summed version of the task events at a given time point. A design matrix specifies which events are active at any time-point. The events in the design matrix are convolved by a canonical hemodynamic response function. These convolved regressors are related the observed BOLD response, allowing for scaling by the weights of the GLM for each event. The GLM model, however, does not take into account the relationships between multiple voxels. Neural representation and processing are likely to not solely rely on the average firing rate changes (related to the idea of frequency coding) but potentially also to the joint activity pattern of distributed neural populations (related to the idea of population coding). Whereas GLM analysis draws on a voxel's or region's signal amplitude being either higher or lower for one condition than another, newer statistical models such as multi-voxel

pattern analysis (MVPA), utilize unique contributions of multiple voxels within a voxel-population to a measureable pattern of activity. MVPA employs machine-learning approaches and is usually implemented by training a classifier to distinguish trials of different conditions within a subset of the data. The trained model is then tested by predicting the conditions of the remaining (independent) data.

All of these approaches are now becoming fairly standardized in young, healthy human volunteers. However there may be reasons to use fMRI in animals, such as rats or rhesus macaques. One goal of such studies can be to compare brain function across species. Other researchers attempt to combine fMRI with intracranial recordings (e.g. single unit recordings, fast-scan cyclic voltammetry), stimulations or circumscribed brain lesions. Moreover animals can be scanned repeatedly and over long periods of time for longitudinal studies, which is often not feasible for human subjects. Experimenters may also be able to control long-term environmental variables (such as the social network size (139)) in animals, which is impossible to do in humans. However fMRI in animals remains challenging for several reasons. Even head-fixed animals move considerably more in the scanner than human volunteers. Note that movements not only of head and brain but also of the body relative to the (fixed) head may cause increased artefacts (e.g. ghosting - displaced reduplications of the image in the phase encoding direction). Moreover the fact that many animals have smaller brains than humans may increase heart-rate and breathing related noise, movement-related artefacts and noise due to field inhomogeneities (from air-filled sinuses) relative to the size of the brain and the areas generating the signal of interest. Finally much of the pre-processing and analysis although conceptually similar to human fMRI is still less well established even without brain lesions or stimulation devices.

FMRI is also routinely used in clinical settings, for example in order to plan neurosurgery (for example, to determine the location of eloquent cortex to avoid severe disability when surgical lesions are made). An interesting area of clinical research is the study of diseased or injured

brain functioning to disentangle causes of pathological behaviour and compensatory mechanisms. However it may seem that clinical use of fMRI still lags behind use in cognitive neuroscience. This might be because patients are often more difficult to scan than young healthy volunteers. Moreover brain pathology (tumours and lesions) may render pre-processing and statistical analysis more difficult. Finally "normal" behaviour and brain function may be less variable than pathological brain function, further obstructing statistical inference. However there is reason to believe that fMRI can be a powerful tool for understanding "functional" brain pathology and maladaptation as well as compensation and recovery.

FMRI experiments are often conducted either as block-designs, event-related designs or with the help of computational models. In a block design, two or more conditions are alternated in series of trials. Each block will have a specific duration of a certain number scans and present one single task condition. The different blocks are designed to differ only in the cognitive process of interest, and can then be contrasted to reveal "activation" or "de-activation" relating to this cognitive process. Block designs offer considerable statistical power, but less flexibility particularly when studying temporally unfolding processes such as learning. Event-related designs allow for more flexibility when investigating complex cognitive processes. However, the statistical power of event related designs is usually smaller. Both block and event-related designs are based on the assumption that cognitive processes are recruited selectively in different conditions and that the associated neural activity (and blood flow) add linearly. Rather than defining the regressors for fMRI analyses based on task-context and behavioural variables some researchers have used mathematical models (either fitted to behavioural data or just based on optimal mathematical solutions) to derive "hidden variables", which in turn can be regressed against fMRI data. Not surprisingly this approach has been particularly successful in neuro-economics, which as discussed above proposes hidden variables such a value-representations to guide behaviour (i.e. choices).

All of these different designs are somewhat susceptible to the choice of baseline – the condition in which a certain cognitive process of interest is thought not to be present. It is tempting to use rest (i.e. absence of any task) as a safe baseline, but as I will discuss in Section 1.2.1.4 the brain is never completely at rest and some regions have been reported to be more active during supposed "resting conditions". Moreover all of these designs usually assume a similar temporal profile of all neural processes and the respective hemodynamic response. But even if hemodynamic response to increased neural firing is similar across regions (and evidence suggests this is not necessarily true) the time-course of processing or the temporal intervals over which regions integrate may be vastly different. I have briefly touched upon similar ideas when discussing the role of areas V5/MT and LIP in perceptual decision-making. I also suggested that adaptive learners should change their learning rate to the volatility of the environment: in very changeable environments they should emphasise only very recent outcomes whereas in stable environments more distant past outcomes might also be considered. It is plausible to think that reward history may be represented with different time constants in the brain and presuming the same temporal profile of neural processing and the same hemodynamic response function for all regions may therefore be inappropriate. A study by Murray et al (CITATION Murray Wang a hierarchy of intrinsic) suggests that different brain areas are not only specialised for distinct types of information and computations but also for different temporal ranges from short time-scales in perceptual areas to longer time-scales in prefrontal cortex. Another study by Neta et al (CITATION Neta Petersen temporal activity) identified different temporal response profiles to feedback in a wide array of tasks, suggesting different neural time-courses for feedback processing in the brain.

I have discussed task-fMRI here as a method to investigate brain function from a "modular perspective". It is not uncommon to interpret the results of fMRI studies as showing that one specific brain region carries out a well-defined cognitive process or represents a presumed computational variable (CITATION O'Reilly and Mar Grek Letters). However sometimes

methodological approaches that built on fMRI data arrive at “network-accounts” of cognitive processes. One example are meta-analyses of fMRI data. Some critics of fMRI studies often raise concerns about problematic statistical analyses based on low-power and small-sample sizes. It is hoped that some of these concerns may be addressed by running meta-analyses on big databases of imaging results such as the “brainmap database” (which contains 2638 papers, 12661 experiments, 100 paradigm classes, 52524 subjects and 100914 locations) and the “neurosynth” initiative (containing 386455 activations reported in 10903 studies). Some meta-analyses of specific cognitive concepts using such databases often turn out a number of different regions. For example a meta-analysis on “valuation” as carried out by McGuire and Kable (70) suggests a role not only for vmPFC but also of posterior cingulate and perigenual cingulate cortex in valuation and value-comparison processes. This may be taken to imply that there is not a single value-comparison module in the human brain but that a network of brain regions engaged in this process.

Another example are analysis techniques applied to fMRI data to study how “functional connectivity” (i.e. the functional connectivity denotes the symmetrical statistical association or dependency between elements of the system (140)) between brain regions is affected by cognitive contexts. Psycho-Physiological Interaction analysis (PPI) is a brain imaging analysis that tries to estimate the functional connectivity between a brain region and the rest of the brain with relation to the performance of a particular psychological task. Dynamic causal modelling (DCM) tries to go even further in trying to test different graphical models of effective connectivity between cortical or subcortical nodes in order to infer their causal interactions during cognitive tasks.

FMRI currently dominates the field of neuroimaging and, perhaps more generally, the wider field of cognitive neuroscience. Since its invention at least 300,000 studies have been published using fMRI. The advantages of fMRI are apparent: The technique is non-invasive (does not even require injection of contrast agents), reasonably cheap and very powerful at

imaging the whole-brain at high spatial (although limited temporal) resolution. Nevertheless some researchers have questioned whether fMRI has told us anything about the brain (141). A fraction of fMRI studies have taken established paradigms from cognitive psychology and tried to carefully localise differences in blood flow during different task contexts. They may be criticized for not telling us anything of fundamental interest to cognitive neuroscience – how a specific cognitive process is mechanistically implemented by neural tissue. They may merely manage to reveal some "anatomical details". Another set of studies has tried to devise very complex psychological paradigms and accompanied them with very sophisticated computational models to then loosely relate them to fMRI data, because blood flow changes somewhere in the brain may verify such models. Again they can be criticised because although they may derive a mechanistic account of human cognition they fail to offer fundamental insights into how they are implemented in the brain. One may be forced to admit that decision neuroscience has produced both types of studies. However there is reason to be optimistic that fMRI, with its ability to image the whole brain non-invasively with ever increasing quality and resolution will yield further fundamental insights into how cognition is actually implemented in the brain in the near future.

1.2.2 Methods for studying brain connectivity

I have argued that advances in neuroanatomy as well as associationist theories of representation of information have fuelled a "connectionist turn". In this section I will introduce some neuroanatomical methods that allowed the study of connections. I will then turn to two methodological approaches that are currently used to discern connectivity and brain networks: DW-MRI and resting state fMRI. These two techniques are of particular importance for this thesis.

Even before Wernicke's school of "disconnectionism", researchers investigated the organization of nerve pathways. The father of modern anatomy – Andreas Vesalius –

dissociated the superficially located grey matter from the harder, white substance below. In gross dissections he noted the adhesion of the two hemispheres via the corpus callosum (see Figure 1-11) and described several central white matter structures such as the corona radiata, the internal and external capsule and the cerebellar peduncles (142).

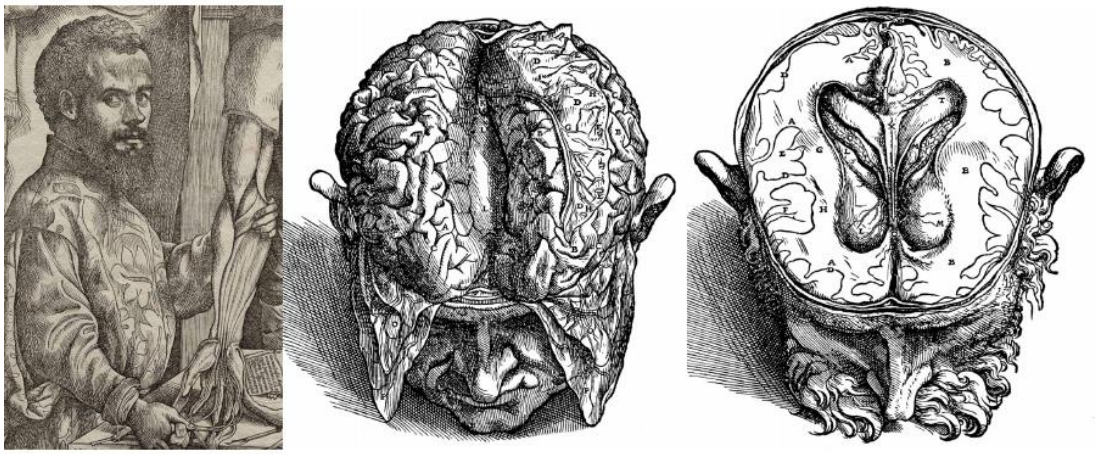


Figure 1-11: Andreas Vesalius conducted gross dissections of brains and defined some of the major white matter structures in “De Humani Corporis Fabrica” (1543).

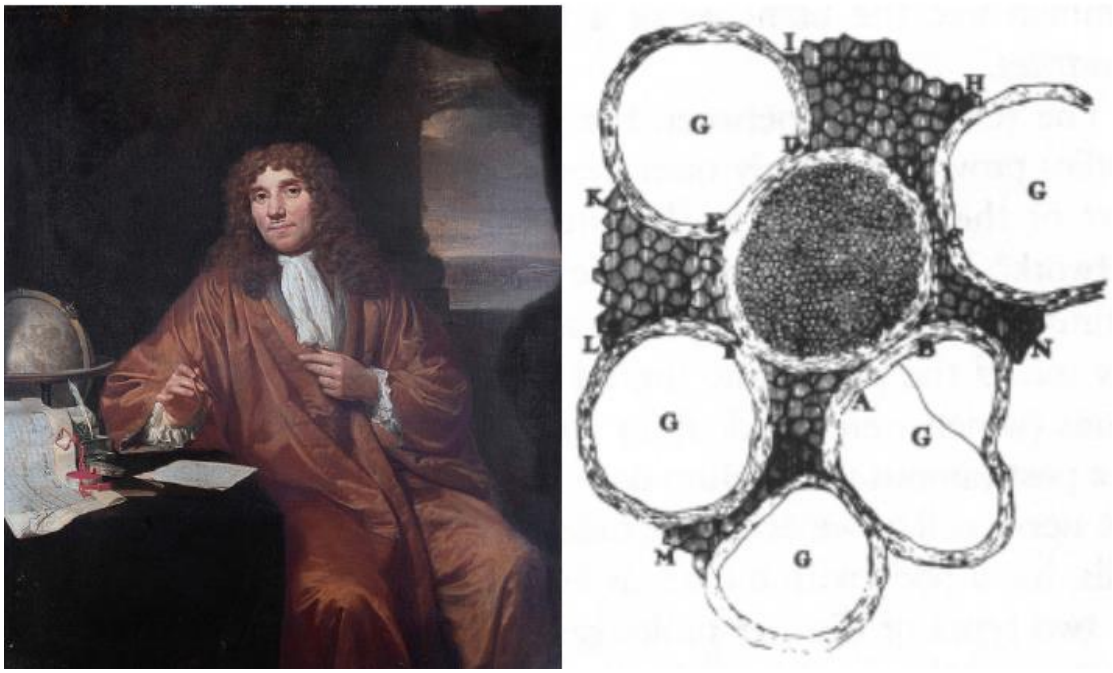


Figure 1-12: Antonius van Leeuwenhoek was among the first to study the microscopic structure of nerve fibres. He described his discoveries in a letter to the Royal Society if London in 1674.

Marcello Malpighi and Antonius van Leeuwenhoek were among the first to study the microscopic structure of nerve fibres. In the 18th century several developments facilitated the microscopic investigation and tracing of white matter fibres: Improved tissue fixation techniques (through e.g. chromic acid and formaldehyde), the invention of the microtome for precise tissue sectioning, the invention of new stains (such as myelin stains) and improved microscopes provided new opportunities for tracing major fibre pathways in the brain. These advances allowed for example Paul Emil Flechsig to describe the developmental patterns of fibre myelination (using myelin stains). Wernicke, in addition to his contributions to neurology, published several atlases of the human brain based on myelin-stained materials. However, even with myelin stains the tracing of small and complex fibre pathways through areas of crossing fibres remained difficult and poorly myelinated and non-myelinated fibres were very difficult or impossible to trace (142).

Figure 1-13: The development of better microscopes (not shown), microtomes (upper left, Caldwell automatic microtome by Cambridge Scientific Instrument Co. in 1884), new tissue fixation techniques (not shown) and staining techniques gave the microscopic study of white matter fibres a new edge. Myelin Stains selectively stain myelin (lower left, note the difference in the amount of myelinated fibres in rodents, chimpanzees and humans). Camillo Golgi (upper

right) developed a silver stain (lower right) that allowed the staining of neurons and their extensions.

It is quite surprising to learn that the existence of nerve fibres was an accepted concept even before the discovery of the neuron soma itself and that the question of whether the two were actually associated entities was debated among neuroscientists. Camillo Golgi's silver staining technique solved this puzzle, because it allowed the staining of neuronal cell bodies together with their dendritic and axonal extensions. This paved the way for Ramon y Cajal's famous neuron theory (142).

Augustus Waller was the first to demonstrate anterograde (orthograde) degeneration – the disintegration of an axon and its myelin sheath distal to a lesion of that axon. Later, von Gudden described retrograde somal degeneration – the decomposition and death of an axon and its cell body proximal to a lesion. Trans-synaptic anterograde and retrograde degeneration describes the death of cells that lie downstream and upstream of an affected neuron, respectively. The principles of anterograde and retrograde degeneration are illustrated in Figure 1-14 and Figure 1-15 .

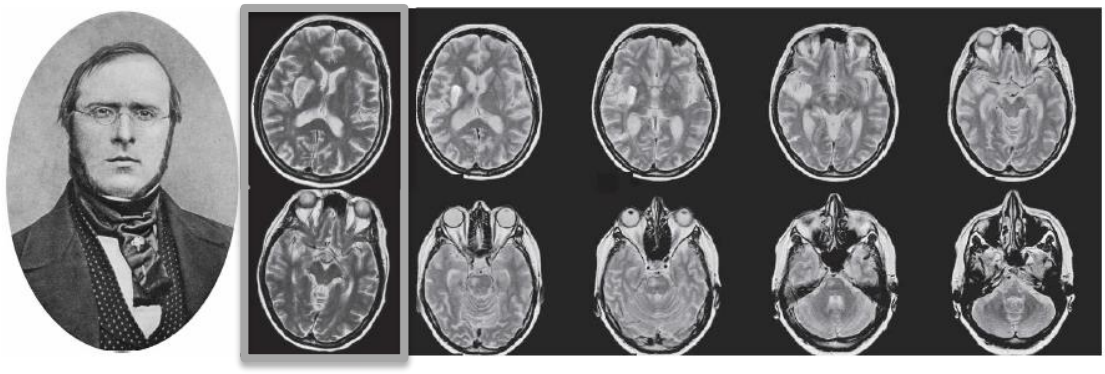


Figure 1-14: August Waller (left) was the first to describe anterograde (now often called "Wallerian") degeneration. The case on the left illustrates Wallerian anterograde degeneration: In the acute stage (the two leftmost images, grey frame) of this basal ganglia / internal capsule stroke there is signal enhancement only in the ischemic area (at the level of the internal capsule – upper image – but not on the level of the crura cerebri – lower image). Three months later affected fibres have disintegrated and signal hyperintensity can be found all the way down to the brainstem.

Following the discovery of anterograde and retrograde degeneration in the central nervous system, anatomists started to use experimentally induced brain injury in combination with microscopic delineation of myelin sheath decomposition as an experimental tracing technique. This approach required induction of a localised lesion to a specific brain area followed by a survival period to allow for degeneration to take place. Tissue fixation and myelin staining allowed for the precise localisation of the degeneration of myelin and therefore for tracing microscopically verified fibre pathways over long distances. In the 20th century the development of stains that specifically stain degenerating axons allowed a further refinement of this approach.

However these techniques were largely replaced, when contemporary tract tracing techniques using retrograde or anterograde intraneural transport of compounds such as horseradish peroxidase or radiolabelled amino-acids were introduced in the 1960s and 1970s. These approaches capitalize on the neuron's own axoplasmatic transport system which is usually used for the transport of organelles, vesicles, macromolecules, enzymes and neurotransmitters. Anterograde transport can operate on a rate of 100-450mm/day and retrograde transport can operate at 100-200mm/day. An anatomical tract tracing experiment is usually carried out in the following way: The skull of the experimental animal is opened and the tracer is injected into the area of interest. The compound is taken up by the neurons in that area via endocytosis and then transported either anterograde or retrograde during a post-injection survival period (usually on the order of 2-5 days). After that time the tissue is fixated and stained, and the tracers are made visible using for example chemical, immunohistochemical or autoradiographic processing. The development of a number of different fluorescent tracing compounds has allowed for multi-tracer experiments using up to five different tracers in a given experimental brain. Figure 1-15 depicts an example of such combined tract tracing experiments.

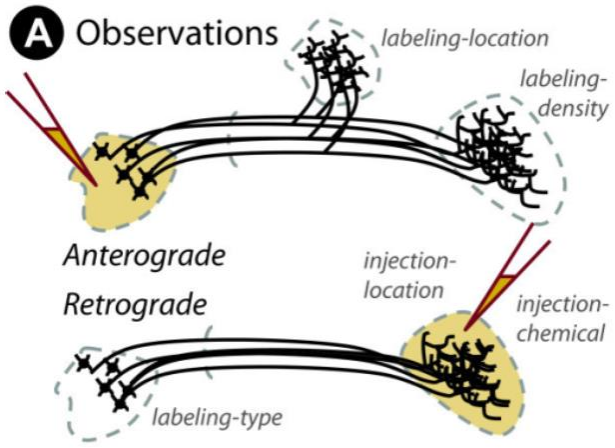


Figure 1-15: Basic principles of tract tracing experiment. Tracer is injected in brain region of interest. It is then taken up by neuron somas and transported anterograde (upper part) or by axons and axon terminals and transported retrograde (lower part). It is therefore possible to trace afferent or efferent projections of a given brain region.

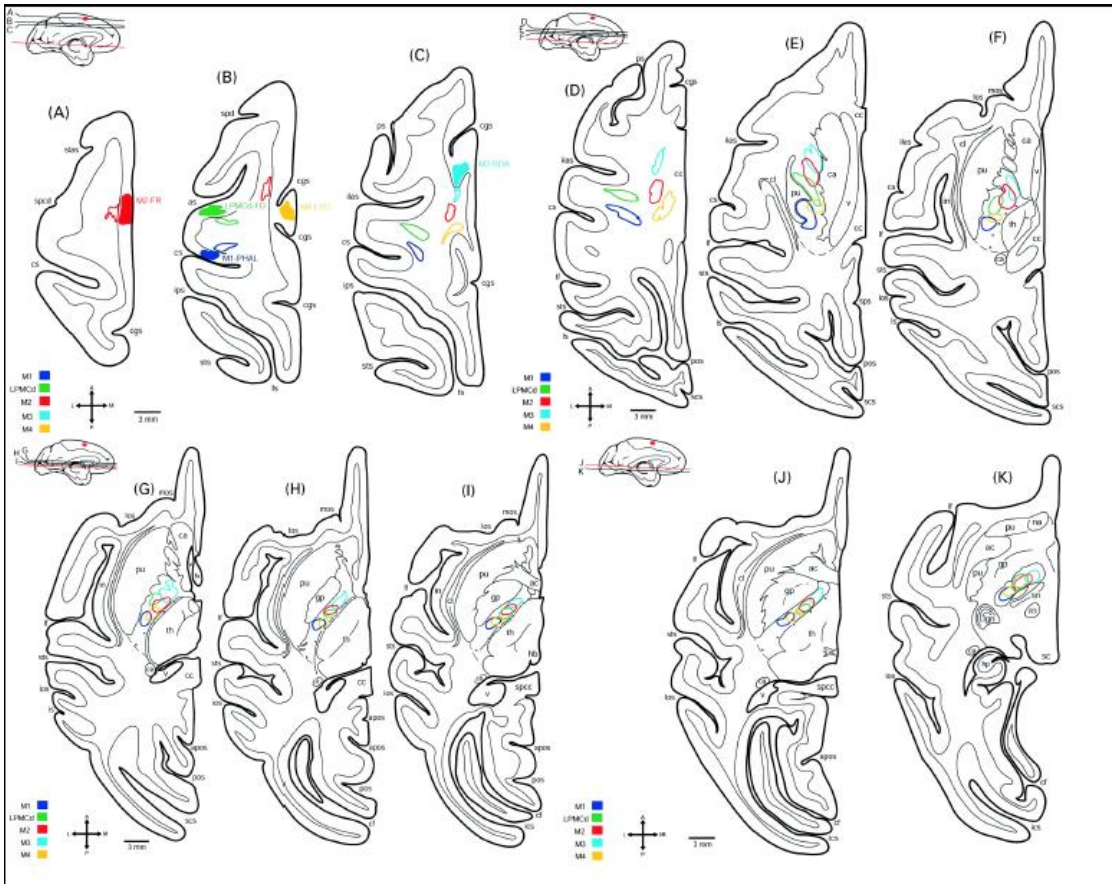


Figure 1-16: This complex tract tracing experiment by Morecraft et al. (143) traces from five arm representations in M1, pre-PMd, SMA, RCZ and CCZ.

The advent of transneural tracers using viruses (such as herpes viruses or the rabies virus) in the 1980s and 1990s opened up new possibilities for even more complex, long range trans-synaptic tracing experiments. Contrary to conventional tracers these viruses replicate in recipient neurons and therefore allow intense transneural labelling.

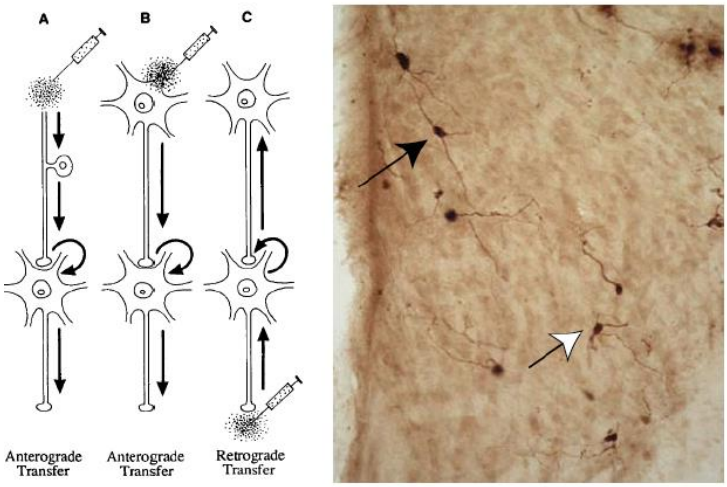


Figure 1-17: Transneural tracing using viruses. The sketch on the left shows anterograde trans-synaptic tracing starting from pseudo-unipolar or multipolar neurons and retrograde trans-synaptic tracing. The picture on the right shows trans-synaptically traced neurons in the brain stem using the rabies virus (144).

In 2007, Lichtman and Sanes proposed a new technique for “staining” neurons and their extensions in a way that renders them easily discernible from neighbouring neurons. This technique was termed “brainbow” and relies on random expression of different ratios of red, green, and blue derivatives of green fluorescent protein in individual neurons. This allows flagging each neuron with a distinctive colour. It is hoped that this technique will yield major contributions to the study of neural connections in the brain.

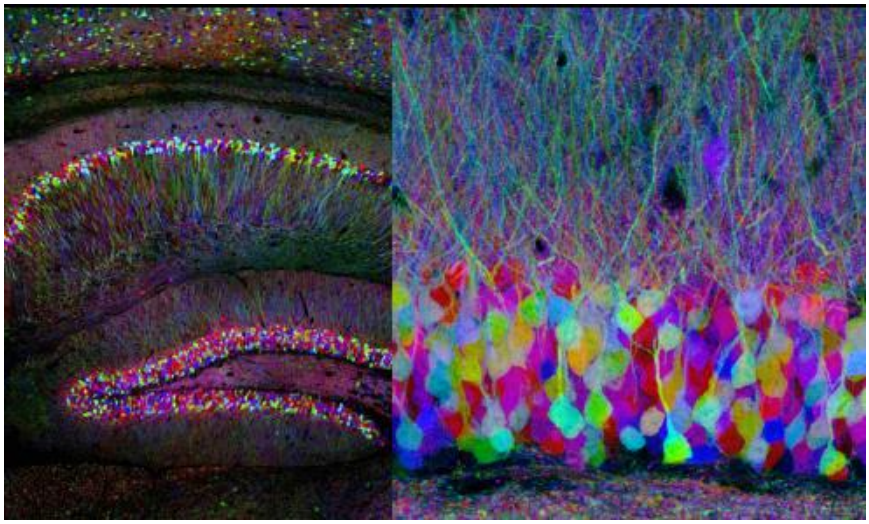


Figure 1-18: Brainbow-stained neurons in the hippocampus (left), the dentate gyrus (right) (145, 146).

Although tracing techniques I have described seem very powerful tools for discerning the organization and layout of major white matter fibre pathways down to the level of specific brain areas or even neurons, they all rely more or less on optical microscopy. This in turn implies that most of them are not able to make any strong conclusions about true neural connections – via electrical or chemical synapses. Only electron microscopy, which uses a beam of accelerated electrons as a source of illumination allowing for much higher resolutions of up to 50 picometre, can therefore make such strong claims. It took researchers more than a decade to map out the complete connectome of the small nematode worm *Caenorhabditis elegans* with its 302 neurons and all its synaptic connections. Researchers now hope to be able to scale up these efforts to more complex nervous systems using highly automated and paralleled slicing and imaging techniques as well as advanced image recognition software (in order to delineate the structures on all the images obtained).

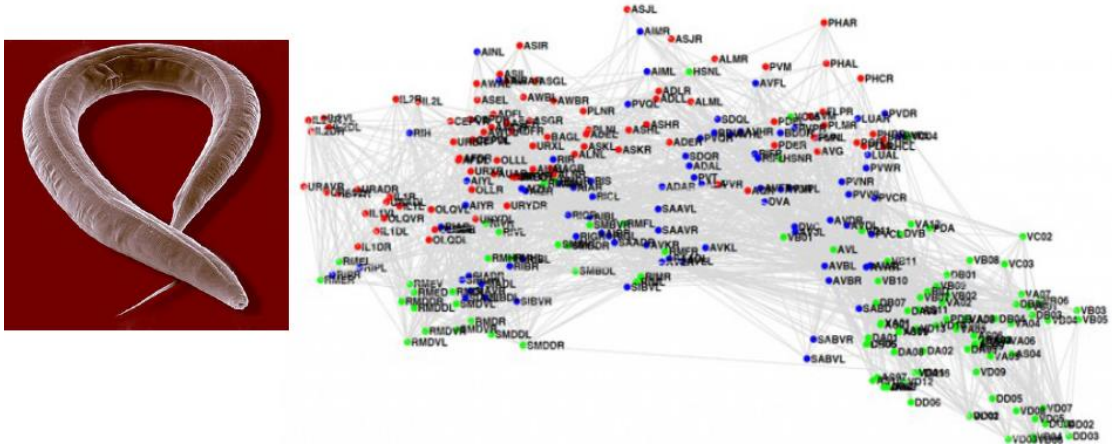


Figure 1-19: Graph depicting the whole connectome of C. elegans.

Finally, MRI based techniques are increasingly being used to study brain connectivity on a macroscopic scale. Cortical thickness correlation relies on the covariance of cortical thickness of highly connected areas across a large number of structural MRI scans. DW-MRI measures the diffusivity of water molecules in the brain and capitalizes on the fact that water diffuses more easily along major white matter fibre bundles as opposed to perpendicular to them. This allows the study of white matter architecture and connectivity via major fibre bundles on a single subject level. Resting-state fMRI derives from the observation that even during rest the there are fluctuations of the BOLD signal. Areas in which the fluctuations measured at rest are correlated are argued to be exhibiting “functional connectivity”. There is evidence that functional connectivity reflects anatomical connections between regions although in some cases the connections may not be direct (147).

When briefly skimming through the history of brain connectivity investigation, several patterns become apparent. Studying structural connectivity has become less less invasive and more refined in its spatial focus. Gross anatomical dissections of major white matter structures, such as the ones carried out by Vesalius and other anatomists, usually cut away other anatomical structures and eventually destroyed much of the brain’s normal architecture in order to follow a single pathway. Degeneration-based tracking approaches still induced large cortical lesions, whereas tracer injection studies caused only limited damage to the tissue. However, both

methods relied on neurosurgery and resulted in the death of the experimental animal. MRI-based approaches are non-invasive and can therefore be carried out in living humans, thereby allowing for the first time a correlation of connectivity with behaviour and cognition. Another observation is that recent methods to study connectivity are increasingly allowing for "hypothesis-free" or data-driven investigations of connections and networks. Gross anatomical dissections require a decision about which major white matter structure to dissect. Tract tracing using degeneration-based approaches or chemical tracers still necessitates a decision on which region to lesion or inject. The investigation of networks of connections or "connectomes" using brainbow, electron microscopy or MRI-based methods could allow for less hypothesis-driven and more data-driven ways of studying the brain and its networks and connections. I will return to this idea in the Chapter 4, which attempts a data-driven way of parcellating the major association fibres of the cerebral cortex. I will now turn to two non-invasive MRI-based methods for studying macroscopic connectivity as they are most important for this thesis.

1.2.1.3 DW-MRI

Diffusion (derived from the Latin diffundere for "to spread out") describes a mechanism of mass transport with net movement of a substance (e.g. an atom, ion or molecule) from a region of high concentration to a region of low concentration. Fick's first law relates the diffusive flux in a fluid to the existing concentration field stating that the magnitude of flux is proportional to the concentration gradient:

$$J = -D \Delta C$$

with J being the net particle flux, ΔC being the particle concentration gradient and D being the diffusion coefficient, which is determined by size of the diffusing molecule, temperature, and

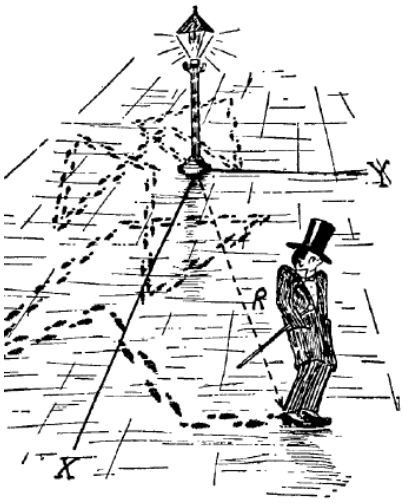
microstructural features of the environment. These microstructural features are of course what many DW-MRI experiments are interested in.

Fick's law describes the diffusive flux of particles along concentration gradients in a manner that is analogous to how heat flows from areas of high temperature to areas of low temperature. On a molecular level diffusive flux solely derives from collisions of particles. This suggests that even without a concentration gradient and in thermodynamic equilibrium diffusion should still occur (if the temperature is higher than absolute zero). This was quite surprising to some, who held Fick's law to imply that diffusion should stop under these conditions. Although net flux vanishes, microscopic motions of molecules still persist and are a prerequisite for DW-MRI.

The Scottish botanist Robert Brown had described the random motions of pollen grains under the microscope and had attributed these to some life force that was causing them. He was disappointed to see the same random movements when studying dust and other dead matter. These random movements, however, are still called "Brownian motion".

Figure 1-20: Adolf Fick (left) described with his “first law” the phenomenon of diffusion. Robert Brown (top left) observed random motions of grains and pollen under the microscope. These random motions are still called “Brownian Motion”.

A mathematical description of this process – also referred to as the “random walk” - was missing. Karl Pearson (1905) asked in a letter to the journal Nature for help with the solution to “The Problem of the Random Walk”: Rather than going home after a long evening in the pub a man starts from a lamppost walking x yards in a straight line in one direction, then turns around through any angle and walks another x yards in a second straight line. Pearson asks for a probabilistic framework to describe the distance of this man from the lamppost after a given amount of time / walks.



The Problem of the Random Walk.

CAN any of your readers refer me to a work wherein I should find a solution of the following problem, or failing the knowledge of any existing solution provide me with an original one? I should be extremely grateful for aid in the matter.

A man starts from a point O and walks l yards in a straight line; he then turns through any angle whatever and walks another l yards in a second straight line. He repeats this process n times. I require the probability that after these n stretches he is at a distance between r and $r+\delta r$ from his starting point, O.

The problem is one of considerable interest, but I have only succeeded in obtaining an integrated solution for *two* stretches. I think, however, that a solution ought to be found, if only in the form of a series in powers of $1/n$, when n is large.

KARL PEARSON.

The Gables, East Ilsley, Berks.

Figure 1-21: In a letter to the journal “Nature” Karl Pearson asks for help or insight to solve “The Problem of the Random Walk”.

Coincidentally at the same time Albert Einstein was seeking evidence that would imply the existence of atoms and turned to “Brownian motion” (although he was unaware of Brown’s observations). Einstein developed exactly such a probabilistic framework in order to describe the motion of a collection of particles undergoing diffusion, introducing the concept of “displacement distributions”. The displacement distribution is a Gaussian function whose

width is determined by the diffusion coefficient mentioned previously. It quantifies the fraction of particles that will traverse a certain distance within a particular timeframe:

$$|x^2| = 2\,D\,\Delta$$

with $|x^2|$ being the mean squared displacement, D being the diffusion coefficient, and Δ being the diffusion time. This framework is still the foundation for how we often model DW-MRI data, for example when fitting a so-called "diffusion tensor", which can be thought of as a three-dimensional displacement distribution of water molecules or their protons.

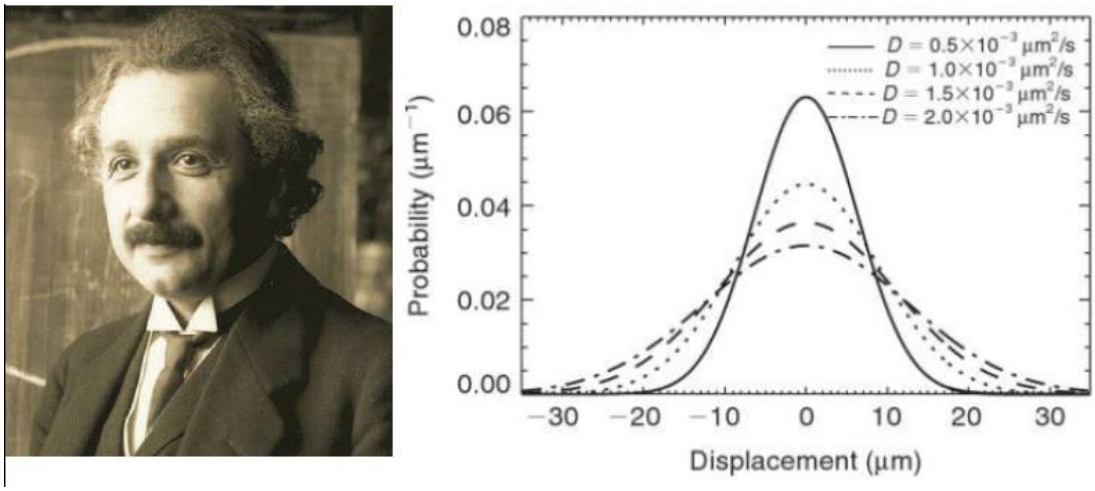


Figure 1-22: The young Albert Einstein addressed the problem of random walks or diffusion by introducing the concept of displacement distributions. The right graph plots displacement distributions for various values of the diffusion coefficient. The diffusion coefficient is exactly the same as in Fick's first law and depends on temperature and size of the diffusing molecules. The time allowed for diffusion was 40ms (142).

MRI allows us to quantify diffusional characteristics in a wide range of samples, and because these molecular diffusional processes are influenced by the geometrical structure of the environment, we can measure this structure indirectly using DW-MRI. When protons (such as the ones in water molecules in our brain) are put in a strong magnetic field (such as in an MRI scanner, which usually has a 1.5 to 9 Tesla field strength) they align parallel or antiparallel to

this magnetic field. Because parallel alignment outweighs antiparallel alignment we have some degree of “longitudinal magnetization” at rest. The protons do not just calmly lay aligned parallel or antiparallel to the magnetic field lines, but they spin around their own axis. This spinning is called precession and the angular frequency of this precession is called “Larmor frequency”, described by the Larmor equation:

$$\omega=\gamma B$$

with ω being the precession frequency, γ being the gyromagnetic ratio (in water the hydrogen nucleus has 2.68 x 108 rad/s/tesla), and B being the magnetic field. If a radiofrequency pulse (RF pulse) is applied to these protons two things happen: some parallel aligned protons will be lifted to a higher level of energy now being aligned antiparallel. Because this changes the ratio of parallel / antiparallel alignment, the net longitudinal magnetization decreases. Moreover, the RF pulse causes the protons to precess in phase, establishing a transversal magnetization, which moves around with the precessing protons. Over time protons lose this phase coherence, therefore losing transversal magnetization. They also switch back to the parallel-aligned low energy state leading to the reappearance of stronger longitudinal magnetization.

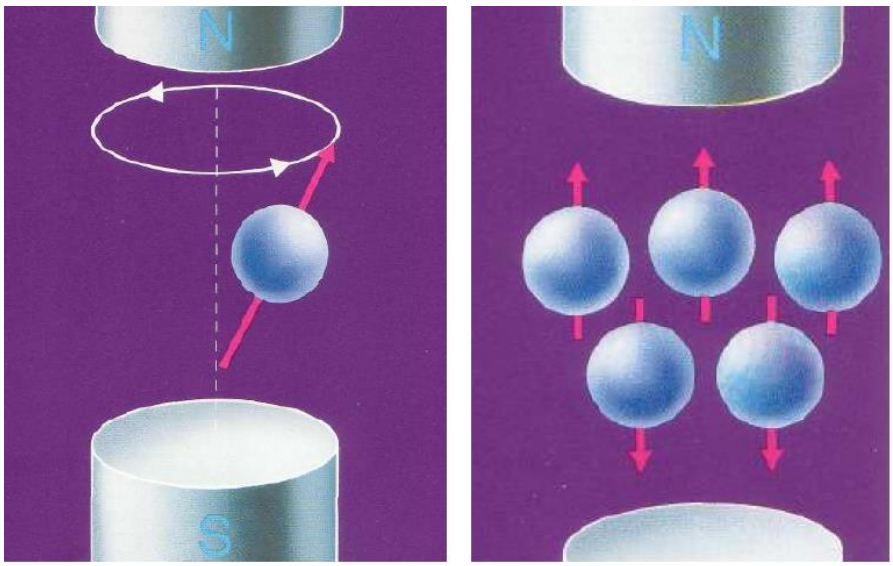


Figure 1-23: Protons in a strong magnetic field align themselves parallel or antiparallel to it and "precess" with a frequency proportional to the strength of the magnetic field. Most of them align parallel to it (in the right depiction it is a ratio of 3/2, in reality in the brain it is thought to be closer to 10,000,007 / 10,000,000) leading to a longitudinal magnetization.

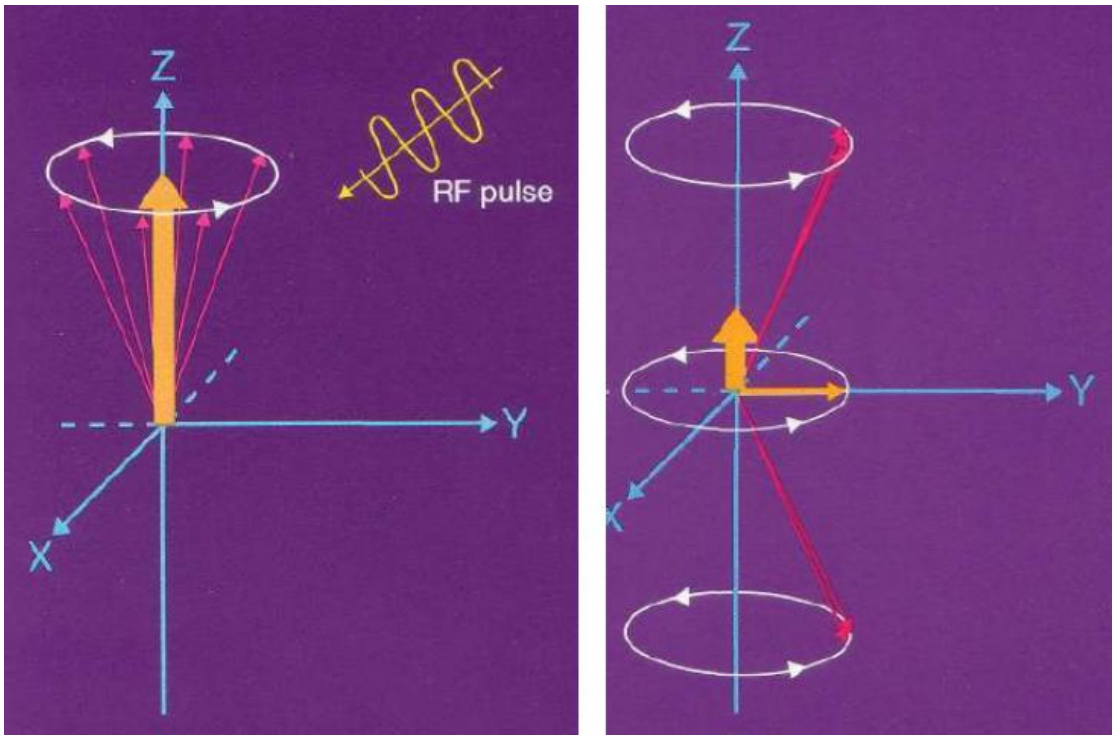


Figure 1-24: With the application of an RF pulse the "mostly" parallel aligned and out-of-phase precessing protons now align parallel and antiparallel (leading to a loss of longitudinal magnetization – z-axis) and start precessing in phase (causing transverse magnetization – y-axis)

Because precession is proportional to the strength of the magnetic field (as stated by the Larmor frequency) within an inhomogeneous magnetic field, the displacement of water molecules can expedite phase coherence loss (i.e. the loss of transversal magnetization). In other words because the strength of the magnetic field in region (A) is different from that in region (B), protons in region (A) precess at a different frequency compared to protons in region (B). Displacement of protons in region (A) to region (B) will lead to quicker phase coherence loss, causing a reduction of the phase-coherence related spin echo signal.

It is easy to imagine how we can now use strong magnetic gradients in a given direction to "artificially" create strong magnetic field "inhomogeneities", which then allow us to indirectly measure the degree of displacement of water molecules along the direction of this magnetic gradient. We might use this to determine the diffusivity in liquids or tissues regardless of the direction of the magnetic field gradient. Using the previously mentioned framework, the displacement distribution will be a Gaussian function and can be visualised as a sphere with the diameter of this sphere reflecting the "amount of diffusivity". In the early years of diffusion MRI it was recognised that (unidirectional) diffusivity of water molecules could change depending on brain microstructure and cellular alterations and could therefore indicate brain ischemia even when conventional scans (T1-weighted, T2-weighted) could not. Thus the introduction of diffusion-weighted imaging was met with enthusiasm as another MRI-based technique gaining new contrast in the brain and complementing T1-weighted and T2-weighted imaging.

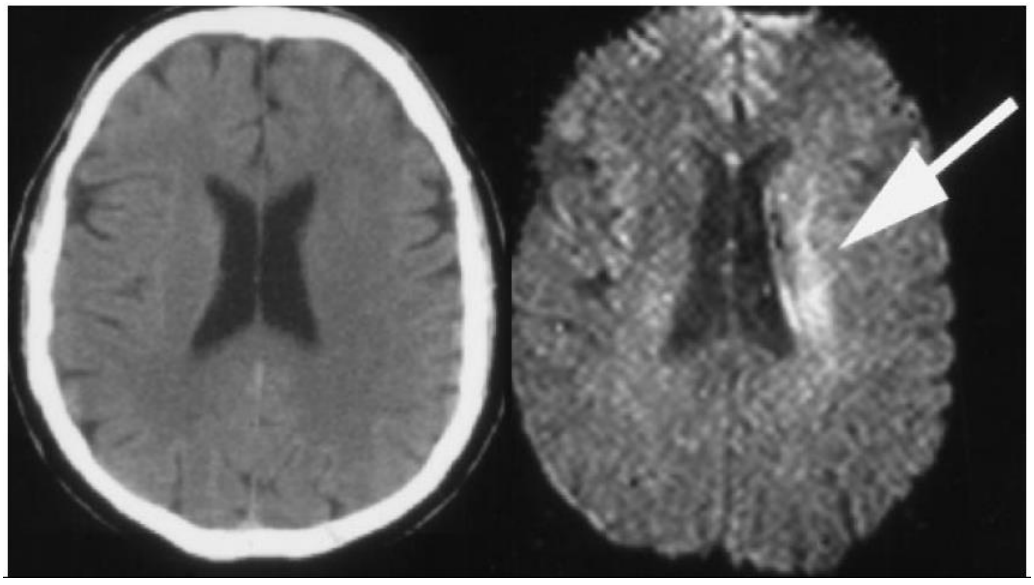


Figure 1-25: Selected neurological images of a 55-year old male bus driver, who had a low velocity accident associated with sudden right hemiparesis and aphasia. The CT scan (left) does

not show any abnormalities 1.5 hours after the accident. The DW-MRI image (right) shows diffusion restriction in the basal ganglia (148).

If the mobility of particles is asymmetrically hindered by certain properties of the surrounding tissue, diffusion in some directions might be “faster” than diffusion in others (e.g. in the direction of white matter tracts) and the displacement distribution might no longer be spherical. Such interactions of water molecules with (brain) tissue micro-architecture can be measured with diffusion MRI.

At about the same time as the finding that diffusivity was reduced in brain ischemia, it was recognised that in certain parts of the brain, especially in the white matter, diffusivity depended strongly on the direction in which the magnetic field gradient was applied. For example, in the corpus callosum water diffusivity was found to be more in line with the x-axis (MNI) but less in line with the y-axis. Whereas in some brain structures diffusivity was found to be similar along all three major axes (e.g. in ventricles, but also in grey matter) and therefore more isotropic, diffusivity in other areas (white matter) was found to be highly directional and anisotropic.

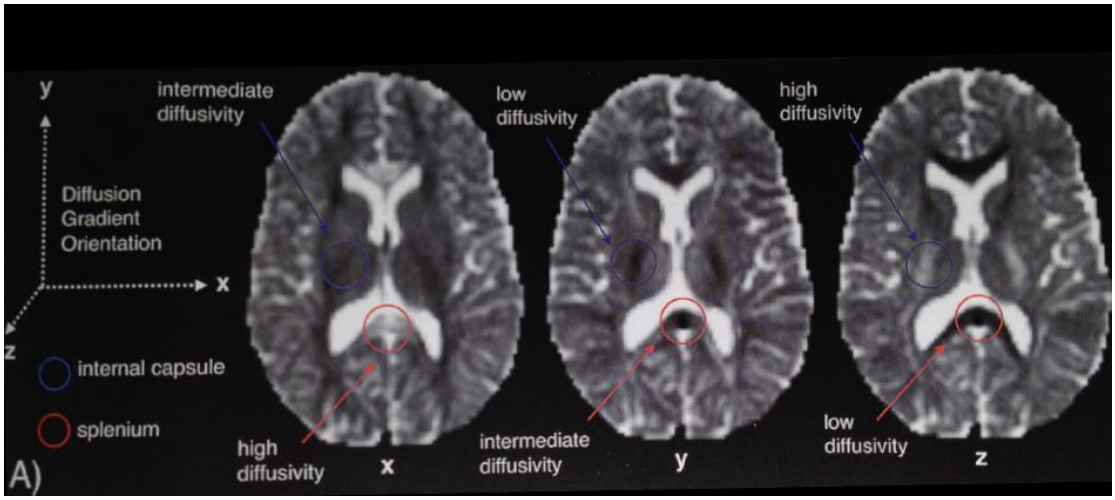


Figure 1-26 shows that diffusivity in some structures largely depends on the direction it is measured in. In the corpus callosum diffusion is high (bright voxels) if measured from left to right (left image) but hindered if measured from anterior to posterior (middle) or top to bottom (right). The internal capsule shows the highest diffusion when measured from top to bottom (right).

To characterise diffusive displacement profiles under anisotropic condition, diffusivity is measured in several different directions, and the results are modelled with a diffusion tensor model (149) (see Figure 1-26). Three so-called eigenvectors (έ1, έ2 and έ3) and their eigenvalues (λ1, λ2 and λ3, see Figure 1-27) are used to describe each voxel's diffusion displacement profile. The tensor can be visualised as a diffusion tensor ellipsoid (see Figure 1-28). The main diffusion direction (έ1) is the direction of the eigenvector with the largest eigenvalue. The fractional anisotropy (FA) is orientation-independent and allows us to draw conclusions on the degree of directionality (anisotropy) in the tissue. This in turn is thought to reflect some functionally relevant aspect of "white matter integrity", such as density of axonal packing, axonal diameter or myelination (see Figure 1-28). This measure of white matter integrity is often used in group-comparisons or parametric studies of white-matter architecture using for example voxel-based morphometry (VBM)-based approaches. Certain white matter micro-structural patterns were reported to be associated with handedness, gender (150, 151), experience and expertise in populations of professional musicians (152, 153), with performance in bimanual motor tasks (154), performance in simple cognitive paradigms (155) and even with measures of functional connectivity (156).

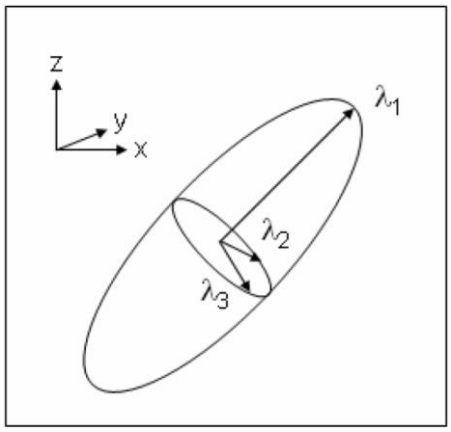


Figure 1-27: Three so-called eigenvectors (έ1, έ2 and έ3, arrows) and their eigenvalues (λ1, λ2 and λ3) describe each voxel's diffusion displacement profile and can be visualised as a diffusion tensor ellipsoid. Picture taken from "Introduction to MRI" at the homepage of the NMR research group at the Department of Cardiovascular Physiology, Heinrich Heine University Düsseldorf (www.nmr.uni-duesseldorf.de/sets/theory.html).

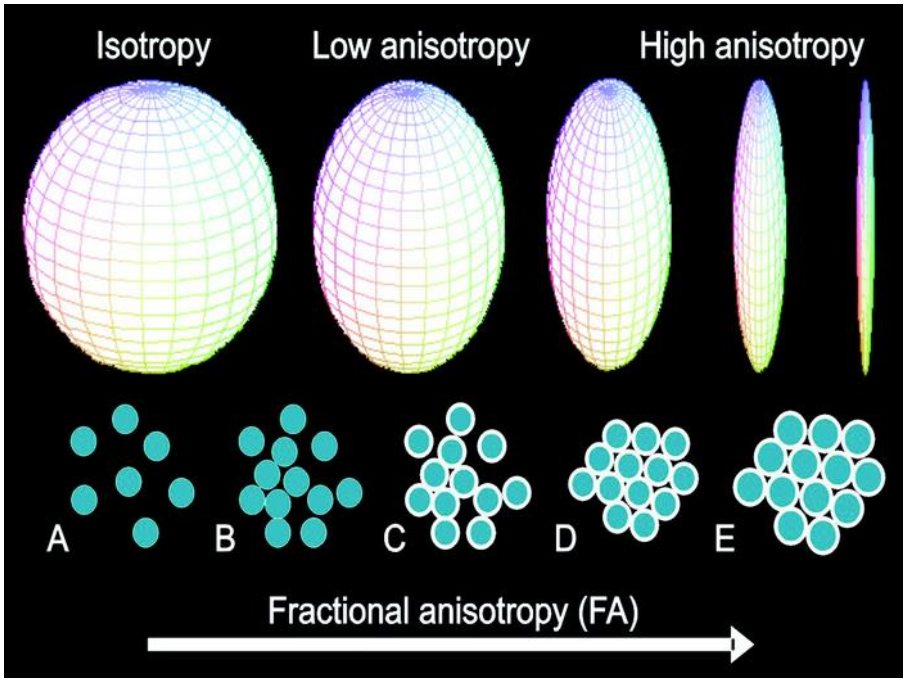


Figure 1-28: FA is thought to reflect microstructural features of the white matter. FA increases from left to right, therefore potentially reflecting denser axonal packing and increased myelination.

Studies comparing white matter properties across subjects crucially rely on accurate whole-brain voxel-wise analyses of DW-MRI data. Therefore, issues of cross-subject registration (particularly of the major white matter tracts), white matter delineation, statistical modelling and multiple comparison correction as well as the question of how to interpret positive and negative results need to be solved. An analysis tool called "Tract-Based Spatial Statistics" (TBSS) tries to address these problems. It attempts to combine ideas from (1) tractography-based and (2) VBM-style cross-subject comparison approaches {Smith, 2006 #450;Smith, 2007 #449;Jbabdi, 2010 #2584}. It is now one of the most commonly used techniques to compare measures of white matter architecture (such as FA or mean diffusivity) across subjects in a voxel-wise fashion.

It relies on first registering all subjects' images to one common template. Secondly this template is "skeletonised", i.e. only the template's "core" white matter – the area with maximum FA – is kept. Subject's individual FA maxima are projected onto this white matter skeleton and then compared using statistical models. This powerful approach might be described as being "anatomically blind" because the white matter skeleton does not relate in a straight-forward way to the known major association, projection or commissural fibre systems. Instead it reflects the locations of white matter regions that are most consistent across subjects. Therefore the resulting TBSS clusters of significant differences are not always easy to interpret in terms of their location and role in connecting grey matter areas. It is possible to envisage how a probabilistic atlas of these major white matter fibre systems could be used to localise significant TBSS clusters and their respective grey matter projection areas. One imagine how a probabilistic atlas of patterns of surface projections of major white matter fibre systems could be used to identify fibres based on their projection patterns in a given subject. One might then skeletonize these anatomically meaningful tracts to compare them across subjects (Chapter 4 puts forward a proposal for implementing such an approach).

It is possible that such approaches might help to tackle questions such as whether behaviourally relevant inter-individual variations in white matter microstructure are due to nature or nurture. If they were mainly formed by experience and learning, we should then expect these inter-individual variations not just in single white-matter pathways but in distributed functionally related networks and therefore might find intra-individual correlations between functionally related pathways. But if these white matter pathways are mainly determined by nature, i.e. by genes, we might find variations in single white matter tracts independent of functionally related pathways. Conversely, these "single tract variations" might then give rise to individual behavioural patterns. One might envisage a technique similar to the above mentioned cortical-thickness correlations, but in combination with white matter parcellation approaches (see Chapter 4), could tackle such problems.

Another powerful way of analysing DW-MRI data is to run "tractography". Tractography allows the 3D reconstruction and modelling of white matter tracts {Behrens, 2009 #2585}. Deterministic tractography streamlines along the main diffusion direction (orientation of the eigenvector έ1) in each voxel starting from a given seed voxel. It is powerful in delineating the layout of major white matter fibre bundles particularly through areas of unambiguous fibre orientation (such as the corpus callosum) and is therefore easy to interpret. However, it may stop in regions of high uncertainty (e.g. low FA) and therefore fail to find any connections through regions of, for example, crossing fibres. Moreover, it accumulates errors over long streamlining distances. The goal of probabilistic tractography is, instead of stopping tracking in areas of high uncertainty or deriving tracts with potentially high amounts of accumulated errors, to develop a representation of the uncertainty with any eventual statement about connections and tracts. Probabilistic tractography allows tracking through regions of high uncertainty (such as in areas of crossing fibres) and then tries to acknowledge and quantify the degree of uncertainty. The result of such tracking approaches however is not a single tract or connection, but rather a probability distribution. This renders the interpretation of results

more difficult and allows for statements that are potentially different from the ones anatomists would like to make (159-161).

Needless to say DW-MRI has its limitations and therefore does not constitute an alternative to tract-tracing techniques but rather a complement {Jbabdi, 2011 #2583}. Diffusion MRI does not inform us about the direction of a connection and the number of its synapses. Additionally the spatial resolution is limited and reliably modelling crossing fibres within a voxel remains challenging. Nevertheless, diffusion MRI is the only technique that allows noninvasive *in-vivo* imaging of white matter architecture in humans. It could help to understand the physiological and behavioural consequences of microscopic white matter structure and therefore allow important insights in structure, function, development and disease of the human CNS (142, 162).

1.2.1.4 Resting State fMRI

As mentioned above most human-fMRI studies have attempted to link mean increases (and decreases) or patterned changes in BOLD to the execution of cognitive tasks. However recently researchers have tried to investigate the random fluctuations of the BOLD signal at rest, which previously had been viewed as noise or baseline-activity. Why should one be interested in these “noisy” baseline fluctuations?

On theoretical grounds it has been argued that such “random” fluctuations of activity must fulfill a purpose because otherwise they would be a considerable waste of energy. Although the human brain represents only 2% of total body mass, it consumes 20% of the body’s energy. Most of this energy is fueling ongoing neuronal signaling (163, 164). Task-related increases in neuronal metabolism are usually on the order of 5%. That implies that when compared with large resting energy consumption fMRI might be investigating (energetically speaking) the tip of an iceberg.

Another argument in favor of investigating the resting fluctuations of the BOLD signal was mentioned earlier and relates to the problem of an "adequate" baseline. Rest-conditions are sometimes used as baseline condition for task-fMRI. However the fact that BOLD undergoes these substantial random fluctuations and some brain areas appear to be even more active during rest as compared to many task conditions (164) may make researchers wonder what they are comparing their task-fMRI signals to. The fact that "imagination" of certain "tasks" recruits the same regions as the tasks themselves (165) may be taken to suggest that even at "rest" humans are engaging in some "imaginary" tasks. Ongoing random fluctuations have even been shown to predict subsequent task performance and have therefore been suggested to reflect an ever active brain predicting future events and task-requirements in order to minimize surprise or optimize task-performance (166).

Correlations of "random" fluctuations are to some degree reflective of structural connectivity (as for example measured with DW-MRI) (167) and functional relations (co-activation in tasks) (168, 169). In addition such coupling has been shown to be predictive of task-performance (170, 171). This is why "coupling" of these fluctuations in BOLD has attracted considerable attention. The coupling of a region's fluctuations can be studied using normal pair-wise correlations for regions of interest or partial correlations. Seed-based correlation analyses allow for brain-wide estimates of correlation-values to one seed. Independent-component and hierarchical clustering analyses try to delineate "networks" of co-fluctuation in an unbiased and data-driven way.

The fact that resting-state fMRI scans are easy to acquire has made them a valuable tool for clinical studies. Researchers have tried to link differences in coupling between regions or within networks with disease, disease severity or progression. For example it has been argued that different forms of dementia disrupt distinct "resting state networks" in a consistent manner (172). Such discoveries may be of substantial diagnostic value – one can envisage resting-state fMRI to be used to augment early diagnosis of neurological and psychiatric

disorders. Moreover these findings may even suggest targets for therapeutic intervention. Resting-state networks and their intrinsic coupling strengths can be affected by non-invasive cortical stimulation protocols (133).

However there is some disagreement as to how resting-state fMRI network patterns should be compared between human subjects or subject groups. One elegant approach is called “dual-regression”. It is premised on ideas that in some ways parallel those that underpin the “TBSS” approach described above. Dual-regression first tries to identify a set of resting-state networks on the group-level using group (single-subject’s concatenated data) ICA (this step may be thought of as somewhat parallel to the group FA-skeleton). Then the spatial distribution of these group-ICA-networks is regressed against each individual subject’s resting state data in order to obtain their time-courses (spatial regression). These time-courses are then regressed against the subject’s data again in order to obtain the individual spatial distributions of the “canonical” group-ICA patterns (temporal regression – it is therefore called “dual-regression”, this somewhat parallels the projection of individual skeletons onto the group-skeleton). As with TBSS this approach may be considered as an elegant compromise between keeping as much data from each individual-subject on the one hand and making group-comparisons feasible on the other. But just as with TBSS, a potential weakness is that dual-regression is largely agnostic about the anatomical meaning of its “networks”; instead network-identification via ICA as its strength. Just as I have described ways in which TBSS might be improved by being more anatomically informed so one might suggest more anatomically informed versions of dual-regression analyses. For example, One might try to regress group-level resting-state coupling patterns from known anatomical regions onto individual subject’s data in in a dual-regression manner. This might allow for the investigation of group-level differences or parametric variability in coupling patterns of known anatomical regions.

One old puzzle in the field of resting state fMRI is the role of the so-called default-more network – a network of strongly coupled regions in the medial prefrontal cortex, posterior

cingulate cortex, precuneus, posterior temporo-parietal junction area, temporal pole and hippocampus. These regions are also known to be more “active” during rest as compared to many task-contexts. It has therefore been associated with “tasks” that humans (and other primates) may choose to engage in during rest – mind wandering, memory retrieval and prospective planning, recursive thinking, or self-referential cognition. Indeed parts of this set of areas have been implicated in cost-benefit decision-making (32), social cognition (173) particularly recursive social inference (28) and the susceptibility for (social) patterns (174) as well as prospective and retrospective memory (175, 176), and reality awareness / source monitoring (177). The fact that these regions are more active during rest has been taken to suggest that humans may engage in any one or a combination of these when placed in the scanner and not required to do anything. However one might consider an alternative (not mutually exclusive) account: Value representation and comparison, for example, is known to be related to a “peculiar” pattern of “activity” in medial prefrontal cortex (one of the nodes of the default mode network). When a choice is presented to subjects overall BOLD in medial prefrontal areas is suppressed and it is only on top of this suppression that the well-established positive valuation and value-comparison signal emerges. This may be due to the “special” nature of processing in these regions (178, 179), which is argued to rely on an intricate excitation-inhibition balance (180). However this may suggest that equalizations of increased activity and increased normal processing during rest in these regions should be considered with caution.

2 Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex

2.1 Abstract

Human ventrolateral frontal cortex (vlFC) is identified with cognitive processes such as language and cognitive flexibility. The relationship between it and the vlFC of other primates has therefore been the subject of particular speculation. We used a combination of structural and functional neuroimaging methods to identify key components of human vlFC. We compared how vlFC areas interacted with other brain areas in 25 humans and 25 macaques using the same methods. We identified a core set of eleven vlFC components that interacted in similar ways with similar distributed circuits in both species and in addition, one distinctively human component in ventrolateral frontal pole. Fundamental differences in interactions with posterior auditory association areas in the two species were also present – these were ubiquitous throughout posterior human vlFC but channeled to distinct frontal regions in monkeys. Finally there were some differences in inter-regional interactions within vlFC in the two species.

2.2 Introduction

The vlFC is identified with cognitive processes pre-eminent in humans such as language and cognitive flexibility. Broca's area is a part of left vlFC and is associated with language (181), while other vlFC areas, sometimes in the right hemisphere, have been linked to cognitive control – high-level top-down control of behavior – by influencing processing in other brain regions (182-186){Procyk, 2008 #2594}.

Despite vlFC's identification with such aspects of human cognition, key features of its neuroanatomy, such as cytoarchitecture, seem homologous in human and non-human

primates such as the macaque (187). Apparent homology in neuroanatomy is puzzling when macaques lack cognitive skills that humans possess. There is evidence that macaque vlFC is involved in auditory processing (188), orofacial motor-control (189), and gesture recognition (190), which might relate to processes necessary for language. However, emphasis has also been placed on macaque's vlFC's role in multimodal sensory integration (188, 191), the selection of environmental features for attention, and their subsequent use to guide flexible decision-making and action-selection (191). How the macaque brain areas responsible for these processes relate to areas in human vlFC areas implicated in attention and task control remains uncertain.

An alternative hypothesis is that there are new areas in human vlFC which support uniquely human cognitive abilities. A recent cytoarchitectonic and receptor-density based investigation identified previously unreported vlFC regions in locations associated with language and cognitive control (113). It is possible that no corresponding regions exist in monkeys but no test has been conducted. Alternatively, another aspect of vlFC may differ between species – the interactions that regions have with each other and with the rest of the brain. Such interconnections and interactions determine the information that brain regions have access to and the influence they exert.

In the current study we used diffusion-weighted magnetic resonance imaging (DW-MRI) parcellation techniques to test the former hypothesis: the potential existence of new areas in humans. DW-MRI identifies areas corresponding in position and extent to those identified by histological examination (192). Another non-invasive MRI-based technique, resting-state functional MRI (fMRI), was used to explore the latter hypothesis – that there are inter-species differences in the functional networks vlFC participates in ["functional connectivity" (140)]. Functionally connected networks observed at rest often overlap with networks observed during performance of cognitive tasks (193). Such inter-regional coupling-patterns often partly reflect anatomical connectivity (147). Previous studies have shown dorsal prefrontal and

parietal cortex participate in similar networks in macaques and humans (192, 194-196). Here we test whether this is true in vlFC.

2.3 Results

The first goal was to identify component parts of the entire human ventrolateral prefrontal cortex (Figure 1-1a). There has recently been interest in the existence of specialized brain areas near the border between ventrolateral prefrontal and ventral premotor cortex (PMv). To ensure the possibility of investigating them in our study we took care to include PMv as well as ventrolateral prefrontal cortex within the vlFC region of interest (ROI, see Figure 2-2a). Posteriorly we delimited the ROI by the posterior end of the precentral gyrus to include the crown of the precentral gyrus but not the rostral bank of the central sulcus. Posterior to the posterior end of the inferior frontal sulcus, the ROI was dorsally delimited by what had previously been established to be the border between PMv and dorsal premotor cortex (PMd) (197). In addition to PMv we also examined all areas in the pars opercularis, pars trianguris and pars orbitalis of the inferior frontal gyrus, the whole of the inferior frontal sulcus and the deep frontal operculum. Hence our ROI was delimited dorsally by the fusion of the middle frontal gyrus and the inferior frontal sulcus to include the whole inferior frontal sulcus. More anteriorly we included all the tissue ventral to a line extending anteriorly from the anterior tip of the inferior frontal sulcus to the most anterior point in the brain. This line ran along the surface of the brain approximately in the axial plane in the standard Montreal Neurological Institute (MNI) coordinate system. Thereby we probably included part of the frontal pole in this most anterior area, and for reasons of completeness we also included a more dorsal region into our ROI identified as frontal pole by Sallet et al., (2013). The ROI included both medial and lateral frontal pole and was medially delimited by the paracingulate sulcus. Ventromedially the ROI was delimited by the lateral orbital sulcus and, more posteriorly at the level of the deep frontal operculum by the circular sulcus. The ROI could be defined in an

unbiased and easily reproducible way in all subjects by using the automated grey matter-white matter segmentation tool FAST v4.1 (FMRIB's Automated Segmentation Tool) (198) on the MNI standard brain template (MNI_152 template) in combination with these gross anatomical landmarks. This standard-space ROI was then registered to each individual brain using non-linear registration (FMRIB's Non-linear Image Registration Tool FNIRT).

Because of potential differences between right and left vlFC we carried out parallel investigations of both regions defined by the same anatomical criteria. However, we judged results to be similar enough to focus on right vlFC and to present left vlFC results only in Supplementary Information (SI).

To identify component vlFC regions we used DW-MRI based tractography in 25 healthy human participants. We estimated the connections of every vlFC voxel in each subject (199) (Figure 2-1b). The parcellation approach identifies voxels within vlFC (Figure 2-1a) with similar estimated profiles of connectivity with the rest of the brain. The connectivity matrix (Figure 2-1b) is used to generate a symmetric cross-correlation matrix (200) reflecting the correlation in connectivity pattern between all vlFC voxels. This cross-correlation matrix is then regrouped using K-means clustering (Figure 2.1c) to identify voxels sharing connectivity profiles (192, 201). The spatial locations of the voxels in each parcel are then established; voxels cluster into anatomically coherent parcels (Figure 2-1d).

With the correct number of regions in human vlFC being unknown, we carried out a series of parcellations into two to 25 regions. The parcellation into ten distinct regions was particularly consistent across subjects and showed the smallest variance of the average cross-correlation suggesting it is the most coherent clustering into similarly correlated clusters (SI Methods 6.1.2 discusses determination of regions). Similar ten-cluster parcellations were obtained in both left and right hemispheres. This specific parcellation solution of ten different regions was further corroborated with a second series of parcellations using an iterative procedure (SI Methods 6.1.4) similar to the approach used by Beckmann et al., (201). This iterative

parcellation procedure yielded similar ten cluster results. Each of these ten regions were subsequently subjected to further sub-parcellation to test if a more detailed parcellation of vlFC could be obtained. Two of the ten vlFC regions could reliably be parcellated further into two regions. Therefore, in summary, we identified twelve similar vlFC regions in all our human subjects.

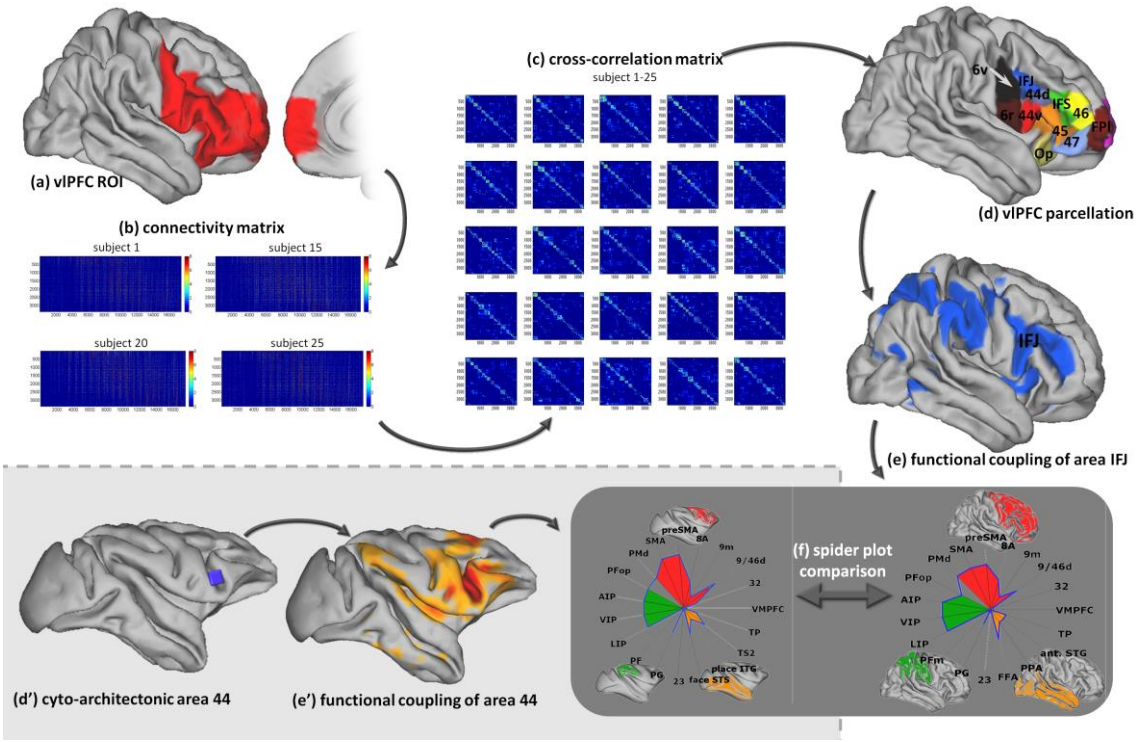


Figure 2-1: Overall approach of the study: We used DW-MRI based tractography in 25 human participants to divide vlFC ROI (a) into regions with consistent connectivity profiles with the rest of the brain. Probabilistic tractography was performed from each voxel in the ROI and yielded a connectivity matrix between all VLFC voxels and each brain voxel for each participant (25 connectivity matrices, 4 example connectivity matrices in (b). These matrices were then used to generate a symmetric cross-correlation matrix, which was then permuted using k-means segmentation for automated clustering to define different clusters (yielding 25 clustered cross-correlation matrices, (c)). These clusters were projected back onto the individual participant's brain and overlayed onto the MNI152 standard brain (d). FMRI analyses in the same 25 participants determined functional connectivity between these distinct vlFC regions and the rest of the brain (example functional connectivity pattern for IFJ in blue in (e)). We related the functional connectivity fingerprints of the different vlFC regions to the resting-state functional MRI based connectivity profiles from different cyto-architectonically defined areas in vlFC in 25 macaques (example functional connectivity pattern for macaque's area 44, (e')).

The names given to each vlFC region need to be considered with caution. One strategy is to avoid naming areas but we found this approach renders our findings difficult to communicate. Instead we used names largely based on cross species similarities in the areas' connectivity profiles. In several, but not all cases, they can be related to previous cytoarchitectonic and receptor-density based parcellation schemes (113).

Following parcellation, we used fMRI from the same 25 humans to determine functional connectivity profiles for each vlFC region. We took the BOLD signal from each area and established which other voxels had correlated BOLD time courses (Figure 2-1e shows the example case of areas where the BOLD signal was correlated with the inferior frontal junction (IFJ) BOLD signal). We then summarized each vlFC area's pattern of coupling in a spider plot (Figure 2-1f). The spider plot shows the relative degree of coupling between each vlFC area and a set of target regions. These target regions were chosen because they are thought, on the basis of anatomical and functional evidence, to correspond between humans and macaques and because different vlFC areas were expected to have diagnostically different coupling patterns with different sets of these target regions. Similar spider plots were calculated for a set of macaque vlFC regions (Figure 2-1d') and the corresponding target areas in other parts of the brain (Figure 2-1e'). Finally (Figure 2-1f) we compared the dissimilarity or "distance" between each human frontal area's coupling fingerprint and the coupling fingerprint associated with each frontal area of the macaque (196, 202).

2.3.1 Regions in ventral premotor cortex

The human precentral gyrus territory overlapping with PMv (197) was subdivided into dorsal and ventral regions (brown and black, Figure 2-2 and

Figure 2-3; SI Methods 6.1.3) with centers of gravity at [x=51 y=1 z=33] and [48 5 12] in MNI standard space, respectively. Given their topographical location, these areas probably

correspond to the cytoarchitectonically defined agranular 6v and dysgranular 6r areas although the ventral region extended into the operculum and therefore probably included parts of op6 (113).

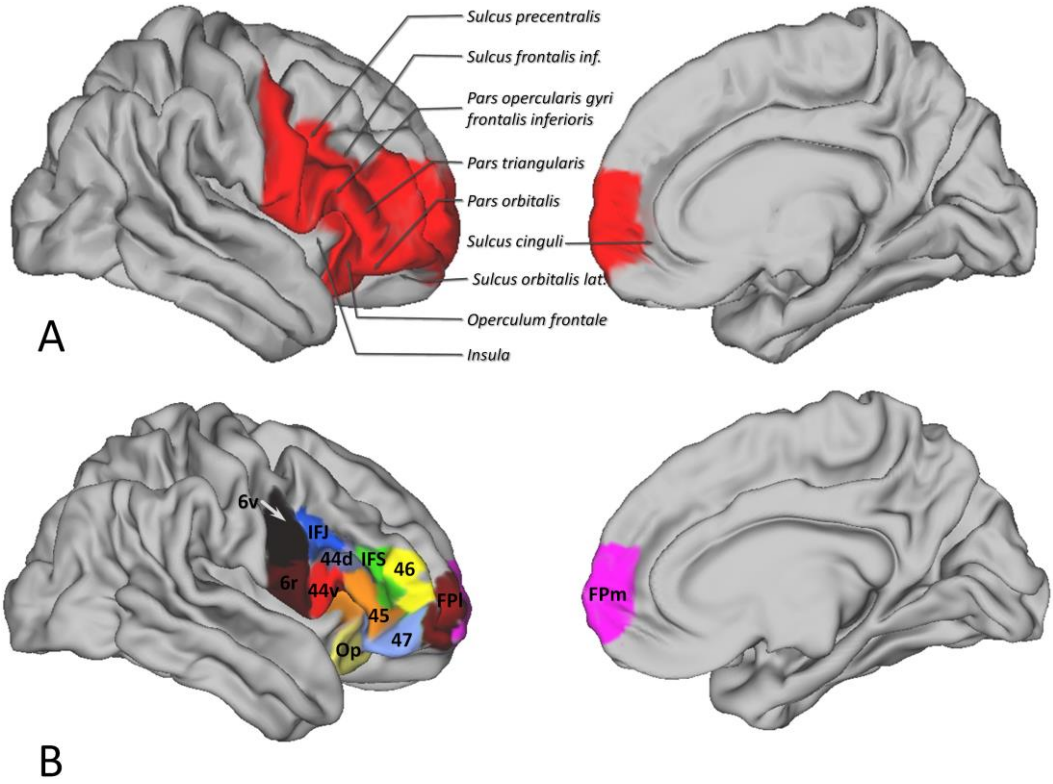


Figure 2-2 (A) shows the right vlFC ROI. Dorsally it included the inferior frontal sulcus and, more posteriorly it included PMv; anteriorly it was bound by the paracingulate sulcus, ventrally by the lateral orbital sulcus and the border between the dorsal insula and the opercular cortex. (B) shows a schematic depiction of the result of the twelve cluster parcellation solution using an iterative parcellation approach. We subdivided PMv into ventral and dorsal regions (6v, 6r, purple and black). We delineated the IFJ area (blue) and areas, 44d (grey) and 44v in lateral pars opercularis (red). More anteriorly we delineated areas 45 (orange) in the pars triangularis and adjacent operculum and IFS (green) in the inferior frontal sulcus and dorsal pars triangularis. We found area 12/47 in the pars orbitalis (light blue) and area Op (bright yellow) in the deep frontal operculum. We also identified area 46 (yellow), and lateral and medial frontal pole regions (FPl and FPm; ruby colored and pink). See also SI Figure 6-1, SI Figure 6-2, SI Figure 6-3 and SI Figure 6-5.

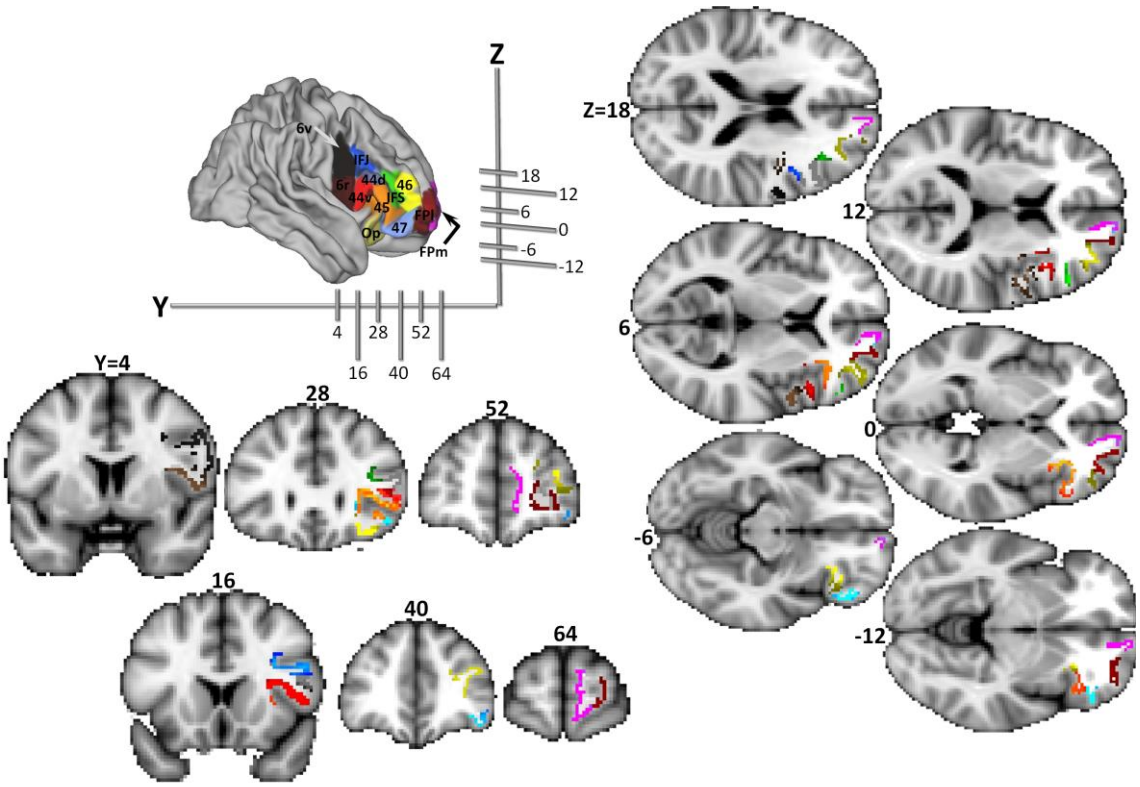


Figure 3

Figure 2-3 shows six coronal and six axial slices through the MNI152 standard brain as provided by FSL with the twelve vlFC sub-regions obtained by the tractography based parcellation overlayed. The Z and Y coordinates for the respective slices are depicted next to the grey axes which show the approximate location these slices were taken from.

Human 6v and 6r (Figure 2-4) were similar in their functional connectivity. Both showed strong coupling to motor cortex, primary somatosensory cortex (S1), supplementary motor area (SMA), cingulate motor areas (CMAs) and PMd. Area 6r had stronger coupling with visuomotor areas including 7m, intraparietal sulcus (IPS), anterior inferior parietal lobule (IPL), whereas 6v had stronger coupling with somatosensory regions such as S1 and anterior parietal operculum, and with PMd. Neither coupled with inferior temporal lobe suggesting their visual information is mainly derived from parietal cortex.

There has been speculation about potential similarities of these regions in humans and macaques (113, 190, 203). We found that functional connectivity patterns of macaque areas F5c and F5a resembled those of human areas 6v and 6r, respectively and so below we refer to them collectively as 6v/F5c and 6r/F5a.

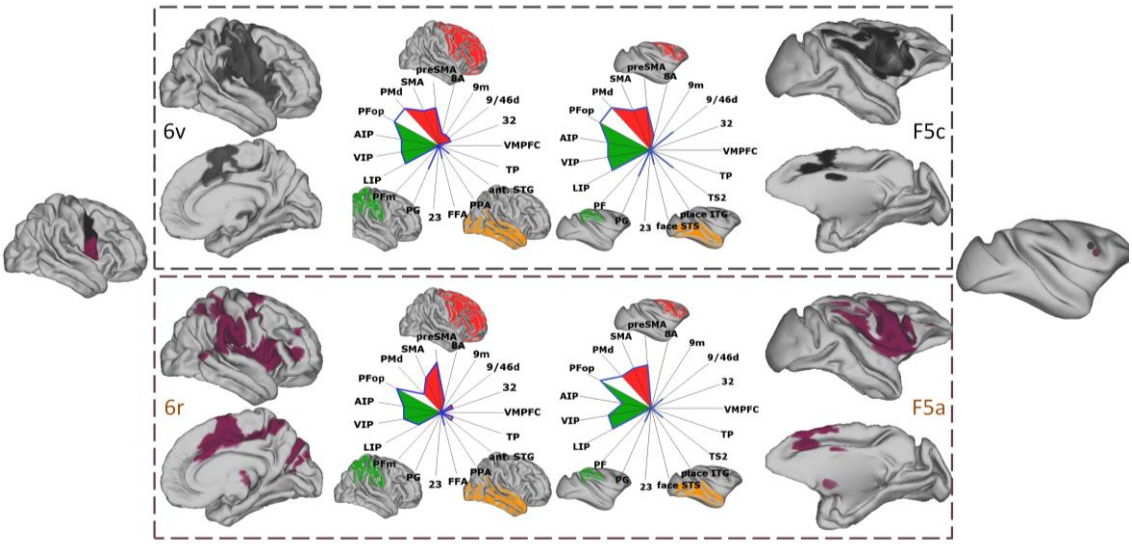


Figure 2-4: Resting state fMRI derived-functional connectivity patterns of human (left panels) areas 6v (black) and 6r (brown) which resemble those of macaque (right) areas F5c (black) and F5a (brown). In the middle we show the intensities of these human and macaque resting state fMRI-derived functional coupling patterns in a selected number of target ROIs that can be easily matched between the two species plotted on a spider plot. These spider plots were used to match human and monkey vlFC sub regions. See also SI Figure 6-4 and SI Figure 6-6.

2.3.2 Pars opercularis and inferior frontal junction

We identified four sub-divisions directly anterior to the precentral gyrus. The first occupied the posterior inferior frontal sulcus at its junction with the precentral sulcus (blue, Fig 2B) and had a center of gravity at [41 12 25]. This subdivision overlaps with the territory that has been called IFJ (185). Although our analyses demonstrated that IFJ can be reliably distinguished from adjacent vlFC areas, it remained a possibility that our IFJ region was not truly a separate region but just the most ventral extension of a more dorsal frontal area such as area 8A. To test this possibility we created a ROI composed of IFJ and of area 8A [as defined by Sallet and colleagues (2013)] and subjected it to further parcellation attempts. In all subjects we found that this larger ROI could be parcellated into two regions corresponding to IFJ and to area 8A suggesting that the two areas are distinct (SI Figure 6-5).

Another more ventral region was located in lateral pars opercularis with a center of gravity at [41 19 10] (red, Figure 2-2B). It posteriorly bordered F5a/6r. It therefore probably corresponds to area 44 (113). Unlike almost all other areas that we investigated, we found that this area could be reliably sub-parcellated into dorsal and ventral subcomponents with centers of gravity at [45 23 19] and [39 20 10] in all subjects. On the basis of their location we suggest they correspond to 44v and 44d (113) although 44v extended into the operculum and therefore probably included parts of op8.

A fourth region was established in the middle section of the inferior frontal sulcus at [43 32 15] (green, Figure 2-2B). It bordered IFJ and 44d posteriorly, area 46 anteriorly, 44v inferiorly, and area 45 in the pars triangularis. This region is located in a similar region to the inferior frontal sulcus regions IFS1 and IFS2 (113) and so we refer to it here as IFS.

Although our analyses demonstrated that IFS can be distinguished from other vlFC areas, it is still possible that IFS is not a truly separate region but just the most ventral part of a dorsal frontal cortical area such as area 9/46v. To test this possibility we created a ROI composed of IFS and of area 9/46v [as defined by Sallet and colleagues (2013)] and subjected it to further parcellation attempts. In all subjects we found that this larger ROI could be parcellated into two regions corresponding to IFS and area 9/46v suggesting that the two areas are distinct. We were not, however, able to reliably subdivide IFS into smaller component areas (SI Figure 6-4).

Examination of functional coupling suggested two important differences in the networks in which these four more anterior regions IFJ, 44d, 44v, and IFS participated, which contrasted with F5c/6v and F5a/6r network patterns. A related pattern of differences was found in macaque vlFC (Figure 2-5). Although all four regions remained coupled with F5c/6v and F5a/6r, they were also coupled with first, a number of dorsolateral prefrontal areas and second, visual association areas in occipito-temporal cortex.

Although the occipito-temporal areas do not appear in the spider plots in Fig 5 this coupling pattern is apparent in the rest of Figure 2-5. Occipito-temporal areas are absent from the

spider plots because the spider plots focused on regions for which there is clear evidence of human-macaque correspondence (see SI Table 6-3 for detailed account and references to cross species homology of spider plot target areas); obviously it is only the coupling with such areas that can be used to precisely decide how similar a vlFC area's functional network is in the two species, but such correspondences remain uncertain in occipito-temporal cortex. In macaque these areas included V4, TEO, TE, and TPO.

The coupling of IFJ in humans suggests it occupies a transitional location between premotor and prefrontal cortex, distinguishing it from the more caudal premotor cortex that has little coupling with prefrontal cortex. To date, there has been no direct comparison of these human precentral subdivisions and areas in the macaque using the same technique. We found a macaque vlFC region could be identified with a similar coupling-profile – area 44 in the fundus of the inferior limb of arcuate sulcus. In contrast, the pattern of coupling of human 44v, with comparatively little interaction with parietal cortex, most closely resembled that of macaque area ProM. The coupling pattern of region 44d, which sits in between IFJ and 44v, bore broad similarities with the coupling pattern of IFJ. So we suggest that this region, like IFJ, may correspond to area 44 in the macaque. Nevertheless it was notable that some aspects of 44d's coupling pattern were also reminiscent of 44v. Finally, the functional connectivity profile of human IFS resembled that of macaque area 45B.

The same cognitive control and language processes have been associated with IFJ and 44v (181, 182, 185), but our parcellation and functional connectivity results suggest that they have quite separate identities. To better understand the subdivisions we identified here, we related the functional coupling patterns of IFJ, 44v, and 45 (see below) – three regions that have particularly been implicated in language and cognitive control – to meta-analysis and derived co-activation patterns for several different cognitive tasks (6.1.6 and SI Figure 6-9a). Consistent with their different coupling profile, the meta-analysis suggested that these regions were associated with quite different functional roles, with area 44v involved in lower level control

processes seen in typical inhibition tasks such the go/no-go task, and IFJ involved in higher-level processes such as task switching. Another region sometimes associated with inhibitory processes, area 45, was not associated with either of these functions, but rather implicated in social and language processes. Thus, the distinct coupling patterns of the human IFJ, 44v and 45 accord with the distinct roles they play in language and cognitive control.

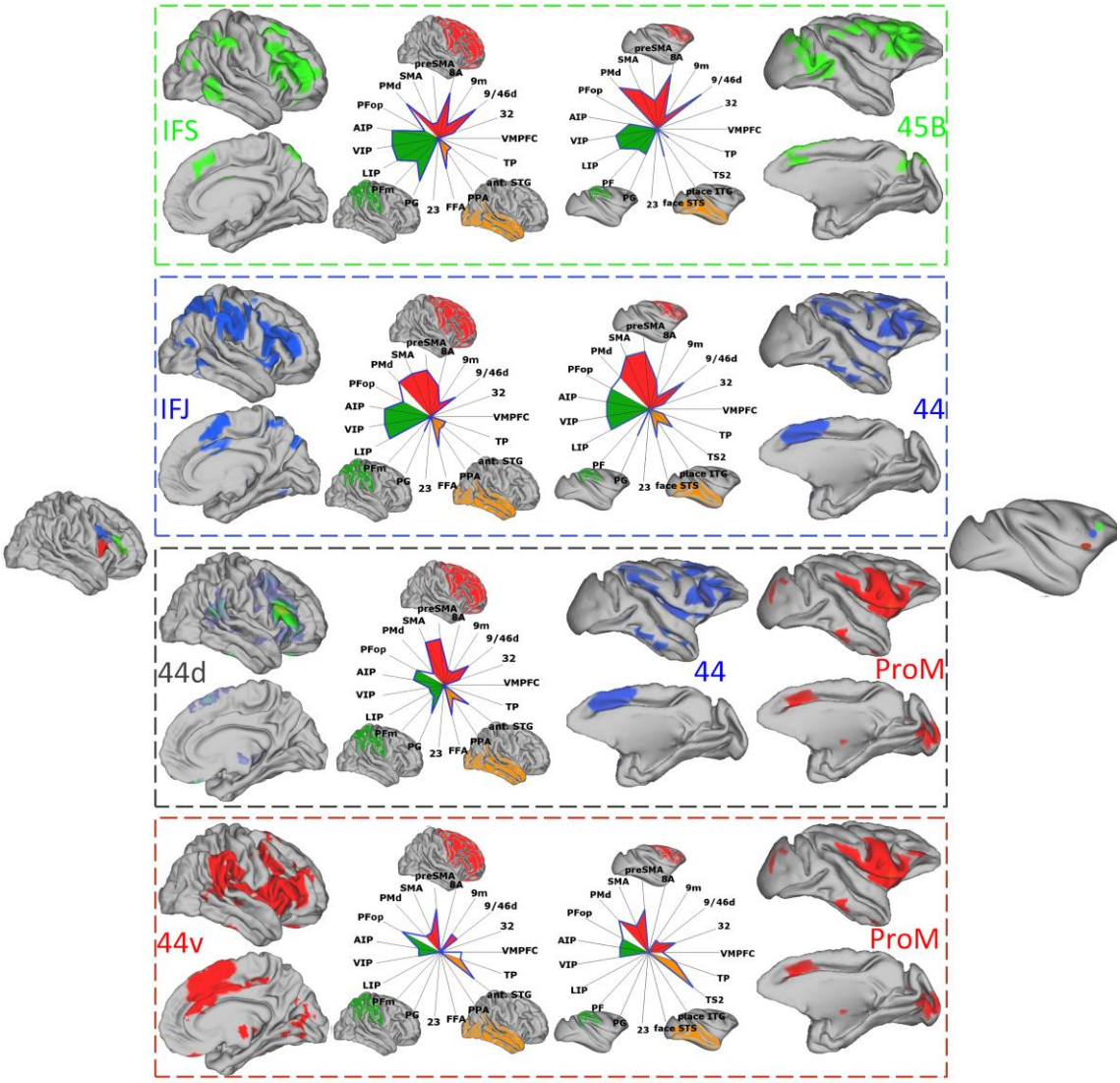


Figure 2-5: Resting state fMRI-derived functional connectivity patterns of human areasIFS (green), IFJ (blue), 44d (green-grey) and 44v (red), and resting state fMRI-derived functional coupling patterns of the proposed corresponding areas in macaque: areas 45B (green), 44 (blue) and ProM (red). Human area 44d resembled both macaque area 44 and ProM. In the middle we show spider plots of these regions. Conventions as for Fig 2. See also SI Figure 6-10.

2.3.3 Pars triangularis and orbitalis

A more anterior vlFC region covered much of pars triangularis and the adjacent operculum with a center of gravity at [38 29 -1] (orange, Figure 2-2). This subdivision seems to partly overlap with territory previously identified as 45 (204) or, perhaps more precisely, 45A. We subsequently call this area 45 although it does not perfectly overlap with BA45 as defined by (205). However its location with respect to the other areas and its coupling pattern (see below) resemble area 45 (206). Another area was found in pars orbitalis and adjacent deep frontal operculum. This region is the second one in which it was possible to conduct further sub-parcellation into areas that had a consistent location in all subjects. One area was located on the gyrus of the pars orbitalis (cyan, Figure 2-2) with a center of gravity at [42 34 -7] and therefore probably corresponding to area 47 or 47/12 (187) while the other area was in the deep frontal operculum with a center of gravity at [31 26 -15] (bright yellow in Figure 2-2). This area is adjacent to areas that Amunts and colleagues (2010) refer to as Op9 and Op10 and partly overlaps with a region referred to as the frontal operculum by Higo et al (2011). We use "Op" when referring to it in our human subjects.

In their functional coupling profiles human area 45 and 47/12 most closely resembled macaque 45A and 47/12, respectively. Functional coupling patterns of human and macaque areas 45/45A and 47/12 shared common features such as coupling with other prefrontal areas and with temporal areas (Figure 2-6Figure 2-6). In this they resembled the immediately posterior regions (IFJ/44d (human)/44 (macaque) and 44v/ProM). However, there was little evidence of coupling with action-selection regions in parietal and premotor cortex. The main difference between 45 and 47/12 was that the latter was more closely linked to anterior parts of both temporal and prefrontal cortex in both species. The functional coupling profile of human Op most resembled the functional coupling profile of the lateral agranular insula (Ial and Iapl in Carmichael and Price, 1994) in the macaque. Both human Op and macaque Ia were

coupled with inferior and superior temporal cortex, perirhinal cortex and temporal pole, amygdala, and anterior dorsomedial frontal cortex.

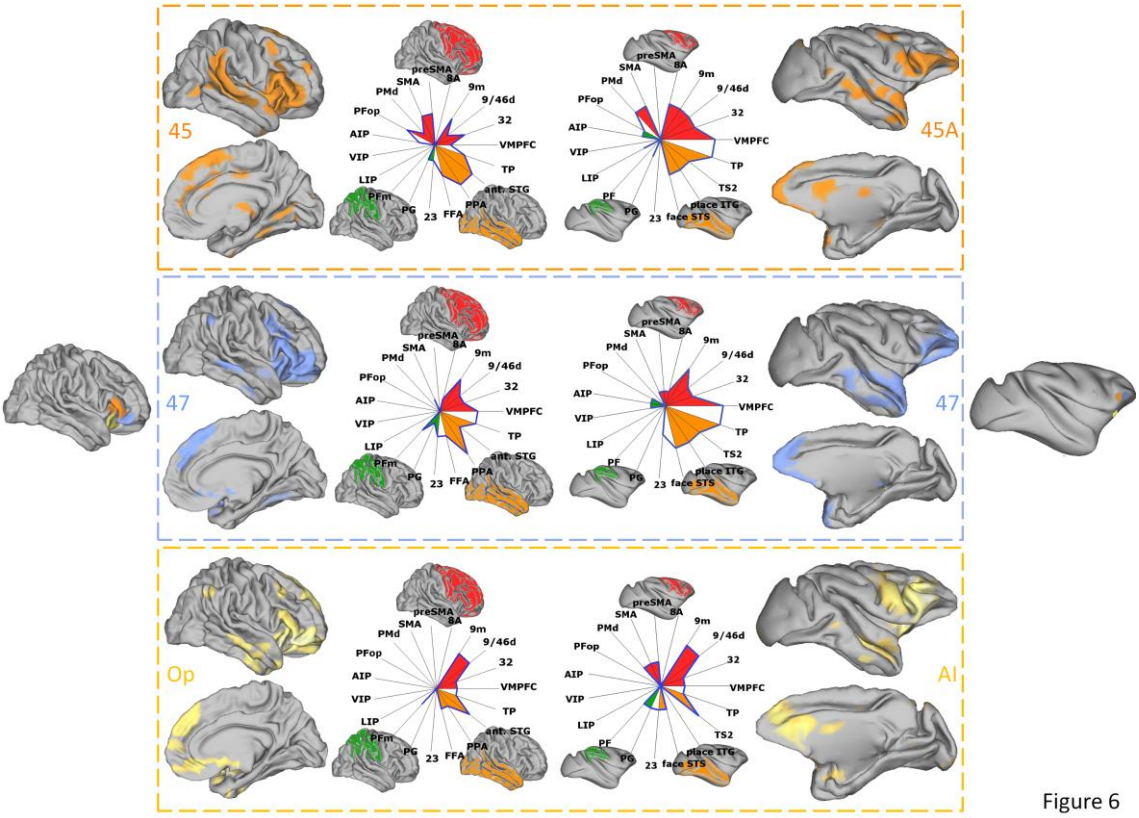


Figure 6

Figure 2-6: Resting state fMRI-derived functional connectivity patterns of human areas 45 (orange), 12/47 (light blue), and Op (light yellow), and resting state fMRI-derived functional connectivity patterns of the proposed corresponding areas in macaque (right): areas 45A (orange), 12/47 (light blue), and Op (light yellow). In the middle we show spider plots of these regions. Conventions as for Figure 2-2.

2.3.4 Human-macaque differences in posterior temporal cortex and intrinsic vlFC connectivity

Thus far, we have reported striking similarities between species. In both species posterior ventral frontal cortex is coupled with sensorimotor regions important for action-selection while anterior vlFC is coupled with temporal visual association-cortex. Such a combination suggests that, in both species, vlFC is well placed to categorize, select and determine the flexible and often context-dependent influence visual representations have over decision-

making and action-selection. Despite evidence for these similarities we also observed major differences, most prominently in coupling with auditory areas. In humans, there was coupling between auditory association areas in posterior superior temporal cortex and many vlFC areas (Figure 2-7, SI Figure 6-6 and SI Figure 6-7). This coupling was not limited to areas traditionally implicated in language, but was apparent throughout 6v, 6r, 44v, IFJ, 44d, 45, and IFS. In macaques this coupling was meager.

It might be argued that for some reason our scanning acquisition or analysis procedures were insensitive to long range coupling between auditory association cortex and frontal lobes in macaques. This, however, cannot be the case as we identifed coupling of the same macaque posterior auditory association regions with medial prefrontal and anterior cingulate cortex (ACC) regions that are concerned with social interaction (207, 208) (Figure 2-7). In fact there is a "double-dissociation" in posterior superior temporal-frontal lobe coupling in the two species; in macaques posterior superior temporal coupling was strongest with medial frontal cortex while in humans it was strongest with vlFC.

We also looked at vlFC intrinsic functional coupling to establish if there was any between-species difference in the coupling of regions within vlFC (see SI Figure 6-8 and SI Methods 6.1.8). We noted less coupling between the two regions situated in PMv (6v/F5c – 6r/F5a) for the human as compared to the macaque. We also found less coupling between 6r/F5a and 44v/ProM in human compared to macaques. However we found stronger coupling between 44v and IFJ in the human compared to ProM and 44 in the macaque. We also found stronger coupling for 45B/IFS with 45, 45 with 47 and 44v/ProM with 46 in the human compared to the corresponding areas in the macaque (see SI Figure 6-8 with SI Figure 6-8A showing the functional coupling in human vlFC; – white asterisks indicate significantly less, black asterisks significantly more functional coupling in human as compared to macaque). In summary, inter-regional coupling within prefrontal areas appeared stronger in humans than macaques but the opposite was true for ventral premotor areas.

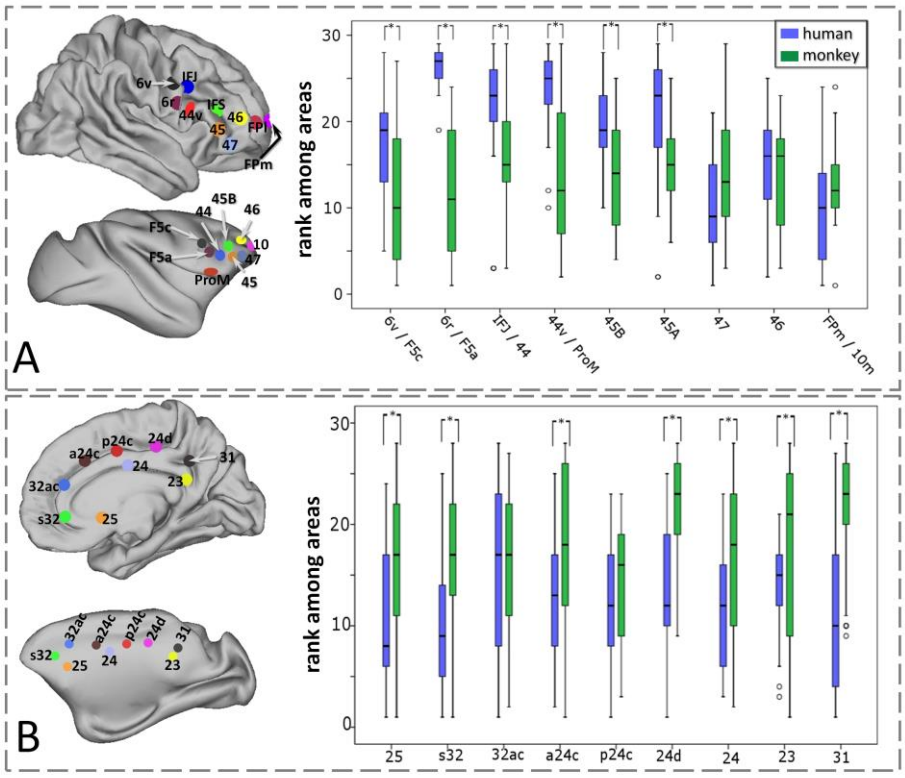


Figure 7

Figure 2-7: fMRI-derived functional connectivity of a posterior auditory association area in the human and macaque (area Tpt) with region vlFC areas (top panel) and cingulate cortex (bottom panel). Graphs show the ranking of functional connectivity with the and cingulate areas in macaques and humans among all ROIs presented in spider plots in Figure 2-4, Figure 2-5 and Figure 2-6. The average intensity in the respective ROI from the seed-based correlation analysis z-maps was rank-ordered from lowest to highest values (i.e. high ranks reflect strong functional coupling) in two separate ranking analyses (one for auditory – vlFC coupling strength and a second for auditory – cingulate coupling strength). Asterisks mark significant between-species differences (Mann–Whitney–Wilcoxon rank-sum test; $p<0.05$). See also SI Figure 6-7, SI Figure 6-8 and SI Figure 6-9.

2.3.5 Areas in the anterior prefrontal cortex

Three areas were identified in anterior human prefrontal cortex. One, on the edge of the investigated region, was actually in dorsolateral prefrontal cortex and corresponds to area 46 (yellow, Figure 2-2) with a center of gravity at [33 47 13] and has also been described elsewhere (196). Two areas were delineated in the frontal pole (ruby and pink, Figure 2-2) with centers of gravity at [13 59 2] and [26 54 0]. We refer to them as medial and lateral frontal pole (FPm and FPl).

Human area 46 matched macaque area 46 in its functional connectivity. Like other vlFC areas, both were coupled with the frontal-parietal networks but the coupling was more limited in this case. For example, in both humans and macaques, parietal coupling was strongest in IPL as opposed to IPS. The functional connectivity patterns of FPm and FPl (Figure 2-8) are distinct and only FPm corresponded in direct and simple ways to a macaque prefrontal brain region – the macaque's frontal pole region 10 – while FPl did not. In both humans and macaques FPm/10 were dramatically different to area 46 in that they showed little coupling with parietal cortex but instead, and unlike 46, they coupled with temporal pole, posterior cingulate cortex and amygdala. These clear differences in coupling-patterns of FPm/10 and 46 resemble differences in their connection patterns known from track-tracing experiments (209).

The coupling-pattern of FPl bore little resemblance with that of FPm/10 but it rather partly resembled that of area 46; like area 46 it was coupled with IPL, and unlike FPm/10 it was not coupled with amygdala or temporal pole. However, FPl's coupling-pattern was not identical to that of area 46; its IPL coupling was much more restricted than area 46's and indeed most of PFl's coupling was just with other prefrontal regions. In summary, human prefrontal cortex contains a region, FPl, which lacks simple correspondence with any in macaque prefrontal cortex. FPl does not show those coupling-patterns that are usually distinctive features of FP cortex, but it instead exhibits coupling-patterns reminiscent of dorsolateral prefrontal cortex.

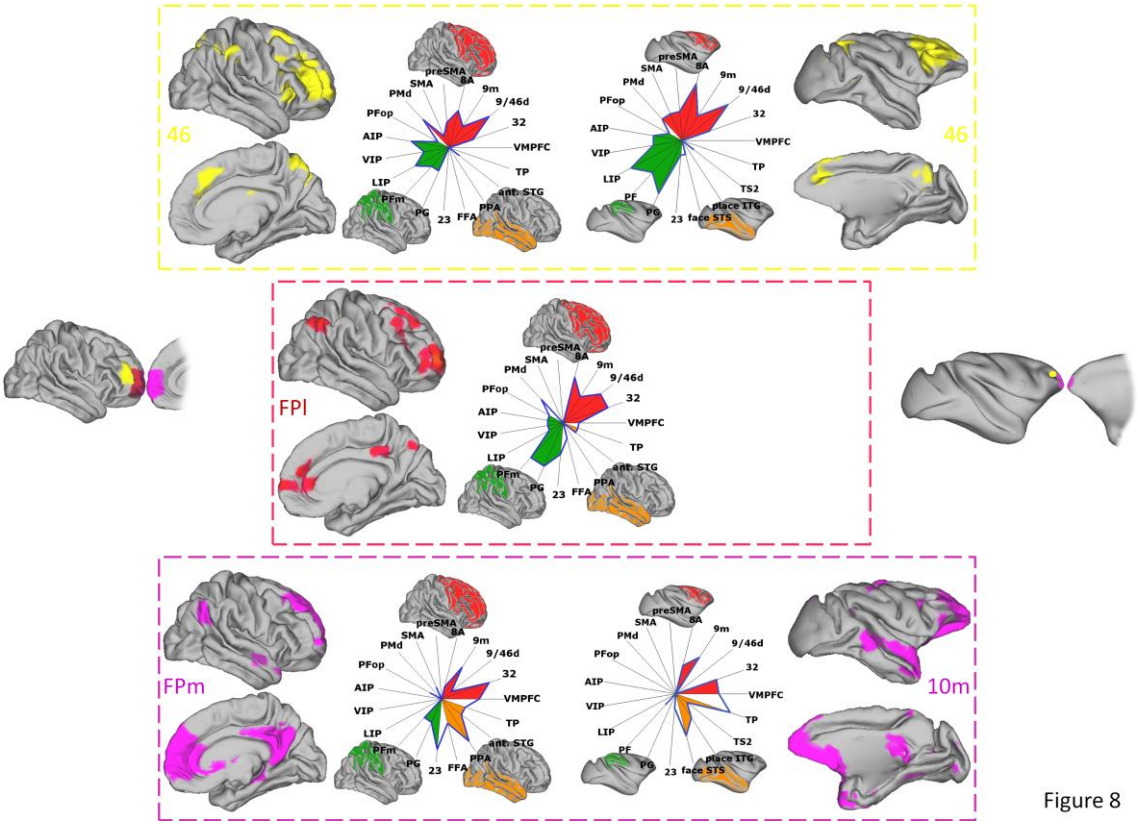


Figure 8

Figure 2-8: Resting state fMRI-derived functional connectivity patterns of human (left) areas 46 (yellow), FPl (ruby) and FPm (pink), and resting state fMRI-derived functional connectivity patterns of the proposed macaque correspondents (right): area 46 (yellow), and area 10m (pink). Area FPl could not easily be matched to any macaque vlFC region but had some features of area 46. In the middle we show spider plots of these regions. Conventions as for Figure 2-2.

2.4 Discussion

We identified fundamental similarities but also striking differences between monkeys and humans in the way vlFC regions link with the rest of the brain. We discuss these results in four parts, focusing briefly on (1) PMv, then (2) posterior vlFC areas associated with cognitive control, (3) vlFC areas and their hypothesized role in language processing in humans, and (4) areas in the most anterior part of the prefrontal cortex.

2.4.1 Ventral premotor cortex

Previous studies have dissociated PMd from PMv using DWI-tractography based estimates of connectivity (197). Here we show further dissociation between superior and inferior regions within human PMv that we refer to as 6v and 6r. A related distinction has also been proposed by Schubotz et al., (210). Their functional connectivity patterns resemble those of macaque areas F5c and F5a, respectively. There was functional coupling between both areas in both species and coupling with anterior intraparietal area (AIP), pre-SMA, and parts of ventral prefrontal cortex. Connections between these areas have previously been reported in tracer injections studies (211, 212). Further subdivision of area 6v, which lies on the edge of the ROI we investigated, may be possible; the existence of coupling between 6v and ventral intraparietal area (VIP) and SMA resembled the anatomical connections of macaque area F4 that are known to exist (211).

Macaque F5 hosts “canonical neurons” encoding specific object-oriented action plans and affordances as well as “mirror neurons” which are active both when monkeys perform object-oriented actions and when they observe the same actions (203). More posterior F5c may represent actions performed in a context-dependent way, while F5a may code actions at a more abstract level (213). Such differences may be a consequence of differences in connectivity with parietal and other sensorimotor areas (214), which were also apparent in our study.

2.4.2 Posterior vlFC and cognitive control

Parts of posterior vlFC have been associated with cognitive control processes, including inhibitory motor control (182), cognitive task control (186), cognitive flexibility (185), and information updating (215). One of the first important results was that DWI-tractography parcellation reliably distinguished five areas - IFJ, 44d, 44v, Op, and IFS - from the surrounding

cortex, this is from both the more posterior PMv and the more anterior and dorsal parts of prefrontal cortex. There has been interest in the possibility that regions of vlFC differ in terms of their involvement in various tasks of cognitive flexibility (216, 217). The parcellation, functional coupling, and meta-analyses we report here all suggest the regions participate in different cognitive processes.

For all five areas we identified areas in macaque vlFC that they resembled (Figs 5 and 6). Neurophysiological studies of cognitive control in macaques have focused on the principal sulcus region, but our results suggest that investigation of areas 44, Ia, and ProM might also shed light on the mechanisms of cognitive control.

For both species the functional coupling patterns of these five regions participated in partly similar networks as premotor areas 6v/F5c and 6r/F5a, and indeed most were also coupled with 6v/F5c and 6r/F5a. From the similarity between the coupling patterns of these five areas and areas 6v/F5c and 6r/F5a their involvement in action-selection may be inferred. However, the additional and distinct coupling of IFJ, 44d, 44v, Op, and IFS with visual association areas in occipito-temporal cortex and dorsolateral prefrontal cortex suggested important differences in function. We propose that these five areas have more flexible control over action-selection – by reference to information from dorsolateral prefrontal cortex about higher order goals and sequential information (218) and with access to richer descriptions of visual features from the temporal lobe that go beyond the action affordances encoded in parietal visuomotor neurons.

2.4.3 vlFC areas associated with language

Although macaques lack language we found regions in macaque vlFC that showed a pattern of functional coupling similar to human vlFC regions implicated in language. Human areas 45 and 47/12, which play roles in semantic processing, resemble macaque areas 45A and 47/12. We also found similar subdivisions in the human left and right hemispheres. Although Broca's area

is only present in the left hemisphere, similarities in cytoarchitecture and receptor-densities in the two hemispheres have been reported before (113).

The development of abilities such as language may, therefore, have exploited pre-existing mechanisms for sensory-motor mapping (190), sequential motor learning (219), or multimodal sensory integration (188, 191).

Although we found similar coupling patterns with premotor and parietal cortex, we also noted pronounced species differences in how most vlFC regions coupled with posterior auditory association areas. Whereas macaque auditory association areas in the posterior temporal cortex coupled strongly with ACC areas known to play a role in social cognition (207), human posterior auditory association areas coupled more strongly with almost all vlFC regions.

These striking discrepancies might relate not just to the prominence of auditory information in human language but even more generally to the poor ability of macaques to use auditory information for a range of purposes. For example, macaques perform poorly in auditory working memory tasks (220) and while they are readily able to learn to use visual information to guide arbitrary and flexible patterns of decision-making, they find it difficult to use auditory information in the same way (221). The ability that humans possess in using auditory and verbal codes to guide decision-making and in short term memory may depend on the coupling between posterior auditory association cortex and ventrolateral prefrontal cortex that only they, and not macaques, possess.

2.4.4 Anterior vlFC

Frontopolar cortex has been associated with representing future and alternative courses of action (27, 222). By contrast its more medial part has been implicated in social cognition (223). In the present study we subdivided it into lateral, FPl, and medial, FPm, subdivisions. Both were distinct from dorsolateral prefrontal area 46. Whereas human area 46 and FPm could be matched to macaque areas 46 and 10 on the basis of their functional connectivity, human area

FPl could not easily be matched to any of the macaque prefrontal cortex (PFC) regions. This supports the notion that human anterior PFC contains a region not found in the macaque and suggests it may support distinctive cognitive abilities. One way in which new regions arise during speciation is as a consequence of connections normally associated with one area invading an adjacent area (224). Human FPl's coupling-pattern may reflect such a process. The existence of a distinctive human FPl region may explain Mantini's and colleagues' (225) recent finding of distinctive components in human fMRI activity patterns.

2.4.5 Comparison with previous studies

The ROI we investigated was defined on the basis of sulcal landmarks but it corresponds approximately to the frontal and opercular component of the fronto-parietal task control and the cingulo-opercular networks or the frontoparietal and ventral attentional networks that have been described previously (193, 226).

Parcellation studies have previously identified areas 44 and 45 within human vlFC (204, 227). In this study we were also able to identify ten other regions beyond areas 44 and 45 including a new frontopolar region FPl. With the exception of FPl we were able to relate them to areas in the monkey brain.

Previous studies of human areas 44 and 45 have focused on estimating the probability that the areas are connected with a limited number of major white matter fascicles. By using similar resting-state BOLD coupling measures in both species, we estimated interaction strengths between each of our twelve vlFC regions and 19 other brain areas in specific regions of temporal, parietal, premotor, and cingulate cortex. This made it possible to show that human and macaque vlFC regions are fundamentally similar in their interactions with specific parietal, premotor, and cingulate regions and with visual association regions in temporal cortex but not with superior temporal auditory association cortex. VlFC circuits in humans and macaques differ in the manner they interact with an important part of auditory association cortex, but

they greatly resemble each other in the manner they interact with visual, action-selection and decision-making areas, thereby emphasizing the role of human vlFC, even human vlFC areas such as 44 and 45, in functions other than language.

2.4.6 Limitations

The DWI-tractography parcellation technique identifies reproducible distinctions between brain areas (192, 228, 229) and areas that correspond with histologically defined brain areas, both in terms of their positions in standard MNI space and with respect to sulcal landmarks (192, 230, 231). This suggests that areas we identify here correspond to separable anatomical entities within vlFC that have a reliable location in standard MNI space. However, it is important to note that probabilistic DWI-tractography only *estimates* the strength of evidence for a connection and that there are both false positives and negatives in the connections it estimates (162, 232). However, the correspondence between DWI-tractography parcellation and histologically defined brain areas reflects the fact that the parcellation approach draws on estimates of connectivity between the ROI and every other voxel in the brain rather than just on estimates of connectivity with a single region or pathway which might be inaccurately characterized.

Other laboratories have attempted to parcellate human cortex on the basis of differences in functional connectivity measured with fMRI and there are similarities, albeit not complete ones, in the schemes that have been proposed using this approach (193, 195, 226) and between such approaches and schemes employing DWI-tractography based parcellation approaches (192, 233). This probably reflects the fact that functional coupling between any two brain areas is a function of the anatomical interconnections between the areas; functional coupling is reduced when direct anatomical connections between the areas are removed but only abolished when indirect anatomical connections are also removed (147). Functional connectivity measurements have the advantage that they are less sensitive to inter-areal

distance than are DWI measurements. By drawing on both DWI-tractography and fMRI connectivity we have attempted to combine the strengths of each method but note that it cannot replace detailed anatomical investigation both in animal models and human tissue acquired post-mortem (e.g. 113, 209). Nevertheless, we hope that our results contribute to an understanding of how data derived from animal models can be used to aid understanding of human brain function.

2.5 Methods

2.5.1 Diffusion-weighted tractography-based parcellation

Diffusion-weighted images were acquired in 25 healthy right-handed participants on a 3 T Siemens Magnetom Verio MR scanner using standard DW-MRI protocols (SI.1). Analyses were performed using tools from FSL (Functional MRI of the Brain Software Library), the Human Connectome Project Workbench, and custom made software written in Matlab (MathWorks). DW-MRI data were pre-processed in a standard way as previously described (192; SI.2).

For each participant, probabilistic tractography was run from each voxel in the right vlFC ROI (the same steps were also performed for the left vlFC; SFig 3) to assess connectivity with every brain voxel (whole brain "target" was down-sampled after tractography to 5 mm isotropic voxels for the connectivity matrix to be manageable, however the whole vlFC ROI was tracked in original FA space, see Figure 2-1), using a model accounting for multiple fibre orientations in each voxel (199). A connectivity matrix between all right vlFC voxels and each brain voxel was derived and used to generate a symmetric cross-correlation matrix of dimensions (number of seeds x number of seeds) in which the (*i,j*) element value is the correlation between the connectivity profile of seed *i* and the connectivity profile of seed *j*. The rows of this cross-correlation matrix were then permuted using k-means segmentation for automated clustering to define different clusters (Figure 2-1c and SI Figure 6-1). The goal of clustering the cross-

correlation matrix is to group together seed voxels that share the same connectivity with the rest of the brain. The parcellation into *ten* distinct regions was particularly consistent across subjects and showed the smallest variance of the average cross-correlation suggesting the most coherent clustering into similarly correlated clusters (SI Methods 6.1.3 discusses determination of regions and concept of "brain region"). We carried out an additional analysis in which we iteratively parcellated vlFC into smaller and smaller regions and identified the same ten regions (SI Methods 6.1.4 and SI Figure 6-2). Using this approach, however, we showed that two of the areas could be subdivided bringing the total areas identified to twelve.

2.5.2 Human resting-state fMRI data acquisition, preprocessing, and analysis

Human resting-state fMRI data were collected for the same group of 25 healthy volunteers in the same session using the same scanner. Participants were instructed to lie still and keep their eyes open and fixated at a cross. Resting state fMRI data acquisition and pre-processing were carried out in a standard way as previously described (192; SI.5)(SI Methods 6.1.5).

To establish the functional connectivity of each vlFC regions, we created ROIs from every parcel in each individual subject's parcellation. As representation of each parcel we took the 5% most likely voxels from each parcel's whole vlFC certainty maps (which in MNI-space are always 170 voxels for a 3400 voxel vlFC ROI) from the fuzzy-clustering derived parcel-specific maps. The fuzzy-clustering algorithm gives a certainty of each voxel belonging to each of the clusters and hence provides a "certainty map" of the whole vlFC ROI for each cluster. Note however that we also tried extracting the time-courses from each parcel as a whole from the "normal" k-means clustering and obtained largely similar resting state functional connectivity patterns. Moreover we drew cubic regions of interest (3x3x3 voxels, i.e. 6mm isotropic) at the centre of gravity of each individual subject's parcel and also at the centre of gravity of each parcel in the group maps and obtained qualitatively similar results. Individual ROIs were

registered from each subject's DW-MRI space into fMRI space using FLIRT. Then the major Eigen time series representing activity in each of the vlFC clusters was calculated.

Individual statistical maps were then calculated using a seed-based-correlation analysis which is part of FSL (fsl_sbca) as previously described (192, 234) in order to infer the functional connectivity of these ROIs with the rest of the brain. For each vlFC subdivision we created a model consisting of the first Eigen time series of that region and the confounding time series representing head movement (six regressors resulting from motion correction using MCFLIRT) and the Eigen time-series of white-matter and cortico-spinal fluid (CSF). The results of each individual subject's ROI-specific seed-based-correlation analysis were then entered into a general linear model (GLM) analysis. The resulting Z-statistical images were thresholded using clusters determined by $Z > 4.0$ and a (corrected) cluster significance threshold of $P<0.05$. These thresholded group Z maps were projected onto the CaretBrain as provided by the Human Connectome Project Workbench using the "surf_proj" algorithm as implemented in FSL and then visualised using the Human Connectome Project Workbench. Unthresholded Z maps were quantified by extracting the average intensity of each cluster's functional connectivity Z-map in a number of cortical regions of interest (SI Table 6-1, SI Table 6-2 and SI Table 6-3). We drew 3x3x3 voxels, i.e. 6mm isotropic, ROIs centered on the coordinates mentioned in Suppl. Table 1 and then averaged the Z-value from the unthresholded seed-based-correlation analysis derived Z-maps within these ROIs. These values for all ROIs and all vlFC areas were then binned into four equally sized bins. These were then displayed on a spider plot (Figure 2-4, Figure 2-5, Figure 2-6 and Figure 2-8).

2.5.3 Macaque resting-state fMRI data acquisition, preprocessing and analysis

Resting-state fMRI and anatomical scans were collected for 25 healthy macaques (*Macaca mulatta*) (four females, age: 3.9 years, weight: 5.08 kg) under light inhalational anesthesia with isoflurane (for detailed information on anesthesia protocol, monitoring of vital signs, data

acquisition and pre-processing see SI Methods 6.1.7). Protocols for animal care, MRI, and anesthesia were performed under authority of personal and project licenses in accordance with the United Kingdom Animals (Scientific Procedures) Act (1986).

The goal of this part of the study was to test for similarities between the functional networks previously established for human vlFC and vlFC regions in the macaque. We therefore aimed to map the resting state functional connectivity networks of all cyto-architectonically described areas in macaque vlFC and adjacent ventral premotor and insular areas. Functional connectivity networks of several regions [the lateral area 10, medial area 10, ventral area 10, area 46, area 9/46v, area 47/12, area 45A, area 45B, area 44, area 8A (187, 235); motor proisocortex (ProM), the anterior insula (Ai), and the gustatory area (236); areas F5a, F5c, and F5p (203)] were identified using a seed-based-correlation, analysis which is part of FSL (fsl_sbca), as in the human subjects, to infer the functional connectivity of these ROIs with the rest of the brain. The seed regions were specified in stereotactic atlases of the macaque monkey brain (237, 238) and then drawn on the standard brains of the respective atlas. They could then be registered between atlases using FLIRT.

2.5.4 Comparison of resting-state functional connectivity of macaque and human vlFC

We identified the macaque regions most comparable to human vlFC regions in terms of functional connectivity. The functional connectivity between each vlFC area and 19 areas in the rest of the brain (each of which corresponded in both humans and macaques: SI Table 6-2), was plotted on spider plots for both species. The ROIs were 6mm isotropic (coordinates in SI Table 6-1) for humans and 3mm isotropic for macaques identified using stereotactic atlases of the macaque monkey brain and then drawn on the standard brains of the respective atlas. We then determined which macaque vlFC region corresponded most closely to each human vlFC.

A formal comparison between human and macaque vlFC coupling patterns was performed by calculating the summed absolute difference [the "Manhattan" or "city-block" distance (196, 202)] of the normalized coupling scores after binning into four equally sized bins. This yielded a summary measure (SI Figure 6-6) of the difference in coupling patterns for each pair of areas in the two species. The summary measure can then be used to compare the functional coupling pattern of each human vlFC region with those of all vlFC regions in the macaque.

2.5.5 Functional coupling of visual and auditory association regions with regions in vlFC and medial frontal cortex

To compare functional coupling between vlFC and auditory association cortex in posterior superior temporal cortex in humans and macaques we rank-ordered the strength of functional coupling (from lowest to highest) between area Tpt and ten vlFC areas that corresponded in both humans and macaques (average z-value in cubic ROI 6x6x6mm in human and 3x3x3mm in macaque) among the 19 previously described target regions (same regions as in the spider-plots). In other words, we assessed how strong the coupling between area and Tpt and each vlFC sub-region ranked in comparison to coupling between Tpt and each of the other 19 areas (VMPFC, area 32, 9/46d etc.). We then performed a Mann–Whitney–Wilcoxon rank-sum test (significance level $p<0.05$) comparing the rank areas assigned to the coupling strength between each vlFC region and Tpt in the two species (Figure 2-7, SI Figure 6-8 presents the same analysis not only for area Tpt [left part of Fig] but also for anterior auditory association area TS2, SI Figure 6-9 presents the same analysis for the left hemisphere).

Another analysis compared the functional coupling of area Tpt (Figure 2-7; for area TS2 see SI Figure 6-8 and SI Figure 6-9) and nine cingulate cortex areas (Beckmann et al., 2009). This means that there were always 29 (vlFC) or 28 (cingulate) ROIs in each of the two separate analyses. The average z-values in the target ROIs representing coupling of area Tpt with vlFC or

cingulate areas were ranked for each individual and compared between species with a Mann–Whitney–Wilcoxon rank-sum test ($p<0.05$ significance level). A rank of 29 (vlFC analysis) or 28 (cingulate analysis) was the highest possible rank (strongest coupling of an ROI among all 29/28 ROIs).

3 Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex

3.1 Abstract

Reward-guided decision-making depends on a network of brain regions. Among these are the orbitofrontal and the anterior cingulate cortex. However, it is difficult to ascertain if these areas constitute anatomical and functional unities, and how these areas correspond between monkeys and humans. To address these questions we looked at connectivity profiles of these areas using resting state functional MRI in 38 humans and 25 macaque monkeys. We sought brain regions in the macaque that resembled ten human areas identified with decision making and brain regions in the human that resembled six macaque areas identified with decision making. We also used diffusion-weighted MRI to delineate key human orbital and medial frontal brain regions. We identified twenty-one different regions many of which could be linked to particular aspects of reward-guided learning, valuation, and decision-making and in many cases we identified areas in the macaque with similar coupling profiles.

3.2 Introduction

As humans we make decisions by taking into account different types of information, weighing our options carefully, and eventually coming to a conclusion. We then learn from witnessing the outcome of our decisions. Human functional magnetic resonance imaging (fMRI) has had a major impact on elucidating the neural networks mediating decision-making and learning but key insights can only be obtained in neural recording, stimulation, and focal lesion studies conducted in animal models such as the macaque. Combining insights from human fMRI and animal studies is, however, not straightforward when there is uncertainty about basic issues such as anatomical and functional correspondences between species (37). For example, while there are many reports of decision value-related activity in human ventromedial prefrontal

cortex (vmPFC) (239, 240) it is unclear whether they can be related to reports of reward-related activity either on the ventromedial surface of the frontal lobe (241, 242), in the adjacent medial orbitofrontal sulcus (243), or indeed to any macaque brain area. It is claimed other areas implicated in reward-guided decision making and learning, such as parts of anterior cingulate cortex (ACC), are not found in macaques (244) but such theories have never been formally tested.

In addition, there is simply uncertainty about the basic constituent components of decision-making and learning circuits. To return to the example of vmPFC, while this region is often contrasted with similarly large subdivisions of frontal cortex such as lateral orbitofrontal cortex (lOFC) and ACC (47) it is unclear whether, and if so how, it should be decomposed into further subdivisions. Moreover, there are sometimes fundamental disagreements about how brain areas contribute to decision-making and learning. For example, it has been claimed both that ACC does (44, 54, 72) and does not (245) contribute to reward-based decision-making and that it is concerned with distinct processes for task control, error detection, and conflict resolution (186, 246). Reliable identification and location of ACC subcomponent regions could assist the resolution of such debates.

In the current study we formally compared brain regions implicated in reward-guided decision-making and learning in humans and monkeys and attempted to identify their key sub-divisions in relation to function (Figure 3-1). We used fMRI in 25 monkeys and 38 humans to delineate the functional interactions of “decision-making regions” with other areas in the brain while subjects were at rest. Such interactions are reliant on anatomical connections between areas (147) and determine the information an area has access to and the way it can influence other areas and thereby behaviour. Each region of the brain has a defining set of interactions, a connectional or interactional “finger-print” (247), that can be compared across species (196, 202, 248). We focused on areas throughout the entire medial and orbital frontal lobe including ACC, lOFC, vmPFC and frontal pole (FP) that have been related to decision making in humans

and monkeys. The results suggested areal correspondences between species as well as finer functional fractionations within regions than previously assumed. In a second step we used a complementary technique, diffusion-weighted magnetic resonance imaging (DW-MRI), to confirm the existence of 21 distinct component regions within human medial frontal and orbitofrontal cortex. The results suggest every day human decision-making capitalises on a neural apparatus similar to that supporting decision making in monkeys.

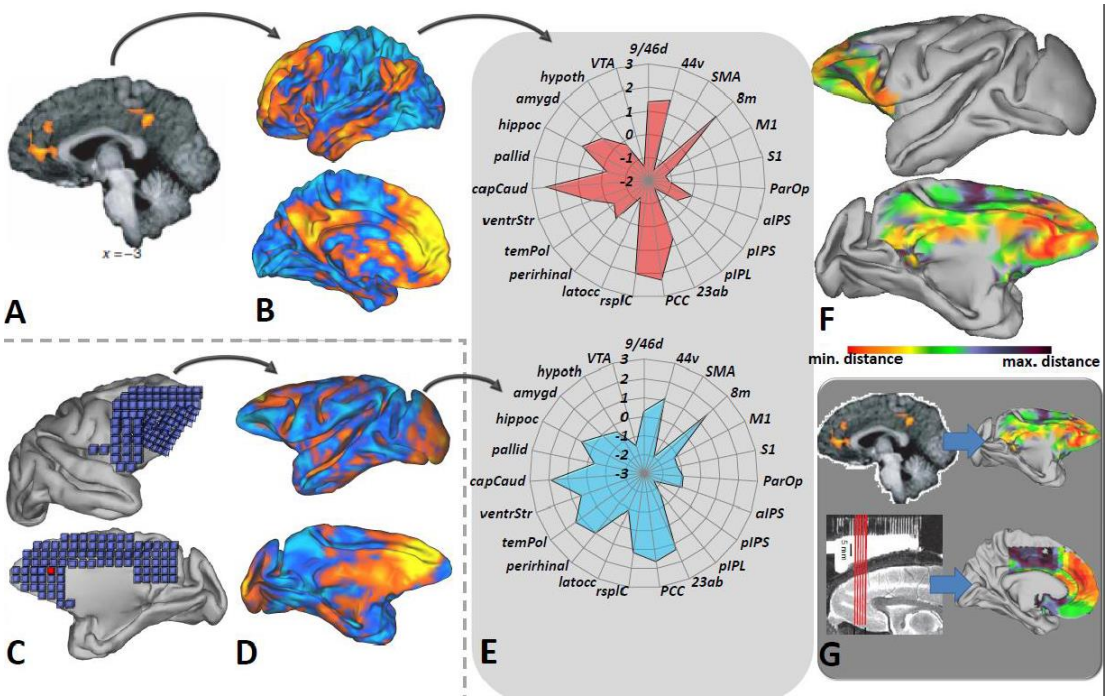


Figure 3-1: Overall approach of the study. FMRI analyses in 38 humans and 25 macaques were used to establish the whole brain functional connectivity of regions in medial and orbital frontal cortex identified with reward-guided learning and decision making in the two species. The example shows the macaque brain regions that have a similar coupling profile to a human vmPFC region (A) identified in a decision making study (32). (B) Each region's functional connectivity with 23 key regions was then determined and (C) summarized as a functional connectivity fingerprint. (D) Once the functional connectivity fingerprint of a human brain area was established it was compared with the functional connectivity fingerprints of 380 ROIs in macaque orbital and medial frontal cortex (one example is shown here) by calculating the summed absolute difference [the "Manhattan" or "city-block" distance (196, 202, 248) of the coupling scores. (E) Examples of the functional connectivity fingerprints for a human (blue) and a monkey (red) brain area. Most monkey ROIs matched human areas relatively poorly and extremely good and extremely bad matches were relatively rare. We used two standard deviations below the mean of this distribution of summed absolute differences as a cut-off to look for "significantly" good human to monkey matches. (F) A "heat map" summarizing the degree of correspondence between the functional connectivity patterns of each voxel in the macaque and the human brain region shown in panel a. Warm red areas indicate macaque

voxels that correspond most strongly. (G) Complementary parts of the investigation started with the functional connectivity fingerprints of both human (top) and macaque (bottom) brain areas involved in reward-guided learning and decision making and then compared them with the functional connectivity fingerprints of areas in the other species.

3.3 Results

The first goal was to relate human frontal cortical regions implicated in decision making and learning in neuroimaging studies to monkey brain regions (Figure 3-1F). We also did the reverse; we identified human brain regions resembling monkey frontal cortical regions implicated in decision making and learning.

We took the following approach: each time we looked at a specific human region we compared its functional coupling pattern to those associated with a set of 448 different monkey regions of interest (ROIs) within ACC, lOFC, vmPFC and FP trying to find one that would constitute the closest monkey match to the human region. We represented the closeness of the match at each voxel as a heat map. In most analyses several warm regions were apparent in the map but in each case the hottest area was found in an orbital or medial frontal region and it was this region that was taken as the best candidate for interspecies correspondence and highlighted with an arrow in Figure 3-2 – 2-5. Notably, the other warm areas in the map tended to be ones with which the frontal area being investigated was connected; in other word the approach identified areas in similar functional circuits in the other species but in each case it particularly highlighted one potential homologue in frontal cortex. We also did the reverse by matching monkey regions to 417 different human ROIs trying to find the best human match for any given monkey area. The comparisons were based on the frontal regions' activity coupling with 23 cortical and sub-cortical ROIs already identified as comparable in the two species (SI Table 7-3).

3.3.1 VmPFC, perigenual ACC, and subgenual ACC

Human vmPFC activity has been linked to subjective values of objects and choices (56, 249, 250). Positive correlation between activity and the values of chosen options and negative correlations with values of rejected options suggests a role in decision-making (56) or attentional selection (251). However, there is also evidence vmPFC tracks values of items even in the absence of any decision or when watching somebody else choosing (252, 253). Some studies have suggested these value representations are independent of reward type (money, food) (31) and reflect the impact of other factors affecting valuation such as delay (32) indicating a "common currency" value representation. However, other studies have hinted at a posterior-to-anterior gradient of increasing abstractness in value representation (254). Here we show that this diversity of activity patterns is a consequence of the existence of several distinct component regions within human vmPFC that can each be linked to different regions in the macaque.

We took the peak Montreal Neurological Institute (MNI) coordinates from a study implicating vmPFC in decision-making (56) and established its functional coupling pattern in humans (Figure 3-2A). Similar activations have been reported nearby (54, 255, 256). It had positive functional coupling with OFC, retrosplenial, lateral occipital, inferior temporal, posterior temporo-parietal junction (TPJp) and perirhinal cortex, as well as considerable coupling with the amygdala, hypothalamus, and ventral striatum. However, it is negatively coupled with lateral FP (FPl) cortex, ACC, and mid-inferior parietal lobule (IPL). A similar coupling pattern was seen when we took a coordinate from a food valuation task reported by McNamee and colleagues (254). A coupling pattern very similar to this was found in the monkey on the medial gyrus rectus near area 14m as defined by Mackey and Petrides (111).

We next examined the coupling pattern associated with a more anterior vmPFC region implicated especially in representing values of abstract goods and choices (254) (Figure 3-2B). Like adjacent regions it exhibited some degree of coupling with amygdala, hypothalamus,

ventral striatum, medial temporal cortex, and temporal pole but had stronger coupling with posterior cingulate cortex, precuneus, and the head of the caudate. It did not show negative coupling with the FPl, but instead with supplementary motor area (SMA), pre-SMA, M1 and IPS. This region's fMRI coupling pattern matched that of the monkey's anterior gyrus rectus and frontal pole in areas 11m and 10m (111, 238).

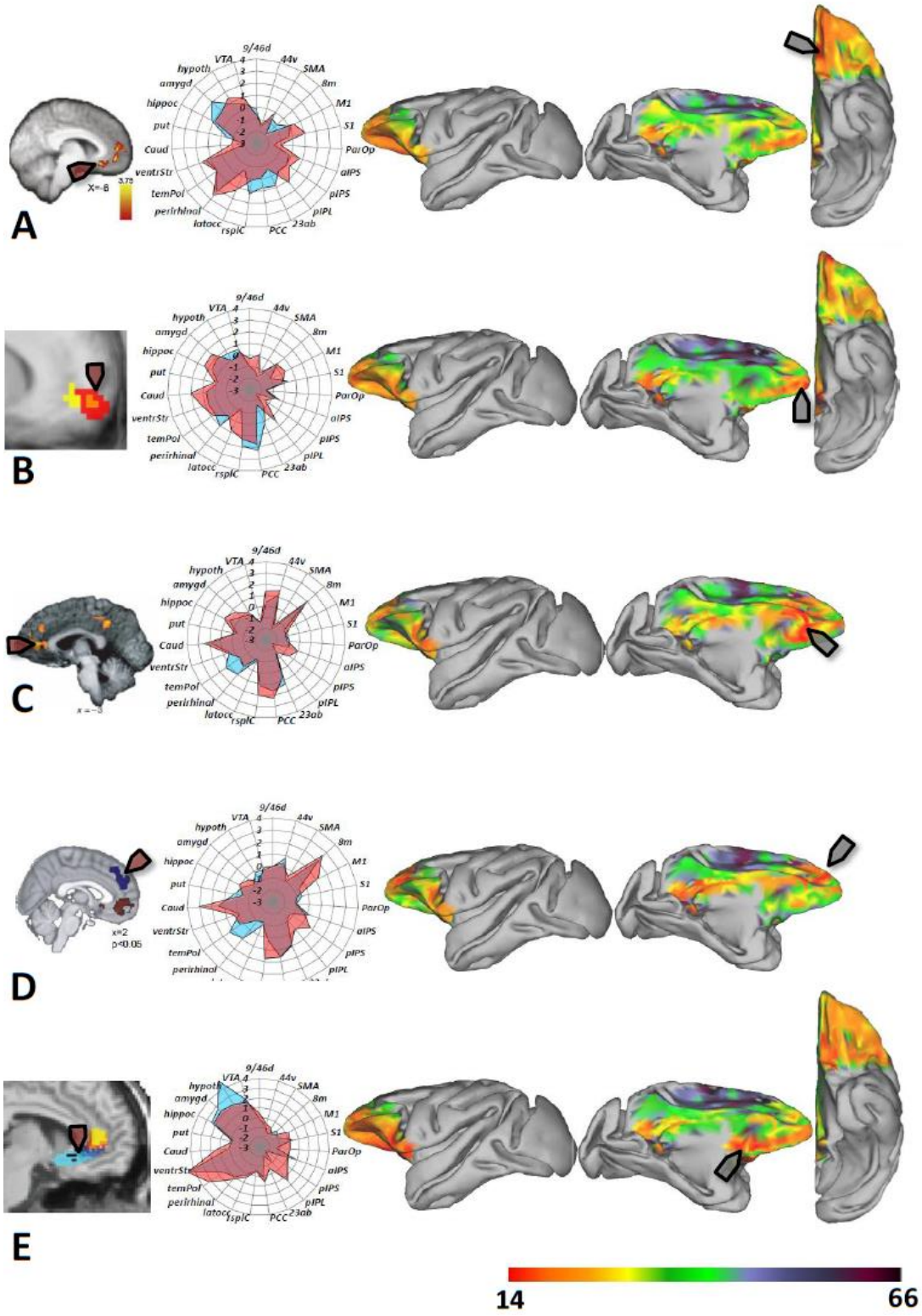


Figure 3-2: Human medial frontal regions (left) linked with (A) reward-guided decision making (56), (B) more abstract reward-guided decision-making (254), (C) cost-benefit valuation (32), (D) imagining the reward outcomes of others (257), and (E) self-valuation and depression (258), could all be linked to macaque brain regions (right) via similarities in their coupling patterns (center, blue = macaque, red = human). However, a human brain area (D) associated with reward outcome imagination did not correspond in a simple way with any area in the macaque.

We looked at a region identified by Kable and Glimcher (32) (Figure 3-2C) that has activity that closely tracks the subjective values of delayed monetary rewards. The region is more dorsal and closer to the genu of the corpus callosum than the value-comparison regions discussed so far (Figure 3-2A, B). One suggestion is that it has a more direct role in tracking values rather than in making value guided decisions (259-262). This region's coupling pattern was also distinct. Although it coupled with temporal pole and TPJp, it also coupled with posterior cingulate cortex, precuneus, dorsolateral prefrontal cortex (dlPFC), FPl and the head of the caudate. It correlates negatively with SMA, dorsal premotor area (PMd), M1 and IPS. In a further analysis we examined an even more dorsal perigenual region linked to individual variation in cost benefit decision making (fig.4d in 54, and in the supplementary information in 60) and found that it was associated with a similar pattern of coupling although now there was less coupling with medial temporal regions such as the amygdala and with dorsal and ventrolateral prefrontal cortex and more coupling with premotor areas such as SMA. Voxels with similar coupling patterns were also found in the macaque in an arc of perigenual cortex corresponding to cingulate area 24 and perhaps part of area 32 (263). It is interesting to note that another region showing high similarity is in retrosplenial areas 30 and 29. This reflects the fact that regions that are connected, as is the case for pgACC and retrosplenial cortex (264), tend to have similar connections with other regions across the brain. Importantly, the best matching area of human pgACC is macaque area pgACC.

This region is not identical with the one previously mentioned to be related to value-comparison in vmPFC. We formally tested the region's coupling patterns for significant differences using permutation testing and cluster-mass thresholding (SI Figure 7-7A). We established that whereas vmPFC is more strongly coupled with OFC, inferior temporal and temporal pole areas perigenual ACC was more strongly coupled with posterior cingulate and retrosplenial cortex. These significant differences were largely similar across species and are

consistent with the notion that these regions have different roles to play in value representation and reward-guided decision-making (259-262).

Activity has also been reported in a more dorsal and anterior region (35). Nicolle and colleagues (35) found activity here when subjects were asked to imagine what values delayed monetary rewards would have for another person. A similar region was active when people imagined how much they would like new and unexperienced food items which were nevertheless composed of familiar components (30). The resting fMRI coupling pattern associated with this region resembled that associated with the two perigenual regions reported by Kable et al and Kolling et al (Figure 3-2E), however, the comparative weakness of coupling with amygdala and temporal cortex areas meant that it most closely resembled a swathe of tissue in the macaque that extended from the anterior cingulate sulcus through area 9 on the dorsal convexity to the principal sulcus. These dorsomedial prefrontal cortex areas (dmPFC) can be reliably dissociated from vmPFC areas. We conducted a formal test for significant differences (SI Figure 7-7B) in coupling patterns between dmPFC, (areas 32d and 9 from the connectivity based parcellation, see below) and vmPFC (areas 14m and 11m). dmPFC was significantly more coupled with cingulate motor areas, pre-SMA, dlPFC, and posterior ventrolateral prefrontal cortex (vlPFC) as well as the head of the caudate, whereas vmPFC was significantly more coupled with other parts of OFC, inferior temporal and lateral occipital, as well as precuneus and ventral striatum. These significant differences were largely similar across species.

Another more posterior and subgenual vmPFC region is implicated in the altered pattern of self-valuation associated with depression (265) (Figure 3-2D). This area is the target of deep brain stimulation in patients with treatment resistant depression (258). We looked at the coupling pattern of this region and found it resembled other vmPFC regions; it shared strong positive functional coupling with OFC, posterior inferior temporal and perirhinal cortex but had stronger coupling with ventral striatum, amygdala and hypothalamus. In addition it had strong

coupling with medial temporal cortex and temporal pole. It was strongly negatively coupled with FPl, dlPFC and dmPFC. A similar coupling pattern was observed in the monkey in subgenual cingulate voxels in area 25 (111). This region's coupling pattern however could be distinguished from that of the more anterior vmPFC identified with value-comparison and decision making. When a formal statistical comparison of the coupling patterns associated with these regions in mid vmPFC and subgenual vmPFC (Figure 3-2A,B) was made it was clear that there was a significant difference (SI Figure 7-8A). Moreover a three way comparison of subgenual vmPFC, anterior vmPFC, and pgACC confirmed that all three regions were robustly separable (SI Figure 7-8B).

Next, we wanted to relate areas identified in macaques back to human frontal cortex (Figure 3-3). Only a small number of studies have recorded in this area of the macaque brain. Monosov and Hikosaka (242) report two sub-regions in macaque vmPFC: a ventral sub-region contained neurons persistently more active when monkeys experienced appetitive stimuli while a more dorsal sub-region was more active when monkeys perceived aversive stimuli or "punishment". By establishing the functional coupling of these two adjacent monkey vmPFC sub-regions we were able to match them, respectively, to a more antero-ventral (Figure 3-2A) and a more postero-dorsal region within human vmPFC (Figure 3-3B). The coupling patterns associated with the more antero-ventral region resembled the region linked to simple value-guided choices (Figure 3-2A) while the coupling pattern associated with the postero-dorsal area resembled that associated with the subgenual region linked to altered self-valuation and depression (Figure 3-2D).

Amemori and Graybiel (72) have reported neurons in a dorsal perigenual cingulate region that play an important role in cost-benefit decision making. The fMRI coupling pattern associated with this region was distinct to that associated with the areas studied by Monosov and Hikosaka (242) but it resembled to some degree that associated with voxels in human dorsal perigenual cingulate cortex (Figure 3-3C) near regions implicated in cost-benefit integration in

humans (32) (Figure 3-2C) especially the more dorsal ones studied by Kolling and colleagues (54, 60).

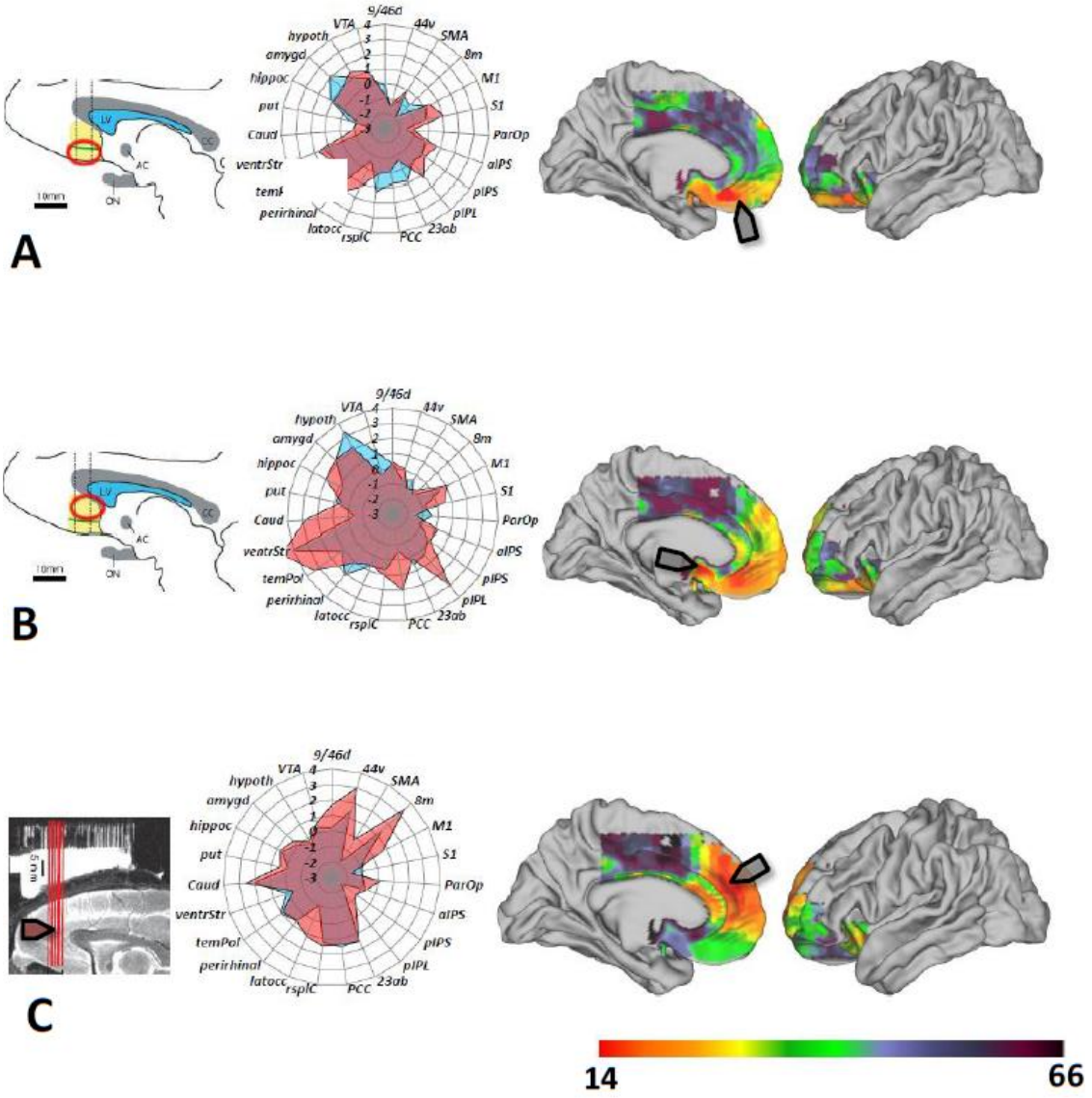


Figure 3-3: Macaque medial frontal regions (left) associated with (A) positive reward expectations (242), (B) negative outcome expectations (242), and (B) cost-benefit decision making (72) could all be linked to human brain regions (right) via similarities in their coupling patterns (center).

3.3.2 Dorsal ACC

Dorsal ACC has been linked with executive control, inhibition, task switching, conflict resolution, and motor planning (186, 215, 266, 267). Here we tried to establish whether a so-

called “task positive” region in human ACC (186) (Figure 3-4A) could be found in monkey ACC. In humans this region is consistently “active” in several different cognitive control tasks and activity recorded here is sometimes referred to as being in the posterior rostral cingulate zone (RCZp) (268). Based on its strong coupling with sensorimotor areas (M1, S1, SMA, pre-SMA, PMd, PMv), areas in intraparietal sulcus (IPS, anterior IPS, and posterior IPS), moderate coupling with dlPFC, and strong negative coupling with temporal lobe areas (TPJp, temporal pole, and STS) we were able to match it to a monkey ACC region. This region was located in the middle of the cingulate sulcus close to CMAv (103).

Other human neuroimaging studies have implicated ACC in learning and updating values of choices {Walton, 2004 #99;Behrens, 2007 #153;O'Reilly, 2013 #100;Procyk, 2014 #2593}. Activity in this region is sometimes described as being in RCZa {Picard, 1996 #98;Amiez, 2013 #2592;Amiez, 2012 #1370} (Figure 3-4B). This region’s strong positive functional coupling with dlPFC, parietal operculum, insula and pallidum and strong negative coupling with vmPFC, ventral striatum, amygdala, posterior IPL, and temporal lobe areas made it similar to a monkey region in the anterior cingulate sulcus near CMAr (103).

Both of these ACC regions, RCZa/CMAr (Figure 3-4B) and RCZp/CMAv (Figure 3-4A) had quite distinct coupling patterns to the perigenual ACC regions discussed in the previous section (Figure 3-2C, Figure 3-3C). In addition they were distinct to other even more posterior ACC regions that correspond to the caudal cingulate zone/dorsal cingulate motor area (CCZ/CMAd) and to more ventral parts of area 23a (area 23/ab) (SI Figure 7-2). This was confirmed with a formal statistical comparison between these three region’s coupling patterns (SI Figure 7-9A). However the three cingulate motor areas CCZ, RCZp and RCZa also shared fundamental similarities in their coupling patterns: All three showed some degree of coupling with pre-SMA, SMA, dlPFC and premotor as well as anterior, IPL caudate and putamen. We looked for a significant gradient of change in functional coupling from anterior to posterior cingulate motor regions and showed that moving anteriorly increased coupling with dlPFC, and parts of vlFC

and insula, whereas moving posteriorly led to increased coupling with posterior cingulate cortex and precuneus, sensorimotor and parietal regions (SI Figure 7-9B). Finally, another statistical comparison between the coupling patterns associated with RCZa/CMAr, RCZp/CMAv, and pgACC further confirmed the existence of differences in coupling between the most anterior cingulate motor area and the dorsal part of the pgACC (SI Figure 7-9C).

Finally we sought to identify human brain regions resembling the ACC region in which neurons have been recorded that encode reward prediction errors (271, 272) (Figure 3-4C). This region most closely resembled the human RCZa (Figure 3-4B).

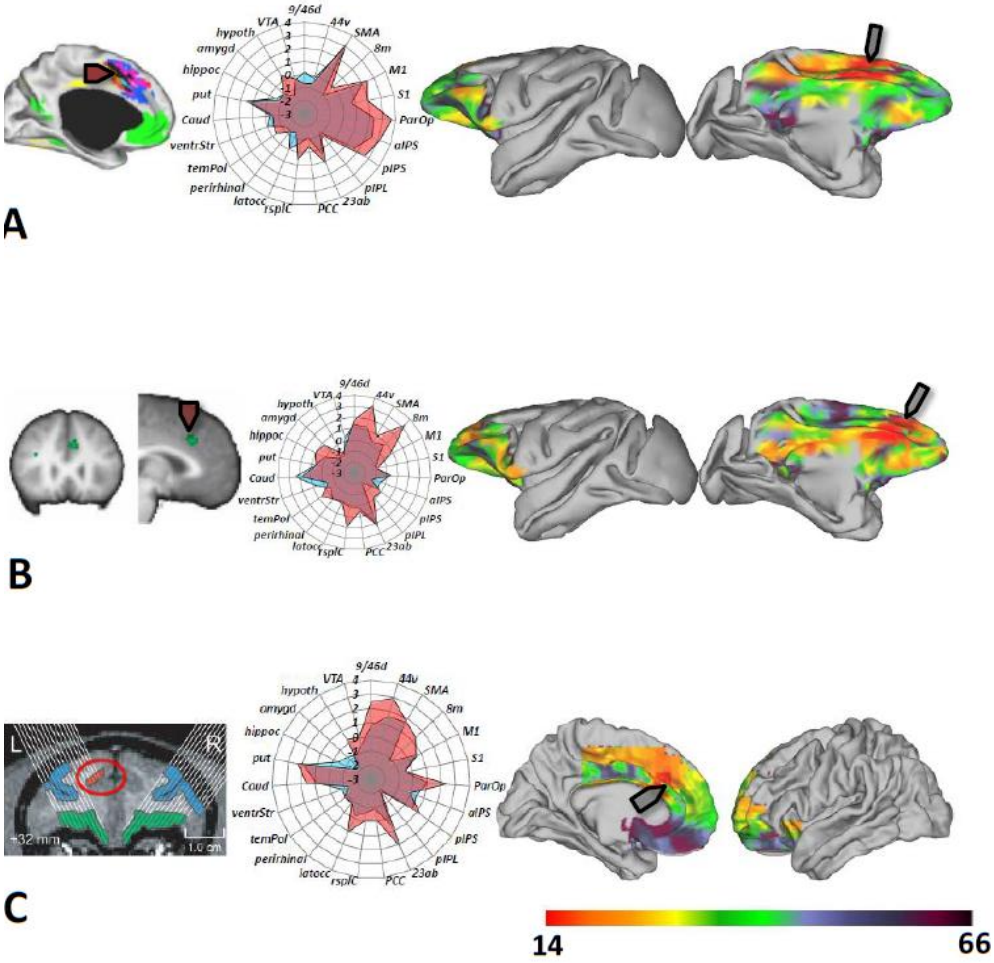


Figure 3-4: Human ACC regions (left) linked with (A) cognitive control (186) and (B) reward-guided action selection and behavioral updating (82) could be linked to macaque brain regions (right) via similarities in their coupling patterns (center). (C) A macaque ACC region (left) linked with reward-guided behavioral updating (271) could be linked to a similar human ACC region to that shown in panel B.

3.3.3 OFC and FP

In macaques neurons in the medial orbital sulcus in central OFC have been associated with context-independent value representation (273) whereas lOFC has been associated with credit assignment (57). There is also evidence of a link between human lOFC and credit assignment (274). Human FPl has been implicated in coding the value or relevance of alternative choices or strategies that are not pursued at the moment but which might be in the near future (47).

We attempted to match human posterior lOFC and anterior lOFC/FP regions with regions in the monkey. There has been particular interest in the possibility that human lOFC, especially posterior lOFC, is important when adverse outcomes, including punishment, lead to behavioural change (259, 275). However, it has proved difficult to identify a particular region of macaque OFC in which neurons are specialized for responding to punishment and instead reward-related neurons have been found across the medial-lateral extent of OFC (276-278). One possibility is that there is a major difference between human and macaque OFC but an alternative hypothesis is that lOFC, even in humans, is not just concerned with punishing outcomes but that it is concerned with using any outcome, positive or negative, to guide behavioural change; for example positive outcomes might lead to behavioural changes in learning situations. Consistent with this idea some recent studies have suggested that the same lOFC region previously highlighted as responsive to punishers (259, 275) has a more general role in credit assignment and the linking of stimuli to reward outcomes (274, 279) (Figure 3-5A). By contrast more anterior lOFC areas in humans are active during decision making under ambiguity (280) (Figure 3-5B).

The human posterior lOFC region was, in terms of its strong positive functional coupling with vmPFC, ventrolateral prefrontal cortex (vlFC), ventral striatum, and temporal cortex and its negative coupling with ACC, sensorimotor, and parietal areas, most similar to a region in monkey lOFC. Such a correspondence is consistent with this region being concerned with the

mediation of behavioural change as a result of a general role in learning stimulus reward associations (48, 274, 279). However, we were unable to match the more anterior human lOFC region linked to ambiguity (Figure 3-5B), that positively coupled with mid-IPL, FPl, and dlPFC, and negatively coupled with ACC, vmPFC, precuneus, ventral striatum, and temporal cortex, to any of the 448 ROIs in the monkey frontal cortex. Formal statistical comparisons of the coupling patterns associated with human anterior and posterior lOFC confirmed the existence of differences in both species (SI Figure 7-10A).

Next we sought human regions corresponding to areas investigated in the macaque. Neurons in monkey medial orbitofrontal sulcus encode context-invariant values of goods (273). This region showed strong positive coupling with the rest of OFC, regions in the temporal lobes (temporal pole, parts of the hippocampus, superior and inferior temporal gyri, perirhinal cortex), lateral occipital cortex and ventral striatum and negative coupling with sensorimotor areas and premotor areas (Figure 3-5C). It matched a region with a similar coupling profile in the posterior part of human central OFC. This part of the human brain is not the same as the more medial vmPFC regions that have been the focus of investigations of value encoding in the human brain (Figure 3-2) but it is similar to a region in which activity has been related to the identity of reward outcomes (279). In this central OFC region, unlike in lOFC, activity is related to the identities of rewards rather than the association between rewards and predictive cues. We conducted a formal statistical comparison between the coupling patterns of a vmPFC/medial OFC region and a lateral OFC region (SI Figure 7-10B). This comparison confirmed significant differences between these two regions, which have previously been linked with processing of positive and negative outcomes (259) or value comparison and credit assignment (47), respectively. Statistical comparisons of the coupling patterns associated with vmPFC (Figure 3-2A), cOFC (Figure 3-5C), and lOFC (Figure 3-5A) confirmed that they each participated in different neural circuits in both species (SI Figure 7-12A). This suggests that

there are not only important differences between vmPFC/medial OFC and lateral OFC (259, 275), but that a central OFC region is also distinct.

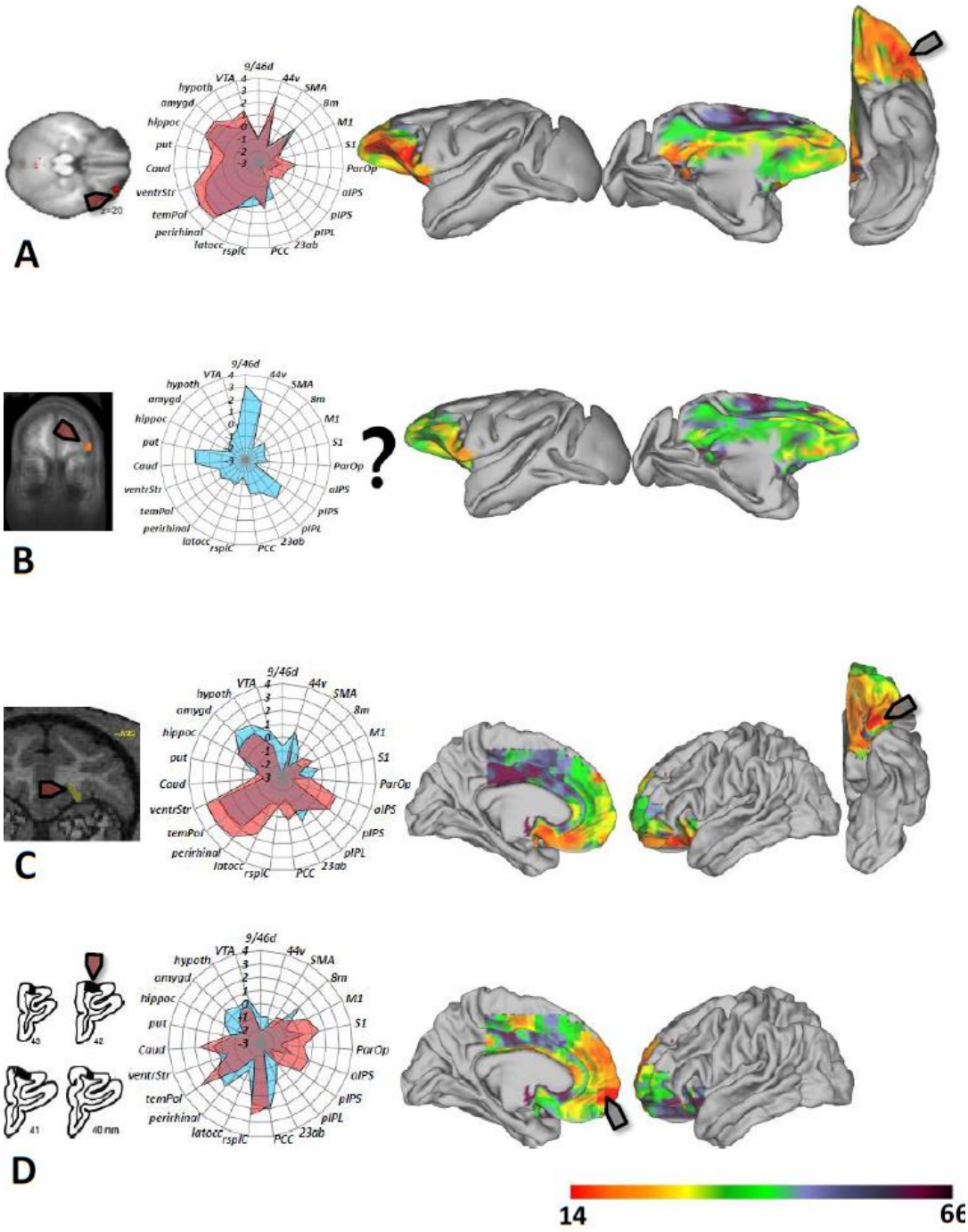


Figure 3-5: Human OFC regions (left) linked with (A) stimulus-reward association (274) and could be linked to a macaque lOFC (right) via similarities in the regions' coupling patterns (center). (B) A human brain concerned with decision-making under ambiguity (280) did not correspond strongly to any macaque brain region (C) Macaque regions (left) linked with (C) context independent and outcome identity dependent value signals (273) and (D) decision outcome monitoring (281) could be linked to human brain regions (right) via similarities in coupling patterns (center).

A region in monkey FP (Figure 3-5D) implicated in linking choices with their outcomes (281) exhibited strong positive coupling with PCC, hippocampus, temporal pole, and head of the caudate and negative correlation with the cingulate motor areas, insula and IPS (Figure 3-4D). It therefore matched a region in human FPm.

3.3.4 DWI-MRI parcellation of medial and orbital frontal cortical areas for decision-making

So far our comparison of fMRI coupling patterns in macaques and humans has suggested several correspondences but it has also suggested a more fine-grained fractionation of vmPFC, OFC, ACC, and FP than is often assumed. In the next part of our investigation, a different imaging modality and analysis approach, DW-MRI tractography, was used to provide an independent test of the parcellation of human frontal cortex suggested by the fMRI coupling pattern analysis.

DW-MRI can be used to estimate the structural connectivity of each MRI voxel and this information can then be used to group together voxels sharing similar profiles of connectivity with the rest of the brain (282). Despite the technique's limitations (162) the areas previously identified using the DW-MRI approach correspond in their spatial position to areas identified in cytoarchitectural analysis (192, 230, 231, 248, 283). A previous parcellation of human ACC on the basis of DW-MRI tractography has been published (201) but the region investigated in that study excluded much of vmPFC, all of OFC, the area between ACC and FP, and because it did not establish dorsal boundaries of ACC areas it left open the possibility of additional areas. Here we investigated a much larger ROI comprising vmPFC, ACC, FP, and OFC in their entirety and extending into adjacent posterior cingulate cortex and dorsal frontal cortex.

We ran probabilistic tractography from each voxel in the ROI (Figure 3-6A) in 38 right-handed human participants in both left and right hemispheres with and without paracingulate sulci. In

general the paracingulate sulcus is prominent in the left but not always in the right hemisphere (284) and so results are shown for the left hemisphere in cases with a paracingulate sulcus (Figure 3-6B) and in the right hemisphere in cases lacking a paracingulate sulcus (SI Figure 7-3The results were, however, broadly similar. We correlated the pattern of structural connectivity of every voxel in the ROI with all other voxels in the same ROI and obtained a symmetric cross-correlation matrix. We then used k-means clustering to group together voxels with similar connection patterns. We used this approach recursively, dividing the initial ROI into two smaller sub-divisions and subsequently subdividing the resulting areas further in three to four parcellation steps. We stopped when results ceased being consistent across subjects. In a final step we returned to the fMRI data and established the functional coupling patterns of all of the DW-MRI parcellation-derived sub-regions and compared them with the coupling fingerprints of the functional areas investigated in the first part of the study (Figure 3-6C) and with the 448 monkey frontal ROIs aiming to find the closest match for each (SI Figure 7-5 and SI Figure 7-5).

We found 5 clusters in dorsal ACC bordering medial aspects of M1, SMA, pre-SMA and 8m (which were in turn also identified by this parcellation but were already discussed elsewhere (285)). Three overlapped with clusters 4, 5, and 6 as proposed by Beckmann et al (201) and which correspond to the areas Picard and Strick (268) call RCZa, RCZp and CCZ (Figure 3-6B, red, orange and brown) with left hemisphere centres of gravity [-9, 20, 34], [-8, 7, 40], and [-11, -26, 42] respectively. The functional coupling patterns of RCZa and RCZp resembled coupling patterns of the two cingulate areas in the first part of the study (Figure 3-4). More ventrally we delineated two regions that, together, overlapped with Beckmann's (201) cluster 7. These were termed 23a/b and 24a/b (Figure 3-6B, green and dark-yellow) with centres of gravity [-4, 19, 23], and [-6, -15, 35]. They were coupled to each other, to premotor cortex and SMA, to the IPL, and to some degree with dlPFC and with negative coupling with temporal lobe areas (SI Figure 7-4). There was a posterior-to-anterior gradient of decreasing sensorimotor

and pallidal coupling and increasing dlPFC and caudate head coupling. Similar coupling patterns are found for monkey cingulate gyrus areas such as 24 and 23d (SI Figure 7-4) (286). The functional coupling pattern of the more anterior part of 24 suggested a correspondence with the perigenual ACC area (32) implicated in cost benefit decision making (Figure 3-6C; Figure 3-3C).

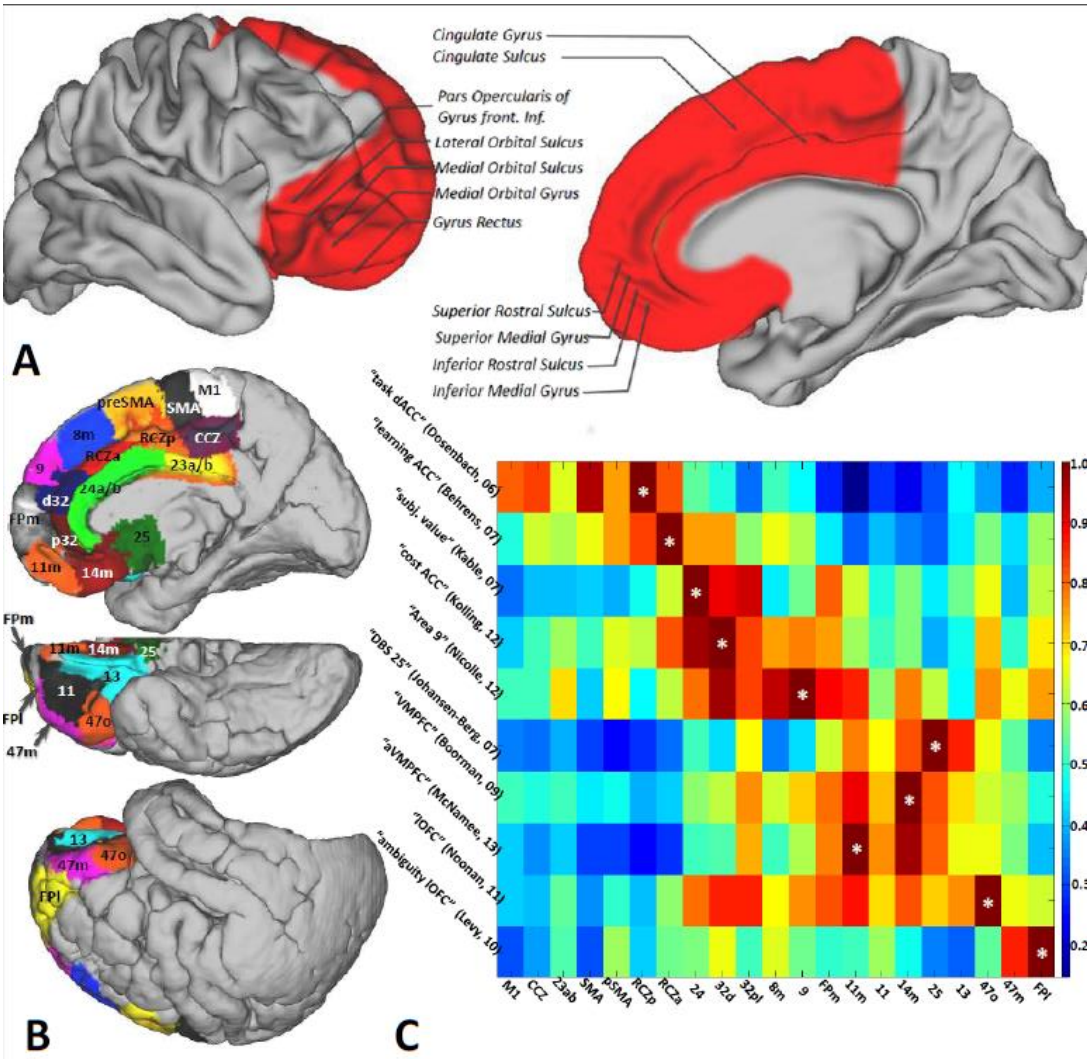


Figure 3-6: (A) The human medial and orbital region investigated and (B) the sub-regions that could be identified on the basis of differences in DW-MRI-estimated connectivity. (C) Correspondences between the fMRI-based functional coupling maps associated with decision making areas (Figure 3-1 – 2-5) and orbital and medial sub-regions identified via DW-MRI parcellation (B). Warm red colors indicate similarities in the functional coupling maps and asterisks indicate areas with highest spatial correlation (see SI Methods 7.1.1.1 for more information).

More anteriorly on the medial surface we delineated seven clusters (Figure 3-6B). A subgenual area (Figure 3-6B, green) overlapped with Beckmann's (201) cluster 1 and the region described by Johansen-Berg and colleagues (258) as a DBS-target-site in depression. Its centre of gravity [-4, 5, -8] suggested correspondence with area 25 (111). Two further clusters lay in vmPFC. The centre of gravity [-9, 23, -19] of the more rostral (Figure 3-6C, orange) suggested a resemblance with area 11m (111). The more posterior overlapped in position with Beckmann's (201) cluster 2 (Figure 3-6B, dark-red) and its centre of gravity [-9, 23, -18] linked it to area 14m (111). The fMRI coupling patterns of these two regions indicated correspondence with vmPFC decision-making regions (Figure 3-6C; compare Figure 3-2A,B). We identified two regions (Figure 3-6B, dark blue and dark red) with centres of gravity [-9, 36, 28], and [-11, 47, 4] suggesting correspondences with subdivisions of area 32: d32, and p32 (287). The fMRI coupling pattern associated with the more dorsal of these, area d32, resembled the coupling pattern of the dorsal perigenual ACC area linked to cost-benefit decision making during foraging (54). A medial area 9 and a medial FP cluster were also found (Figure 3-6B, pink and grey) that overlapped with area 9 identified by Sallet and colleagues (196) and FPm identified by Neubert and colleagues (248) at [-10, 51, 29] and [-11, 60, 4]. Area 9's coupling pattern resembled that of the area linked to imagination of other peoples' values and unexperienced rewards (256, 257) (Figure 3-6C). FPm's functional coupling pattern was similar to the monkey frontal polar region investigated by Tsujimoto and colleagues (281) (Figure 3-6C; Figure 3-5D).

On the lateral surface we delineated three clusters (Figure 3-6B, yellow, pink and orange) with centres of gravity at [-31, 50, 60], [-30, 50, -8], and [-32, 24, -11] which we relate to FPl (248, 283) and two distinct components of area 47/12 we refer to as 47/12m (anterior) and 47/12o (posterior), respectively (111). FPl, which resembled the ambiguity area described by Levy and Glimcher (280) (Figure 3-5B, Figure 3-6C) , exhibited little correspondence with areas in the macaque brain (248). 47/12m exhibited only a limited degree of correspondence with

macaque 47/12. 47/12o, which resembled the lOFC region concerned with credit assignment (274) (Figure 3-5A, Figure 3-6C), resembled the posterior part of 47/12 in macaque.

Next we divided the central part of the OFC into areas corresponding to clusters 2 and 3 of Kahnt and colleagues (288) (Figure 3-6B, bright blue and black) with centres of gravity at [-13, 22, -21] and [-22, 30, -17]. These two central OFC areas had very distinct coupling patterns: The posterior region showed strong positive coupling with ventrolateral frontal cortex, vmPFC, perirhinal cortex, the temporal pole, amygdala and the ventral striatum and so resembled area 13 in the macaque (111) and the region studied by Padoa-Schioppa (273). By contrast the more anterior OFC region, we refer to as 11, coupled strongly with regions such as dlPFC, vlPFC, vmPFC, temporal pole and ventral striatum and resembles a monkey anterior central OFC region.

In a final set of analyses, we sought to investigate whether the human brain areas in cingulate and orbitofrontal cortex that we identified and the macaque brain areas we identified as homologous occupy a similar position within the brain networks of the two species (see also SI Methods 7.1.4). First, we calculated the dissimilarity matrices of all 23 homologous frontal regions with one another within each species and performed a correlation between the two species' matrices. The two dissimilarity measures were highly correlated between species (ρ = .79, p<0.001). Second, we calculated two measures of the centrality of the medial and orbital frontal regions within the network of target areas. Centrality measures indicate how important a node is within a network. Again, the correlations between species in these measures were highly significant (degree centrality: ρ = .57, p = .005; eigenvector centrality: ρ = .56, p = .005). Third, to provide another line of evidence regarding the "relational similarity" of the frontal areas under investigation, we performed a hierarchical clustering analysis for both species on the whole-brain functional coupling of all areas derived from the tractography-based parcellation of the human medial and orbital cortex (SI Figure 7-13A) and their proposed monkey equivalents (SI Figure 7-13B). The hierarchical clustering analyses did not reveal

identical relationships between areas in the two species when they were examined at the finest level, but it can be seen that at the broader level clustering into families of regions is largely similar across species. Five broad groups of areas are identifiable in both species: (1) the cingulate motor areas, (2) the perigenual and medial frontopolar regions, (3) the vmPFC and anterior central OFC regions, (4) the posterior medial and central OFC regions, and (5) lateral orbitofrontal cortex. This supports the idea that despite fine grained differences between species it is possible to identify broad similarities in the way that frontal areas interact in circuits in the two species.

3.4 Discussion

Despite the apparent sophistication of human reward-guided decision-making and learning it appears to depend on medial and orbital frontal cortical circuits that are similar to ones that can be identified in macaques. Moreover careful inspection of the patterns of inter-regional interaction revealed a finer level of parcellation between component areas within ACC, vmPFC, and OFC than is often assumed in decision-making investigations.

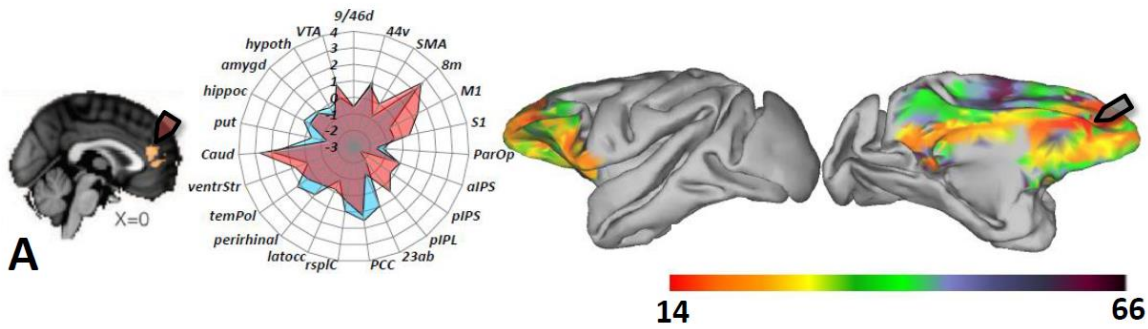
In a previous study of parietal cortex it was shown that areas with distinctive DW-MRI-estimated connectivity profiles corresponded to distinct cytoarchitectural regions (192, 230, 231). Although DW-MRI investigation is no substitute for more detailed investigation of cytoarchitecture it may guide such investigations and it provides information about the approximate extent of anatomical areas on a scale that corresponds with the majority of human neuroimaging studies. Human vmPFC is one of the most frequently reported areas of activity in investigations of reward-guided decision-making but it was possible to show that it consists of sub-regions each exhibiting a distinctive pattern of activity coupling with the rest of the brain. Moreover in most cases the regions were associated with distinctive DW-MRI-estimated connectivity profiles.

A region near area 14m has been linked to decisions or attentional selection of choices (56, 251, 255, 256). In comparison to ACC and OFC it was, in both species, distinguished by strong positive coupling with hypothalamus, ventral striatum and amygdala (289). Reward-related activity has been reported in a similar area in the macaque (242) and lesions here disrupt reward-guided decision-making (290). Two more anterior areas, 11m and 11, were linked with more abstract choices (254) but they nevertheless bore a resemblance to areas 11m and 11 in the macaque (111). In monkeys lesions that include area 11 and area 13 disrupt decision-making guided by contrasts in outcome identity and reward-specific satiety and not just reward amount (291).

In monkeys neurons in central OFC encode the value of a specific item, regardless of the value of the other items it is presented with (273). The functional coupling of this region resembled that of a part of human central OFC, area 13, but it was different to the more medial vmPFC areas that have been the focus of investigations of value representations in the human brain (Figure 3-1). It did, however, correspond to the human region identified by Klein-Flügge and colleagues (279) as coding reward identity. This pattern of results is broadly consistent with the scheme suggested by Rudebeck and Murray (292) in which value based decisions and reward identity based decisions are mediated by medial and central orbital regions respectively. As discussed below, however, an even more lateral OFC region can be distinguished that is concerned with establishing stimulus-reward associations and credit assignment (57, 274, 279).

A dorsal medial frontal area in humans, area 9, was linked to decision-making when the outcome had to be imagined or modelled in some way (256, 257) and its activity coupling pattern resembled area 9 (235) in the monkey medial frontal cortex. A slightly more ventral area, d32, resembled the region active when decisions have to be made about whether rewards are worth the cost of foraging (54). Although it has been argued that no exact homologue of human d32 exists in the macaque (287) it was notable that its activity coupling

pattern resembled that seen in the dorsal part of macaque area 32 which is in the anterior cingulate sulcus – a region sometimes called 32(s) (293). A region with a similar coupling pattern in the monkey has also been associated with cost-benefit decision-making (72). The coupling pattern for this region could be distinguished from, but nevertheless resembled, a distinct region, pgACC, that is active when participants trade magnitudes of rewards against their respective delays (32). This cost-benefit-comparison or derivation of a "common currency" appears to be a distinct process to the comparison of values associated with different choices in order to make a decision; this latter process is associated with more ventral parts of the medial surface in areas 11m and 14m (Figure 3-2, 2-3, 2-6, SI Figure 7-1).



Human d32 and monkey 32(s) are very distinct in their coupling patterns from more posterior dorsal ACC areas that have also been implicated in decision making, learning and cognitive control. The more anterior RCZa/CMAr is responsive to feedback about motor strategies especially when this is relevant to changes of behaviour in the future both in humans (54, 82, 266, 270, 294) and macaques (271, 295). It seems more active whenever there is evidence that the current mode of action is suboptimal and alternative behaviours should be considered. The strong positive coupling of RCZa with areas involved in cognitive control - such as dlPFC, supramarginal gyrus, pre-SMA, inferior frontal gyrus and subthalamic nucleus - might support this function of seeking for alternative behavioural strategies. The more posterior RCZp area has been called a task positive region that is active in situations of conflict (69), when subjects need to engage in various different tasks (186) or during low-frequency responding (266).

Generally speaking this area seems to be more active whenever there is a higher demand placed on motor control, be it conflict, effort or infrequent motor behaviour. Its strong coupling with sensorimotor areas (M1), IPS and areas that have been implicated in top-down motor control (SMA, PMd, ventral premotor area) might support this function. One might think of the different role of RCZp and RCZa as paralleling the role of areas 45 and 47/12: Whereas area 45 seems to select certain objects and visual features for attention (296), the more anterior area 47/12 might be more involved in learning new visual objects–reward associations (47). RCZp and RCZa might do something similar in the action domain: Whereas RCZp might help selecting and attending specific motor plans, RCZa might acquire evidence for switching behavioural strategies away from the current mode of action.

A correspondence was identified between lOFC area 47/12o in macaques and humans. In contrast to medial OFC and vmPFC and central OFC, monkey lOFC is necessary for associating specific outcomes with specific choices (57). Similarly, in humans, it is more active in situations when specific rewards can be associated with specific outcomes (274, 279). Its coupling pattern -- strong positive correlation with vlFC, middle and inferior temporal lobe areas as well as perirhinal cortex would put it in an excellent position to access visual information of varying degrees of abstractness while its strong coupling with ventral striatum, amygdala and the vmPFC link it with reward and value-guided choice networks. It is therefore ideally placed for credit assignment and learning about which sensory features in the environment are rewarding.

When decisions have to be made in highly ambiguous situations a more anterior lOFC/FPl region tends to be active (280, 297). Moreover FPl has been proposed to track values of alternative choices (56, 60) and stores information about alternative strategies or task sets when they should not be used at the moment but might be returned to later (298, 299). In short this region shows activity correlated with the number of alternative options, strategies and task sets, or the amount of evidence that alternative options might be worth pursuing.

This would put it in a good position to signal uncertainty about the effectiveness of our current behaviour, and therefore allow critical evaluation and meta-representation of our decisions (300). FPl coupled strongly with mid-IPL, dlPFC, and pre-SMA and was not matched to any monkey ROI. Note that the functional coupling pattern of human FPl was largely different from the coupling pattern of monkey FPm as described by Tsujimoto and colleagues (281). Our results suggest Tsujimoto and colleagues (281) recorded from an area more similar to human FPm.

The similarity in areas' coupling patterns between species provides a strong case that these areas perform similar functions in the two species. This is further demonstrated by the comparison of different measures of the positions that the cingulate and orbitofrontal regions occupy within the larger brain network. The coupling with predefined target regions, the different measures of the centrality of the areas, and the grouping of areas into sub-networks based on their whole-brain functional connectivity all provide evidence for a largely conserved organisation of decision making areas in these two primates.

To conclude the fMRI coupling patterns of areas in medial and orbital frontal cortex in humans and macaques suggest several meaningful subdivisions can be identified even within areas, such as vmPFC, that are often treated as unitary. Moreover, many correspondences can be identified between the species although they are not always those that are widely assumed. A small number of areas may be found only in humans and not in macaques.

3.5 Methods

3.5.1 Human participants

Resting state BOLD functional MRI (rs-fMRI) and diffusion-weighted and T1-weighted structural images were acquired in 38 healthy right-handed (according to Edinburgh Handedness Inventory: mean +/- SD, 0.84 +/- 0.19) participants (20 female; age range: 20– 45

years; mean age +/- SD,30.7 +/- 10.1 years) on a 3 T Siemens Magnetom Verio MR scanner in the same session using standard DW-MRI and rs-fMRI protocols (SI Methods 7.1.1). All participants gave written informed consent in accordance with ethical approval from the local ethics committee. Participants lay supine in the scanner, and cushions were used to reduce head motion. Participants were instructed to lie still and keep their eyes open and fixated at a cross. Twelve participants had no paracingulate sulcus in their left hemisphere (noparacingulate_left), 20 had a prominent paracingulate sulcus in their left hemisphere (paracingulate_left), and 22 had no paracingulate sulcus in their right hemisphere (noparacingulate_right). For both, the rs-fMRI based functional coupling analyses and the DW-tractography based parcellation we calculated group-results for the two most common patterns (paracingulate_left; noparacingulate_right) groups trying to establish whether any inter-individual differences could be explained by the two factors hemisphere (left v.s. right) or paracingulate sulcus (prominently present or absent). Both DW-tractography based parcellation and rs-fMRI results yielded largely similar results for all subjects.

3.5.2 Human resting-state fMRI data acquisition, preprocessing, and analysis

Analyses were performed using tools from FSL (Functional MRI of the Brain Software Library), the Human Connectome Project Workbench, and custom made software written in Matlab (MathWorks). Resting state fMRI data acquisition and pre-processing were carried out in a standard way as previously described (192) (SI Methods 7.1.1). To establish the functional connectivity of each region, we created ROIs based on exemplary functional imaging studies with human participants that related a specific aspect of decision making to “activity” in a circumscribed part of one of these areas (SI Table 7-2) We drew a cubic ROI (3x3x3 voxels, i.e. 6mm isotropic) centered on the peak MNI coordinate of a given study. These ROIs were registered from MNI-space to each subject’s individual rs-fMRI space via the T1-weighted

structural image using FNIRT and brain boundary-based registration (BBR). Then the major Eigen time series representing activity in each of the ROIs was calculated.

Individual statistical maps were then calculated using a seed-based-correlation analysis which is part of FSL (fsl_sbca) as previously described (192, 234) in order to infer the functional connectivity of these ROIs with the rest of the brain. For each ROI we created a model consisting of the first Eigen time series of that region and the confounding time series representing head movement (six regressors resulting from motion correction using MCFLIRT) and the Eigen time-series of white-matter and corticospinal fluid (CSF). The results of each individual subject's ROI-specific seed-based-correlation analysis were then entered into a general linear model (GLM) analysis. The resulting Z-statistical images were projected onto the CaretBrain as provided by the Human Connectome Project Workbench using the "surf_proj" algorithm as implemented in FSL and then visualized using the Human Connectome Project Workbench. Unthresholded Z-maps were quantified by extracting the average intensity of each ROI's functional connectivity Z-map in a number of cortical and sub-cortical regions of interest that we refer to as target regions to distinguish them from the orbital and medial frontal regions that were the focus of our investigation (SI Table 7-3). These regions were chosen, first, because they are known to be interconnected with particular orbital and medial frontal regions in the macaque and so the functional connectivity patterns of different orbital and medial frontal areas are likely to be distinguishable on the basis of their coupling with these areas. Second, the areas were chosen because their homology in humans and macaques has already been established. We drew ROIs (3x3x3 voxels, i.e. 6mm isotropic) centred on the coordinates listed in SI Table 7-3 and then averaged the Z-value from the unthresholded seed-based-correlation analysis derived Z-maps within these ROIs. These values were then displayed on a spider plot (Figure 3-1 – 2-5). On the basis of these "coupling fingerprints" the human decision making areas were then compared to 448 different ROIs in the monkey frontal cortex to establish the monkey ROI with the most similar coupling pattern (see below). Note that the

spider plots illustrate the simple correlation between a frontal lobe area and the target areas as opposed to partial correlations because: 1) it makes the link between the spider plots and the functional connectivity z-maps in Figure 3-1 – 2-5 transparent; 2) it prevents under-estimation of the coupling between a frontal area and a target area if they are both also coupled to a second target area (recall that the target areas were chosen because they were likely to be connected to the frontal areas being investigated). Comparisons between the coupling patterns of different frontal areas to a given target area become transparent even if only one of the frontal areas and the target area being examined are both coupled to another target area; 3) it ensures that judgments about similarities in the coupling patterns of frontal areas in humans and macaques are not influenced by interspecies differences in the way that target areas are interconnected with one another. However, it is important to remember that although it is the case that such coupling patterns reflect monosynaptic connections they do not do so exclusively (147). In addition a complementary analysis based on partial correlation is shown in SI Figure 7-6.

In another analysis monkey regions from different neurophysiological recording studies were compared, based on their functional coupling patterns, to 417 ROIs in human frontal cortex. These 417 different ROIs (3x3x3 voxels, i.e. 6mm isotropic) were drawn in equal distance to one another (6mm) in order to cover the medial and orbital frontal and frontopolar cortex. The region covered everything that has been referred to as ACC, OFC or FP (Figure 3-5A). It therefore comprised the whole region investigated by Beckmann and colleagues (201), except the two most posterior clusters (8 and 9). It included the cingulate gyrus and sulcus (including the dorsal bank of the paracingulate sulcus if present) and extended posteriorly to include all cingulate motor areas as delineated by Beckmann and colleagues (201) and Amiez and Petrides (301). It, therefore, covered the most anterior tip of the marginal sulcus but excluded the precuneus and posterior cingulate cortex. Anteriorly it included subgenual ACC and vmPFC. Moreover it contained areas 9, FPm and FPl, as delineated by Neubert and Sallet and

colleagues (196, 248), the OFC, and the orbital part of the inferior frontal gyrus (pars orbitalis gyri frontalis inferioris). These 417 different ROIs were then registered from MNI-space to each subject's individual rs-fMRI space. In order to infer the functional connectivity of these 417 ROIs with the rest of the brain we used exactly the same procedure as described above for the decision-making study-based ROIs. In this way we were able to obtain 417 different group Z-statistical images which were subsequently used to generate 417 different "coupling fingerprints" using exactly the same 23 cortical and sub-cortical target ROIs as mentioned above. These 417 different "coupling fingerprints" could then be compared to each of the monkey areas from neurophysiological recording studies (see below).

3.5.3 Macaque resting-state fMRI data acquisition, preprocessing and analysis

Resting-state fMRI and anatomical scans were collected for 25 healthy macaques (Macaca mulatta) (four females, age: 3.9 years, weight: 5.08 kg) under light inhalational anesthesia with isoflurane (for detailed information on anesthesia protocol, monitoring of vital signs, data acquisition and pre-processing see SI Methods 7.1.2). Protocols for animal care, MRI, and anesthesia were performed under authority of personal and project licenses in accordance with the United Kingdom Animals (Scientific Procedures) Act (1986).

The goal of this part of the study was to test for similarities between the functional networks of areas implicated in decision making in human and monkey frontal cortex. We therefore aimed to map the resting state functional connectivity networks of areas derived from exemplary neurophysiological recording studies of macaque monkeys which had related neuronal activity measures to specific aspects of decision making. We drew cubic ROIs (3mm isotropic) centered on the center of gravity of the recording site of each study. These ROIs were registered from standard-space (302) to each monkey's individual rs-fMRI space using FNIRT. Then the major Eigen time series representing activity in each of the ROIs was calculated and seed-based-correlation analysis (fsl_sbca) was used, as in the human subjects,

to infer the functional connectivity of these ROIs with the rest of the brain. As in the human subjects unthresholded group Z maps were quantified by extracting the average intensity of each ROI's functional connectivity Z-map with 23 cortical and sub-cortical target regions of interest (SI Table 7-3). These coupling fingerprints of monkey "decision making areas" were then compared to 417 different ROIs in the human orbital and medial frontal cortex to establish regions with the best corresponding functional coupling pattern (see below).

For the reverse comparison (matching human neuroimaging derived "decision making regions" to areas in the monkey frontal cortex) we drew 448 equally spaced cubic ROIs (3mm isotropic) to cover the whole cingulate gyrus and sulcus, as well as areas 9, 10, 12, 45, 47/12, 14, 11, 13, and Iai as defined by (238). These 448 different ROIs were then registered from standard-space to each monkey's individual rs-fMRI space using FNIRT. In order to infer the functional connectivity of these 448 ROIs with the rest of the brain we used exactly the same fsl_sbca-based procedure as described above. In this way we were able to obtain 448 different group Z-statistical images which were again used to generate 448 different "coupling fingerprints" using exactly the same 23 cortical and sub-cortical target areas as mentioned above. These 448 different "coupling fingerprints" could then be compared to each of the human areas from neuroimaging studies (see below). We also conducted diffusion-weighted tractography-based parcellation of the same orbital and medial frontal region (Figure 3-5A) in our human participants (see below) and matched the functional coupling profiles of each of the different parcels to these 448 monkey ROIs using the same "coupling fingerprint" matching approach.

3.5.4 Comparison of resting-state functional connectivity of macaque and human decision making areas

A formal comparison between human and macaque coupling patterns was performed by calculating the summed absolute difference [the "Manhattan" or "city-block" distance (196,

202, 248) of the coupling scores. This yielded a summary measure of the difference in coupling patterns for each pair of areas in the two species (e.g. for human "task-positive ACC" compared to monkey ROI number 267 the summed absolute difference between coupling profiles is 34.78). The summary measure can then be used to compare the functional coupling pattern of each human region with those of all 448 regions in the macaque and vice versa. The Manhattan distance has previously been used to compare the coupling patterns of brain areas across species (196, 202, 248). It is a useful metric for comparing connectivity because it is summarizes the whole pattern of coupling for an area. It is the whole pattern of connectivity rather than any particular connection that distinguishes areas and so this metric is appropriate. In addition, unlike some other possible approaches, it is less sensitive to any idiosyncrasy in the estimate of any one connection. This summary measure was then back-projected onto the monkey brain for each of the 448 ROIs to show regions with low absolute difference in red and regions with high absolute difference in brown/black (Figure 3-1 – 2-5). Thus the right hand side of Figure 3-2 – 2-5 can be thought of "heat maps" in which warm red colors indicate the regions in the brain of one species that best correspond with the area highlighted from the other species on the left hand side.

3.5.5 Diffusion-weighted tractography-based parcellation of orbital and medial frontal ROI

DW-MRI data were pre-processed in a standard way as previously described (192) (SI Methods 7.1.3 and 7.1.4). For each participant, probabilistic tractography was run from each voxel in the orbital and medial frontal ROI (Figure 3-6A, SI Figure 7-2A) in three groups (noparacingulate_right, noparacingulate_left, paracingulate_left) to assess connectivity with every brain voxel (whole brain "target" was down-sampled after tractography to 5 mm isotropic voxels for the connectivity matrix to be manageable, however the whole orbital and

medial frontal ROI was tracked in original FA space, using a model accounting for multiple fiber orientations in each voxel (199). A connectivity matrix between all orbital and medial frontal voxels and every other brain voxel was derived and used to generate a symmetric cross-correlation matrix of dimensions (number of seeds x number of seeds) in which the (*i,j*) element value is the correlation between the connectivity profile of seed *i* and the connectivity profile of seed *j*. The rows of this cross-correlation matrix were then permuted using k-means segmentation for automated clustering to define different clusters (Figure 3-6B, SI Figure 7-3). The goal of clustering the cross-correlation matrix is to group together seed voxels that share the same connectivity with the rest of the brain.

We used a recursive or iterative clustering procedure here similar to the one used by Beckmann and Neubert and colleagues (201, 248). In this way the orbital and medial ROI was parcellated into subregions via three to four parcellation steps. Parcellation was stopped when the resulting parcel could not be further sub-divided in a similar way in all subjects into either two, three or four sub-regions (with preference to two over three and three over four). A subdivision was considered reliable if the topography of the different clusters was the same in all subjects of a particular group (for example noparacingulate_left).

4 Connectivity based parcellation of principal systems of association fibers in human and monkey cerebral cortex

4.1 Abstract

Association fibres are connections between areas within a cortical hemisphere. Long-range association fibre systems are functionally relevant. They constitute long-distance connections between higher-order association cortices and may thus be relevant for information integration across different sensory and cognitive domains. Moreover they have been implicated in a number of neurological and psychiatric syndromes called “disconnection” syndromes.

I demonstrate a novel data-driven method of dividing the white matter into its major association fibre systems in both the human and the macaque monkey. I show that this can be used in order to link the topography of white matter association fibres with their respective pattern of cortical projection. I also suggest that this can be used to compare the skeleton of connections and the cortical distribution of their projections across a wider range of species in order to test hypotheses about brain evolution. Finally I demonstrate a new way of assessing potential disconnection syndromes or hodological pathologies more generally.

4.2 Introduction

4.2.1 Modular and connectionist views of brain structure and function

Ever since the early days of modern neuroscience in the 19th century, two major frameworks have shaped the way that we think about the brain and its structure, function and dysfunction. Early clinicopathological correlation studies such as Broca’s reports on aphasia emphasised localisation of function in discrete and highly specialised cortical areas (303). Support came

from early microscopic-anatomical studies that attempted to divide the cortex into a distinct set of sub-regions based on for example cell types or cytoarchitecture. It was hoped that these (micro-) structurally defined areas would relate to functionally specialised modules (304). On the other hand the study of so-called disconnection-syndromes such as Wernicke's report on conduction aphasia invoked ideas of distributed and parallel processing in networks of brain regions connected via white matter tracts (102).

Early studies of the brain's structural connections distinguished between projection fibres (connect cortex with subcortical regions; e.g. the fornix), commissural fibres (connect cortical regions between hemispheres; e.g. corpus callosum), and association fibres (connect cortical regions within hemisphere; e.g. the arcuate fascicle). These association fibres were thought to be of particular importance for cognition and brain dysfunction (104, 305, 306).

An early principle of connectivity - "Flechsig's rule" (after psychiatrist Paul Flechsig) – held that major association systems never connected two "*primordial zones*", i.e. areas that were already fully myelinated at birth and largely corresponded to primary sensory and motor areas (102). Rather association fibres were taken to be long-range connections between *association cortices*, allowing for the integration of higher-order information across domains. Their disconnection in turn would lead to the impairment of such "integrative processing" whilst sparing early sensory or motor functions.

Conduction aphasia was among the earliest disconnection syndromes proposed (102). It was linked with disconnections between Broca's and Wernicke's area in the dominant hemisphere whilst sparing these cortical regions. It therefore leaves both, speech production and comprehension intact but results in a word repetition deficit. Some forms of visual agnosia were also argued to be disconnection-syndromes. Lesions sparing visual cortices but disconnecting them from higher order areas in for example the temporal cortex may cause complex agnosias whilst leaving visual perception largely intact. Apraxia and pure alexia have

also been studied as disconnection syndromes during the classical “associationist” era in the late 19th century (102, 307, 308).

4.2.2 Geschwind’s theory on disconnection-syndromes

Norman Geschwind revived and broadened the idea of disconnection-syndromes in the 1960s (309, 310). In line with “Flechsig’s rule“ he suggested that in humans primary sensory and motor regions are never directly connected to one another but only via association-cortices (102). Disconnection syndromes may thus arise either from lesions to these obligatory relay areas or their respective connections (one may be tempted to call them “dis-association” rather than “disconnection” syndromes). This theory relates to “connectionist” theories about the evolution of higher cognitive functions. Geschwind suggested that in lower mammals primary sensory and motor regions are connected directly. However as one moves up the phylogenetic scale connections come to be made between newly developed regions of “association” cortex “inserted” between the “older” areas. Long-range fibres (association and commissural) then in turn connect these association areas. Geschwind suggests that in the monkey these associative connections are still largely relayed via the limbic system. However in humans the development of new “tertiary” association regions (e.g. in the inferior parietal lobule) “freed” inter-modality connections from the limbic system.

The below proposed method of dividing the major association fibres based on their cortical projection patterns aims to test such hypotheses. It allows for a data-driven way of delineating the long-range association fibres in the human brain in order to verify such general “principles” of connectivity, e.g. with abundant connections between association cortex and absence of connections between primary areas. Moreover I hypothesise that this approach can be used in different species allowing for the cross-species comparison of general motifs of connectivity.

4.2.3 New MRI-based methods for studying brain connectivity and graph-theory

The advent of new brain-imaging techniques such as PET, fMRI and TMS in the late 20th century again supported modular theories of brain-function. In contrast to "associationists' " believes that complex cognitive functions could only emerge through the interaction of brain areas, a growing body of literature succeeded in linking higher order cognitive processes (such as planning, problem solving, decision making, cognitive control) to distinct areas of the cerebral cortex. On the other hand elegant tracer-injection techniques and the advent of DW-MRI sparked new enthusiasm for connectivity and the study of the brain as an assembly of "networks" (142). The emergence of graph-theoretical analyses of connectivity data in recent years may have furthered this enthusiasm (311).

These new MRI-based imaging methods offer some crucial advantages. Both task-fMRI and DW-MRI are less invasive than previous methods (e.g. neurophysiological recordings or tracer injections). Their ability to image the whole brain at once may allow for less hypothesis-driven investigations of both, modular processing and connectivity. Most importantly however they may contribute to an integration of modular and connectionist views of brain structure and function and may hence allow for a reassessment of topological (topos = place, i.e. "localised") and hodological (hodos = path, i.e. connection-based) accounts of brain dysfunction (102, 104, 312).

One of the early data-driven approaches for studying the network architecture of the brain came from graph-theoretical analyses of connectivity data (140, 311). Graph theory is a mathematical tool that describes networks as consisting of nodes and edges (their connections) and analyses them in order to derive their general properties. One early suggestion of graph-theoretical analyses was that the human brain shares similarities with so-called "small world networks" (140). Such networks are more structured than "random graphs" but have shorter average shortest-path-lengths than perfectly structured lattices. Simply put most connections in the brain appear to connect neighbouring brain regions but a

reasonably small number of long-range connections provide short-cuts and therefore reduce the average shortest-path between any two brain areas substantially. Vaguely put this may allow for just the right balance between modular processing and information integration.

Another early graph-theoretical finding was that different brain regions had a different number of connections with a small group of regions (so-called "hubs") being particularly highly connected (140, 313-315). Later on these views were refined to incorporate the observation that brain areas are often grouped together in "families of areas" or modules. These may in turn be connected to "provincial hubs". Long-range connections then may connect these provincial hubs with global hubs (314). These ideas of regional and global hubs may remind of Flechsig's and Geschwind's theories of connections via secondary and tertiary association cortex. Moreover the long-range "short-cuts", which decrease average shortest-path length, may be identified with the major association fibre systems. Interestingly recent graph-theoretical analyses have implicated "hubs" as particularly vulnerable targets for neural degeneration (313), which may be in line with Geschwind's broader notion of "disconnection syndromes".

4.2.4 Limitations of graph-theoretical accounts

Graph-theory is a powerful data-driven way to study brain connectivity and network properties. However some graph-theoretical approaches have been "anatomically agnostic" in that they treated all areas and connections equally as "nodes" and "edges", respectively {Jbabdi, 2013 #2586}. For example some graph-theoretical accounts ignore whether two edges / connections are part of the same or of different major white matter fibre systems. However the exact three-dimensional layout of a path and its topography particularly in the case of long-range association fibres may determine the efficacy and cost of a connection, the speed of conduction or the "ontogenetic profile" of a connection, i.e. how it changes with development and aging (316, 317). Hence two connections may be more likely to share

similarities in developmental and aging trajectories, risk to lesions, or co-vary in their myelination or connectivity strength across subjects if they are part of the same major white-matter fibre system. If this is true they may link with behavioural co-variation across subjects in tasks that are supported by different brain regions but the same major white matter fibre systems. Moreover one may think about hubs as regions, which participate in a wider set of major association fibre pathways rather than regions, which only have a higher degree of connectedness, because such “multi-pathway hubs” might be likely to have access to very diverse information from largely distinct networks.

Major association fibre systems may also shape topological configurations and motifs of connectivity and even function. For example the “onion-shape” of frontal-parietal connectivity with more anterior / posterior frontal regions being connected to more anterior / posterior parietal regions (318, 319) may be largely reflective of the topological organisation of the fibres within the superior longitudinal fascicles (SLFs). The parallel anterior-to-posterior hierarchies of ventral/orbital frontal and anterior temporal connections with anterior / posterior ventral frontal areas being connected to anterior / posterior temporal areas (320, 321) may be linked to the organisation of fibres within the extreme capsule fibre complex (EmC) and the uncinate fascicle (UF).

Similar to the idea of co-development and aging the major white matter association systems may be places for the spread of plasticity and degeneration (322). However some graph-theoretic approaches may be blind to these effects since they discard the exact topographical layout of connections when distilling network-models as consisting of nodes and edges. We therefore aimed to find an approach that is on the one hand data-driven like graph-theoretic approaches but on the other hand faithful to topographic- anatomical information.

4.2.5 A novel technique of dividing white matter into association fibre systems

We therefore devised a novel technique of dividing the white matter into the major association fibre systems based on similarities in cortical projection patterns. We ran probabilistic tractography from all voxels in several coronal slices through the white matter. Using k-means clustering we grouped together white-matter voxels with similar grey-matter projection patterns. We were thus able to reconstruct all known major association fibres of the human cerebral cortex and their respective pattern of grey-matter projection. In this manner we obtained two probabilistic atlases: One atlas denotes the topography of all major white matter association fibre systems. The other atlas is a probabilistic depiction of (surface) patterns of cortical projections for each of these association fibre systems. The same analyses were carried out in the rhesus macaque allowing for the cross-species validation of this approach.

We hope that this data-driven way of identifying connections and white matter fibre systems may proof useful in several different ways:

(1) First a probabilistic atlas of fibre systems in the white matter together with their probabilistic cortical projection patterns may allow us to link “grey-matter” and “white matter” imaging data more effectively. As mentioned above a growing body of imaging literature using for example PET and fMRI have linked rapid changes in “activity” in specific cortical regions with defined cognitive processes. Another set of studies has linked between-subject differences of white-matter microstructure as measured with DW-MRI to more stable (trait-like) behavioural differences. Learning and neuroplasticity has been associated with changes in both, grey and white matter. One may be able to use our linked grey and white matter atlas in order to reconcile these patterns of transient grey-matter activation with patterns of white matter differences across subjects to obtain an integrated view of cognitive processing.

(2) Secondly we show here that this methodological approach can be used in both human and macaque monkey brains. Hence this technique may offer a data-driven approach for studying

connections and network architecture across different species. It may permit the investigation of cross-species differences in both the strength and distribution of projections. Moreover it may allow for the discovery of new connections or even new connectional principles such as the above-mentioned principle of primary sensory / motor areas being either connected directly (in lower mammals), connected via association areas and the limbic association cortex (monkeys) or via secondary and tertiary association areas.

(3) Thirdly this methodical approach may offer ways to re-evaluate above-mentioned "disconnectional" or hodological accounts.

4.2.6 A tool for a new hodotopic framework for clinicopathological correlations

In recent years the notion of connectional pathologies has broadened significantly. Catani and Ffytche (102) suggest to also consider syndromes of "hyper-connectivity" as part of the "hodotopic" framework for clinicopathological correlations. They propose that connectivity-related dysfunctions may therefore range from classical disconnection syndromes, which emerge from pure white matter lesions, such as conduction aphasia, apraxia and agnosia to the combination of fronto-frontal hyper-connectivity with frontal disconnection from other brain regions in autism (323, 324). However patterns of connectivity may even change during the course of a disease. For example it has been suggested that prodromal and early stages of Alzheimer's disease are associated with hyper-connectivity in certain regions, which may later turn into hypo-connectivity (325, 326).

More generally the concept of "disconnection syndromes" has been invoked with respect to several neurodegenerative diseases. Seeley et al have linked the atrophy pattern of different neurodegenerative disorders such as Alzheimer's disease, fronto-temporal dementia, semantic dementia, progressive non-fluent aphasia and corticobasal syndrome to distinct networks of highly interconnected brain regions (327). In a follow-up study this group suggested that impairments to specific networks caused by a given neurodegenerative disease may be

compensated for by enhanced connectivity in alternative networks (328). These findings have been explained with reference to anterograde and retrograde trans-synaptic degeneration. Although for example in the case of Alzheimer's disease early pathology is thought to be localised in perirhinal cortex (329), it has also been shown that connected regions such as the posterior cingulate (PCC) exhibit a decrease in metabolic activity early in the course of the disease (330). This may be explained in terms of neurons in PCC degenerating as a result of being deprived of inputs (anterograde trans-synaptic degeneration) or losing trophic support from target regions (retrograde). Other theories have linked these "networks" of degeneration with correlated gene-expression in highly connected areas (331). Finally more controversial theories have compared the trans-neural spread of degeneration and accumulation of fibrillary proteins such as amyloid-β, α-synuclein and tau with prion-diseases (e.g. Creutzfeldt-Jakob Disease) (332).

Future research will need to clarify whether some neurodegenerative diseases can be understood from a hodological perspective. However caution should be exercised: it may be unwise to broaden the notion of disconnection-syndromes to capture any distributed pathology. Lesions or dysfunctions in circumscribed cortical areas are likely to go beyond the affected cortical site and include other directly connected cortical regions. These phenomena may not be considered disconnection syndromes. Rather we may want to think of hodological pathologies as arising primarily from aberrant connections with at least initially spared cortical processing. Such increased or decreased conduction may particularly disrupt behaviour that relies on the simultaneous cooperation of cortical regions or the rapid transfer of outputs from one area to another.

I suggest here that the derived probabilistic maps of connections and projections may be used in order to delineate novel sets of hodological pathologies. One way to achieve this could be via the triangulation between (A) behavioural abnormality, (B) brain pathology and the (C) probabilistic maps of connections and projections. For example we might wonder whether a

certain neurological or psychiatric disorder such as neglect or autism can be thought of as a disconnection syndrome. (A) Linking the behavioural pathology (attentional reorienting in the case of neglect (333)) and its "normal" cortical network (the intraparietal sulcus [IPS] and lobule [IPL] and frontal eye fields [FEF] (334)) with (C) the probabilistic surface projection patterns already suggests that directing attention heavily relies on the cooperation of cortical regions connected via the superior longitudinal fascicles 2 and 3 (SLFs). We may now link (B) brain lesions causing neglect with (C) the probabilistic atlas of association fibres and again delineate SLFs 2/3 as the main focus of overlap for lesions causing neglect (335, 336). This may support the notion of neglect being a disconnection syndrome. However note that the same triangulation procedure may work in cases where macroscopic white matter lesions cannot reliably be found (e.g. in the proposed hyper-connection or mixed syndromes) by replacing (B) white matter pathology with patterns of grey matter change (e.g. atrophy), hyper/hypo-metabolism, changes in functional coupling or any other profile of grey matter change and in turn linking these with (C) probabilistic projection maps of association fibres.

To summarise in the following chapter I will demonstrate a novel data-driven method of dividing the white matter into its major association fibre systems in both the human and the macaque monkey. (A) I show that this can be used in order to link the topography of white matter association fibres with their respective pattern of cortical projection. (B) I also suggest that this can be used to compare the skeleton of connections and the cortical distribution of their projections across a wider range of species in order to test hypotheses about brain evolution. (C) Finally I demonstrate a new way of assessing potential disconnection syndromes or hodological pathologies more generally.

4.3 Methods

4.3.1 Human data acquisition and pre-processing

Human in-vivo diffusion MRI data were provided by the Human Connectome Project (HCP), WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University (337). The minimally pre-processed datasets of the first 15 subjects (9 female, age range 26-35 years) from the Q2 public data release were used. Data acquisition and pre-processing methods are detailed in Ugurbil et al. (338), Sotiropoulos et al. (339) and Glasser et al. (340). In brief, 1.25 mm isotropic resolution data were collected across the entire brain on a customized 3T Siemens Skyra scanner using a monopolar Stejskal-Tanner diffusion scheme with a slice-accelerated EPI readout. Sampling in *q*-space included 3 shells at *b* = 1000, 2000, and 3000 s/mm^2. For each shell 90 diffusion gradient directions and 6 *b* = 0's were acquired with reversed phase-encoding direction for TOPUP distortion correction (341).

We used BedpostX to fit a crossing fibre model to the data (199). A multi-shell extension was used to reduce overfitting of crossing fibres due to non-mono-exponential diffusion decay (342). Up to three fibre orientations per voxel were allowed. This produced voxel-wise posterior distributions of fibre orientations that were subsequently used in probabilistic tractography. In addition, we used a 32k surface created using FreeSurfer on the T1-weighted image (acquired using an MPRAGE sequence at 0.7 mm isotropic resolution) and aligned to the diffusion space as part of the HCP's minimum pre-processing pipeline (340) to display the data. All data were pre-processed and analysed and visualized using tools from the FMRIB Software Library (343), custom code in Matlab (The Mathworks) belonging to the in-house MR Comparative Anatomy Toolbox (Mr Cat) and the Connectome Workbench (344).

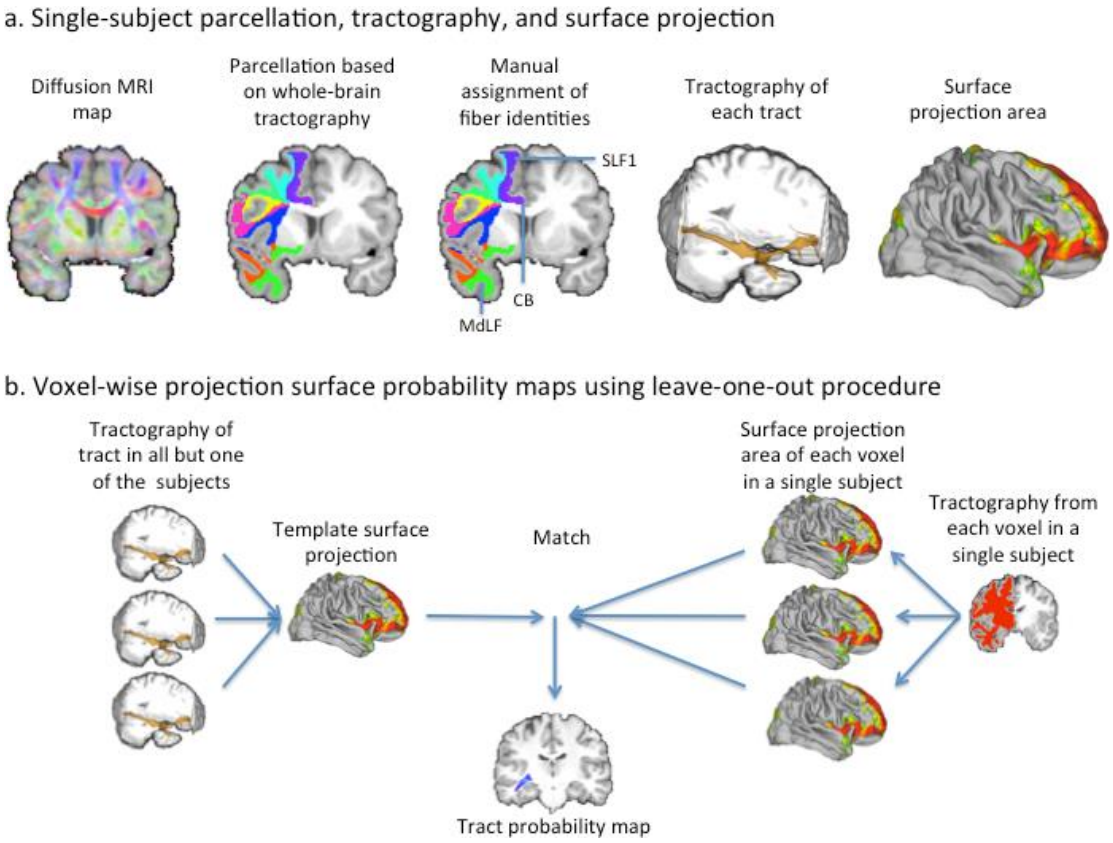


Figure 4-1 depicts the overall methodological approach of the first part of the study on human association fibre systems. Investigation of monkey association fibre systems was conducted largely in parallel to this. Part (a) depicts the steps of the parcellation of four coronal white-matter ROIs (only one representative ROI shown here) in 15 individual subjects (only one representative subjects shown here) from left to right. DW-MRI maps were used to parcellate coronal ROIs into different clusters based on each voxel's pattern of projection to the cortical surface as estimated using probabilistic tractography. Probabilistic tractography was run from all identified clusters and summed across cluster-labels (CingB, SLF1, 2, 3, AF, EmC, UF, MdLF, ILF) and subjects to obtain group-representation of all tracts (only one representative tract – EmC – shown here). Part (b) depicts the creation of white-matter probability maps. Correlating all white-matter voxel's surface projection patterns of each subject with the group-derived surface projection patterns of each tract, but leaving the specific subject out, created these maps.

4.3.2 Human white matter parcellation and tractography

Figure 4-1 depicts the overall methodological approach of this study. For each subject, four right hemisphere coronal regions of interest were created incorporating all the white matter at the level of the (1) head of the caudate (mean y=19), at the level of the precentral gyrus (y=3),

the planum temporale (y=-13), and the splenium of the corpus callosum (y=-29). A cortical grey-matter target mask was obtained by using FSL's automated tissue-segmentation tool FAST and by removing all subcortical grey matter manually. For each coronal slice, probabilistic tractography was performed from each white matter voxel to each voxel in the whole cortical grey matter of the brain. To reduce the computational load, the target cortical grey-matter mask was down-sampled to 5 mm isotropic for this analysis. Tractography was unconstrained except for a two coronal waypoint masks 6 mm anteriorly and posteriorly of the coronal region of interest to enforce longitudinal tracking for the first few steps. Moreover in order to exclude any projection fibres (i.e., fibres projecting from the neocortex to basal ganglia, brainstem structures, cerebellum and spinal cord) we used an exclusion mask that comprised the basal ganglia, the thalamus, the brainstem and the cerebellum. This exclusion mask was based on the Harvard-Oxford Subcortical Structural Atlas and was manually refined in each individual subject. Finally a sagittal exclusion mask at z=0 was used to exclude any commissural fibres (corpus callosum and anterior commissure). Probabilistic tractography was run from each voxel of the white-matter ROI to the down-sampled cortical grey-matter mask using FSL's "probtrackx" with the following settings: a curvature threshold of 0.2, 2000 steps with a step length of 0.3125, 5000 samples, a fibre-threshold of 0.01, no distance threshold, random sampling of the initial fibres above fibre-threshold (randfib = 1), using loop checks and correcting the path distributions for the length of the pathways.

The resulting data were stored in a matrix of dimensions (number of coronal white matter voxels × all voxels in the rest of the brain) that was subsequently used to create a cross-correlation matrix of dimensions (number of coronal white matter voxels × number of coronal white matter voxels) indicating the similarity in connectivity profile of each of the voxels in the coronal region of interest. The rows of the cross-correlation matrix were permuted using the *k*-means clustering algorithm to group together voxels with a similar connectivity profile. In this algorithm the number of clusters needs to be set by the experimenter. We expected to find at

least nine clusters belonging to the longitudinal fibres, but it is possible to find clusters of voxels that do not belong to any of those because they may correspond to projection fibre pathways or short u-fibres. We therefore conducted k-means clustering of the cross-correlation matrix multiple times obtaining solutions for nine to sixteen clusters. The silhouette measures (345) of the different solutions were compared to determine the number of clusters with lowest silhouette value. We also included a small distance constraint of 0.1 in the k-means clustering to increase the chances of obtaining continuous clusters.

In the next step we labelled the clusters in each individual subject's "optimal" parcellation solution. This labelling was done based on each cluster's topography (location relative to other grey and white matter structures, extend, shape) with reference to two previous atlases of the white matter (346, 347). More specifically clusters that were located at least partly in the cingulate gyrus or the cingulate sulcus in the frontal and parietal lobes were labelled "CingB"-cluster (because they resembled the topography of the cingulum bundle). Clusters that were located in the vicinity of the parahippocampal gyrus in the temporal lobe were also labelled "CingB". Clusters that were located in the superior frontal gyrus and in the superior parietal lobule were labelled cluster "SLF1" (because they resembled the topography of the superior longitudinal fascicle 1). Clusters in the middle frontal gyrus and in the vicinity of the inferior parietal sulcus were labelled "SLF2". Clusters in the dorsal aspect of the inferior frontal gyrus and in the inferior parietal lobule were labelled "SLF3". Clusters in the ventral aspect of the inferior frontal gyrus and in the vicinity of the frontal, parietal and posterior temporal operculum were labelled "AF" (because they resembled the arcuate fascicle). Clusters that lay between the insular cortex and the claustrum were labelled "EmC" (because they resembled the extreme capsule). Clusters in the superior temporal gyrus and the vicinity of the superior temporal sulcus were labelled MdLF (resemblance of middle longitudinal fascicle). Clusters in the middle and inferior temporal gyrus were labelled ILF (resemblance of inferior longitudinal

fascicle). Clusters in the lateral orbitofrontal and temporopolar white matter were labelled UF (resemblance of uncinate fascicle).

We then performed tractography from each of the assigned clusters in each coronal slice. Probabilistic tractography was run using FSL's probtrackx2 both in 3D-diffusion space and to a 32k surface created using FreeSurfer. This step used the same settings as the tractography for the parcellation (see above), but the target brain was not downsampled. In addition to the masks described above, we used all other clusters in the same slice as exclusion masks, ensuring that our tractography is not contaminated with fibers going through neighbouring clusters. Since tractography of longitudinal fibers have a tendency to find pathways back-projecting (for instance, AF fibers from the posterior to frontal cortex often enter the extreme capsule going back through the temporal lobe), we also added a coronal exclusion masks excluding dorsal (fronto-parietal) white matter tracts on the level of the anterior commissure for the ILF, MdLF, and EmC and a coronal exclusion mask excluding ventral (temporal) tracts on the level of the anterior commissure for the superior longitudinal and arcuate fascicles. We added a full coronal slice at MNI y=-12 for the UF to prevent it projecting back through the EmC.

For each subject, the tractograms from the different clusters belonging to a single fibre (e.g., SLF1 tractograms resulting from clusters identified in the different coronal slices) were log transformed, averaged, and normalized to the maximum value in the average tractogram. Subsequently, for each tract, the subject-specific tractograms were averaged and normalized to the maximum value to create the tractogram for the population. To make the results available to the scientific community, we created an atlas of each tract's course compatible with FSLview MRI data display tool that will be made available online.

4.3.3 Human Surface analyses

In the tractography step above, tracts that hit the grey/white matter border, as defined by a FreeSurfer segmentation of each individual's structural scan as part of the Human Connectome

Project's minimal pre-processing pipeline (340) were recorded in the 32k vertex space. To create surface projection maps of each tract, these surface projections were analysed in a manner analogous to the tractograms described above. For each subject, the surface projections from the different clusters belonging to a single fibre were log transformed, averaged, and normalized to the maximum value in the average surface projection. Subsequently, for each tract, the subject-specific surface projections were averaged and normalized to the maximum value to create the projection surface for the population.

Projecting the data in surface space allowed us to perform statistical comparisons of the surface projections of each tract. We created a model containing nice regressors containing the surface projection of each tract for each subject, as well as regressors to explain subject-specific variance. Contrasts comparing the surface projections of each tract in the group where then performed using permutation testing as implemented in FSL's "randomise" tool (348). Statistical tests were performed in vertex space and the results were thresholded at a level FWE-corrected level of $p<0.05$.

To further quantify how each tract contributes to the connectivity profile of individual areas we created connectivity fingerprints (349) of specific areas with all nine tracts. We used surface projections of previously published probabilistic atlases of dorsal frontal areas 9/46d, 9/46v, and 46 (196), ventrolateral frontal areas 6v, 6r, 44d, 44v, 45, and 47 (350), and five area along the inferior parietal lobule (192). Each area's mask was projected to the 32k FreeSurfer surface and thresholded at 50% of the population. This allows us to examine the specific combination of tracts that form the connectivity fingerprint of each cortical area. The mean projection strength of each tract for each subject was recorded and normalized to the maximum and minimum projection strength of each tract. Note that this thus created a connectivity fingerprint showing the relative contribution of each tract to the connectivity profile of each area, but that the values of tracts cannot be compared across areas.

4.3.4 Human probabilistic projection map and leave-one-out analysis

In the next step we aimed to derive a probabilistic map of cortical surface-projections of all tracts. Moreover we verified the results of our white-matter parcellation by running a "leave-one-out" comparison between all white-matter voxel's projection patterns for each particular subject and the tract surface atlas but leaving the specific subject out from the atlas of tract-projection patterns.

More specifically we created the probabilistic surface atlas of tract-projection patterns from the probabilistic tractograms on the 32k vertex FreeSurfer surface obtained by running probtrackx2 after the parcellation (see above). The tractography-derived surface-maps were as before log-transformed and averaged within each subject across ROIs for every particular cluster-label (CingB, SLF1, 2, 3, AF, EmC, UF, MdLF, and ILF). They were then normalized within each subject to values between 0 and 1. Then they were summed across subjects to obtain a probabilistic representation of all tract-labels. The resulting group-derived probabilistic surface maps were once more normalized to their maximum value and multiplied with 100 so they may be used more intuitively. The values on the cortical surface in each vertex may reflect the "probability" of a given tract projecting to this "patch of cortex".

For the leave-one-out analysis we again used probtrackx2 to run tractography to the "cortical surface" (i.e. the 32k vertex surface) this time from each voxel in every subject's white matter. The settings for the probabilistic tractography, the waypoint and exclusion masks were exactly as in the initial tractography for the parcellation. In the next step we obtained the spatial correlation on the cortical surface between every voxel of each subject's white-matter projection pattern (i.e. approximately 200,000 voxels) and the group-derived probabilistic surface projection patterns of the nine different tracts (CingB, SLF1, 2, 3, AF, EmC, UF, MdLF, ILF). However we compared a given subject's white-matter projections only to the probabilistic atlas that was created leaving this particular subject out. For example we spatially correlated each white-matter voxel's surface projection pattern of subject 7 with the nine tract's

probabilistic atlases created from subjects 1-6 and 8-15. Each white matter voxel of subject 7 will be assigned nine values – reflecting the spatial surface correlation with each of the nine probabilistic tract-atlases. These correlation values are then combined to one volume for each subject and each of the nine tracts. They were then normalized across the volume to obtain probabilities between 0 and 1. Each subject's nine probabilistic maps were then registered to the MNI152_1mm template using FLS's FNIRT tool. For each subject we then determined in each voxel, which tract had the highest normalized correlation value. We then calculated the voxel-wise percentage of subjects of the group which showed a particular tract as having the highest correlation value. We hence created another probabilistic atlas in the white matter. This atlas depicts each voxel's probability of yielding a certain surface projection pattern with probabilistic tractography. White matter renderings of these results were created using custom-made software written in MATLAB.

4.3.5 Monkey data acquisition and pre-processing

In-vivo diffusion MRI data were obtained from six rhesus monkeys (*Macaca mulatta*, 3 female, average age 5.12 years) using a 3T whole-body scanner at the University of Oxford. Protocols for animal care, magnetic resonance imaging, and anaesthesia were carried out under authority of personal and project licenses in accordance with the UK Animals (Scientific Procedures) Act (1986) using similar procedures to those that we have described previously (351, 352). During scanning animals were kept under minimum anaesthetic using isoflurane and kept under continuous veterinary observation. A four-channel phased-array radio-frequency coil in conjunction with a local transmission coil was used for data acquisition (Windmiller Kolster Scientific, Fresno, CA, USA). Diffusion MRI scanning was performed using a twice-refocused diffusion-weighted spin echo sequence (353) with TE/TR: 102 ms/8.3 sec; resolution: 1 x 1 mm; slice thickness: 1 mm; 60 isotropically distributed diffusion directions; b = 1000 s/mm^2. In order to correct for image distortion, images were acquired with both left-right

and right-left phase encoding direction and subsequently combined (341) using the FSL TOPUP tool. Six datasets were acquired in a single session and averaged to create a single set of data. Voxel-wise model fitting of diffusion orientations and tractography was performed in the same way as in the human datasets. We used BedpostX to fit a crossing fibre model to the data (354).

Ex-vivo diffusion MRI data were obtained from one rhesus monkey (*Macaca mulatta*, male, age at death 4.03 years) using a 7T magnet with a Varian DirectDrive^{tm} (Agilent Technologies, Santa Clara, CA, USA). The brain was perfusion fixed with formalin and subsequently fixed in agar. Seven non-diffusion-weighted (b = 0 s/mm^2) and three diffusion-weighted (b = 4000 s/mm^2) datasets were acquired using a single line readout, 2D Stejskal-Tanner pulse sequence (355) (TE/TR: 25 ms/10 sec; array size: 128 x 128; resolution 0.6 mm x 0.6 mm; 128 slices; slice thickness: 0.5 mm; receiver bandwidth: 100 kHz). All co-registered b = 0 s/mm^2 and b = 4000 s/mm^2 datasets were averaged to create a single set data. The diffusivity of ex-vivo tissue is generally reduced due to death and fixation, while diffusion anisotropy is largely preserved (356). This change in diffusivity necessitates the use of larger b-values to achieve equivalent diffusion contrast to in-vivo data, as is achieved here by increasing the diffusion coefficient from b = 1000 to 4000 s/mm^2 (357, 358). Again the same overall approach was used for voxel-wise model fitting of diffusion orientations and tractography.

4.3.6 Monkey white matter parcellation and tractography

The white-matter parcellation in the monkey was carried out largely paralleling the steps for the human white-matter parcellation. It was performed solely on the monkey in-vivo diffusion MRI data. In each subject, six right hemisphere coronal regions of interest were created incorporating all the white matter. As in the human we drew ROIs at the level of the (1) head of the caudate (mean y=10 according to the macaque brain template of (302)), at the level of

the precentral gyrus (y=0), the planum temporale (y=-10), and the splenium of the corpus callosum (y=-17). In the case of the monkey we added two more regions of interest, one at the level of the posterior extend of FEF (y=6) and one at the level of the central sulcus (y=-5) in order to make clustering less susceptible to the choice of ROIs in this smaller group of subjects. A cortical grey-matter target mask was obtained by manually segmenting a macaque brain template (302) into the different tissue-components (grey and white matter, CSF) and removing all subcortical grey matter manually from the grey-matter mask. This cortical grey-matter mask was then nonlinearly registered to each individual monkey's DW-MRI space using FNIRT and then manually adapted to comprise all cortical grey matter voxels. For each coronal slice, probabilistic tractography was performed from each white matter voxel to each voxel in the whole cortical grey matter of the brain. To reduce the computational load, the target cortical grey-matter mask was downsampled to 5 mm isotropic for this analysis. As for the human analysis tractography was unconstrained except for a two coronal waypoint masks 3 mm anteriorly and posteriorly of the coronal region of interest to enforce longitudinal tracking for the first few steps. Again in order to exclude projection fibers we used an exclusion mask that comprised the basal ganglia, the thalamus, the brainstem and the cerebellum. This exclusion mask was manually drawn in each monkey's DW-MRI space. Finally as in the human a sagittal exclusion mask at z=0 was used to exclude any commissural fibers. Probabilistic tractography was run from each voxel of the white-matter ROI to the down-sampled cortical grey-matter mask using FSL's "probtrackx" with the following settings: a curvature threshold of 0.2, 2400 steps with a step length of 0.25, 5000 samples, a fibre-threshold of 0.01, no distance threshold, random sampling of the initial fibres above fibre-threshold (randfib = 1), using loop checks and correcting the path distributions for the length of the pathways.

As was done for the human white-matter parcellation the resulting data were stored in a connectivity matrix, which was subsequently used to create a cross-correlation matrix. The cross-correlation matrix was permuted using the *k*-means clustering algorithm to group

together voxels with a similar connectivity profile. We again used the lowest silhouette measure (345) of the different nine to sixteen clustering solutions in order to determine the number of clusters for the best clustering solution.

As for the human data we labelled the clusters in each individual monkey's "optimal" parcellation solution. This labelling was done based on each cluster's topography (location relative to other grey and white matter structures, extend, shape) with reference to two previous atlases of the white matter (346, 347) using exactly the same labels as for the human data. Please note that this should not be taken to suggest that these white-matter tracts are identical across the two species. In fact one of the goals of this experiment was to allow for testing whether they are similar in terms of their size, topography, projection patterns and tract integrity. However we used the parallel labels largely in line with previous literature of human and monkey white matter anatomy. Criteria paralleled those for the human labelling.

Again largely paralleling the methods used for the human we performed tractography from each of the assigned clusters in each coronal slice with the same settings as in the parcellation, but the target brain was not downsampled. In addition to the masks described above, we used all other clusters in the same slice as exclusion masks and added a coronal exclusion masks excluding dorsal (fronto-parietal) white matter tracts on the level of the anterior commissure for the ILF, MdLF, and EmC and a coronal exclusion mask excluding ventral (temporal) tracts on the level of the anterior commissure for the superior longitudinal and arcuate fascicles. We added a full coronal slice at y=-7 for the UF to prevent it projecting back through the EmC.

For each monkey, the tractograms from the different clusters belonging to a single label were log transformed, averaged, and normalized to the maximum value in the average tractogram. Subject-specific tractograms were averaged and normalized to the maximum value to create the tractogram for the population.

4.3.7 Monkey probabilistic projection map, leave-one-out analysis, and comparison with high-resolution ex-vivo data

As for the human white-matter parcellation we aimed to derive a probabilistic map of cortical surface-projections of all tracts. We created the probabilistic surface atlas from aforementioned tractograms obtained by running tractography from all clusters after the parcellation. These average tractograms were projected onto a 10k vertices monkey- surface template using the "surf_proj" algorithm as implemented in FSL. Because the parcellation was carried out in both hemispheres we could average across hemispheres and subjects to obtain a probabilistic representation of all tract-labels based on 12 monkey hemispheres (from six subjects). These added tractograms were again normalized to their maximum value and multiplied with 100 obtaining a probabilistic depiction of surface projection of the association-fibre labels.

We ran a leave-one-out analysis paralleling the methods used for the human data to obtain the spatial correlation on the cortical surface between every voxel of each monkey's white-matter projection pattern and the group-derived probabilistic surface projection patterns of the nine different tracts. We compared a given monkey's white-matter projections in a given hemisphere to the probabilistic atlas that was created leaving this particular subject and hemisphere out (one hemisphere's voxels are matched to projection patterns from 11 hemispheres). As in the human these correlation values were again averaged across subjects and another probabilistic atlas in the white matter was obtained. This atlas depicts each voxel's probability of yielding a certain surface projection pattern with probabilistic tractography.

In order to verify the white-matter parcellation results obtained in six in-vivo monkey DW-MRI datasets we ran a similar spatial correlation analysis on the cortical surface between the previously obtained probabilistic projection atlas with surface projections from every white-matter voxel of one high-resolution ex-vivo DW-MRI dataset. Tractography was run from each

white-matter voxel of the ex-vivo sample to the cortical grey matter using parallel tractography methods and settings. The target cortical grey matter mask was a manually drawn. As before we used two coronal waypoint masks 3 mm anteriorly and posteriorly of the coronal region of interest to enforce longitudinal tracking for the first few steps. Again in order to exclude projection fibers we used a manually drawn exclusion mask that comprised the basal ganglia, the thalamus, the brainstem and the cerebellum. A sagittal exclusion mask at z=0 was used to exclude any commissural fibers. Probabilistic tractography was run from each voxel of the white-matter using FSL's "probtrackx" with the following settings: a curvature threshold of 0.2, 3840 steps with a step length of 0.125, 10000 samples, a fibre-threshold of 0.01, no distance threshold, random sampling of the initial fibres above fibre-threshold (randfib = 1), using loop checks and correcting the path distributions for the length of the pathways. Each voxel's tractogram was projected onto the 10k vertices monkey- surface template using the FSL's "surf_proj". They were then spatially correlated with the probabilistic surface atlas of tract-projections from the 12 in-vivo hemispheres. As before each white-matter voxel was assigned nine correlation values reflecting spatial correlation with the nine tract-labels. White matter renderings of these results were created using custom-made software written in MATLAB.

4.3.8 Meta-analyses

Next we wanted to qualitatively compare the different association fibre's surface projection patterns with patterns of activity or grey-matter density change that are linked with potential "disconnection syndromes". The latter were obtained in part by performing meta-analyses of VBM and fMRI studies using the BrainMap database performing likelihood estimations (359) of structural and functional brain activity. The BrainMap database (www.brainmap.org) is an online depository of published functional and structural neuroimaging experiments with coordinate-based results (x,y,z) in Talairach or MNI space.

The BrainMap-database was queried on June 24th, 2015, when the database contained 2177 papers, 83 paradigm classes, 40934 participants, 10330 experiments, and 82135 locations. Neurosynth was searched on June 24th, 2015 when it comprised 386,455 activations and reported in 10,903 studies and meta-analyses of 3342 terms.

In a first meta-analysis we wanted to investigate the network of brain regions that had repeatedly been implicated in speech repetition in the literature. This specific cognitive ability is thought to be impaired in patients with conduction aphasia. We thus searched the BrainMap database for experimental paradigms that probed "overt repetition" (*Experiments: Paradigm Class is Recitation / Repetition (Overt)*) looking for activations (*Experiments: Activation is Activation only*) using fMRI (*Experiments: Imaging Modality is fMRI*) in healthy participants (*Subjects: Diagnosis is Normal*). Such a search yielded 17 papers, with 48 of 85 experiments matching criteria.

In a second set of meta-analyses we turned to cortical networks that had been implicated in spatial attention and hemispatial neglect. We attempted to find a network of regions reported for spatial attentional control more generally. We thus searched the BrainMap database for experimental paradigms that probed "visuospatial attention" (*Experiments: Paradigm Class is Visuospatial Attention*) looking for activations (*Experiments: Activation is Activation Only*) using fMRI (*Experiments: Imaging Modality is fMRI*) in healthy participants (*Subjects: Diagnosis is Normal*). Such a search yielded 102 papers, with 329 of 478 experiments matching criteria.

In a third set of meta-analyses we turned to reading, an ability that is impaired in patients with dyslexia. A recent study argued dyslexia can be understood as a disconnection syndrome (360). We searched the BrainMap database for experimental paradigms that probed "Reading (overt/covert)" (*Experiments: Paradigm Class is Reading (overt/covert)*) looking for activations (*Experiments: Activation is Activation Only*) using fMRI (*Experiments: Imaging Modality is fMRI*) in healthy participants (*Subjects: Diagnosis is Normal*). Such a search returned 61 papers, with 242 of 361 experiments matching criteria.

Next we wanted to obtain a representative map of cortical regions implicated in memorizing. Memory-impairments are the hallmark of almost all forms of dementia and several other neuro- degenerative diseases. Recent studies have implicated that some neurodegenerative diseases and their microscopic correlates (e.g. amyloid deposits and neurofibrillary tangles in the case of Alzheimer's disease) propagate across the brain along major fibre-connections and may thus resemble patterns known from disconnection-syndromes (327, 328, 332). The BrainMap database was searched for studies that had probed "memory" more generally (*Experiments: Behavioural Domain is Cognition = Memory*) looking for activations (*Experiments: Activation is Activation Only*) using fMRI (*Experiments: Imaging Modality is fMRI*) in healthy participants (*Subjects: Diagnosis is Normal*). We obtained 379 papers, with 1413 of 1883 experiments matching criteria.

Next we turned to attention deficit hyperactivity disorder (ADHD) – a neurodevelopmental psychiatric disorder, which has recently been linked with decreased connectivity in pathways connecting prefrontal and parieto-occipital areas (361). We searched the BrainMap database for studies involving ADHD-patients (*Subjects: Diagnosis is Attention Deficit / Hyperactivity Disorder*) and found 17 papers, with 52 of 83 experiments matching criteria. Development stuttering has also been linked to changes in white matter integrity (362). Searching the BrainMap database for studies involving patients with developmental stuttering (*Subjects: Diagnosis is Developmental Stuttering*) yielded 19 papers, with 129 of 163 experiments matching criteria. Finally we attempted to link patterns of grey-matter change to patterns of surface-projection from the parcellation-derived atlas. A considerable body of literature has implicated schizophrenia as a disconnection syndrome (363) particularly with reference to temporo-frontal connections. We searched the BrainMap database this time for voxel-based morphometry studies involving schizophrenia-patients (*Subjects: Diagnosis is schizophrenia*) and found 129 papers, with 355 of 391 experiments matching criteria. Similarly we searched the BrainMap database for VBM studies involving patients with frontotemporal dementia

(*Subjects: Diagnosis is schizophrenia*) and found 15 papers, with 60 of 95 experiments matching criteria.

These meta-analysis derived maps were projected onto the cortical surface using FSL's surf_proj tool. We then compared them to all nine probabilistic surface-projection atlases derived from the connectivity-based parcellation using spatial correlation on the surface.

In a final analysis we used the individual subject's whole-brain summary-maps from the task-data of the HCP from the language processing (story) task and the social cognition (theory of mind) task. We conducted a regression analysis predicting the distribution of activity in these tasks on the cortical surface of each individual subject with the same subject's surface-projection patterns of the nine major association fibre bundles.

4.4 Results

Using DW-MRI data of 15 healthy human subjects and six macaque monkeys, we parcellated coronal sections of the cortical white matter and identified anterior-posterior fibers belonging to nine major longitudinal association tracts. We performed probabilistic tractography from each of the seeds, showing the tract course of each fibre throughout the white matter. We recorded the number of streamlines that reached the boundary of the white and grey matter and created a surface - projection map of each tract. We will first describe the projection maps obtained from the parcellation in human and monkey in-vivo data. We will then present the tract course of each fibre throughout the white matter as obtained from the leave-one-out analysis in both species. After that we present a parcellation in an ex-vivo DW- MRI data-set from one macaque monkey. We will then discuss the relationship of these findings with tentative disconnection-syndromes.

4.4.1 Connectivity-based tractography and surface analyses

We performed connectivity-based parcellations of coronal slices through white matter to identify the bodies of nine major longitudinal fibers in the human and monkey brain and subsequently performed tractography from each of these clusters towards the cortex to create a projection map of each tract. The resulting projection maps are displayed in Figure 4-2. We will discuss the details of these maps in detail below.

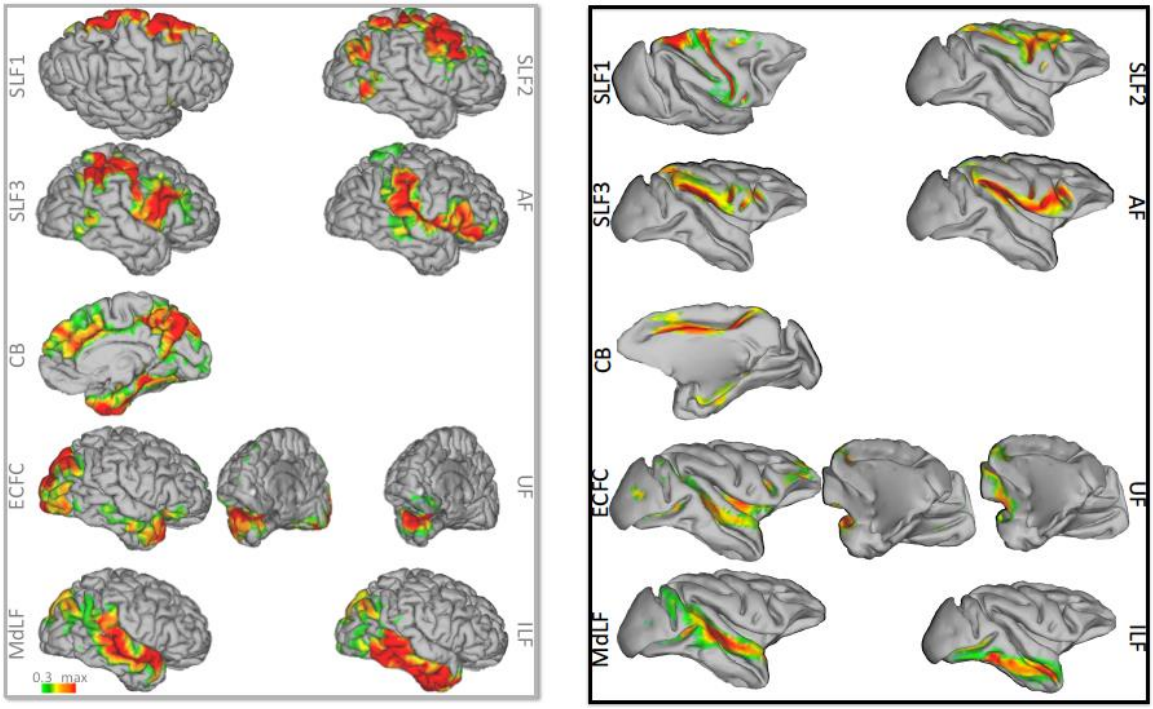


Figure 4-2: Surface projection maps of all fibers. Normalized average projection maps of each of the nine tracts thresholded at 0.3 projected onto the grey-matter surface. On the left panel the human projection maps are plotted onto the grey-matter surface rendering of subject 1. On the right panel the monkey projection maps are plotted onto the F99 surface template.

4.4.2 Frontal-parietal fibers

In the dorsal part of the cortex we aimed to identify the three separate branches of the superior longitudinal fascicle and the arcuate fascicle (SLF1, 2, 3, AF in Figure 4-3). Previous studies often group at least some of these fibre bundles together and the relative contribution of each of them to the connections of distinct frontal cortical areas is often left open. Here,

using the high-resolution data from the Human Connectome Project in combination with a data-driven parcellation of the white matter itself, we are able to isolate each of the different fibers in each individual subject. Moreover, using surface-based statistics, we can quantitatively compare the projection areas of these different fibers.

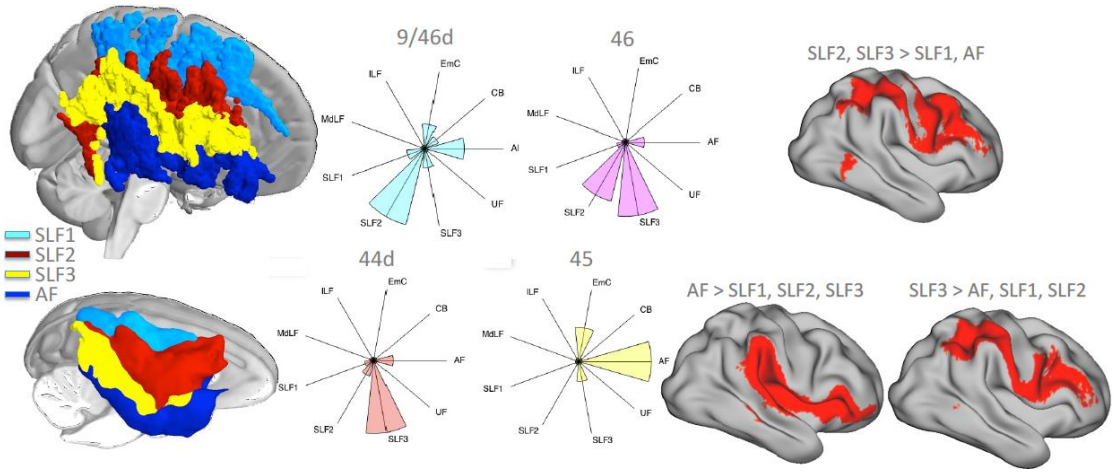


Figure 4-3: The left panels show 3D-renderings of all dorsal tracts (SLF1 [light blue], 2 [red], 3 [yellow], and AF [dark blue]) in both humans and monkeys. The middle panels show connectivity fingerprints of four frontal areas (9/46d, 46, 44d, 45). The right panels show results of statistical comparisons of surface projections between these dorsal fibres: Areas where SLF2 and SLF3 project significantly more across the group of subjects when compared to SLF1 and AF are shown in the right upper panel. Areas where AF's projections are significantly stronger than those of SLF1, 2, 3 are depicted in lower left. Areas where SLF3 is stronger than AF, SLF1 and SLF2 are shown on lower right.

As described previously (305, 364-366) the first branch of the superior longitudinal fasciculus (SLF I) runs in the superior parietal lobule and the superior frontal gyrus and connects areas in the superior parietal lobule and precuneus (areas 5, 7, dorsal aspects of 1, 2, 3), to the superior frontal cortex (dorsal aspects of M1 and PMd, SMA, pre-SMA, areas 8, 9, 32). We also observed a side-branch of the SLF1 running from the dorsomedial-to-ventrolateral thus connecting medial SFG (SMA, pre-SMA) with posterior IFG (areas 6v and 44). Such fibres might

be attributable to the previously reported separate short-range association pathway, termed the frontal aslant tract (305) and may thus not be part of the SLF1 in the narrow sense.

The second branch (SLF II) runs in the vicinity of the intraparietal sulcus (IPS), below the hand-area in M1 and reaches the middle frontal gyrus. It connects areas in the IPS (areas AIP, MIP) with lateral somatosensory and motor cortex, premotor cortex, FEF, and parts of dorso-lateral prefrontal cortex (areas 9/46d, 46, 8, 9). The third branch (SLF III) runs through the inferior parietal lobule in the vicinity of the IPS, ventral of the hand-area in M1, towards the inferior frontal sulcus and gyrus. It thus connects the intraparietal sulcus and inferior parietal lobule (areas PF, PFG, PG) to the inferior frontal gyrus (PMv, areas 44, 45, 47). In addition to these previously reported patterns of connectivity we also noted both for SLF2 and SLF3 a branch of fibres deflecting ventrally at the level of the posterior IPS and curving around the STS to reach the posterior inferior temporal sulcus and middle and inferior temporal gyrus, ventral to the projection areas of the AF (see also statistical comparisons in Figure 4-3). Again given the limitations of diffusion MRI tractography, it is important to consider the possibility that these projections are spurious, i.e., a 'rogue' tract entered into by tracts in the posterior part of the dorsal cortex where many tracts converge. However it should be noted that such connections had been reported previously but had been grouped with the arcuate fascicle (305). Moreover previous studies using a different method – namely resting-state fMRI – which does not have these constrains of spurious tracking had reported strong coupling between frontal areas such as FEF, 9/46v, IFJ and IFS with the same inferior temporal cortex regions, which are here shown to be reached by SLF2 and 3 (196, 350).

The arcuate fascicle in turn runs in the posterior half of the superior temporal gyrus and in the vicinity of the superior temporal sulcus, curves around the vertical part of the lateral fissure through the supramarginal gyrus to then run in the vicinity of the frontal operculum touching the circular insular sulcus and the lateral orbital sulcus. It thus connects regions in the posterior superior temporal cortex (areas 41, 42 22) and the temporal and parietal operculum

to regions in the ventral frontal cortex (areas 6r, 44v, 45, 47/12m and 47/12m) and the frontal operculum (Fop, OP4-10 according to (113)).

The specificity of the projections of the different branches of the SLF and the AF is illustrated by the connectivity fingerprint of four areas in the lateral prefrontal cortex: 9/46d, 46, 44d and 45. Although some of these areas are directly adjacent and are often referred to together under the more generic term 'dorsolateral prefrontal cortex', the fingerprints displayed in Figure 4-3 show that the more dorsal area 9/46d almost exclusively receives projections from the second branch of the SLF, whereas SLF2 and 3 reach the more anterior area 46. This difference was confirmed in an ANOVA with the factors AREAS (9/46d, 9/46v [not shown] and area 46) and TRACTS (all nine tracts obtained from the parcellation), which yielded an AREAS×TRACTS interaction ($F(16,224)=9.036$, $p<0.001$). Areas 44d and 45 (see Figure 4-3), often collectively referred to as "Broca's area" in the left hemisphere, largely receive distinct connections via SLF3 and AF, respectively. In support of the 'onion' model of fronto-parietal connectivity, which suggests that more anterior/posterior frontal regions are connected to more posterior/anterior IPL regions (319), we demonstrate that in the inferior parietal cortex, AF reaches more anterior areas, whereas SLF3 reaches more posterior areas. Interestingly, the posterior parts of the inferior parietal cortex are also reached by ventral longitudinal fibers (see below). Another ANOVA on IPL regions as defined by Mars et al (192) with the factors AREAS (PFop, PFt, PFM, PGa, PGp) and TRACTS again yielded significantly different combinations of connectivity with the nine fibers (AREAS×TRACTS interaction: $F(32,448)=19.833$, $p<0.001$).

Finally we ran statistical contrasts between these four fibres' surface projection patterns (see Figure 4-3). All of these tracts showed significantly different projection areas in extended patches of cortex. SLF1 projected significantly more strongly to the medial frontal pole, area 8m, pre-SMA, SMA, medial aspects of M1 and S1 and to areas in the superior parietal lobule and precuneus (areas 5, 7) when compared to SLF2, 3 and AF. SLF2 projected significantly

more strongly to areas in the middle frontal gyrus such as 9/46d and area 8A and to areas in the SPL when contrasted to SLF1, 3 and AF. SLF3 in turn (see Figure 4-3) showed significantly stronger connections to the lateral frontal pole, areas 46, IFS, IFJ, 6v, anterior IPL, IPS and notably an area in posterior middle temporal gyrus when contrasted to AF, SLF1 and 2. The AF (see Figure 4-3) projected significantly more strongly to areas 47, 45, 44v, 6r, the frontal and parietal operculum and the posterior superior temporal gyrus and sulcus when contrasted with projection patterns of SLF1, 2 and 3. When SFL2 and 3 where contrasted with SLF1 and AF (see Figure 4-3) we found significantly stronger projections to middle frontal gyrus and inferior frontal sulcus regions (46, 9/46d, 946v, IFS, 8A, IFJ), premotor cortex, IPS, supramarginal and angular gyrus, as well as posterior middle temporal gyrus.

4.4.3 Temporal and fronto-temporal fibers

In the temporal cortex we aimed to dissociate the two major tracts that run through the temporal lobe – the middle and inferior longitudinal fascicle. Moreover aimed at delineating three major fibre systems that connect mainly temporal cortex areas with frontal regions – the extreme capsule fibre complex (sometimes termed inferior fronto-occipital fasciculus (305)), the uncinate fascicle and the cingulum bundle. Again the connectivity-based parcellation approach was able to reliably delineate these fibre systems and yield significant differences in their patterns of cortical connections.

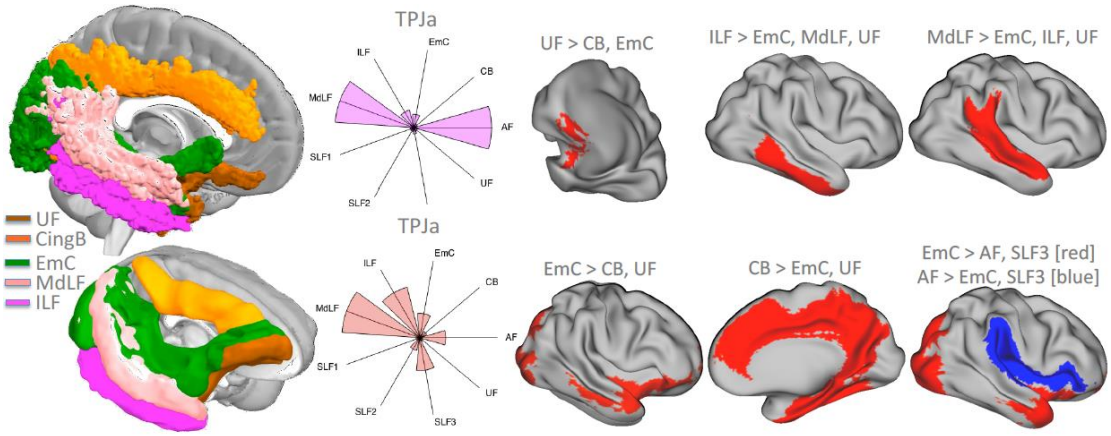


Figure 4-4: The left panels show 3D-renderings of all temporal and fronto-temporal fibers (CingB [orange], EmC [green], UF [brown], MdLF [pink] and ILF [purple]) in both humans and monkeys. The middle panels show connectivity fingerprints of two TPJ areas (anterior and posterior TPJ). The right panels show results of statistical comparisons of surface projections between these temporal and fronto-temporal fibers: Areas where UF projects significantly more across the group of subjects when compared to CB and EmC are shown in the top left panel. Areas where ILF's projections are significantly stronger than those of EmC, MdLF and UF are shown in the upper middle panel. The upper right panel shows regions where MdLF's projections are significantly stronger than those of EmC, ILF and UF. In the left lower panel EmC's cortical projections are contrasted with CB and UF. Areas where CB's projections are significantly stronger than those of EmC and UF are shown in the lower middle panel. The lower left panel finally compares EmC to two tracts that where described in the previous section: areas where EmC is stronger than AF and SLF3 are shown in red, whereas areas where AF is significantly stronger than EmC and SLF3 are shown in blue in the same panel.

We first delineated cingulum bundle – a large c-shaped fibre bundle in the medial frontal, parietal and temporal lobes. Its anterior-dorsal component runs through the white matter of the cingulate gyrus and in the vicinity of the cingulate sulcus. It curves around the genu of the corpus callosum. Its posterior-ventral section runs within the parahippocampal gyrus. Branches of the cingulum bundle leave the tract's core towards the dmPFC, the precuneus, cuneus and inferior temporal cortex. The cingulum bundle thus connects vmPFC (medial area 14), subgenual and perigenual cingulate areas (areas 25, 24, 32), dmPFC (medial areas 9, 8, 6), cingulate motor areas, posterior and retrosplenial cingulate regions (areas 23, 29, 30), paracentral lobule, precuneus and cuneus (areas 7 and 19), lingual (areas 18 and 19), fusiform areas (areas 19 and 37) parahippocampal gyrus (area 34), as well as the presubiculum, the

entorhinal, the perirhinal cortex and the amygdala. In addition to these extensive long-range connections the cingulum bundle is rich in short association fibres that link various cingulate regions with each other.

Secondly we delineated a bowtie-shaped tract in the occipital, temporal and frontal lobes that we term extreme capsule fibre complex because its fibres “squeeze” through the extreme capsule (between the insular cortex and the claustrum, equivalent to the tie of the bow-tie). Note that other authors have referred to similar fibres as “inferior fronto-occipital fasciculus” (305). The EmC derives from the inferior and medial surface of the occipital lobe, touches the dorsolateral wall of the posterior horn of the lateral ventricle, runs in the vicinity of the inferior temporal sulcus and middle temporal gyrus, deflects dorsally between the claustrum and the insular cortex just above the uncinate fasciculus and then curves around the anterior frontal operculum (OP8-10 (113)) to split into dorsomedial and ventrolateral/orbital fibres. It thus connects the inferior and medial occipital areas 17, 18, 19 and posterior medial temporal cortex area 37, lateral occipital and posterior inferior temporal cortex regions, superior temporal sulcus, middle temporal gyrus regions, and anterior superior temporal gyrus regions with insular, frontal opercular, anterior ventrolateral and orbital prefrontal and dorsomedial prefrontal regions.

Thirdly we identified a hook-shaped tract that ran in the anterior temporal, orbital and medial frontal lobe, which we call uncinate fascicle (UF). This tract ran in the vicinity of the temporal pole, the planum polare and the entorhinal cortex and lateral of the amygdala to then curve dorsally and track medial of the insular cortex and ventral of the EmC to reach the medial orbital gyrus. It hence connects temporopolar (temporal proisocortex), inferotemporal and perirhinal regions, anterior STS, parahippocampal areas and the amygdala with medial (areas 14m, FPm), central (areas 13, 11) and lateral OFC.

Finally we identified two major tracts running in the occipital and temporal lobes and to a limited degree in the inferior parietal lobule. One ran largely in the superior occipital and

temporal lobe, as well as in the IPL – the middle longitudinal fascicle. (MdLF). The other ran largely in the inferior occipital and temporal lobes – termed inferior longitudinal fascicle (ILF). The MdLF runs vertically through the posterior division of the supramarginal gyrus then curves around the vertical end of the lateral fissure and runs in the superior temporal gyrus and in the vicinity of the superior temporal sulcus towards the planum polare. It thus connects regions in the posterior SMG such as PFG and PGa with regions in the transverse temporal gyri, planum temporale, superior temporal gyrus and sulcus, as well as the planum polare. The ILF runs in the lateral occipital white matter and the posterior IPL, tracks the middle and inferior temporal gyrus and inferior temporal sulcus to then reach inferior temporopolar regions. It thus connects areas in the lateral occipital and posterior inferior temporal cortex such as V3, V4, V5/MT and LOC, posterior inferior IPL region PGp with regions in the inferior temporal gyrus such as TEO, areas TEp and TEa and inferior temporopolar proisocortex.

The specificity of the projections of the different temporal and fronto-temporal fibers is illustrated by the connectivity fingerprint of two different areas in the temporo-parietal junction (TPJa, TPJp). Although the two regions are often grouped to together as one region implicated in both, attentional reorienting and social cognition, the fingerprints displayed in Figure 4-4 illustrate distinct patterns of connectivity. Whereas TPJa is mostly reached by fibres running in the AF and MdLF, TPJp receives projections via ILF, MdLF and to a limited degree EmC. These distinct connectivity patterns may support distinct functional roles in either social cognition or attentional reorienting.

Finally, we ran statistical contrasts between these five fibre's surface projection patterns (see Figure 4-4). All of these tracts showed significantly different projection areas in extended patches of cortex. Particularly UF's projected significantly more strongly to the planum polare of the temporal pole, the OFC (area 13) and vmPFC (25, 14m, 11m) when compared to CB and EmC (see Figure 4-4). ILF projected significantly more strongly to areas in the posterior STS, the middle and inferior temporal gyrus and ITG when contrasted to EmC, MdLF and UF (Figure

4-4). MdLF in turn (see Figure 4-4) showed significantly stronger projections to the parts of SMG and AG, almost the entire STG and STS but excluding the planum polare when contrasted to EmC, ILF and UF. The EmC when contrasted with CB and UF (see Figure 4-4) projected significantly more strongly to areas in the medial, inferior and lateral occipital cortex including areas V3, V5, V7 and LOC, regions in the middle temporal gyrus, the anterior superior temporal gyrus and the planum polare, the insula, the lateral OFC including areas 47/12 and the medial frontopolar cortex. The CB in turn when compared to EmC and UF (see Figure 4-4) had significantly stronger projections to the entire cingulate gyrus, dmPFC, medial parietal cortex, cuneus and precuneus cortex, the parahippocampal gyrus, the perirhinal and entorhinal cortices and the inferior temporal pole.

A contrast between EmC, AF and SLF3 shows their distinct projection patterns in the temporal and ventral frontal cortex: AF projects significantly more to the entire planum temporale and large parts of posterior STG and areas 6r, 44v and 45, 47/12m when compared to SFL3 and EmC (see Figure 4-4). EmC in turn projects significantly more to anterior STG, STS and temporopolar cortex, and to areas in the anterior frontal operculum, 47/12o, 11 and medial frontopolar cortex. SLF3 (not shown) projects significantly more strongly to SMG, AG and the most posterior part of the STG, as well as 6v, IFJ, 44d and IFS.

4.4.4 Entropy maps

Finally we used the surface-projection patterns as obtained from the white-matter parcellation to calculate the tract-projection uncertainty or “entropy” for each “patch” of the cortical surface (each surface vertex). Figure 4-5 depicts a surface entropy map. Entropy in the information-theoretic sense reflects the average amount of information contained in each part of a message received and thus characterizes the uncertainty about our source of information. We here use the measure to reflect the uncertainty for each vertex on the surface to be projected to via one given tract, i.e. regions with low entropy are reached largely by one tract

with high probability. Conversely areas with high entropy receive projections from a number of tracts with comparatively lower probabilities.

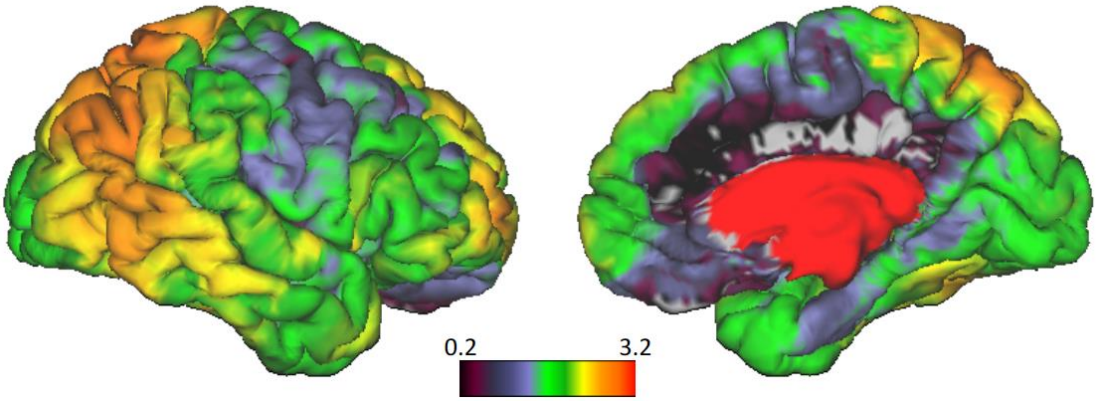


Figure 4-5 depicts a surface entropy map. Colours from black to brown to grey to green to yellow to red depict the entropy value per vertex (region on cortical surface) and thus reflect the uncertainty about being reached by any given tract.

From the entropy-map it is apparent that regions in the cingulate gyrus have comparably low levels of uncertainty. These regions are almost solely projected to via CB. Regions in vmPFC and OFC, but also somatomotor and dorsal premotor regions also have comparably low levels of uncertainty as receiving the biggest number of projections via UF and EmC in the former case and via the SLFs in the latter case. On the other hand dmPFC, vlFC and dlPFC, occipital, as well as STG and anterior inferior temporal cortex have a considerably higher level of uncertainty as most of these are reached by at least three tracts with similar probabilities (e.g. EmC, SLF1 and CB in the case of dmPFC). Finally the highest uncertainty was observed in regions where the projection-patterns of three or more fibre tracts consistently overlapped, such as in posterior inferior temporal lobe (ILF, MdLF, EmC, SLF2, SLF3), angular gyrus (ILF, MdLF, EmC), precuneus (SLF1, CB, EmC) and anterior dlPFC (SLFs, AF, EmC).

These observations are somewhat reflective of previously described “network hubs” in the human brain (315, 367). Such hubs were delineated by graph-theoretical analysis tools and are

taken to be nodes occupying a central position in the overall organization of a network, i.e. a considerable number of “direct” (not in the sense of monosynaptic but in the sense of shortest path) paths run through such hubs. Using a different imaging modality (resting-state fMRI) and a different analysis approach (graph theory, namely degree centrality in the case of (367)) such studies had consistently proposed posterior inferior temporal lobe, angular gyrus, precuneus and posterior cingulate, as well as dmPFC and dlPFC to be cortical hubs. Moreover vlFC, occipital, as well as STG and anterior temporal cortex were implicated to still have considerably high degrees of connectivity. Similar regions were implicated by our analysis to be reached by a considerable variety of different large association fibre systems. This may be surprising because we did not aim to measure the overall degree of connectedness of a given cortical region (or surface vertex) but rather its “participation” in distinct large association fibre systems.

Moreover these findings are of course limited by the fact that this study (as many of the above-mentioned graph-theoretical analyses of connectivity and network-hubs) focused on long-range association fibres but excluded commissural and projection fibres. For example entropy in the cingulate gyrus and lateral part of the primary motor cortex was comparably low. However regions in the mid-cingulate cortex and the primary motor cortex are also reached for example by the pyramidal tracts, thalamo-cortical tracts and cortex-basal-ganglia projections.

4.4.5 Connectivity-based parcellation of white matter and surface projection in monkey

We used a largely parallel approach of performing connectivity-based parcellations of coronal slices through white matter to identify the bodies of nine major longitudinal fibers. We subsequently performed tractography from each of these clusters towards the cortex to create

a projection map of each tract. The resulting white-matter association fibre bundles and cortical projection maps are summarised in Figure 4-6 and Figure 4-7). The results are qualitatively similar to what is reported in the human subjects.

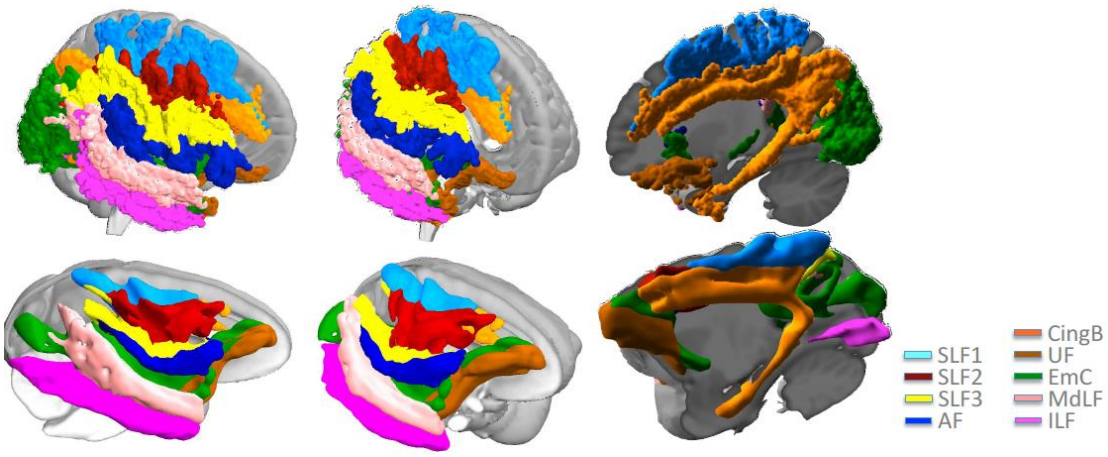


Figure 4-6: The three panels show 3D-renderings of all fronto-parietal, temporal and fronto-temporal fibre systems (SLF1 [light blue], 2 [red], 3 [yellow], and AF [dark blue], CingB [orange], EmC [green], UF [brown], MdLF [pink] and ILF [purple]) in both humans (upper panels) and monkeys (lower panels). The left two panels show a lateral view of the rendered right-hemispheric tracts, the middle panels show an anterolateral oblique view, and the right panels show a medial view with the left hemisphere removed and the right hemisphere partly removed to allow view of the medial aspects of all fibres.

Monkey SLF1 runs in the medial parietal cortex and the superior parietal lobule and in the white matter of the superior frontal cortex dorsal to the cingulate sulcus. Similar to the human results we observed a tentative side-branch connecting dorsomedial frontal areas with ventrolateral frontal cortex similar to the so-called aslant tract. SLF2 in turn runs in the vicinity of the IPS, beneath the hand area of the primary motor cortex and towards the white matter below the arcuate and principal sulci. The SLF3 derives from the IPS, IPL and the parietal operculum to course horizontally through the white matter of the IPL and inferior frontal cortex in the vicinity of the lower bank of the arcuate sulcus. AF is a distinct fibre system which can be delineated from SLF3 and arises from posterior IPL and STG in the vicinity of the most posterior aspect of the lateral fissure to run along the parietal and frontal operculum dorsal of

the external and extreme capsule and the claustrum to then course along the arcuate sulcus and the inferior bank of the principal sulcus.

As in the human the CB rand in the white matter of the cingulate gyrus and the in vicinity of the cingulate sulcus to then run curve around the splenium of the corpus callosum and track in the white matter of the ventral part of the temporal lobe, to reach inferior temporal cortex areas, the perirhinal and entorhinal cortex, as well as the amygdala. The extreme capsule on the other hand run through the occipital lobe and parts of posterior IPL, tracks though the white matter of STG and ITG to then curve dorsally and course through the white matter between claustrum and insular cortex to reach ventrolateral and medial prefrontal cortex. The UF also ran between the ventral part of the claustrum and the insular cortex directly ventral to the EmC and reached areas in the STG, the temporal pole and the vicinity of the amygdala, as well as lateral and central OFC, as well as vmPFC, perigenual ACC and medial frontopolar cortex. Finally MdLF and ILF ran similar to the human tracts in the white matter of the STG and ITG and reached from the lateral occipital cortex to the temporal pole.

Finally we replicated these fibre courses and projection patterns in one single high-resolution ex-vivo data-set. The results of this parcellation are shown in Figure 4-7 and follow a very similar qualitative pattern as those described above.

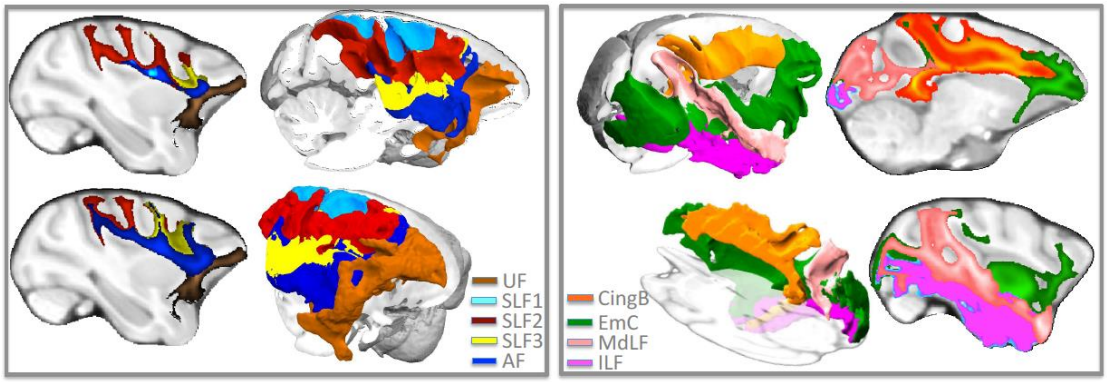


Figure 4-7 shows the results of a connectivity based parcellation of the white matter association fibres in one single high-resolution ex-vivo dataset. The left panel shows 3D renderings and sagittal slices of the four frontal-parietal fibers together with the UF. The right panel depicts the temporal and fronto-temporal fibers. Colour-coding is exactly as in previous figures.

4.4.6 Leave-one-out analysis, probabilistic white-matter atlas and "clinical" relevance

In the previous sections, we investigated the projection surfaces of each of the major longitudinal white matter fibers of the human and the macaque brain. These surfaces reflect a crucial feature of white matter connections that is often ignored in diffusion MRI tractography studies, namely which areas are connected by a particular fibre. This is precisely the information that is relevant to clinicians interested in disconnection syndromes. Although it is good to see where a white matter lesion has occurred, this is ultimately only important if this provides information about the areas that are no longer able to interact with one another. However, although the projections of fibers might be similar across subjects, there is considerable individual variability on exactly where a fibre courses. Therefore, we used the projection surfaces to create a probabilistic map that describes how likely each voxel in the white matter is to have a projection map similar to each of the nine fibers described above. We did this analysis in a "leave-one-out" manner – comparing each white-matter voxels surface projection pattern of a given subject with the surface-based projection atlases of all but this one subject (e.g. comparing subject 1's voxels to the surface atlas of subjects 2-15). This analysis was again carried out in both species.

Figure 4-8 and Figure 4-9 show a summary of the results of this analysis. The resulting probability maps resemble in an intriguing way the parcellation-derived white association tracts described earlier, although these are now based on correlating cortical surface projection patterns.

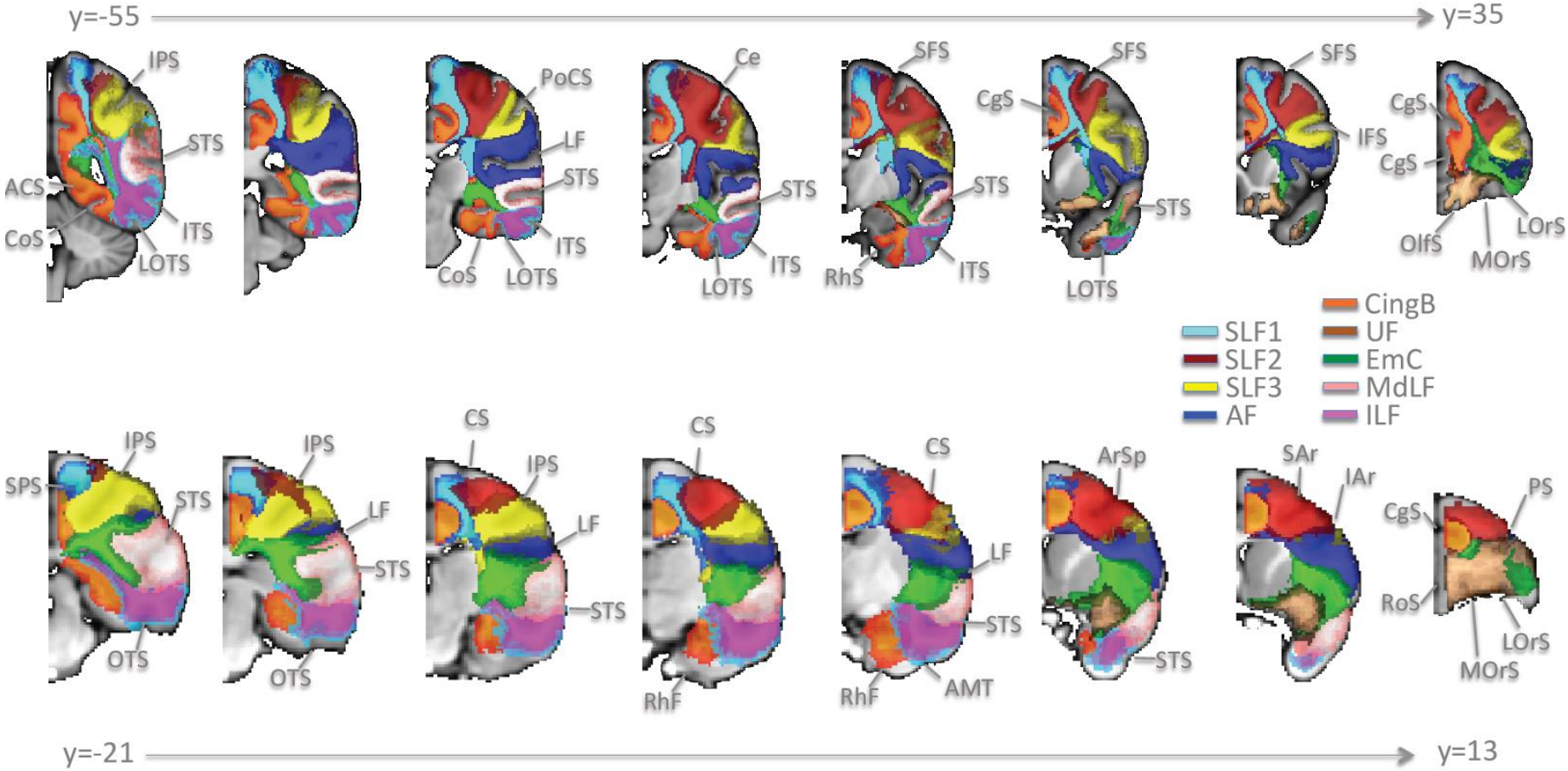

y=-55
y=35
IPS
STS
ACS
ITS
CoS
LOTS
PoCS
LF
STS
ITS
CoS
LOTS
Ce
STS
ITS
LOTS
SFS
CgS
STS
RhS
ITS
SFS
CgS
STS
LOTS
SFS
IFS
CgS
CgS
OlfS
LOrS
MOrS
SLF1
SLF2
SLF3
AF
CingB
UF
EmC
MdLF
ILF
IPS
SPS
STS
OTS
IPS
LF
STS
OTS
CS
IPS
LF
STS
CS
RhF
CS
LF
STS
RhF
AMT
ArSp
STS
SAr
IAr
CgS
PS
RoS
LOrS
MOrS
y=-21
y=13

Figure 4-8 shows coronal slices (y=-55 to y=35 for human and y=-21 to y=13) through the human and monkey white-matter atlases thresholded at 30%. The atlas was obtained by correlating surface projection patterns of each voxel in the white matter with the nine parcellation-derived surface patterns. Colour-coding as in previous figures. Abbreviations: IPS=inferior parietal sulcus; STS=superior temporal sulcus; ITS=inferior temporal sulcus; LOTS=lateral occipitotemporal sulcus; CoS=collateral sulcus; ACS=anterior calcarine sulcus; LF=lateral fissure; PoCS=postcentral sulcus; Ce=central sulcus; SFS=superior frontal sulcus; RhS=rhinal sulcus; CgS=cingulate sulcus; IFS=inferior frontal sulcus; LOrS=lateral orbital sulcus; MOrS=medial orbital sulcus; OlfS=olfactory sulcus; SPS=superior parietal sulcus; OCT=occipitotemporal sulcus; RhF=rhinal fissure; AMT=anterior mid-temporal sulcus; ArSp=arcuate sulcus spur; SAr=superior arcuate sulcus; IAr=inferior arcuate sulcus; RoS=rostral sulcus; PS=principal sulcus

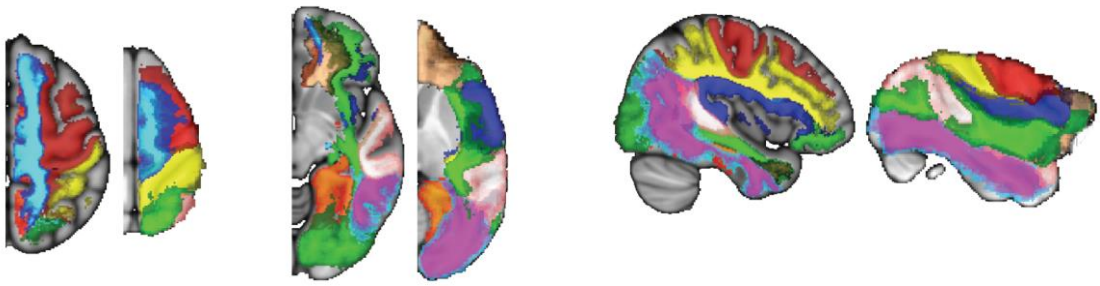


Figure 4-9 shows another set of axial and sagittal slices through the white-matter-atlases. Colour coding as in previous figures.

For example in the most anterior coronal slice (Figure 4-8, y=35) it can be seen that the white matter of the cingulate and paracingulate gyrus has high probabilities to project to the surface in pattern like CB. White matter in the inferior rostral gyrus, the straight gyrus, the medial orbital and lateral orbital gyrus resembles UF's projection pattern. White matter in the orbital part of the IFG and in the lower central white matter resembles EmC. White matter in the vicinity of IFS resembles SLF3, whereas white matter in the MFG and SFG resembles SLF2 and SLF1, respectively.

In a more posterior coronal slice (Figure 4-8, 5th from the left, y=15) it can be seen that the white matter in the CG and in the vicinity of the CS, but now also in the parahippocampal gyrus resembles CB. SFG, MFG and IFG contain white matter with a high probability of resembling the projection pattern of SLF1, 2 and 3, respectively. The white matter underneath the frontal operculum, in the vicinity of the dorsal circular insular sulcus and in the dorsal part of the

extreme capsule resembles AF, whereas white matter in the ventral extreme capsule and in the vicinity of the ventral circular insular sulcus matches EmC. White matter in close to the amygdala and underneath the piriform cortex is likely to belong to UF. White matter in the STG and around the ITS closely resembles the surface projection pattern of MdLF and ILF, respectively.

Finally, when looking at the second-most posterior slice in Figure 4-8 (y=-37) it is apparent that the white matter in the CG and precuneus, as well as in the parahippocampal and parts of the fusiform gyrus resemble cortical projections of CB. Dorsal to the precuneus in the paracentral lobule white matter matches SLF1. Lateral to that in the postcentral gyrus and SMG white matter resembles SLF2 and SLF3. In the parietal operculum and the superior temporal gyrus white matter resembles AF. White matter in STG and the vicinity of STS has a projection pattern similar to MdLF. White matter in inferior MTG, ITG and lateral fusiform gyrus resembles ILF. White matter that touches the dorso-lateral wall of the posterior horn of the lateral ventricle shares similarities in cortical projection with EmC.

The figures also present comparable results from a similar analysis carried out in a “leave-one-out” manner in the monkey. The overall qualitative pattern is largely similar to the results obtained for the human and resembles previously reported studies using tracer-injections (365).

The here-described results can be summarized as an “atlas” of the white-matter in common MNI-standard space. This atlas assigns each white-matter voxel a probability that its pattern of cortical projections resemble any one of the above delineated major association tracts. Such an atlas-tool may be helpful for neuroimaging studies to link specific patterns of cortical activation as obtained with fMRI or PET with changes in white microstructure as obtained for example with DW-MRI derived measures of FA. Moreover, the atlas can be of interest to clinicians, who want to link areas of grey and white matter lesion, atrophy or stimulation. Figure 4-10, Figure 4-11, and Figure 4-12 illustrate how such an atlas may be used.

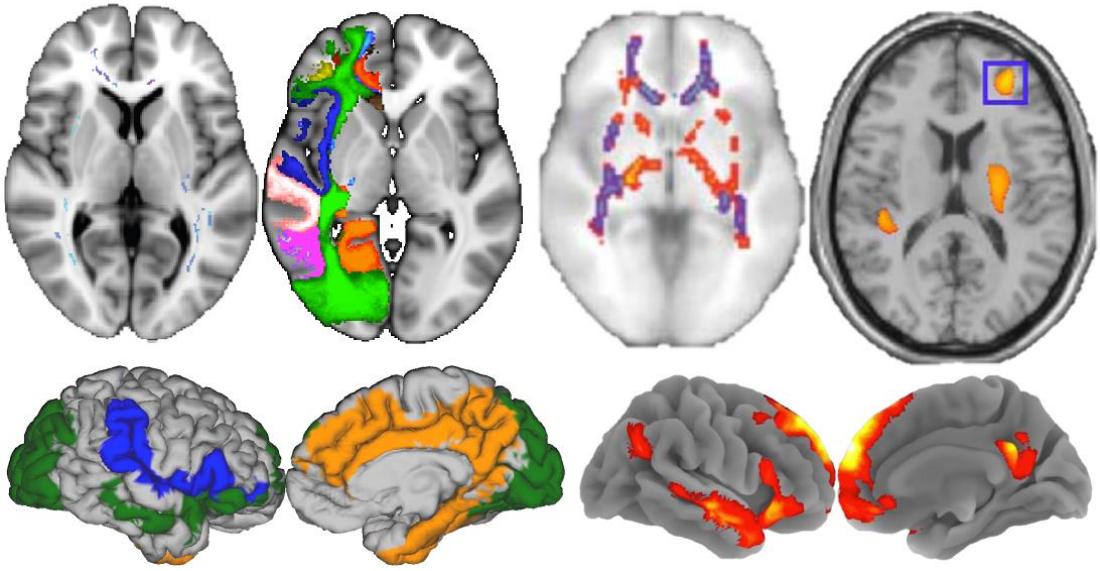


Figure 4-10: Left upper panel a summary image of changes in white-matter integrity relating to the social network size (368). The panel next to it depicts the probabilistic atlas in an axial slice at z=0. The two right panels show changes in white matter integrity reported for patients with autism (369, 370). The two lower panels depict surface projection patterns of EmC (green), AF (blue) and CB (orange) on the left and the canonical set of regions reported for social cognition (as derived from the meta-analysis on the BrainMap-database).

For example, Figure 4-10 shows in the left upper corner a summary image of changes in white-matter integrity relating to the social network size an individual lived in (368). Comparison to the probabilistic atlas depicted next to it reveals that many of the voxels of significantly greater white matter integrity in individuals with bigger social networks are found in AF, EmC and CB. Interestingly, similar white matter regions have been reported to exhibit decreased integrity in patients with autism spectrum disorder (two upper right images) (369, 370) – a neuro-developmental disorder characterized by impaired social interaction. The cortical projection patterns of AF, EmC and CB are plotted in the lower left and can be compared to a canonical set of cortical areas implicated in social cognition and theory of mind (lower left). There is a considerable amount of overlap between the projection patterns of AF, EmC and CB – the tracts that are implicated by social network size and impaired social interaction – and the

network active during social cognition task, particularly in pgACC, FPm, dmPFC, PCC, precuneus, vlPFC, STG and STS and temporopolar cortex.

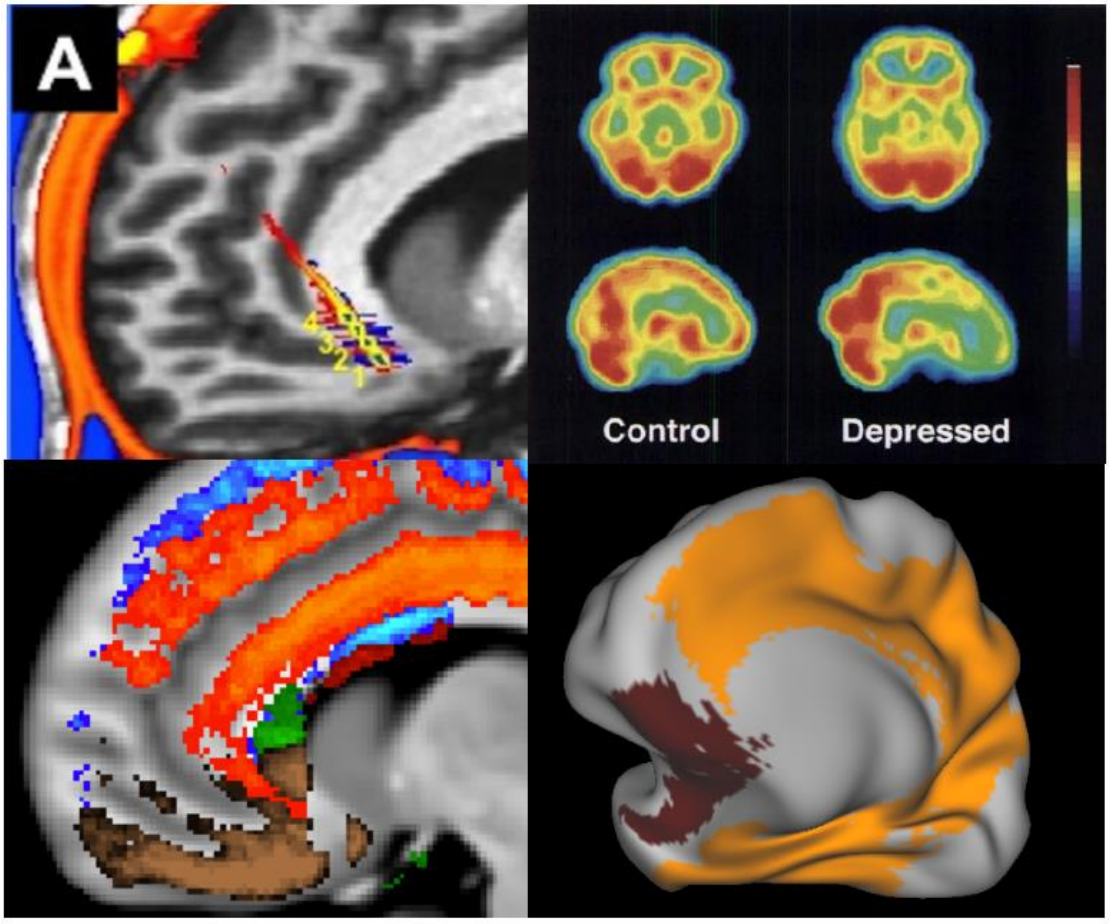


Figure 4-11 depicts in the upper left the white-matter contact locations of deep-brain-stimulation electrodes placed to treat major depressive disorder (371) and areas of hypometabolism in unipolar depression (372). A sagittal slice of the white matter atlas is depicted below together with surface projection patterns of UF and CB.

Another example of how to potentially make use of the white-matter atlas is depicted in Figure 4-11. The upper left image shows a summary of white-matter contact locations of deep-brain-stimulation electrodes close to subgenual ACC placed for the treatment of major depressive disorder (371). Comparison with the atlas depicted below reveals that electrodes are placed in the vicinity of UF and CB (and the corpus callosum and potentially to some degree the EmC). UF's and CB's cortical projection patterns are plotted on the lower right. Interestingly there is a considerable amount of overlap of these projection maps with hypo-

perfused cortical regions as measured with SPECT in patients suffering from unipolar depression, particularly in medial prefrontal and orbitofrontal cortex, cingulate cortex, anterior temporal lobe and amygdala (372).

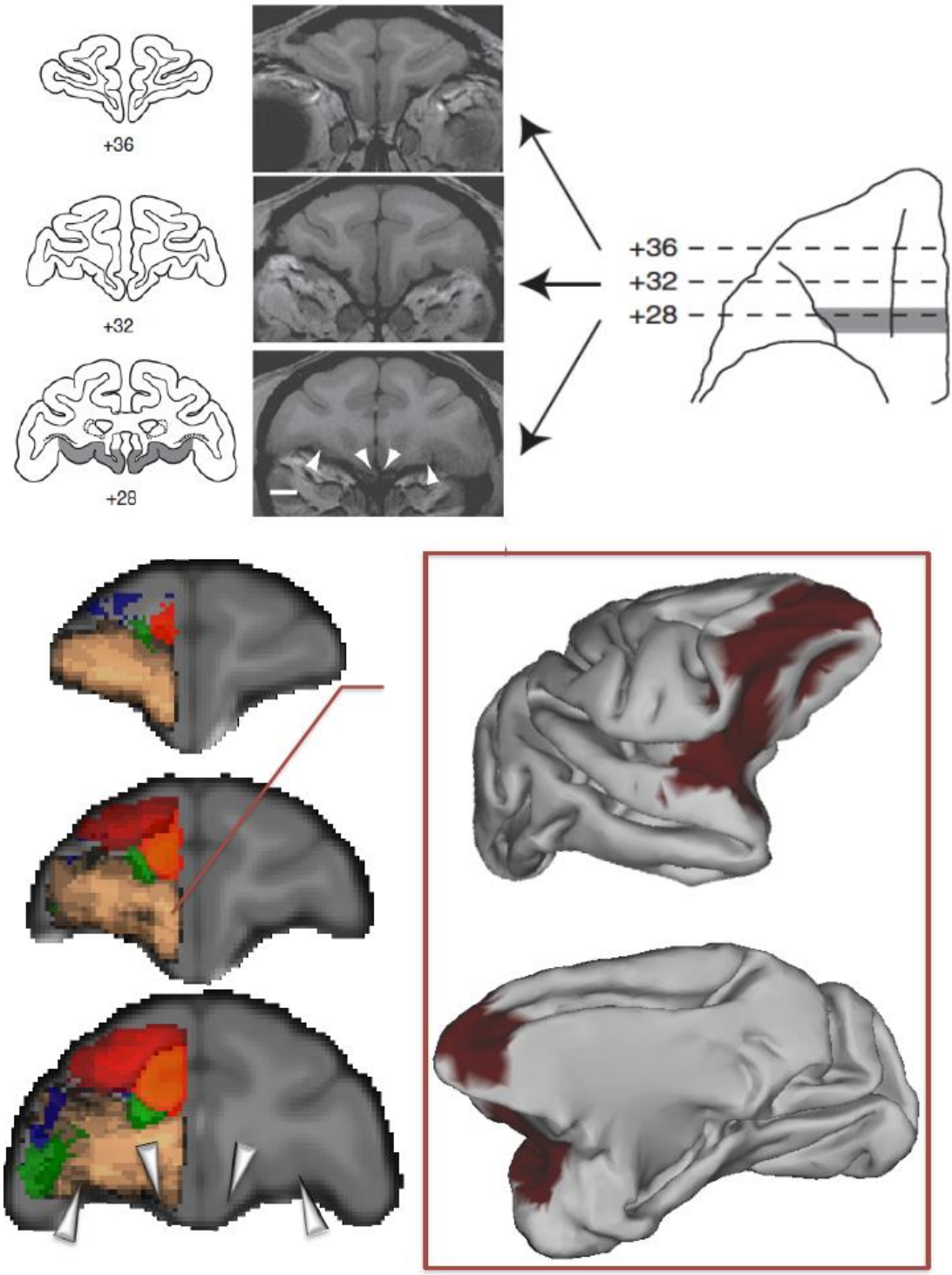


Figure 4-12 the upper part shows a small area in monkey central OFC that was lesioned using an aspiration technique (373). The lower part shows coronal slices of white-matter atlas and a surface depiction of UF.

A final example of how the atlas may be used is depicted in Figure 4-12. The upper figures shows a small area in monkey central OFC that was lesioned using an aspiration technique (373). This lesion destroyed a small strip of cortex involving portions of areas 13l, 13m, 13a and 14c. Such OFC lesions have been shown on the one hand to impair reversal learning and on the other hand to disrupt emotional responses in the presence of the normally anxiogenic objects such as moving snake stimuli. This was often taken to suggest OFC regulates emotion and enhances behavioural flexibility through inhibitory control. However this is at odds with findings that monkeys receiving even more extended bilateral but excitotoxic lesions of central and medial OFC (large parts of areas 11, 13, 14) not showing any impairment in reversal learning and emotion regulation. The difference in the two lesion-paradigms may be that excitotoxic lesions spare white matter fibre bundles whereas aspiration lesions disconnect them. The well-established behavioural effects of the comparably small strip-lesion shown in the upper part of Figure 4-12 may thus not reflect missing cortical tissue but constitute a "disconnection syndrome". The monkey-atlas depicted below suggests that such an aspiration lesion would likely destroy parts of UF. On the lower right we depict the cortical projection pattern of UF. Interestingly there is a considerable overlap between UF's cortical projections and regions implicated in reversal-learning and emotional regulation, particularly in lateral OFC, anterior insula, anterior STG and amygdala.

4.4.7 Comparison to meta-analysis derived neuroimaging maps

The previous section finished by exploring how the probabilistic association-fibre atlas may be used by researchers and physicians to link circumscribed white-matter areas with their respective cortical projection patterns. Changes in confined white-matter areas can accompany specific cognitive abilities or impairments. White-matter areas are also targeted by DBS or affected by multiple sclerosis or stroke related lesions. This final part of the current

study presents some example analyses, which illustrate how our cortical atlas of projection patterns (presented previously) may be used in combination with neuroimaging-derived cortical "activation" maps or maps of cortical atrophy. This may help to establish a formal approach to test for disconnection-syndromes or hodological disorders.

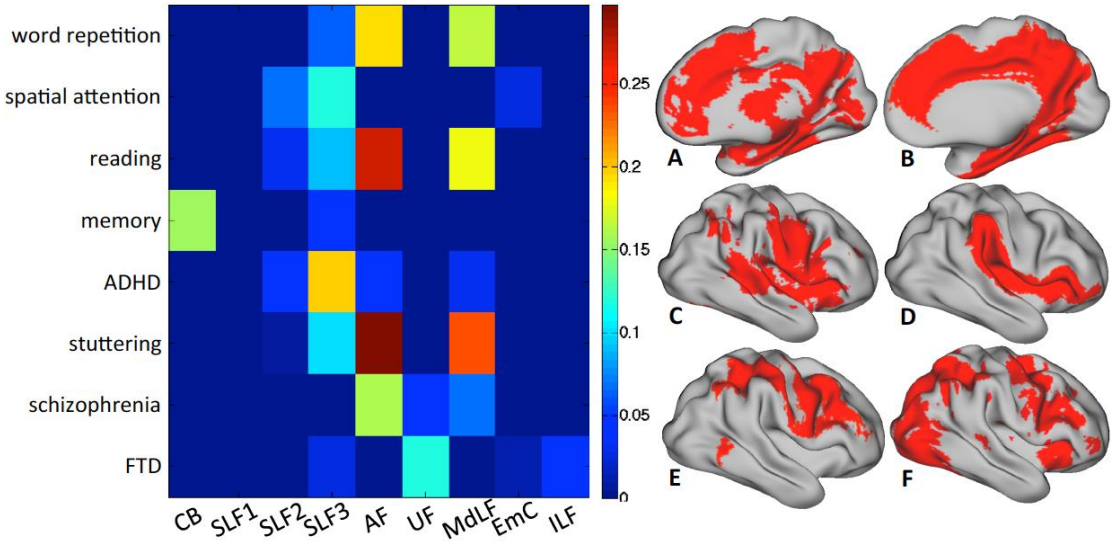


Figure 4-13: The left hand panel shows a correlation matrix summarizing spatial correlations between a set of meta-analysis derived neuroimaging-maps (vertical access) with probabilistic surface projection patterns of nine association fibres (horizontal axis). Three examples of meta-analysis-derived neuroimaging-maps and their arguably matching fibre tracts are depicted on the right ([A] memory, [B] CB, [C] reading, [D] AF, [E] SLF 2+3, [F] spatial attention).

Figure 4-13 summarizes the results of spatially correlating a set of meta-analysis derived neuroimaging-maps with probabilistic surface projection patterns of all nine association fibres in the right hemisphere. Three examples are depicted on the right part of Figure 4-13.

We first ran a meta-analysis on studies probing word-repetition using the Brainmap-database in order to obtain a canonical set of cortical regions involved in speech repetition. The inability

to repeat speech despite intact auditory comprehension and fluent speech production famously characterizes conduction aphasia, the most classical disconnection syndrome. We found that the cortical pattern of word-repetition correlated with the spatial distribution of AF's and MdLF's surface projections. This is in line with notions of conduction-aphasia being the result of a disconnection between regions in the superior temporal lobe and the vlFC.

We secondly ran a meta-analysis on studies probing spatial attentional control. Impairments in spatial attention, attentional reorienting and spatial representation are observed in syndromes like neglect. Neglect has been linked with lesions in posterior parietal, frontal, inferior temporal and lateral occipital cortex in a number of neuropsychological studies. Alternatively the location of the damage in the white matter may be considered the defining feature of neglect and the syndrome be considered the result of a disconnection. We found the cortical network for spatial attention correlated with the surface-projection pattern of SLF3 and to some degree SLF2. The network for spatial attention is plotted next to the projection pattern of SLF2 and SLF3 in Figure 4-13E and F.

Another meta-analysis was conducted on studies probing reading. Impairments in reading are one of the defining symptoms in dyslexia. Recent research had linked dyslexia to impaired connectivity between auditory cortex and left inferior frontal gyrus, despite phonetic representations in primary and secondary auditory cortices themselves being intact. Spatial correlation of the meta-analysis derived cortical map for reading was spatially correlated with projection patterns of all nine association fibres and correlated most strongly with AF and somewhat with MdLF (Figure 4-13C and D). This is consistent with previous suggestions of dyslexia being a disconnection syndrome {Boets, 2013 #2577}.

We also calculated a meta-analysis on considerable number of studies on memory-function. Impairments in memory are observed as early symptoms in a range of different neurodegenerative diseases such as Alzheimer's disease. Recent studies have linked the pattern of propagation of their microscopic correlates (e.g. amyloid deposits and

neurofibrillary tangles in the case of Alzheimer's disease) across the brain, as well as the associated atrophy patterns with long-range connections. Such studies implied that neurodegenerative diseases target specific large-scale human brain networks. Correlating the cortical pattern revealed by a meta-analysis of studies on memory-function with our fibre-projection maps revealed a high correlation with CB (Figure 4-13A and B). The respective meta-analysis based and projection map are depicted on the right side of the same figure. This is consistent with CB connecting to parahippocampal areas, the entorhinal cortex and hippocampal formation – areas known to be crucial for memory formation (e.g. as part of the classic Papez circuit). Moreover these are areas where the earliest histological changes are observed in several neurodegenerative diseases.

In a second step we compared activity patterns associated with a number of psychiatric and neurological conditions with the association-fibre related projection maps. We calculated activation likelihood estimation maps (ALE-maps) on functional imaging studies involving patients with ADHD and developmental stuttering. ADHD is a neurodevelopmental psychiatric disorder and has been linked to changes in fronto-parietal networks. We found that the associated activity pattern spatially correlated with the projection pattern of SLF3. Developmental stuttering is a speech disorder associated with dysfluencies such as hesitations, prolongations and repetitions of speech sounds. Recent studies have implicated underactivity in the ventral premotor, opercular and sensorimotor cortex, as well as auditory areas to occur in these patients. We found spatial correlations of associated meta-analysis results with AF's and to a limited degree MdLF's and SFL3's projection patterns.

Finally we turned to patterns of changes in grey-matter density as observed with schizophrenia and fronto-temporal dementia. Schizophrenia has been conceptualized as a disorder of connectivity for several decades particularly in relation to connections between frontal and temporal areas. We found that patterns of grey-matter change associated with schizophrenia matched AF's projection pattern. This is consistent particularly with the changes observed for

auditory verbal hallucinations – one of the most common symptoms in schizophrenia. Fronto-temporal dementia (FTD) patterns of grey matter change matched most closely UF's projection. This may be taken to be consistent with the profound semantic memory loss in both the verbal and non-verbal.

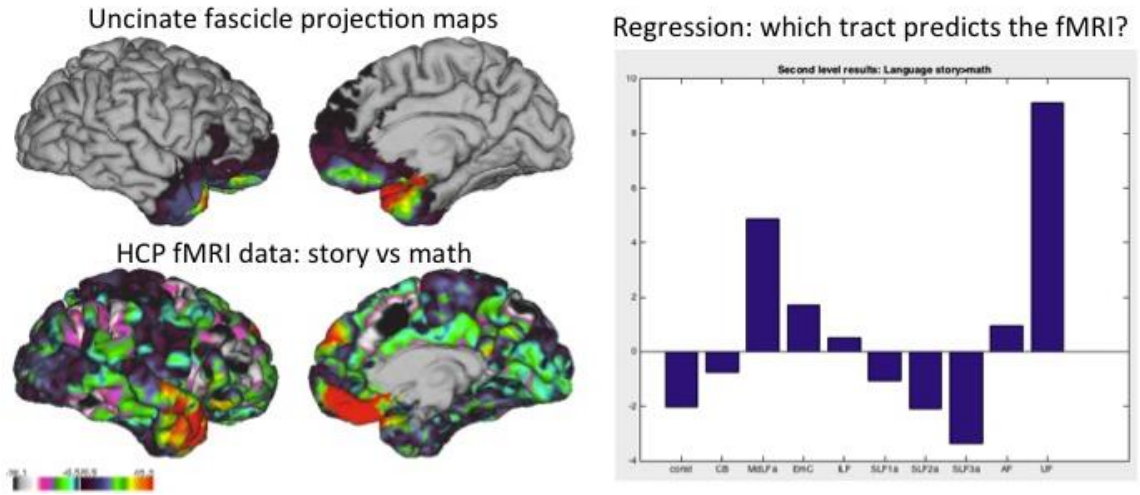


Figure 4-14: The left upper panel shows the surface projection pattern of the uncinate fascicle. The left lower panel shows a group-map from the semantic language (story) task contrasted with the math-task. The right panel shows the result of the regression analysis with average beta-parameters for each tract's surface distribution predicting the fMRI-activity pattern on the individual subject's level.

To explore this further and to validate a final way of linking neuroimaging-derived activity patterns with tract-derived surface projection patterns we conducted a generalized linear model analysis on a single-subject level using the subject-specific fMRI-derived activity patterns from the HCP dataset and the same subject's parcellation derived tract-projection results. This analysis tried to investigate whether any subject-specific activity pattern relating to a semantic language task could be predicted by the projection patterns of its specific white matter association fibres. On the group-level we found that UF's and to a smaller degree MdLF's projection pattern predicted fMRI-activity patterns for the semantic language task (Figure 4-14). This may be linked to the previously described results for FTD.

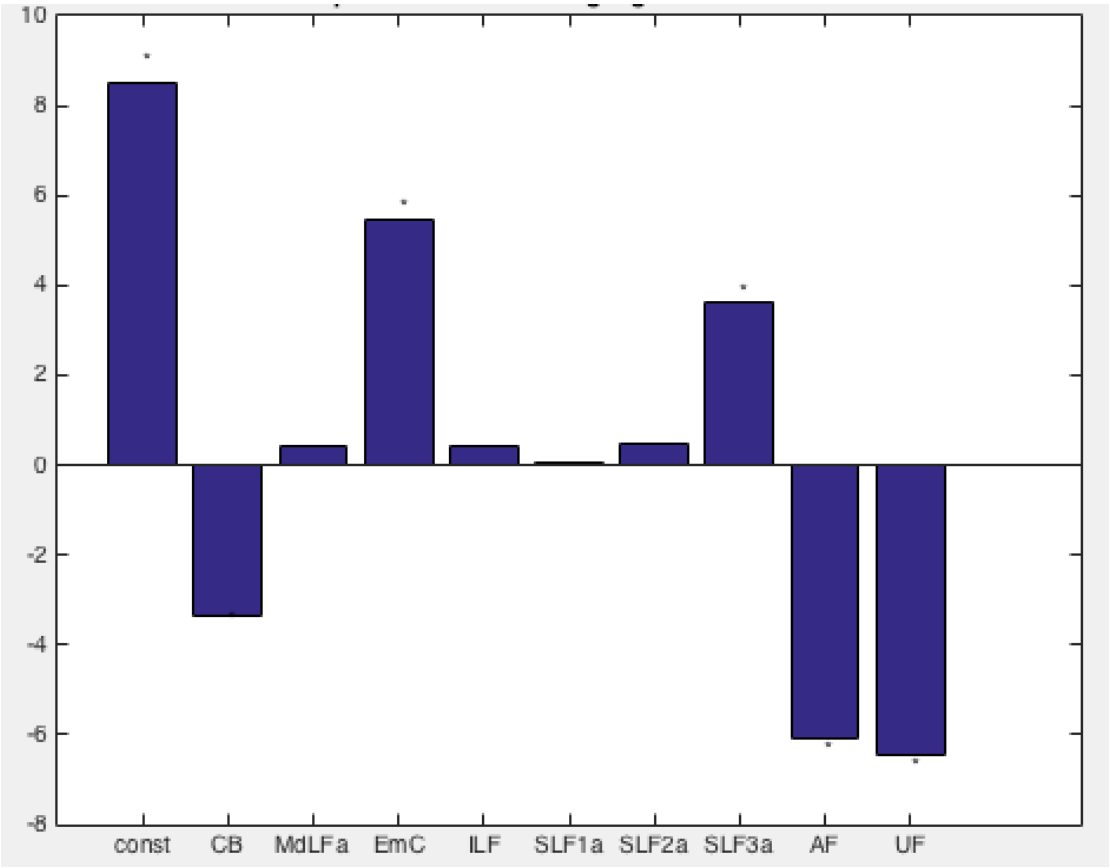


Figure 4-15: shows the result of a similar regression analysis attempting to predict fMRI-activity pattern in the social (theory of mind) task. The graph shows average beta-parameters for each tract's surface distribution predicting the fMRI-activity pattern on the individual subject's level.

A similar analysis was conducted on fMRI-derived activation patterns relating to a theory of mind task revealed that EmC's and to a smaller degree SLF3's projection patterns predicted activity distribution on the cortical surface (Figure 4-15). This is consistent with previously described studies on social cognition, social network size and autism, which were also linked to EmC.

4.5 Discussion

In this chapter I presented a novel technique of dividing the white matter into the major association fibre systems based on similarities in cortical projection patterns. We ran probabilistic tractography from all white-matter voxels in several coronal slices and used k-means clustering to grouped together white-matter voxels with similar cortical projection patterns. In doing this we were able to reconstruct nine major association fibres of the human cerebral cortex and their respective pattern of grey-matter projection.

As a result of this analysis we obtained two probabilistic atlases: One atlas reflects the topography of all major white matter association fibre systems. The other atlas is a probabilistic representation of (surface) patterns of cortical projections for each of these association fibre systems. A largely parallel approach was used in the monkey obtaining qualitatively similar results.

I suggested previously that this approach and the resulting atlases may be used in several different ways and gave examples of some of them. An atlas of the topography of all major white matter association fibre systems may help to identify clusters of changed white-matter integrity as obtained from FA-based VBM or TBSS-style analyses. I previously gave the example of increased white-matter integrity between the insular cortex and the claustrum in people who lived in bigger social groups (368). Comparison with the white-matter atlas allows to consistently label such clusters of significant correlation or difference – in this particular as EmC. Moreover the second atlas of probabilistic cortical projection patterns may now allow comparing EmC's probabilistic surface-projection pattern with fMRI-derived "activity" patterns linked to social cognition. In short we hope a probabilistic atlas of fibre systems in the white matter together with their probabilistic cortical projection patterns may facilitate linking "grey-matter" and "white matter" imaging data more effectively. This link between grey matter activation and white matter change is most often also link of different temporal scales.

Imaging studies using for example PET and fMRI often link short-term changes in neural "activity" in specific cortical regions with defined cognitive processes (374, 375), whereas studies linking between-subject differences of white-matter microstructure as measured with DW-MRI tend to capture more stable (trait-like) behavioural differences or intermediate-term learning effects (142). One may be able to use our linked grey and white matter atlases in order to reconcile these patterns of transient grey-matter activation with stable patterns of white matter differences across subjects to obtain an integrated view of cognitive processing. Future research may try to make use of this anatomical information when comparing white-matter microstructure across subjects. TBSS for example although using a tract-based approach to match white matter voxels across subjects is still largely anatomically blind. It is possible to imagine how a "skeletonisation" of the current white-matter atlas affords an anatomically informed variant of TBSS, where voxel-wise comparisons of specific anatomically defined white-matter fibre bundles can be carried out.

Moreover I showed that the probabilistic surface-projection patterns can be used to calculate a surface-entropy map, reflective of how many distinct white-matter tracts reach any-one particular surface-vertex. This entropy bore some resemblance of previously described maps of "network hubs", even though we here did not measure the overall degree of connectedness of a given cortical region but rather its "participation" in distinct large association fibre systems. Nevertheless our analysis implicated largely similar regions, including the posterior inferior temporal lobe, angular gyrus, precuneus, dmPFC and dlPFC, as well as vlFC, posterior STG and anterior temporal cortex. This may be linked back to the previously-mentioned Flechsig's-rule, which held that long-range association fibres connected secondary and tertiary association cortices but not primary sensory and motor cortices. Indeed in the entropy map it is the higher-order association cortices in posterior parietal, inferior temporal and prefrontal cortex, which have the highest entropy-values, indicating participation in a bigger range of different association fibres. Primary motor and sensory areas on the other hand have

considerably lower entropy. This may indicate that these higher-order association cortices or "hubs" receive connections from a broader range of long-range fibres and potentially play a role in information integration and distribution. These long-range connections may be critical in giving an otherwise fairly modular organized brain its small-world-like architecture.

I also showed that this methodological approach can be used in both human and macaque monkey brains. We did not conduct any formal between-species comparisons in this study. Rather the goal of the current study was to establish this overall approach in high-resolution human in-vivo DW-MRI data. I then went on to demonstrate that it can also be applied to lower-resolution in-vivo monkey data and even high-resolution ex-vivo monkey data. Nevertheless I hope this technique may offer a data-driven approach for studying connections and network architecture across different species. Future research will attempt to use it for the investigation of cross-species differences in (1) the size of a specific tract (the diameter of all its bundles), (2) the strength of a given white matter bundle (e.g. white matter integrity as measured with FA, myelination or average axon diameter) and (3) the size and distribution of a fibres projection-area. Moreover it may allow for the discovery of new connections or even new connectional principles. Moreover there has been considerable interest in hemispheric asymmetries in brain anatomy and structural connectivity in humans and other primates (376-378). Future research may be able to use this technique in order to disentangle whether they are asymmetries of (1) tract's size, (2) white matter integrity or (3) distribution of cortical projection or any combination of these three.

Finally, I proposed that this methodological approach might offer ways to investigate developmental and aging trajectories of white matter tracts and re-evaluate "disconnectional" or hodological accounts of brain disease. I exemplified one way to achieve this via the triangulation between (A) behavioural abnormality, (B) brain pathology and the (C) probabilistic maps of connections and projections.

One of the examples I discussed previously related to the network of white-matter fibres implicated in social cognition (368). A largely similar network was reported to be changed in patients with autism spectrum disorder (369, 370). One of the core-symptoms of autism is impaired social interaction (379) and the cortical projection patterns of the association fibres that were implicated in autism overlapped in large areas with canonical activity patterns for social cognition (52). Another example was the site of DBS in major depressive disorder. DBS in the white matter beneath sub-genual ACC was reported to lie in the vicinity of UF and CB in responders (371). There is a considerable overlap between UF's and CB's cortical projection pattern with known areas of changed metabolism in major depression (372). I also showed that a similar approach may be used in case of monkey white matter lesions beneath the OFC (373).

To further illustrate the feasibility of this approach I presented results from spatial correlations of meta-analysis derived activity patterns reported for speech repetition (an ability impaired in conduction aphasia), spatial attention (hemispatial neglect), reading (dyslexia) and memory (a range of different forms of dementia). In most cases we were found considerable overlaps between cortical networks supporting these tasks and surface-projection patterns of specific association fibres. I finally linked patterns of changes in activity or grey-matter density in patients with ADHD, stuttering, schizophrenia and FTD to the distribution of surface-projections of the association fibres.

The results of these exploratory analyses should not be taken to suggest that all of the above mentioned disease and syndromes are the results of disconnection, although they show that hodological mechanisms would be consistent with the patterns observed. Future research will try to expand on these preliminary findings. As suggested a full "triangulation" would require a complete set of (A) specific behavioural impairment (potentially in combination with network supporting behaviour in healthy brain, e.g. pattern for spatial attention in Figure 4-13E and F), (B) brain pathology (from PET-derived hypo-metabolism maps, atrophy maps etc.) and the (C)

probabilistic maps of connections and projections. In almost all cases I failed to present this complete set of data. Moreover future research will attempt to broaden this perspective by looking at an even bigger range of disorders, potentially including Dejerine's syndrome (also known as pure word blindness) – another classic disconnection syndrome (102, 380, 381), pain asymbolia (a condition in which pain is experienced without unpleasantness) – which may be linked with lesions in CB or EmC (381), visual hypoemotionality (a condition of emotional hyporeactivity to visual stimuli) – which may result from the disconnection of fibres in ILF (382) and primary progressive aphasia (a disease characterized by progressive language and speech difficulties) – which has been linked to EmC and UF (383).

Future research may also try to expand this general methodological approach of linking cortical networks of brain-activity, brain pathology and association fibre projection in order to be able to predict behavioural impairments resulting from precise white-matter lesions. The "art" of linking behavioural symptoms with cortical lesions still lies at the core of clinical neurology (384). The idea of linking disconnection with behavioural anomaly, although dominating neurological localisation in the peripheral nervous system and the spinal cord, may still await further exploration in the case of cortical association and commissural fibre systems.

Similar to the above-mentioned idea of co-development, aging and degeneration being shaped by major association fibre-systems one may speculate that these tracts are a major determinant for flexible cortical reorganisation. Previous studies have noted that cortical lesions or interruption of processing as caused by TMS may be compensated for other cortical regions sustaining "normal" behaviour (114, 129, 385). For example Hartwigsen et al (129) showed that left PMd and SMG complement each other during rapid action reprogramming. It is possible that regions that are tightly structurally connected and share patterns of connectivity to third regions (as shown in the previous Chapter 3) are playing similar functional roles in complex cognition and one of them may thus be able to partially compensate for

interruption of processing in the other. This is interesting as we previously showed that non-invasive stimulation protocols can associative plasticity in corticocortical pathways (386, 387).

Finally it is important to point out that the current study being carried out on DW-MRI data suffers from all its severe limitations. Even though we used high-resolution DW-MRI data here, it is still the case that each voxel in a DW-MRI dataset contains many thousands of axons and modelling fibre orientations particularly in areas of crossing fibres can be a considerable challenge (199). Moreover DW-MRI-based tractography is blind to the directionality of connections and may be susceptible to spurious tracking. Tracer injection studies should therefore still be considered the gold-standard, as they provide definitive evidence for existence of synaptic connections between brain regions and their directionality. In addition it needs to be pointed out that the parcellation method has limitations in itself and a more fine-grained parcellation of certain tracts might be possible, as has previously been demonstrated using tracer data (388).

Finally caution should be noted for cases where I linked neuroimaging-derived (e.g. fMRI or PET) co-activation maps with association-fibre projection patterns. Whereas it is the case that structural connectivity is a major determinant of functional correlation and coactivation (147, 167) co-activation has also been reported to be influenced by synaptic efficacy and regional GABA-ergic tone (387, 389). Therefore changes in coupling as for example observed in MCI and Alzheimer's disorder in the DMN (390) may thus not only relate to disintegration of EmC and CB but also to loss of synaptic efficacy or inhibitory tone. Loss of regional GABA-ergic inhibitory tone in somatosensory cortex was reported in older people and related to decreased tactile acuity. As previously discussed inhibitory tone may play a key role for the specific types of computations carried out by regions in the DMN (180) and the general trend towards loss of inhibitory tone in aging and disease may thus render DMN-areas particularly vulnerable in a wide range of disorders (315, 325, 391) without primarily targeting structural connections.

5 General Discussion

In this thesis every experimental chapter had its own discussion of results and their implications. In this brief general discussion I will take up some of the more general themes and motifs mentioned in the first chapter and turn to them again in light of the findings presented. Particularly I will try to expand on some seeming inconsistencies and breaking points. I will try to briefly reiterate the ideas and points that seem at odds with one another and then explain how they may be reconciled.

Two general ideas have led much of the introduction: (1) Neuroeconomic and neuro-ecological views of decision making and (2) modular versus connectionist views of brain function.

Firstly the theory of decision-making and the neuroscience of choice were discussed and two alleged "turns" in decision neuroscience were mentioned. Neuroeconomics taught us that decisions are not made reflex-like, but with reference to complex representations of the structure of the world (including specific aspects such as expected value, risk, ambiguity). The neuro-ecological turn on the other hand proposed (among other things) that there might be several different decision-making systems most likely working in parallel. Each of these parallel systems evolved during the course of phylogeny and adapted during ontogeny and their specific mode of action may reflect this process of evolution and adaptation. I finished by suggesting that inter-individual variability in choice behaviour and neural representation can partly be understood as arising from adaptation to environmental variables. This may lead to a different understanding of some psychiatric disorders – rather than arising from "broken" brains they might be the result of (mal-) adaptation to potentially extreme individual histories and environments.

Secondly I contrasted modular and connectionist views of brain structure, function and dysfunction. I suggested that connections and interactions of brain regions may be critically involved in achieving complex cognitive functions and that many intricate computations are

distributed in networks of neurons and even brain regions. I specifically stressed the importance of long-range connections between “higher order” brain regions (regions removed from primary sensory input or motoric output) for the integration of information and potentially even for generating a coherent percept of the world.

Some of these ideas may at first seem somewhat incompatible and irreconcilable:

(A) It may initially seem surprising to invoke a stark contrast between modular and connectionist views of brain structure and function and then present studies, which parcellate cortex into distinct modules based on their connections to the rest of the brain. Previous anatomical studies that had tried to establish distinct cortical modules focused on localised microscopic or biochemical aspects such as cytoarchitecture or receptor-density (113, 283, 304, 392). Regional variation in such features was taken to support structural and functional (e.g. computational) modularity. In contrast turning to connections as the defining features of modules could be perceived as counterintuitive. (B) It appears even more astonishing to attempt a parcellation of white-matter into distinct “units” of major association fibre systems using distributed cortical projection patterns. One may be tempted to ask whether such an approach falls into the modular or connectionist school of thought. This is particularly striking given that many of the long-range association fibres seem to connect higher-order association cortex. Different connectivity patterns of neighbouring higher-order association cortex (e.g. in the case of FPm and FPl) was driving the sub-division of a bigger patches of cortex into distinct subregions. At the same time these long-range connections where discussed as allowing for distributed processing and information integration between these higher-level areas {Jbabdi, 2013 #2588}. This is in line with ideas that processing occurs in specialised modules closer to primary sensory and motor areas (e.g. “colour module” in V4 and motion module in V5). However “higher-order” integrative processing is more distributed in networks connected via long-range association fibres. Are the parcellations of mostly higher-order association areas (presented in chapters 2, 3) incompatible with ideas of distributed processing in such higher

order regions connected via long-range association tracts (chapter 4)? (C) Moreover the second and third chapter of the thesis convey the idea of distinct cortical regions with specific patterns of connections constraining inputs and outputs and potentially linking to specific cyto-architecture and computations. However this could be taken to be at odds with the previusly mentioned contextual effects on cognitive processing (see specifically chapter 1.1). Particularly with reference to the "ecological" turn in decision neuroscience I had mentioned that recent research emphasises the influence of contextual variables on decision-making – for example irrelevant and non-choosable options influence valuation and value-comparison. Similarly the number of options or the social context affects value-representation in the brain and subsequent choice behaviour. This seems incompatible with modular views of decision-making and value-comparison and their underlying neural correlates. If a region is a distinct entity with specific cytoarchitecture, specific computational capabilities and a specific set of connections constraining inputs and outputs to match its computational role, why do "irrelevant" contextual variables have an effect on processing? How would they even reach the region if connections were constrained and tailored to a region's computational role? (D) Finally a related question relates to the phylogenetic and ontogenetic adaptation of brain-function to environmental challenges. They could be perceived to be at odds with the notion of distinct and specialised modules. In the thesis I have argued for medial prefrontal and cingulate cortex as well as Broca's area and vlFC being dividable into specific sub-regions based on connectivity. Moreover I have argued these sub-regions may link to functionally relevant sub-units. But is this reconcilable with the notion of adaptation of brain structure and function? Can a region in vlFC or ACC be highly specific in terms of cytoarchitecture, connectivity and computational function, but still adapt to various different environments and evolve to support language and reading or complex long-term financial investment decisions? Does the "recycling" account put forward several times in the thesis, which suggests that complex cognitive abilities re-use an old neural apparatus match with the same regions being

specialised modules with suitable microstructure and connectivity? Either regions are highly specialised and cannot easily be recycled for rather different cognitive abilities or cortical regions are less distinct and processing occurs in diffuse patches of cortex which are not clearly dividable into specific sub-region with distinct functional roles.

I will obviously not be able to address all of these puzzles in great detail here. I will however attempt to briefly outline some thoughts that may proof useful ingredients for a thorough discussion of these problems.

Although this thesis has a particular emphasis on connectivity and co-activation between brain areas it is at the same time tightly related to views of brain function that emphasis structural and functional specialisation in circumscribed patches of cortex {Jbabdi, 2012 #2587}. A parcellation of medial prefontal, OFC and vlFC draws much of its relevance from the belief that different sub-regions have different functional properties. In the case of the white matter-parcellation some of the links with disconnection-syndromes fundamentally imply that even specific connections have certain functional roles and are in that sense somewhat "modular". Such a notion of "modular" connections may seem rather strange. This is particularly the case if one underlies a classic "Fodorian" view of modularity, in which modules are reflex-like, hardwired devices that process narrow types of information in highly stereotyped ways (393). Such a "narrow" view of modular organisation may to some degree be plausible for "peripheral" systems. It is conceivable that visual motion processing in V5/MT is largely reflex-like and highly stereotyped. But several areas discussed in this thesis arguably belong to higher order association cortex and long-range association fibres connect preferably such areas (381). Some researchers have argued for a modular organisation of lower level processing but doubted that higher order association cortex can be subdivided into distinct modules (394). They question that higher level or "central" processes such as reasoning, inference, judgment, semantic processing, grammar and decision making can be generated by inflexible, stereotyped and predetermined computational processes (395). "Narrow" and stereotyped

modular processing is sometimes assumed to be characterised by *access specificity* and *processing specificity (396)*; i.e. only narrow access to very specific types of information and very specific computational processes. The diverse long-range connections observed here for language and decision-making areas in vlFC and mPFC seem at odds with this notion and imply that these computational processes have access to a wider array of information and are thus non-modular.

Broad information access (indirectly measured with our connectivity-profiles) may thus speak against modularity in the regions investigated here (vlFC, mPFC, OFC). Narrow-input modules operate only at the early stages of information processing (e.g., in perception), but it is difficult to see a role for modules at "higher" levels of processing, where information from diverse sources (e.g. via long-range associative connectivity) needs to be integrated. This in turn would question the purpose of dividing association cortex into sub-region after all.

On the other hand several higher-order cognitive functions have been proposed to be modular including language (393) theory of mind (397); decision making (47), spatial orientation (398); number (399); and metacognition (400). It is conceivable that some cognitive mechanisms have access to large and diverse amounts of information but only process information that meets its input criteria. Higher order systems could have wide access but narrow processing criteria. For example systems that generate inferences about the value of items might be sensitive only to information represented in valence format, though information meeting this criterion could be present in many representational systems in the brain (401). This would not invalidate the idea of modularity or render sub-divisions of cortical areas superfluous. For example, one can easily imagine "higher-order" modules that have access to large swaths of central knowledge stores but process information in specialized ways. One example are internet search engines – they are highly specialized to suit one particular purpose but have access to the entire Internet (402). It is possible that processing specificity and access specificity are largely independent features of modules. Higher order association cortices may

be less constrained in their access to information (via a variety of long-range association fibre systems) compared to earlier sensory and motor areas but may on the other hand be more constrained in terms of processing. Future research will determine whether cortical distribution of processing specificity / processing properties resemble maps of access specificity (see Figure 5-1) or are completely independent of them.

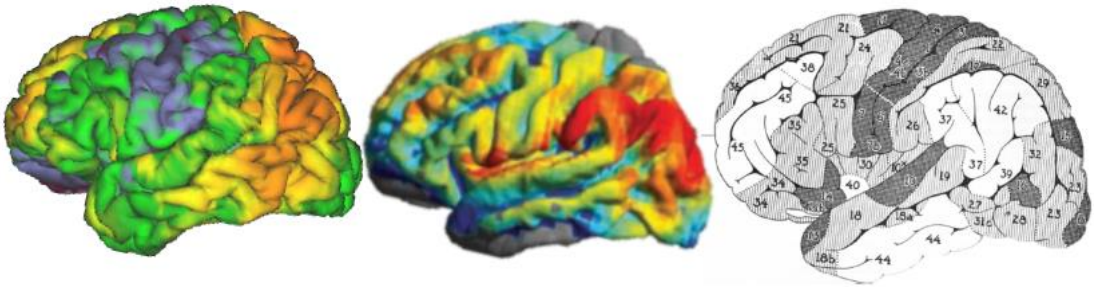


Figure 5-1 depicts three different whole-brain maps: The left panel depicts the previously discussed surface entropy map, which is reflective of the number of distinct major association fibres reaching a given part of the cortex. The middle figure depicts a network centrality map derived from (403) with more central regions ("hubs") in warm colours. The right figure depicts a map based on Flechsig's myeloarchitectonics with primordial areas (dark-grey; overlapping mostly with primary motor and sensory areas); intermediate areas (light grey / vertical lines, parasensory association cortex); and terminal areas (white, association areas for other association areas). There are some overall similarities between the three different maps. In some ways they could be thought of as depicting access-specificity/generality of a given cortical region but may be largely independent of processing properties or processing-specificity/generality.

Moreover one can speculate that these two aspects of modularity – processing specificity and access specificity – are subject to independent evolutionary change and life-time neuroplasticity. It is possible that modules "adapt" either in their processing pattern or in their connectivity and access to information to long-term environmental variables or evolutionary pressures. This is compatible with studies which report circumscribed grey matter changes in association cortex as well as changes in functional connectivity between regions with long-term exposure to specific social variables (83, 139). Modules may be "recycled" for any one of these two reasons or their combination – because they have a specific pattern of information-

access or a specific computational property. Short-term adaptation of a module may lead to the module being able to process slightly different types of inputs than it was initially "designed" by natural selection to process. For example FFA may have evolved to recognize faces of conspecifics because of their great relevance for survival. But FFA can adapt to exert similar types of computations on a broader set of inputs (404). It is conceivable that areas with more access generality (broader access to information, e.g. association cortex) are faster to adapt or evolve to process alternative inputs compared to areas with narrower information access.

Another argument in favour of non-modular processing derives from the following notion: If information was to be processed in highly specialised modules, then there must be meta-modules to distribute information accordingly. For example if it is true that different modules in vmPFC are involved in value comparison of primary reinforcers on the one hand (e.g. foods in posterior vmPFC) and more abstract secondary reinforcers on the other hand (e.g. money, coupons or trinkets in anterior vmPFC) (254) then there needs to be a process of allocating information to these different modules. At least this process would be more domain-general and hence not have the narrow properties of a "module".

However it is not clear that this needs to be the case. Some modular and specialized processing systems do not route information along prespecified channels. For example the endocrine system clearly consists of modules with diverse functions and diverse processing criteria (different receptors and signal transduction cascades for different hormones and transmitters) (395, 396). However they often all have access to a single common pool of inputs (a rich mix of hormones and other molecules) and achieve specialized processing, because receptors have a highly specific recognition site that is capable of selecting its own appropriate inputs ("lock and key" metaphor). This means that each module is sensitive only to its specific inputs and no "metamodule," or routing system, is necessary. Similarly eyes and ears are

exposed to both light and sound, but eyes process only light, and ears process only sound (395).

Another seeming inconsistency that arose from the thesis is that between highly specialised and distinct areas in prefrontal cortex (as proposed particularly in chapters 2 and 3) and the contextual influences on their processing (as discussed in chapter 1). The influence of "irrelevant" contextual information on processing could be taken as evidence against reflex-like and stable modular computation and in favour of flexible, open, general-purpose systems particularly in prefrontal areas. A sub-division of the frontal lobe may thus be only of anatomical but not of functional relevance.

However the notion of areas processing information in a stereotyped way impenetrable to other contextual information is shaped ideas about early sensory systems. For these interactivity might be relatively unnecessary or indeed detrimental. For systems involved in higher order processing, however, such as reasoning, judgment, planning and decision making, there is every reason to expect interactivity and the integration of information from multiple sources. The functional demands on these "higher order systems" would in fact require penetrability of processing.

Rather than conceptualising contextual influences, as external and somewhat "unintended" effects on proper processing in an encapuslated module they should be understood as relevant inputs to the system. The distinction between inputs to processing and outside influences on processing is arbitrary and they could well both be considered inputs to the computation of a system. I have already suggested that broad informational input does not preclude computational specificity in dedicated anatomical sub-regions. As mentioned natural selection may actually favour some modules that integrate from a broad range of information because they would allow for adaptation of behaviour to environmental contexts. This is fully in line with some of the ideas discussed in the first chapter of the thesis – that the influence of seemingly irrelevant contextual information on valuation and decision-making should not be

considered a fallacy but rather be understood as an adaptive feature of human choice behaviour.

Finally in the introduction I had argued in favour of flexible changes in brain structure and behaviour in response to long-term environmental manipulations and challenges. I even argued that some psychiatric disorders should be understood as adaptations of brains to (potentially extreme) circumstances rather than brain damage. However is such a flexible account of cognitive function and underlying brain structure compatible with the idea of circumscribed, highly specialised and evolutionarily conserved brain regions? On the one hand I am arguing for a flexible adaptation of higher-order cognitive processes to specific environmental histories and challenges. On the other hand I discuss fine-grained sub-divisions of cortical areas with consistent profiles of connectivity even across species that diverged about 25 million years ago. How can reliably identifiable and highly conserved brain regions allow humans to handle the myriad of novel challenges of their every-day life?

Evidently humans face and are able to solve challenges their ancestors never faced. They are for example able to read texts in rapid speed. They make long-term investment decisions on the stock market. They are able to use highly complex tools such as robotic surgical systems, they pilot aircrafts and they build skyscrapers. It seems implausible that these challenging and novel tasks should arise from a limited number of specialised modules that are preserved across monkeys and humans.

The here presented high degree of connectional similarity between human and monkey frontal lobe areas precludes any specialised chess, reading or cycling modules. But it does not necessarily rule out that novel cognitive abilities such as playing chess, reading and cycling rely on modular processing in circumscribed cortical circuits. It may only imply that areas and modules cannot easily be identified with their cognitive domain or content. In fact I suggested in chapters 2 and 3 that regions implicated in language and strategic decision-making tasks may in fact not be language and decision-making areas after all. Nevertheless they could still

be highly specialised modules in the sense that they were evolutionarily selected to conduct a specific computational process that proved advantageous (395).

The fact that humans can play chess or read has by some been taken as evidence against evolved specialised systems (405). I would like to argue that it does not preclude chess or reading to be the result of a modular computation that developed during the course of evolution – strategic social cognition systems could be recruited in chess, and systems evolved for identifying objects such as tools or animals could be recruited to identify letters or words in reading (395).

In fact modules with specialised computations carried out on a differently broad set of inputs could allow for the discussed flexibility in adaptation through their dynamic recombination . This is a characteristic of many evolved systems. The human immune system may be taken as another example. Modern humans encounter, fight, and overcome pathogens our ancestors never encountered. This is possible because the adaptive immune system (as opposed to the innate immune system) has a modular architecture (395). Structures on the surface of lymphocytes have specific binding sites that recognize an antigen exactly because it takes only that particular molecular structure as an input. The recognition structures are the result of a recombination of a limited number of building blocks affording the flexibility to detect any potential antigen an individual may ever encounter.

Novelty (novel antigens, novel cognitive tasks) is necessarily part of the adaptive problem each system evolved to solve. Novelty is a potential problem for any architecture that is the product of evolution by natural selection. Every new organism faces an environment that is different from the one faced by its ancestors, at least in the tokens if not the types of entities in the world. Evolution does not prepare the organism for what is to be, only what was. Modular architectures provide flexibility because they allow components to be assembled and combined in novel ways (395).

In this respect, it is interesting to re-iterate that that humans do not always handle novelty all that well, and they operate, in many aspects of modern life, in ways that appear detrimental to fitness or to their own long-term goals. The introductory chapter had therefore proposed to conceptualise these seeming failures with respect to the phylogenetic and ontogenetic history of their underlying modules.

It therefore has to be concluded that the current thesis hinted at distinct and separable sub-regions and modules even in areas that are thought to support higher cognitive function. Many anatomical properties of these areas were conserved across monkeys and humans. Future research will need to show how evidently present adaption to changed and still changing environmental challenges on the behavioural level does actually occur on the neural brain with a large number of different sub-regions that preserved fundamental connectional features throughout a long evolutionary history.

6 Supplemental Information (SI) for "Comparison of human frontal cortex areas for cognitive control and language with areas in monkey frontal cortex" (Chapter 2)

6.1 Supplemental Experimental Procedures and Results

6.1.1 Diffusion-weighted MRI data acquisition

Diffusion-weighted images were acquired in 25 healthy right-handed (according to Edinburgh Handedness Inventory: mean +/- SD, 77 +/- 26.7) participants (14 female; age range: 20– 45 years; mean age +/- SD, 29 +/- 6.6 years) on a 3 T Siemens Magnetom Verio MR scanner.

All participants gave written informed consent in accordance with ethical approval from the local ethics committee. Participants lay supine in the scanner, and cushions were used to reduce head motion. Diffusion-weighted data were acquired using echo planar imaging (65 2.0 mm thick axial slices; field of view, 190 x 190 mm2; and a voxel size of 2 x 2 x 2 mm). Diffusion weighting was isotropically distributed along 60 directions using a b-value of 1.000 $s \times mm^{-2}$. Eight volumes with no diffusion weighting were acquired throughout the acquisition. A structural scan was acquired for each participant in the same session, using a T1-weighted three-dimensional fast, low-angle shot (3D Multiecho MPRAGE) sequence [repetition time (TR), 2530 ms; echo time (TE), 1.69 ms; flip angle, 7.0 °; with elliptical sampling of k space, giving voxel size of 1.0 x 1.0 x 1.0 mm].

6.1.2 Diffusion-weighted tractography-based parcellation

DW-MRI data were pre-processed using tools from FDT v2.0 (Functional MRI of the Brain Diffusion Toolbox; part of FSL 4.1). Eddy-current-distortions were corrected using affine registration of all volumes to a target volume with no diffusion weighting. EPI unwarping was performed using FMRIB's Utility for Geometrically Unwarping EPIs v2.5 (FUGUE). Voxel-wise estimates of the fibre orientation distribution were calculated using Bedpostx, limited to

estimating two fibre orientations at each voxel, because of the b-value and number of gradient orientations in the diffusion data (Behrens et al., 2007).

The vlFC ROI was registered from the MNI152 template brain as provided by FSL to the individual subject's T1-weighted structural image using FNRIT (FMRIB's Non-linear Image Registration Tool) and then registered from the individual subject's structural image into the individual subject's DW-MRI space using FLIRT (FMRIB's Linear Image Registration Tool). For each participant, probabilistic tractography was run from each voxel in the right vlFC ROI to generate a connectivity matrix between all vlFC voxels and each brain voxel as described (200) and used to generate a symmetric cross-correlation matrix for all vlFC voxels. This cross-correlation matrix was then permuted using k-means segmentation for automated clustering to define different clusters. K-means can be sensitive to initialisation, but there are some ways to minimize these effects. In the implementation of k-means used here ("ccops" as implemented in FSL), starting points are chosen as follows: a random point is chosen to be the first cluster centre, then the second cluster centre is the point furthest from the first one, the third the point furthest from the average of the first two, etc until k centres are obtained. Thus, the only randomness is in the position of the first cluster center starting point. To allow the clustering to be fully driven by the connectivity with the rest of the brain no connexity constraint was applied during this clustering procedure (Tomassini et al., 2007). Additionally we clustered the seed voxels in the cross correlation matrix using the fuzzy k-means algorithm. This algorithm, rather than just assigning each voxel in the vlFC seed ROI to one of the clusters, also gives a certainty of each voxel belonging to each of the clusters and hence provides a "certainty map" of the whole vlFC ROI for each cluster. The clusters were then all registered to the MNI152 template by inverting the linear (DW-MRI to T1-weighted structural image) and non-linear (T1-weighted structural image to MNI152_2mm template as provided by FSL) registration matrices / warp-fields that were used before and then matched between subjects for spatial overlap.

6.1.3 Number of clusters of the diffusion-weighted tractography-based parcellation

To determine the optimal number of clusters resulting in consistency across participants as well as coherency within the individual participants, we divided the right vlFC into clusters ranging from two to 25. We emphasize that the goal of this vlFC parcellation into vlFC sub-regions with coherent structural connectivity was not to determine exactly how many sub-regions there are in the vlFC. Such claims can hardly be tenable not only because of the limitations imposed by quality and resolution of the imaging data but also because they depend on what definition of "brain area" is used. Rather, the goal of this study is to test the hypothesis that different sub-regions in the vlFC can reliably be found and that the different structural and functional connectivity of these sub-regions might allow them to exert different types of top-down control onto other brain regions during a variety of cognitive tasks, such as language or cognitive control tasks. Hence the connectivity-based parcellation was limited to identifying sub-regions that are consistent within the group of subjects and then also found to be consistent in a second group of subjects. Similar inter-subject consistency in signal location is a prerequisite for most functional neuroimaging studies.

To ensure that the clustering procedure yields a sensible, reliable and robust parcellation of the vlFC we carried out a series of analyses that each attempted to parcellate vlFC into different numbers of sub-regions. We clustered the probabilistic tracking results into two to 25 clusters and analysed the results in the following way: First, we calculated the average cross-correlation value in the cross correlation matrix for each cluster in each subject (normalised for the overall cross correlation value in the cross correlation matrix) and calculated the variance of these cluster specific cross correlation values for each clustering solution. It is of course the case that the average cross correlation in each cluster increases as the number of clusters increases and the highest possible average cross correlation value would be reached with a clustering into as many clusters as there are voxels in the ROI.

However, we do not expect there to be as many clusters as there are voxels in the vlFC. Within the range of two to 25 clusters a low variance of the average cross correlation should indicate a coherent clustering into similarly correlated clusters without very well or very poorly correlated outlier-clusters. The group average cross correlation values for each cluster in the two to 25 clusters solutions is shown in Supplemental Figure 1B (in red diamonds) together with the variance of the measures (in blue squares). This measure indicates that the ten cluster solution shows the smallest variance in the average cross correlation and might therefore be considered as the solution that provides the most coherent clustering into similarly correlated clusters without very well or very poorly correlated outlier-clusters.

Second, we looked at the spatial overlap between corresponding clusters across the 25 subjects for each of the two to 25 clusters solutions. Supplemental Figure 1C shows in red triangles the average spatial overlap between corresponding clusters across the 25 participants for the different clustering solutions. Two indices of overlap were used. First, we examined the average degree of spatial overlap for a given cluster across all clusters (green dashed line) and, second, the degree of spatial overlap for the cluster with the lowest spatial overlap (cyan dotted line). In both cases parcellation solutions with ten vlFC subregions had maximal spatial overlaps (at least in the range of five to 25 clusters).

Third, in order to cross-validate the vlFC parcellation, we looked at the spatial overlap between the vlFC sub-regions derived from the two to 25 cluster solutions in the 25 participants on a group level with corresponding clusters in an independent group of eight subjects (see Supplemental Figure 1D). This analysis was done to ensure that the proposed vlFC parcellation that was based on a group of 25 subjects also holds true in other subjects. We also wanted to see which of the two to 25 clusters clustering solutions on a group level were reliably replicable in an independent group of subjects. Supplemental Figure1D shows with a red dashed line the average degree of spatial overlap for a given cluster across all clusters and with a blue dotted line the degree of spatial overlap for the cluster with the lowest spatial overlap.

In both cases parcellation solutions with ten vlFC subregions had the maximal spatial overlaps between the group of 25 participants and the independent group of eight subjects (at least in the range of five to 25 clusters).

We also determined the centers of gravity for each of the clusters in each individual participant. Centres of gravity for each of the clusters in each individual participant with ellipsoids representing the 95% confidence limits of the mean group cluster location are plotted on the structural MNI brain for the ten cluster-parcellation (Supplemental Figure 1E and F). Moreover the comparison of our parcellation derived functional connectivity results with functional connectivity of various different cyto-architectonically defined regions in the macaque vlFC further corroborates the plausibility of the parcellation of the vlFC into different sub-regions (for the between species comparison see below SI.6).

It is apparent that an iterative parcellation procedure (see SI.4) yielded largely similar results to the ten cluster parcellation. However using an iterative parcellation procedure it was also possible to further parcellate two of the ten regions (a region that extended from the pars orbitalis to the deep frontal operculum was sub-parcellated into lateral area 47 and the deep frontal operculum area Op; area 44 could be further sub-divided into a dorsal and a ventral part: 44d and 44v).

6.1.4 Recursive parcellation and sub-parcellation attempts

In addition to the ten cluster parcellation solution a recursive or iterative clustering procedure similar to the one used by (201) was conducted. Supplementary Figure 2 illustrates this procedure and shows the results. The parcellation started with the exact same vlFC ROI (leftmost image in SFig. 2) and was conducted in the same 25 subjects. Probabilistic tractography from each voxel in the ROI was used to parcellate the whole ROI into two sub-regions based on differences in the pattern of structural connectivity. These parcels were summed across all subjects to generate a group representation (leftmost figure in the first grey square). This first parcellation (grey arrow with “1.”) divided vlFC into an anterior (blue in

SFig.2 leftmost grey square; including FPl, FPm, 46, 47/12, Op, 45A, IFS) and a posterior (red, including 6v1, 6r1, 44v, 44d, IFJ) region. The group-parcels resulting from this first parcellation into two sub-regions were then used as an ROI for the next parcellation attempt which again parcellated them further into two sub-regions (see SFig.2). In this way the vlFC was parcellated into subregions with 3-4 parcellation steps. Parcellation was stopped if the respective parcel / ROI could not be sub-divided further reliably in all subjects into either two, three or four sub-regions (with preference to two over three and three over four). A subdivision was considered reliable if the topography of the different clusters was the same in all 25 subjects and all parcels could be matched beween all subjects based on spaital overlap.

The BOLD-time courses were extracted from each of the sub-regions on each level of the iterative parcellation and used in a multiple regression analysis using FEAT (FMRI Expert Analysis Tool) in order to infer the major differences in functional connectivity of the different sub-regions with the rest of the brain (controlling for confounding time series representing head movement and the Eigen time-series of white-matter and cortico-spinal fluid). The group results of this analysis were thresholded (z>2.3) and surface-projected (see “Methods”) and are shown next to the group parcellation results (right part of each grey box). The functional coupling maps are plotted in the same colour as the parcellation derived group parcels (e.g., the functional coupling pattern of anterior vlFC parcel is shown in blue and the functional coupling pattern of posterior vlFC in red in the right part of the leftmost grey box).

6.1.5 Human resting-state fMRI data acquisition, preprocessing, and analysis

Whole-brain blood oxygen level-dependent (BOLD) fMRI data was collected for 5 ½ min from each participant, using the following parameters: 44 axial slices; in-plane resolution, 3.0 x 3.0 mm; slice thickness, 3.0 mm; no slice gap; TR, 2.410 ms; TE, 30 ms; 128 volumes. Data were analyzed using tools from FSL. The first six volumes of each functional dataset were discarded and preprocessing was performed as follows: correction of distortion due to magnetic field

inhomogeneities using FUGUE (FMRIB's Utility for Geometrically Unwarping EPIs v2.5), motion correction using MCFLIRT (Motion Correction using FMRIB's Linear Image Registration Tool), non-brain removal using BET (Brain Extraction Tool v2.1), independent component analysis based denoising using MELODIC, spatial smoothing [using Gaussian 5 mm full-width at halfmaximum (FWHM) kernel], grand-mean intensity normalization of the entire four-dimensional dataset by a single multiplicative factor, and high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with high pass frequency cut-off of 100.0 s). Registration of functional images to the skull-stripped structural image was done using brain boundary based registration as implemented in FEAT (FMRI Expert Analysis Tool). Registration of the structural image to the MNI152_T1_2mm_brain MNI template brain in FSL was done using FNIRT (FMRIB's Non-linear Image Registration Tool).

6.1.6 Comparison between functional connectivity patterns of three tractography-based parcellation derived vlFC sub-regions to meta-analysis derived co-activation patterns of various different cognitive tasks

It is important to note that both IFJ and 44 have been associated with various cognitive control and language processes, including attentional reorienting (406), information updating (216, 407), inhibitory motor control (408), syntactic processing (409) and verbal working memory (410). Several recent studies have attempted a functional dissociation of these two areas by exploring their differential involvement in cognitive control (216, 217). IFJ has been proposed to be involved in the processing of infrequent stimuli (216, 411), preparatory cognitive control (217), information updating (412) and feedback processing (413), whereas 44 has been linked with inhibitory control and the flexible adjustment of motor plans (216, 217, 411). Moreover some recent studies have tried to functionally dissociate these two regions by investigating their differential involvement in language and more specifically in different types of syntactic

processing (414). Moreover another set of areas – area 45A and the deep frontal operculum – has also been suggested to be a core region for cognitive control and task set implementation (186, 415). However it is not clear what the different roles of IFJ, 44 and 45A/frontal operculum in cognitive control are. To test whether they are differentially involved in cognitive control and language tasks we carried out a comparison of these areas' functional connectivity profiles with co-activation patterns of functional imaging studies from a meta-analysis database.

We performed several meta-analyses of fMRI studies in order to establish networks of brain regions that have repeatedly been implicated in various different cognitive control and language tasks and therefore performed several meta-analyses of fMRI studies using the BrainMap database (www.brainmap.org) (see table 3 for more information on BrainMap search) and performed activation likelihood estimations (ALE) (416) of functional brain activity associated with these various different tasks. The BrainMap database (417, 418) is an on-line database of published functional and structural neuroimaging results. Reduced data are archived in BrainMap in the form of three-dimensional coordinates in stereotactic space (x,y,z) extracted from the peak locations of reported brain activations in the literature. The peak coordinates of a certain experimental condition were smoothed using a Gaussian distribution (FWHM = 12 mm) to accommodate the spatial uncertainty associated with neuroimaging foci and generate modelled activation images with 2-mm resolution (416). The BrainMap database was queried on 10th July, 2012, when the database contained 2,226 papers, 83 paradigm classes, 42,386 participants, 10,606 experiments, and 84,299 locations, and on 11th August 2012, when the database contained 2. 238 papers, 83 paradigm classes, 42.660 participants, 10.646 experiments. In this series of meta-analyses we tried to establish ALE maps for several different cognitive tasks / cognitive concepts that had been associated with vlFC. We then performed spatial cross-correlations (using fslcc as part of FSL) between the established ALE maps and the networks of functional connectivity derived from the ten different vlFC sub-

regions in order to determine which of the vlFC sub-regions and their respective network of resting state functional connectivity is also likely to be involved in specific cognitive tasks (169) (Supplemental figure 9).

In this way we established that the resting state fMRI derived functional connectivity patterns of IFJ mostly resembled that of co-activation patterns reported for task switching and working memory tasks (such as the n-back task). The resting state network of area 44 showed particular similarity to networks reported in inhibitory / interference control tasks (e.g., go/no-go tasks), whereas area 45/deep frontal operculum activations resembled patterns associated with attentional control (e.g., face discrimination tasks) and semantic retrieval.

Our parcellation and functional connectivity results support the idea that these regions are meaningful sub-regions in vlFC which are involved in different cognitive functions. The very distinct connectivity patterns of human IFJ, 44 and 45A and their suggested macaque homologues 44, ProM and 45A could allow these areas to play distinct roles in syntax processing and cognitive control.

6.1.7 Macaque resting-state fMRI data acquisition and anaesthesia

Resting-state fMRI and anatomical scans were collected for 25 healthy macaques (*Macaca mulatta*) (four females, age: 3.9 years, weight: 5.08 kg) under light inhalational anaesthesia. Anaesthesia was induced using intramuscular injection of ketamine (10 mg/kg), xylazine (0.125– 0.25 mg/kg), and midazolam (0.1 mg/kg). Macaques also received injections of atropine (0.05 mg/kg, i.m.), meloxicam (0.2 mg/kg, i.v.), and ranitidine (0.05 mg/kg, i.v.). Local anaesthetic (5% lidocaine/prilocaine cream and 2.5% bupivacaine injected subcutaneously around the ears to block peripheral nerve stimulation) was also used at least 15 minutes before placing the macaque in the stereotaxic frame. The anesthetized animals were placed in an MRI-compatible stereotactic frame (Crist Instruments) in a sphinx position and placed in a horizontal 3 T MRI scanner with a full-size bore. Scanning commenced ~2 h after induction,

when the ketamine was unlikely still to be present in the system. Anaesthesia was maintained, in accordance with veterinary recommendation, using the lowest possible concentration of isoflurane to ensure that macaques were lightly anesthetized. The depth of anaesthesia was assessed using physiological parameters (heart rate and blood pressure, as well as clinical checks before the scan for muscle relaxation). During the acquisition of the functional data, the inspired isoflurane concentration was in the range 1.0–1.8%, and the expired isoflurane concentration was in the range 0.9–1.7%. Isoflurane was selected for the scans as resting-state networks have been demonstrated previously to be present using this agent (194). Macaques were maintained with intermittent positive pressure ventilation to ensure a constant respiration rate during the functional scan, and respiration rate, inspired and expired CO_2, and inspired and expired isoflurane concentration were monitored and recorded using VitalMonitor software (Vetronic Services Ltd.). In addition to these parameters, core temperature and SpO_2 were monitored throughout the scan.

A four-channel phased-array coil was used for data acquisition (Dr. H. Kolster, Windmiller Kolster Scientific, Fresno, CA, USA). Whole-brain BOLD fMRI data were collected for 53 min and 26 s from each animal, using the following parameters: 36 axial slices; in-plane resolution, 2 x 2 mm; slice thickness, 2 mm; no slice gap; TR, 2000 ms; TE, 19 ms; 1600 volumes. A structural scan (three averages) was acquired for each macaque in the same session, using a T1-weighted magnetization-prepared rapid-acquisition gradient echo sequence (either 0.5 x 0.5 x 0.5 or 0.5 x 0.5 x 1.0 mm voxel resolution). The first six volumes of each functional dataset were discarded, and preprocessing was performed as similarly as possible to the human resting state data preprocessing: non-brain removal, 0.1 Hz low-pass filtering to remove respiratory artefacts, independent component analysis (ICA) –denoising using MELODIC, motion correction using MCFLIRT, spatial smoothing (using Gaussian 3.0 mm FWHM kernel), grand-mean intensity normalization of the entire four-dimensional dataset by a single multiplicative factor, and high-pass temporal filtering (Gaussian-weighted least-squares

straight line fitting, with high pass frequency cut-off of 100.0 s). Registration of functional images to the skull-stripped structural and the MNI macaque template was done using FNIRT.

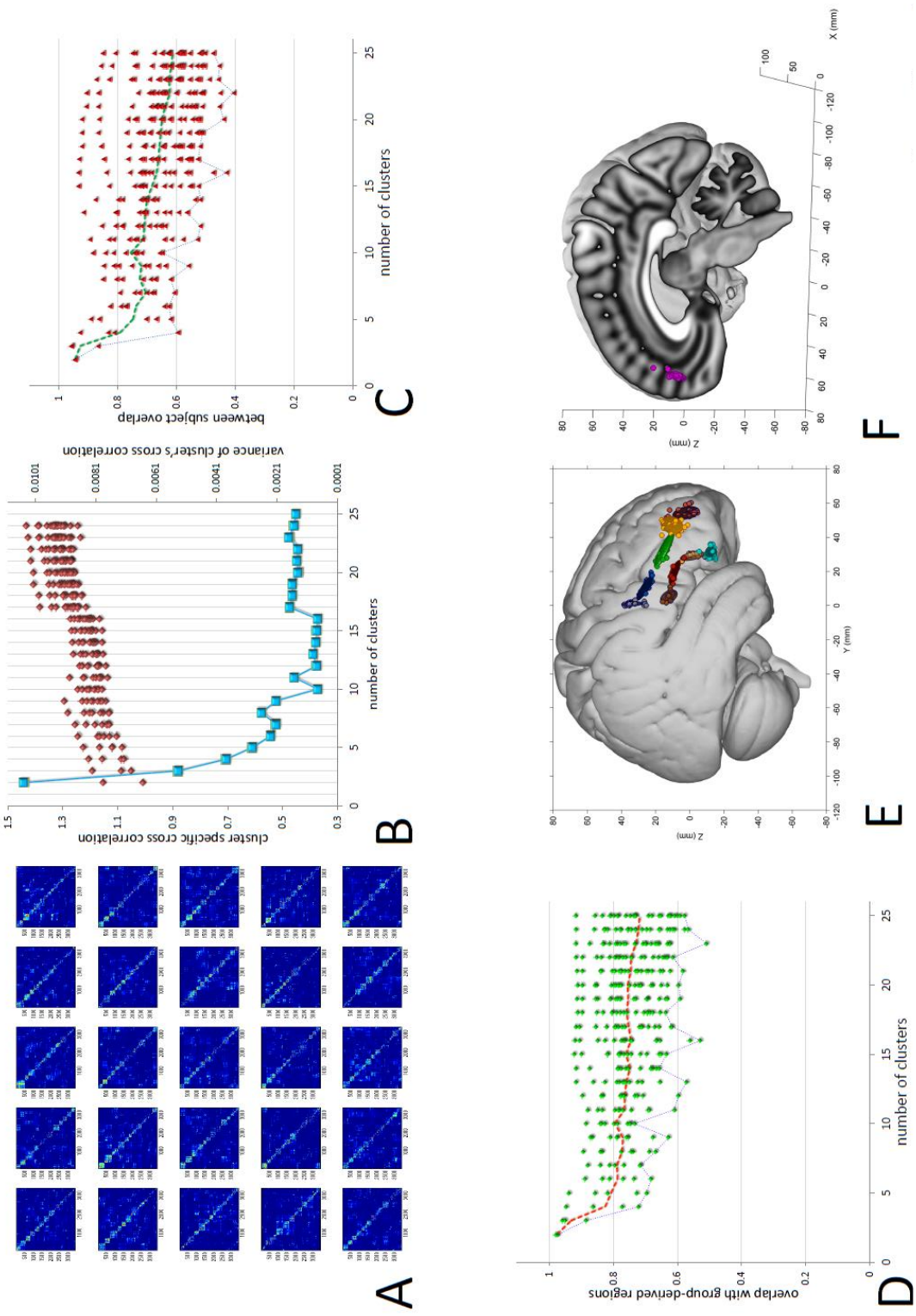
6.1.8 Functional coupling between different vlFC sub-regions.

The goal of this analysis was to determine if there was any between-species difference in the coupling of sub-regions within vlFC. As a representation of each parcel we took the 5% most likely voxels from each parcel's whole vlFC certainty maps (which in MNI-space are always 170 voxels for a 3400 voxel vlFC ROI) from the fuzzy-clustering derived parcel-specific maps. The fuzzy-clustering algorithm gives a certainty of each voxel belonging to each of the clusters and hence provides a "certainty map" of the whole vlFC ROI for each cluster.

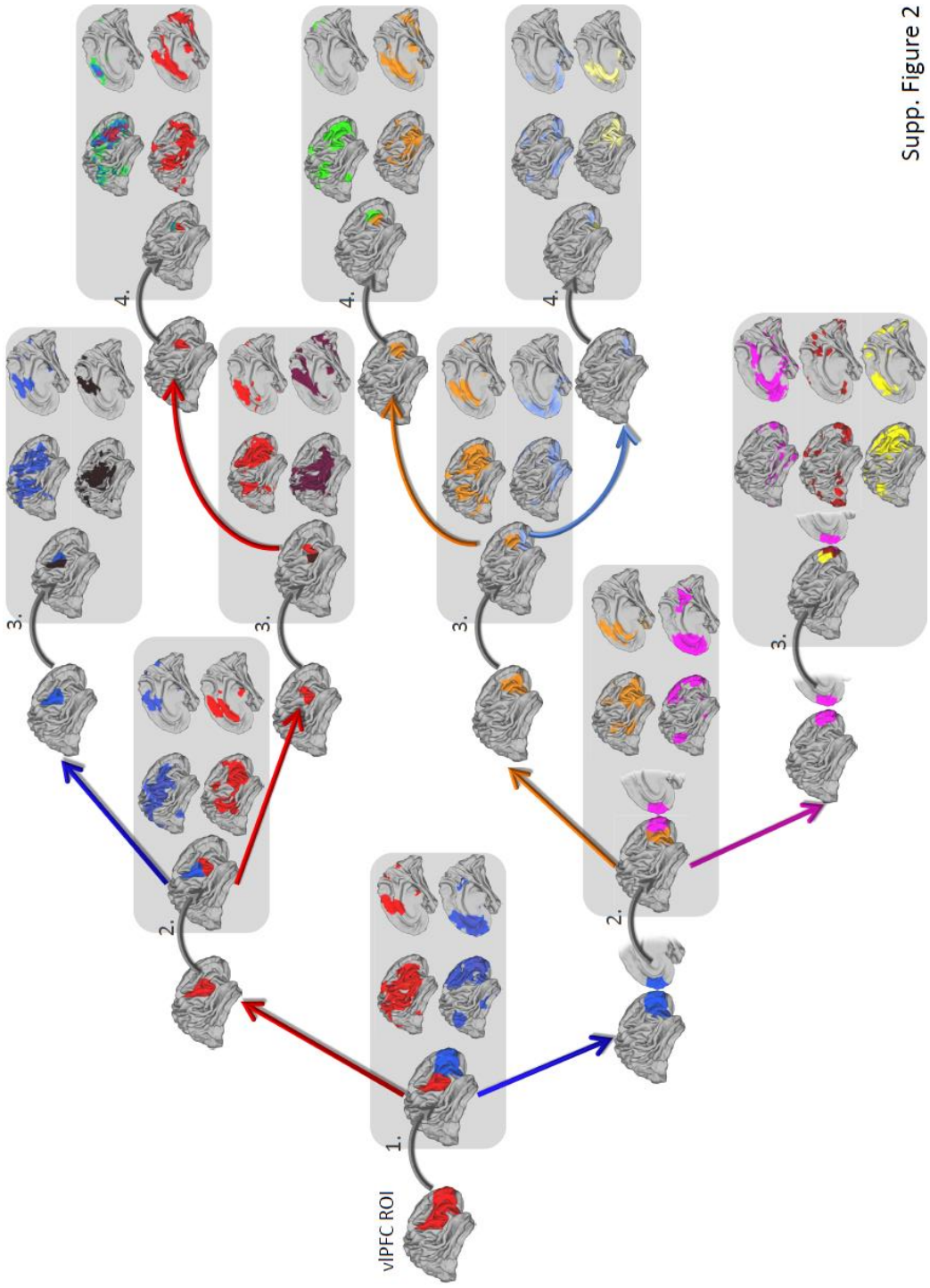
These individual ROIs were then registered from each subject's DW-MRI space into fMRI space using FLIRT. Then the major Eigen time series representing activity in each of the vlFC clusters was calculated. The major Eigen time series is the single time series that best reflects coherent activity across the vlFC region in that it represents the largest amount of variance across the set of voxels within the cluster. As such, the first Eigen time series helps prevent the mixing of signal and noise. We also extracted the average time-course of the whole brain. We then calculated the partial correlation of all vlFC sub-region-specific time-courses for each individual subject. We then Fisher-Z transformed the correlation coefficients. For the macaques we extracted the Eigen time series for each of the vlFC sub-regions that had been shown to be homologous to human vlFC sub-regions. Again the average time-course of the whole brain was extracted as well. We calculated the partial correlation of all monkey vlFC sub-region-specific time-courses and transformed the correlation coefficients using the Fisher-Z transformation (see Supplemental figure 9B). We performed independent samples t-tests between all transformed correlation coefficients and Bonferroni-corrected for multiple comparisons. We found smaller coupling between the two sub-regions situated in PMv (6v/F5c – 6r/F5a) for the human as compared to the macaque. We also noted smaller coupling between the 6r/F5a and

44v/ProM for the human as compared to the macaque. On the other hand we found stronger coupling between 44v with IFJ in the human compared to ProM and 44 in the macaque. We also found stronger coupling for 45B with 45A, 45A with 47 and 44v/ProM with 46 in the human compared to the corresponding areas in the macaque.

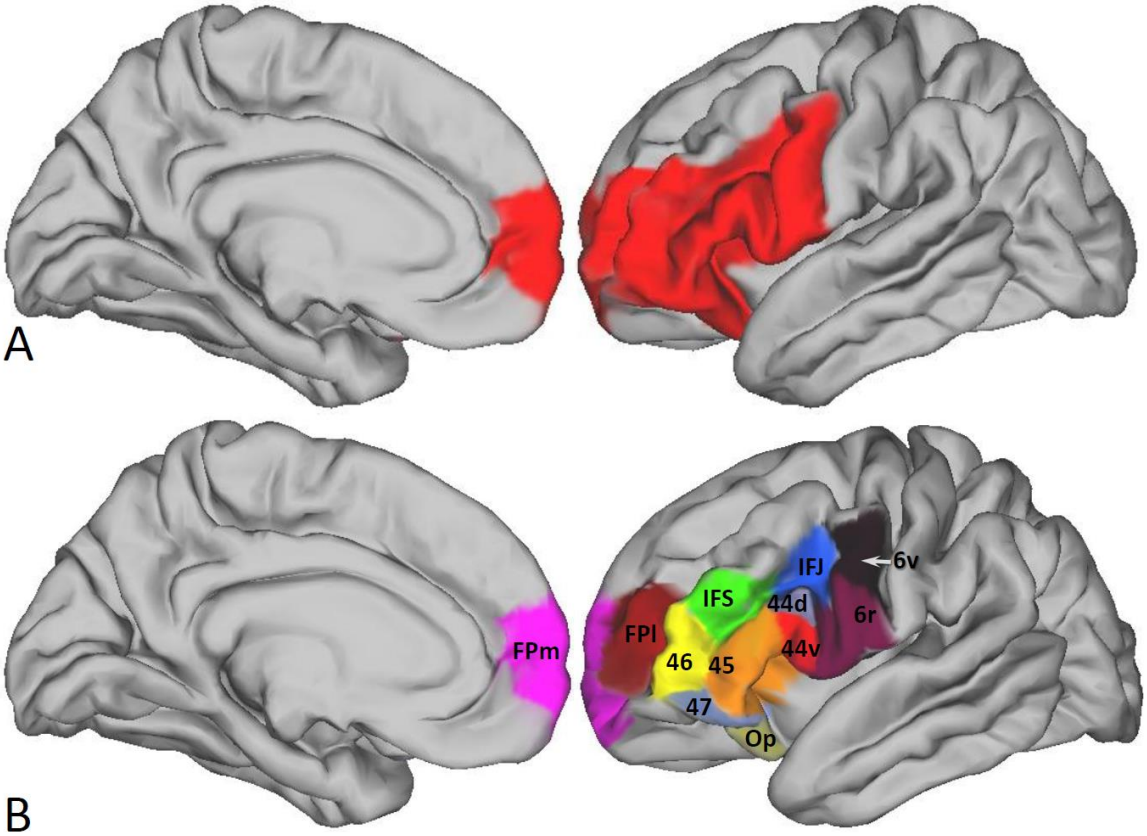
It might be argued that for some reason our scanning acquisition or analysis procedures were causing these differences in coupling within vlFC between macaques and humans. This, however, seems unlikely because we show effects going into both directions: smaller coupling between for example the 6r/F5a and 44v/ProM for the human as compared to the macaque and larger coupling between 44v and IFJ in the human compared to ProM and 44 in the macaque. Moreover in the previous spider-plot comparison analyses we showed the striking degree of similarity in coupling patterns between species including both short-range and long-range connections.



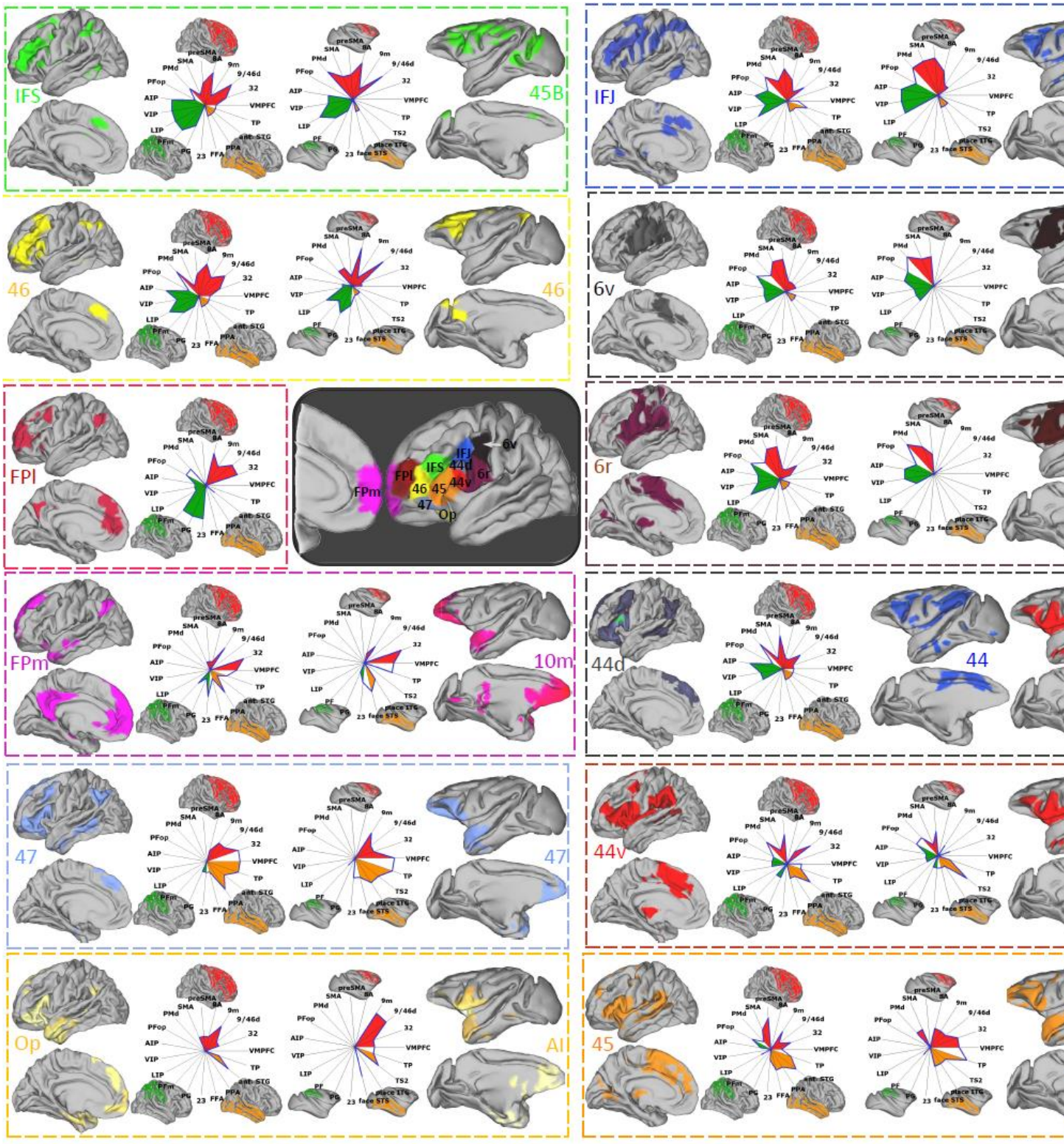
SI Figure 6-1: (A) shows the subject specific cross-correlation matrices for the right vlFC ten cluster clustering solution. The group average cross correlation values for each cluster of the two to 25 clusters solutions is shown in Figure 1B (in red diamonds) together with the variance of the measures (in blue squares). Figure 1C shows the average overlap between corresponding clusters across the 25 participants for the different clustering solutions. 1D shows the spatial overlap between the vlFC sub-regions derived from the two to 25 clusters clustering solutions in the 25 participants on a group level with corresponding clusters in an independent group of eight subjects. 1E and 1F show their respective centers of gravity for the ten way parcellation solution.



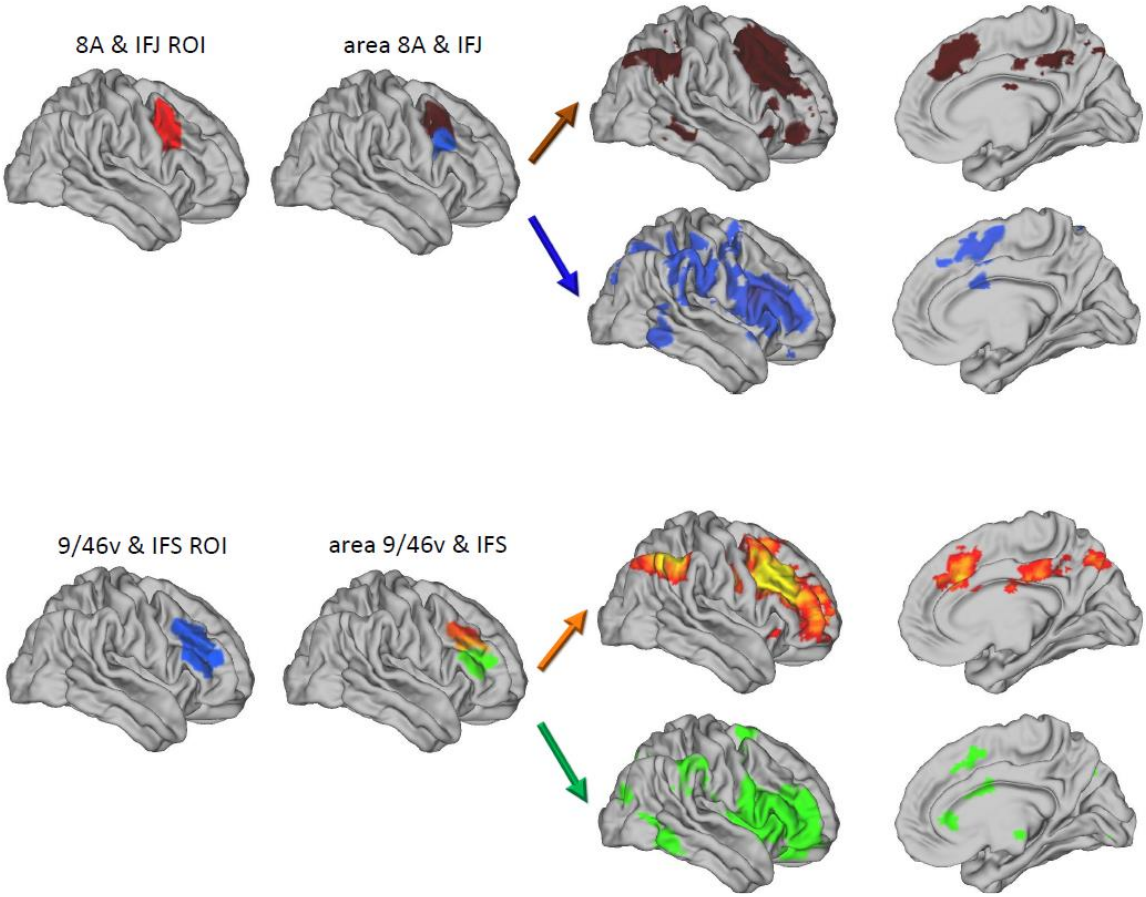
SI Figure 6-2 shows the results of an iterative parcellation of vlFC similar to the one used by Beckmann and colleagues (201) in the anterior cingulate cortex. The parcellation started with the exact same vlFC ROI (red in leftmost panel). The first parcellation (grey arrow with "1.") divided the original vlFC ROI into an anterior (blue in leftmost grey square) and a posterior (red) region. A multiple regression showed, that these two areas have significantly different resting state coupling profiles with the rest of the brain (right part of leftmost grey square). The group-parcels resulting from this first parcellation into two sub-regions were then used as an ROI for the next parcellation attempt which again parcellated them further into two sub-regions (grey arrows with "2."). In this way the vlFC was parcellated into subregions with 3-4 parcellation steps. Parcellation was stopped if the respective parcel / ROI could not be sub-divided further reliably in all subjects.



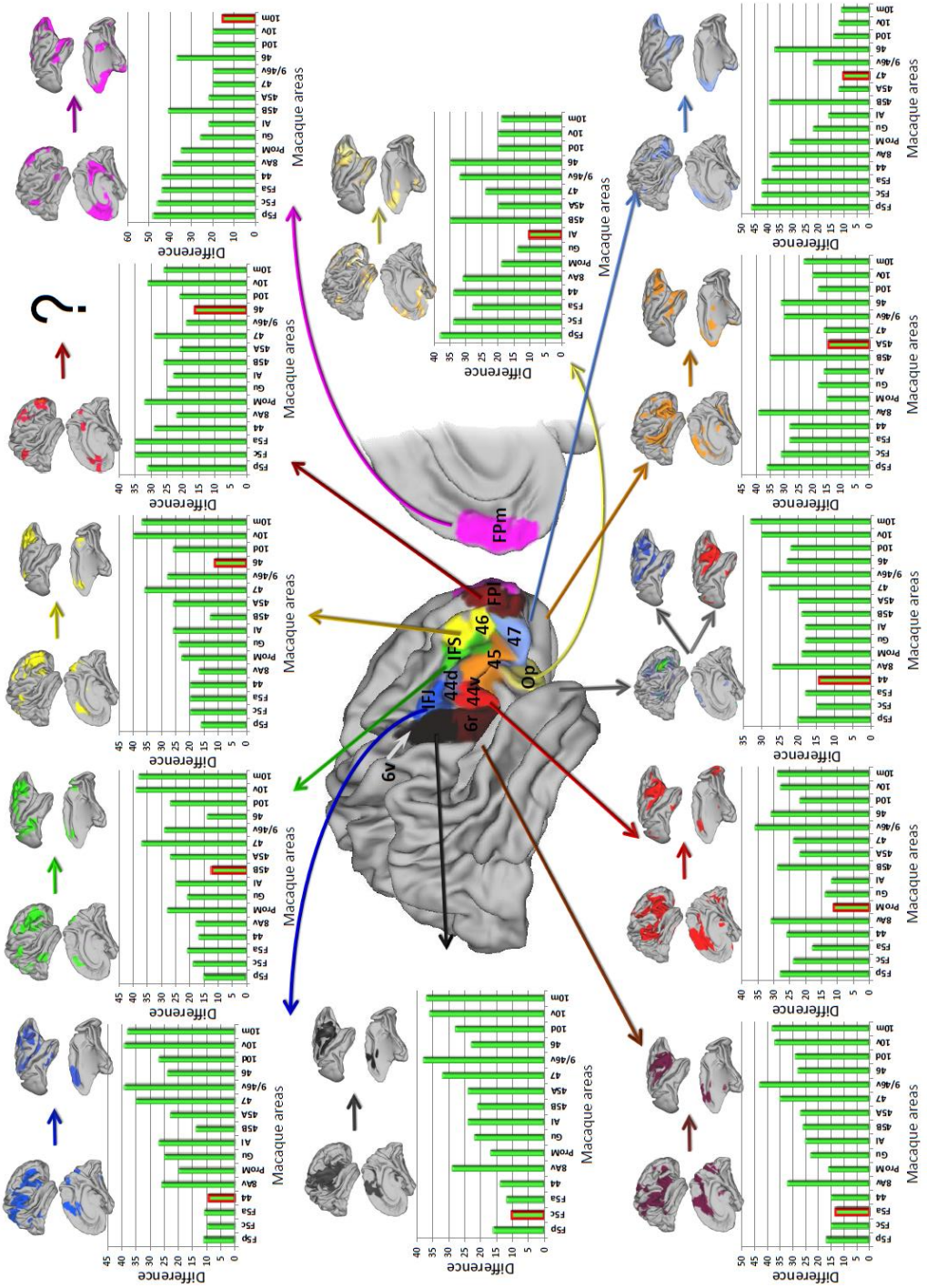
SI Figure 6-3 is the equivalent to Figure 2-2 in the main text for the left vlFC: 2A shows the left vlFC ROI. 2B shows the result of the twelve cluster parcellation.



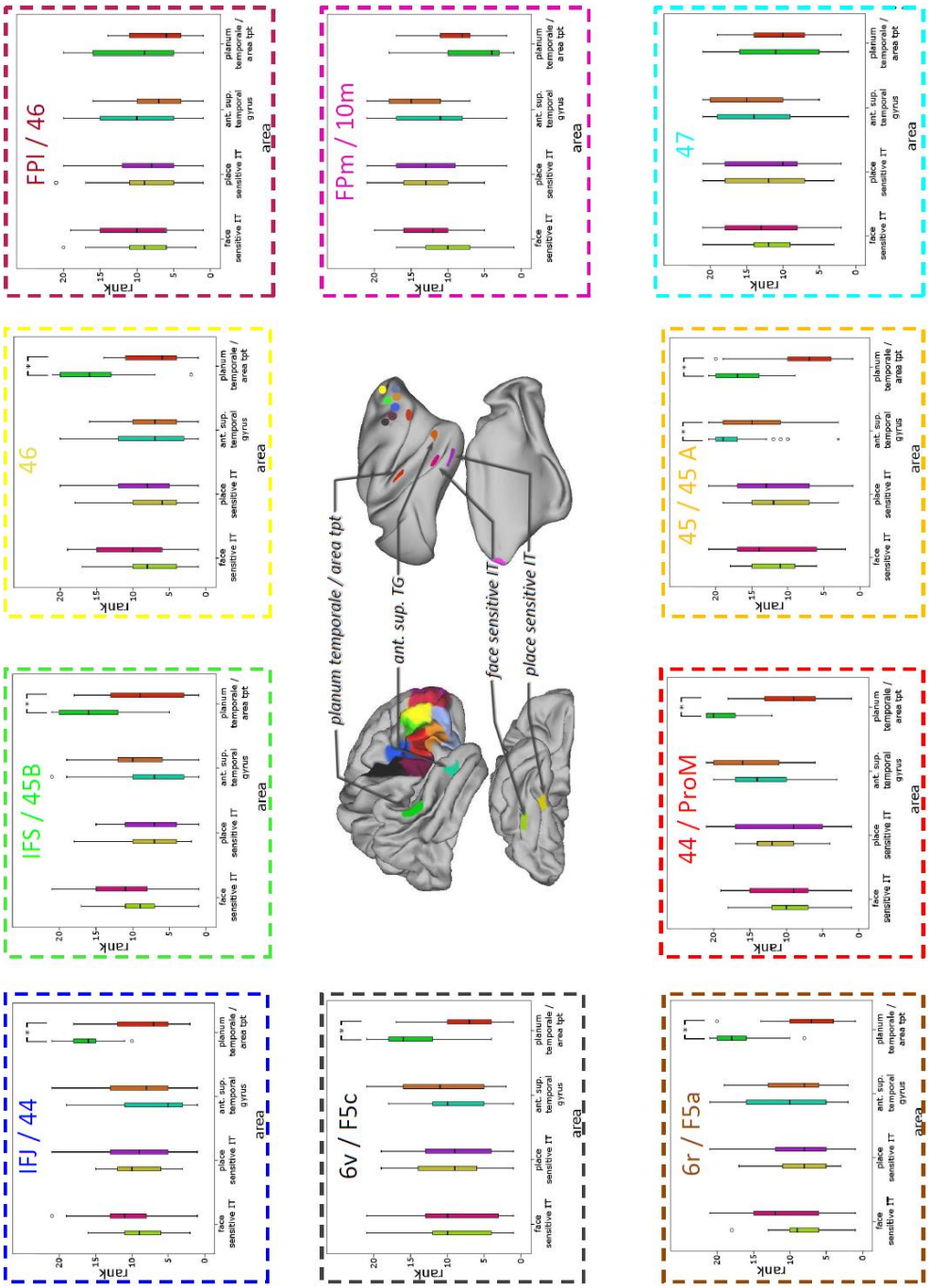
SI Figure 6-4 shows on the left side of each panel the resting state fMRI based functional connectivity patterns of each of the twelve left vlFC sub regions together with their respective spider plots. The right site of each panel depicts the coupling pattern of each suggested homologuous area in the macaque left vlFC together with the respective spider plot.



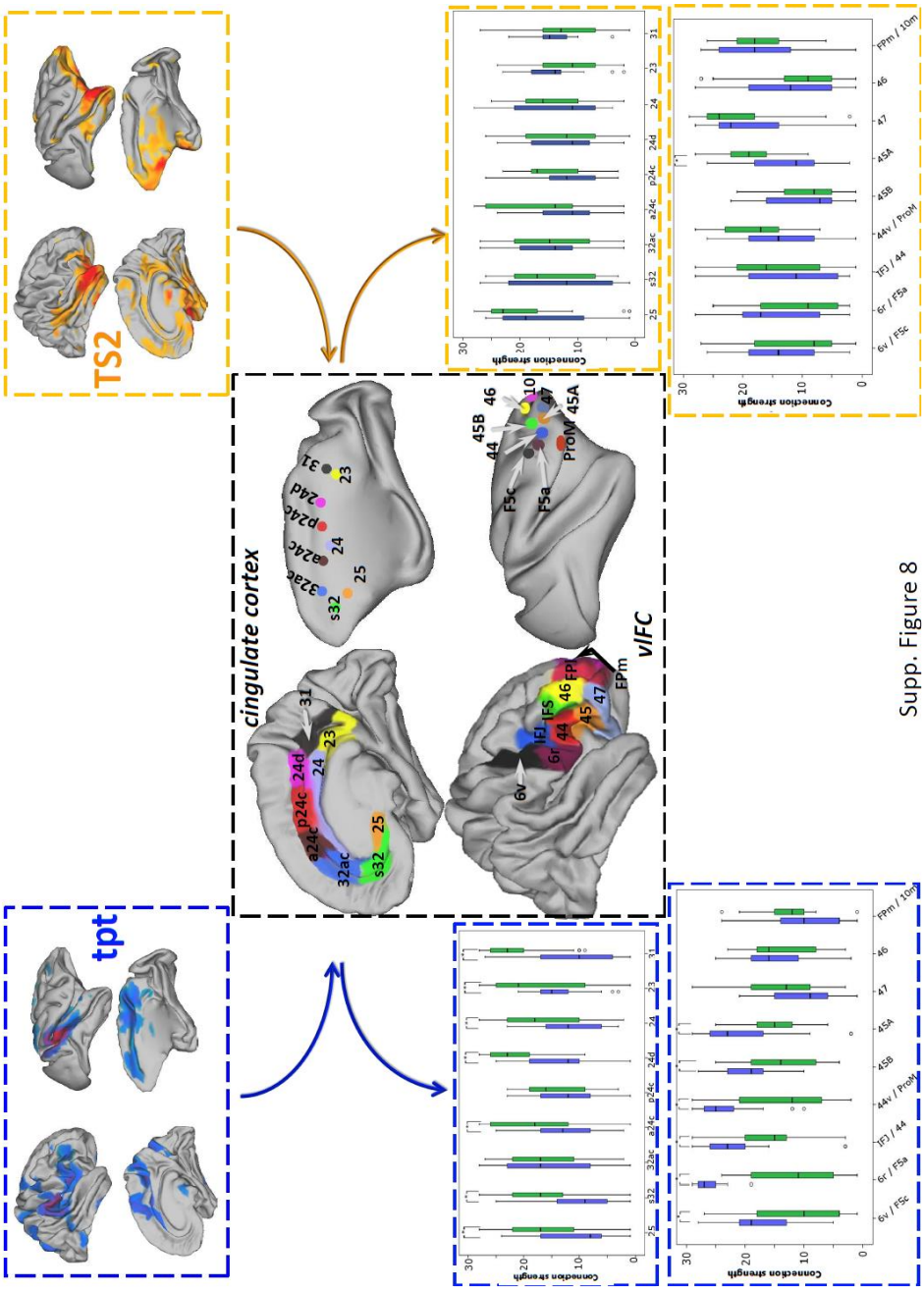
SI Figure 6-5: Additional parcellation analyses were conducted in order to show that IFJ and IFS could reliably be distinguished from 8Av and 9/46, respectively**.** An ROI comprising area IFJ as described here and 8A as described by (196)could be parcellated in all 25 subjects (top left). We established in a multiple regression that these two areas have significantly different resting state coupling profiles with the rest of the brain (top right). We also tested if it would be possible to reliably parcellate a ROI containing the 9/46v region described by (196) and the IFS area into two distinct regions (bottom left). We established, in a multiple regression, that these two areas have significantly different resting state coupling profiles with the rest of the brain (bottom right).



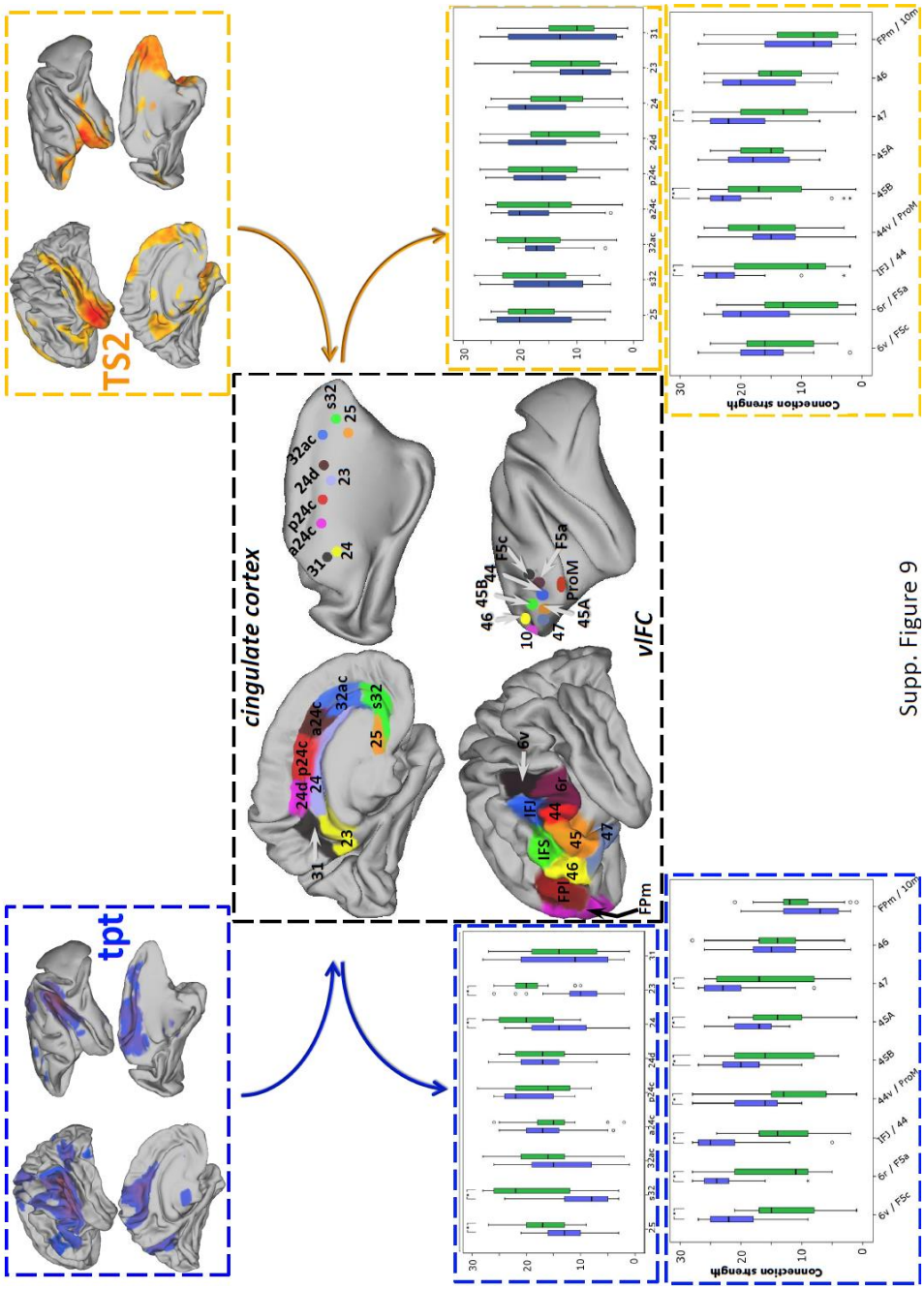
SI Figure 6-6 shows the results of the between species matching of functional connectivity patterns. Each human right vlFC sub-region was compared to all cytoarchitectonically defined macaque vlFC regions according to their coupling patterns. For each human vlFC region we show a bar chart where the height of each bar represents the distance between the connectivity fingerprints of the human vlFC region in question and all macaque vlFC regions. The lowest bar indicates the macaque area that is least different (most similar) to the human vlFC area investigated in terms of its coupling patterns with the rest of the brain. The summed absolute differences between the functional coupling scores of each human vlFC sub-region and 16 areas in macaque vlFC suggested that the human vlFC sub-regions matched the macaque sub-regions as described in the main text and Figure 2-4, Figure 2-5, Figure 2-6 and Figure 2-8.



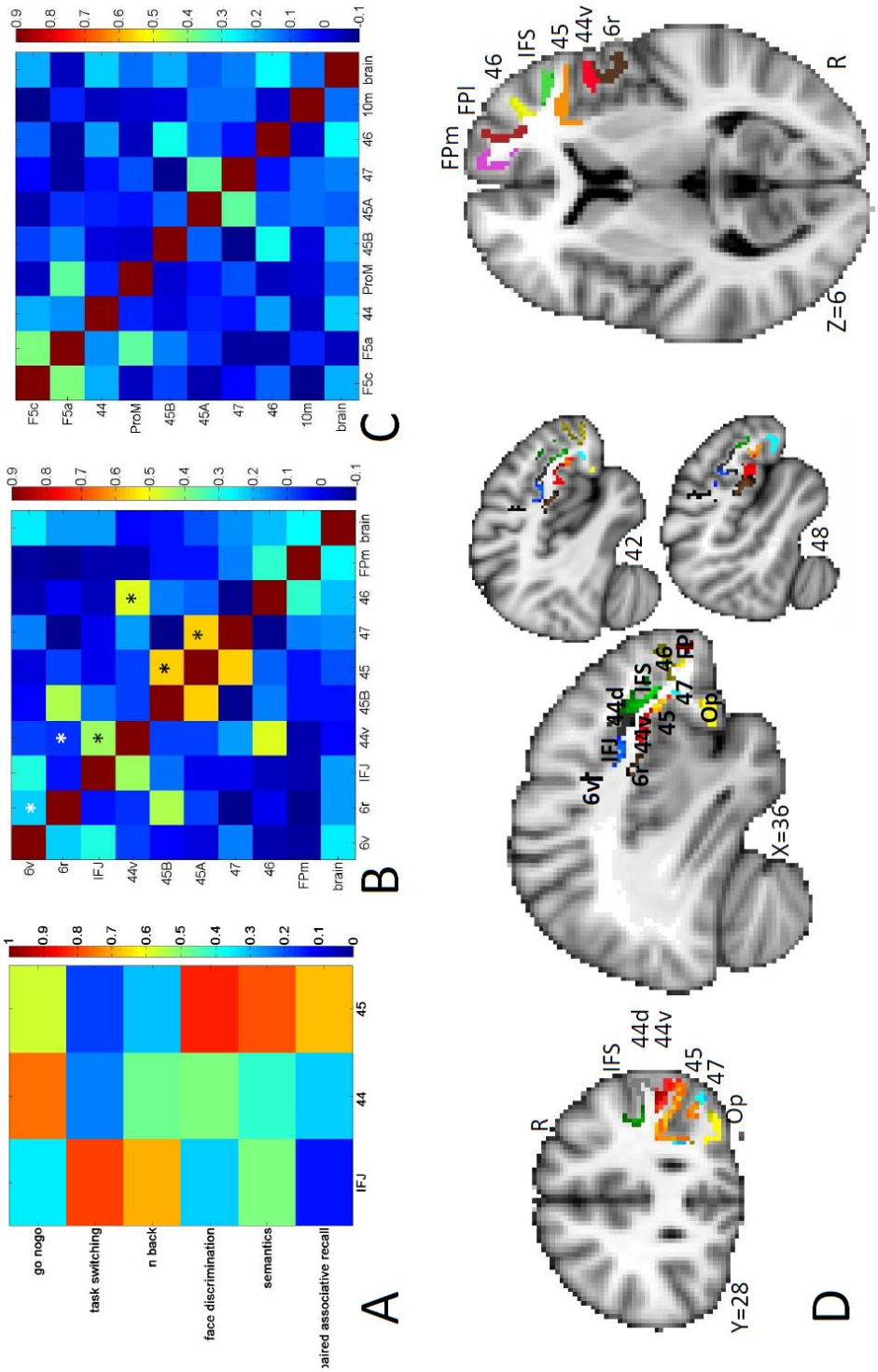
SI Figure 6-7 shows the rank of four regions in the temporal lobe among the 19 regions used to generate the spider plots in main text figures 4, 5, 6, and 8 for humans and macaques. High values mean that the coupling strength (average z-value in the ROI of the parcel derived z-maps) between the respective vlFC region and the temporal lobe region ranks very high among all target regions of interest (e.g., in the top left we can see that the IFJ functional connectivity z-map shows very high coupling with planum temporale in the human [green boxplot], but area 44 [the proposed macaque homologue of IFJ] shows relatively low coupling with area Tpt [the macaque homologue of the planum temporale, red boxplot]; this difference is even significant when tested with a Mann–Whitney–Wilcoxon rank-sum test with $p < 0.05$). "Cold colours" (green, cyan, green-yellow and olive) are used for the human temporo-occipital ROIs and "warm colours" (red, pink, orange, lilac) for the macaque ROIs. The ROIs are depicted on the human and macaque standard brain surface with the exact same colours as used in the boxplots.



SI Figure 6-8 shows fMRI derived functional connectivity patterns of two auditory association areas in the human and macaque (planum temporale and Tpt in blue and a voice-sensitive area in the anterior parabelt region in the human and the macaque in orange) and the functional coupling with areas in the vlFC and cingulate cortex in the left hemisphere. The two top graphs show the functional connectivity of planum temporale / Tpt (blue, left top graph) and anterior parabelt region (orange, right top graph) with the anatomically defined vlFC subregions. The two bottom graphs show the functional connectivity of planum temporale / Tpt (blue, left bottom graph) and anterior parabelt region (orange, right bottom graph) with anatomically defined cingulate sub-regions. Graphs show the ranking of functional connectivity of the auditory association areas with the vlFC and the cingulate ROIs in two separate ranking analyses among all ROIs presented in the spider plots before (see figure 4-6 and 8). Asterisks mark significant between species differences significance according to Mann–Whitney–Wilcoxon rank-sum test with $p<0.05$. The left part of this figure is identical to Figure 7 in the main text.



SI Figure 6-9 shows the same analysis as SI Figure 6-8 for the left hemisphere (planum temporale / Tpt, left TS2 and their interactions with left vlFC and left cingulate cortex).



SI Figure 6-10 (A) shows a matrix of the spatial correlation between the sub-region specific resting state functional connectivity patterns of areas IFJ, 44 and 45 (as shown in main text figures 5 and 6) and the ALE maps for various different cognitive tasks and cognitive concepts established using the BrainMap database (www.brainmap.org). See Supplemental table 4 for more information. B and C show cross correlation matrix of the resting state BOLD time-courses in vlFC regions in human (B) and macaque (C). Human vlFC regions show significantly stronger coupling between some prefrontal regions (black asterisks show significant difference in a between species t-test on the Fisher-transformed correlation coefficient) but also weaker coupling between premotor regions (indicated by white asterisks). D shows the parcellation derived sub-regions plotted on the MNI152 standard brain in a coronal (left), three sagittal (middle) and axial (right) slices.

Area	Abbreviation	MNI coordinates (x, y, z)			Study
Ventro-Medial Prefrontal Cortex	VMPFC	8	44	-14	Cluster 2 in (201) and area 14m in (111); also cf (27)
Area 32	32	8	42	0	(286); also cf (54)
Area 9 medial	9m	9	43	46	Cf cluster 3 in (196) suggested to be area 9
Area 9/46 dorsal	9/46d	38	40	36	Cf. cluster 5 in (196) suggested to be 9/46d
Area 8A / Frontal Eye Field	8A	32	12	56	Cf cluster 8 in (196) and 8Ad in (419) suggested to be 8A
Pre-Supplementary Motor Cortex	PreSMA	6	6	53	Cf. Pre-SMA in (420)
Supplementary Motor Cortex	SMA	4	-8	54	Cf. SMA in (420)
Premotor Cortex, dorsal part	PMd	-30	-2	58	Cf. PMd in (420)
Area PFop	PFop	58	-16	28	Cf. PFop in (318) and red cluster in (192) suggested to include PFop
Anterior Intraparietal Area	AIP	52	-32	44	Cf. blue IPL cluster in (192) suggested to include AIP
Ventral Intraparietal Area	VIP	36	-42	50	Cf. red SPL cluster in (192) suggested to include VIP
Mid-Intraparietal Sulcus	LIP	22	-64	52	Cf. magenta SPL cluster in (192) suggested include be LIP
Mid- Inferior Parietal Lobule	PFm	50	-44	40	Cf. green cluster in (192) suggested to include PFm
Area PG	PG	38	-68	38	Yellow/orange cluster in (192) (suggested to include PGp)
Posterior Cingulate Cortex	23	10	-50	30	Cf (201) and (286)
Fusiform Face Area	FFA	42	-50	-18	Cf. (183)
Parahippocampal Place Area	PPA	28	-48	-16	Cf. (183)
Voice sensitive superior temporal Gyrus	voice	48	6	-12	Cf. (421)
Temporal Pole	TP	34	12	-36	Interior temporopolar region (ITP) in (422)

SI Table 6-1: Human brain areas investigated.

Area	Abbreviation used subsequently	MNI coordinates (x, y, z) (302)		
Ventro-Medial Prefrontal Cortex	VMPFC	1.75	15.5	-1.75
Area 32	32	1.5	15.5	5.25
Area 9 medial	9m	1.5	19	12.5
Area 9/46 dorsal	946d	11.75	12.75	10.25
Area 8A / Frontal Eye Field	8A	13.5	8	13.25
Pre-Supplementary Motor Cortex	preSMA	1.5	11.75	15.25
Supplementary Motor Cortex	SMA	2.25	6.25	18
Premotor Cortex, dorsal part	PMd	12.25	-0.25	17
Area PFop	PFop	21.5	-8	6
Anterior Intraparietal Area	AIP	18.25	-8	11.75
Ventral Intraparietal Area	VIP	11	-14.75	12.75
Mid-Intraparietal Sulcus	LIP	10.5	-19	17
Mid- Inferior Parietal Lobule	pf	21	-11.75	13.75
Area PG	pg	16.5	-15.25	16.25
Posterior Cingulate Cortex	23	1.25	-16	8.75
Fusiform Face Area	Face sensitive STS	23.5	-10	-4
Parahippocampal Place Area	Place sensitive ITG	20	-2	-11
Voice sensitive superior temporal Gyrus	Ts2	22	0	-2.5
Temporal Pole	TP	18.75	7	-7.25
Area F5p	F5p	15.25	4	10.75
Area F5c	F5c	20.75	4	7
Area F5a	F5a	19	6.75	6.5
Area ProM	ProM	20	8	-2.75
Area Area 8Av	8Av	16.5	8.5	10.25
Area 44	44	15.25	8.75	3
Area 45B	45B	14.5	7.5	8.25
Area 45A	45A	16.75	13.25	4.25
Area 47	47	12.5	19.75	6
Area 9/46v	9/46v	13.75	14.25	8
Area 46	46	11	13.5	7.5
Lateral Agranular Insula	AI	15.25	6.25	-4
Area 10d	10d	5.75	22.25	9.5
Area 10v	10v	5	24.25	3.5
Area 10m	10m	1.25	25	5

SI Table 6-2: Macaque brain areas investigated.

Area in Human	Human Area Abbreviation	Monkey Area Abbreviation	Area in Rhesus Monkey	Study / Atlas
Ventro-Medial Prefrontal Cortex	VMPFC	VMPFC	Area 14	(111, 238)
Area 32	32	32	Area 32	(111, 201, 238, 423)
Area 9 medial	9m	9m	Area 9m	(196, 424)
Area 9/46 dorsal	946d	946d	Area 9/46D	(196, 237)
Area 8A / Frontal Eye Field	8A	8A	Area 8A	(196, 238)
Pre-Supplemental Motor Cortex	PreSMA	preSMA	Agranular F6	(196, 238)
Supplemental Motor Cortex	SMA	SMA	Agranular Frontal Area F3	(196, 238)
Premotor Cortex, dorsal part	PMd	PMd	Area F2 6DR 6DC	(196, 238)
Area PFop	PFop	PFop	Area PFop	(192, 237)
Ventral Intraparietal Area	AIP	AIP	Area AIP	(192, 237)
Mid-Intraparietal Sulcus	VIP	VIP	Area VIP	(192, 237)
Lateral Intraparietal Sulcus	LIP	LIP	Area LIP	(192, 237)
Mid- Inferior Parietal Lobule	PFm	Pf	Area PF	(192, 237)
Area PG	PG	Pg	Area PG	(192, 237)
Posterior Cingulate Cortex	23	23	Area 23	(237) (201, 425)
Fusiform Face Area	FFA	Face sensitive STS	Face sensitive superior temporal sulcus	(426)
Parahippocampal Place Area	PPA	Place sensitive ITG	Place sensitive inferior temporal gyrus	(426)
Voice sensitive superior temporal Gyrus	voice	Ts2	Voice sensitive TS2	(427)
Temporal Pole	TP	TP	dorsal temporal pole	(238)

SI Table 6-3: Correspondences between human and macaque areas outside of vlFC.

Search Criteria	papers	participants	experiments	locations
Experiments: Paradigm Class: Go/No-Go AND Activations only	86	2076	301	2743
Experiments: Paradigm Class: Task Switching AND Activations only	42	831	151	1282
Experiments: Paradigm Class: n-back AND Activations only	120	2950	481	3682
Experiments: Paradigm Class: Face Monitor / Discrimination AND Activations only	144	2985	645	4312
Experiments: Behavioral Domain: Cognition: Language - Semantics AND Activations only	253	3990	945	7245
Experiments: Paradigm Class: Paired Associate Recall AND Activations only	43	808	161	1184

SI Table 6-4: Meta-analysis of activation patterns in cognitive tasks.

7 Supplemental Information (SI) for “Connectivity profiles reveal the relationship between brain areas for reward-guided learning and decision making in human and monkey frontal cortex” (Chapter 3)

7.1 Supplemental Experimental Procedures and Results

7.1.1 Human resting-state fMRI data acquisition, pre-processing, and analysis

Whole-brain blood oxygen level-dependent (BOLD) fMRI data was collected for 5 ½ min from each participant, using the following parameters: 44 axial slices; in-plane resolution, 3.0 x 3.0 mm; slice thickness, 3.0 mm; no slice gap; TR, 2.410 ms; TE, 30 ms; 128 volumes. Data were analyzed using tools from FSL. The first six volumes of each functional dataset were discarded and preprocessing was performed as follows: correction of distortion due to magnetic field inhomogeneities using FUGUE (FMRIB's Utility for Geometrically Unwarping EPIs v2.5), motion correction using MCFLIRT (Motion Correction using FMRIB's Linear Image Registration Tool), non-brain removal using BET (Brain Extraction Tool v2.1), spatial smoothing [using Gaussian 5 mm full-width at half maximum (FWHM) kernel], grand-mean intensity normalization of the entire four-dimensional dataset by a single multiplicative factor, and high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with high pass frequency cut-off of 100.0 s). Registration of functional images to the skull-stripped structural image was done using brain boundary based registration as implemented in FEAT (FMRI Expert Analysis Tool). Registration of the structural image to the MNI152_T1_2mm_brain MNI template brain in FSL was done using FNIRT (FMRIB's Non-linear Image Registration Tool).

7.1.2 Macaque resting-state fMRI data acquisition and anesthesia

Resting-state fMRI and anatomical scans were collected for 25 healthy macaques (Macaca mulatta) (four females, age: 3.9 years, weight: 5.08 kg) under light inhalational anesthesia.

Anesthesia was induced using intramuscular injection of ketamine (10 mg/kg), xylazine (0.125–0.25 mg/kg), and midazolam (0.1 mg/kg). Macaques also received injections of atropine (0.05 mg/kg, i.m.), meloxicam (0.2 mg/kg, i.v.), and ranitidine (0.05 mg/kg, i.v.). Local anaesthetic (5% lidocaine/prilocaine cream and 2.5% bupivacaine injected subcutaneously around the ears to block peripheral nerve stimulation) was also used at least 15 minutes before placing the macaque in the stereotaxic frame. The anesthetized animals were placed in an MRI-compatible stereotactic frame (Crist Instruments) in a sphinx position and placed in a horizontal 3 T MRI scanner with a full-size bore. Scanning commenced ~2 h after induction, when the ketamine was unlikely still to be present in the system. Anesthesia was maintained, in accordance with veterinary recommendation, using the lowest possible concentration of isoflurane to ensure that macaques were lightly anesthetized. The depth of anesthesia was assessed using physiological parameters (heart rate and blood pressure, as well as clinical checks before the scan for muscle relaxation). During the acquisition of the functional data, the inspired isoflurane concentration was in the range 1.0–1.8%, and the expired isoflurane concentration was in the range 0.9–1.7%. Isoflurane was selected for the scans as resting-state networks have been demonstrated previously to be present using this agent (194). Macaques were maintained with intermittent positive pressure ventilation to ensure a constant respiration rate during the functional scan, and respiration rate, inspired and expired CO2, and inspired and expired isoflurane concentration were monitored and recorded using VitalMonitor software (Vetronic Services Ltd.). In addition to these parameters, core temperature and SpO2 were monitored throughout the scan.

A four-channel phased-array coil was used for data acquisition (Dr. H. Kolster, Windmiller Kolster Scientific, Fresno, CA, USA). Whole-brain BOLD fMRI data were collected for 53 min and 26 s from each animal, using the following parameters: 36 axial slices; in-plane resolution, 2 x 2 mm; slice thickness, 2 mm; no slice gap; TR, 2000 ms; TE, 19 ms; 1600 volumes. A structural scan (three averages) was acquired for each macaque in the same session, using a

T1-weighted magnetization-prepared rapid-acquisition gradient echo sequence (either 0.5 x 0.5 x 0.5 or 0.5 x 0.5 x 1.0 mm voxel resolution). The first six volumes of each functional dataset were discarded, and pre-processing was performed as similarly as possible to the human resting state data pre-processing: non-brain removal, 0.1 Hz low-pass filtering to remove respiratory artefacts, independent component analysis (ICA) denoising using MELODIC, motion correction using MCFLIRT, spatial smoothing (using Gaussian 3.0 mm FWHM kernel), grand-mean intensity normalization of the entire four-dimensional dataset by a single multiplicative factor, and high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with high pass frequency cut-off of 100.0 s). Importantly, apart from procedures to deal with artefacts due to the externally controlled respiratory frequency in the macaques and the different resolution of scans of the two species' brains, the pre-processing was identical for human and macaque datasets. This is particularly important in light of the influence of certain pre-processing steps on the values of resting state correlations (428). Registration of functional images to the skull-stripped structural and the MNI macaque template was done using FNIRT.

7.1.3 Diffusion-weighted MRI data acquisition and pre-processing

Diffusion-weighted data were acquired using echo planar imaging (65 2.0 mm thick axial slices; field of view, 190 x 190 mm2; and a voxel size of 2 x 2 x 2 mm). Diffusion weighting was isotropically distributed along 60 directions using a b-value of 1.000 $s \times mm^{-2}$. Eight volumes with no diffusion weighting were acquired throughout the acquisition. A structural scan was acquired for each participant in the same session, using a T1-weighted three-dimensional fast, low-angle shot (3D Multiecho MPRAGE) sequence [repetition time (TR), 2530 ms; echo time (TE), 1.69 ms; flip angle, 7.0°; with elliptical sampling of k space, giving voxel size of 1.0 x 1.0 x 1.0 mm].

DW-MRI data were pre-processed using tools from FDT v2.0 (Functional MRI of the Brain Diffusion Toolbox; part of FSL 4.1). Eddy-current-distortions were corrected using affine registration of all volumes to a target volume with no diffusion weighting. EPI unwarping was performed using FMRIB's Utility for Geometrically Unwarping EPIs v2.5 (FUGUE). Voxel-wise estimates of the fibre orientation distribution were calculated using Bedpostx, limited to estimating two fibre orientations at each voxel, because of the b-value and number of gradient orientations in the diffusion data (199).

7.1.4 Diffusion-weighted tractography-based parcellation of human medial and orbital frontal cortex

The ROI for the diffusion-weighted tractography-based parcellation (Figure 3-5A) covered everything that has been referred to as ACC, OFC or FP (Figure 3-5A). It therefore comprised the whole region investigated by Beckmann and colleagues (201), except the two most posterior clusters (8 and 9). It included the cingulate gyrus and sulcus (including the dorsal bank of the paracingulate sulcus if present) and extended posteriorly to include all cingulate motor areas as delineated by Beckmann and colleagues (201) and Amiez and Petrides (301). It, therefore, covered the most anterior tip of the marginal sulcus but excluded the precuneus and posterior cingulate cortex. Anteriorly it included subgenual ACC and vmPFC. Moreover it contained areas 9, FPm and FPm, as delineated by Neubert and Sallet and colleagues (196, 429), the OFC, and the orbital part of the inferior frontal gyrus (pars orbitalis gyri frontalis inferioris). This ROI was drawn in both the left and the right hemisphere.

The ROI was registered from the MNI152 template brain as provided by FSL, to each individual subject's T1-weighted structural image using FNIRT (FMRIB's Non-linear Image Registration Tool). In the individual T1-weighted structural images it was masked by the grey-white matter boundary as defined by FAST v4.1 (FMRIB's Automated Segmentation Tool) (198) and then

registered from the individual subject's structural image into the individual subject's DW-MRI space using FLIRT (FMRIB's Linear Image Registration Tool). For each participant, probabilistic tractography was run from each voxel in the ROI to generate a connectivity matrix between all ROI voxels and each brain voxel as previously described (282) and used to generate a symmetric cross-correlation matrix for all voxels in the ROI. This cross-correlation matrix was then permuted using k-means segmentation for automated clustering to define different clusters. K-means can be sensitive to initialization, but there are some ways to minimize these effects. In the implementation of k-means used here ("ccops" as implemented in FSL), starting points are chosen as follows: a random point is chosen to be the first cluster centre, then the second cluster centre is the point furthest from the first one, the third the point furthest from the average of the first two, etc. until k centres are obtained. Thus, the only randomness is in the position of the first cluster centre starting point. The clusters were then all registered to the MNI152 template by inverting the linear (DW-MRI to T1-weighted structural image) and non-linear (T1-weighted structural image to MNI152_2mm template as provided by FSL) registration matrices / warp-fields that were used before and then matched between subjects for spatial overlap. We used a recursive clustering procedure here similar to the one used by Beckmann and Neubert and colleagues (201, 429). This procedure tries to identify a parcellation solution that is largely consistent across subjects. The ROI is sub-divided into sub-regions with consistent estimated structural connectivity with the rest of the brain. These sub-regions are then further sub-divided until between subjects-consistency breaks down. In this way we delineated 21 main component parts in the ROI. We also determined centres of gravity for each of the clusters. This iterative parcellation procedure yielded largely similar results in the two hemispheres (Figure 3-5B, SI Figure 7-3).

To establish the functional connectivity of each orbital and medial parcel, we created ROIs from every parcel in each individual subject's parcellation. Individual ROIs were registered from each subject's DW-MRI space into fMRI space using FLIRT. Individual functional coupling

maps were calculated using the same seed-based-correlation analysis approach described above. The resulting Z-statistical group images were projected onto the CaretBrain as provided by the Human Connectome Project Workbench using the "surf_proj" algorithm as implemented in FSL and then visualized using the Human Connectome Project Workbench. Z maps were quantified using the same 23 cortical and sub-cortical ROIs mentioned above. These parcel-specific functional coupling fingerprints could then be compared to 448 macaque orbital and medial frontal ROIs by calculating the summed absolute difference of the coupling scores (see above). When all of these parcellation-derived functional coupling fingerprints were matched to any of the 448 monkey ROIs, we obtained a summed absolute difference for each human area – monkey ROI comparison (i.e. 23 x 448 = 9408 summed absolute differences). When plotting these differences in a histogram it is apparent that they are normally distributed (SI Figure 7-2G). Most monkey ROIs matched the human areas relatively poorly, whereas extremely good and extremely bad matches were relatively rare. We used two standard deviations below (note: that small absolute differences were considered good matches) the mean of this distribution of summed absolute differences as a cut-off to look for "significantly" good human to monkey matches. In this way we were able to identify monkey ROIs that matched all the different component parts of human medial and orbital frontal cortex with the exception of area FPl.

In addition to establishing the functional coupling patterns of all parcellation-derived subregions to compare them within and across species, we also conducted several tests for significant differences in coupling patterns between specific subregions.

To test for statistically reliable differences in coupling patterns between area 9 and 32d as well as RCZa, we used randomise permutation testing of the coupling patterns in both species independently. We found that the posterior ACC region coupled significantly more strongly with the other cingulate motor areas, pre-SMA and SMA, an area in dlPFC, posterior vlPFC, insular cortex, striatum and putamen. PgACC on the other hand coupled more strongly with

both parts of the frontopolar cortex, posterior cingulate and retrosplenial cortex, anterior vlFC, posterior TPJ, various parts of the STS, ventral striatum and the head of the caudate. Area 9 coupled significantly more strongly with parts of retrosplenial cortex, and areas in the inferior parietal cortex. These significant differences in functional coupling could support the notion of these areas contribute differently to decision making and learning. This is particularly important in the case of pgACC and the area bordering it immediately posterior – these are often collectively referred to as "ACC" but here we demonstrate that their coupling patterns are significantly different (see SI Figure 7-11A). Moreover we conducted an additional analysis searching for significant differences in coupling between posterior and anterior OFC regions, classifying 47o, 14m and 13 as posterior OFC and 47m, 11, and 11m as anterior OFC. As shown in SI Figure 7-11B this analysis yielded significant differences in coupling between posterior and anterior OFC. We conducted another statistical comparison of the coupling patterns associated with vmPFC (Figure 3-2A), central OFC (Figure 3-5C), and lOFC (Figure 3-5A) and confirmed that they each participated in different neural circuits in both species (SI Figure 7-12A). This suggests that not only are there important differences between vmPFC/medial OFC and lateral OFC (259, 275) but that a central OFC region is also distinct.

To further investigate whether the brain areas in cingulate and orbitofrontal cortex that we identified as homologous between species occupy a similar position within the brain networks of the two species we performed the following additional analyses. First, we calculated the dissimilarity matrices of all 23 homologous frontal regions with one another within each species using the same Manhattan distance measure applied to the connectivity fingerprints of the original matching analyses and performed a correlation between the two. The two dissimilarity measures were very highly correlated between species ($\rho = .79$, $p<0.001$).

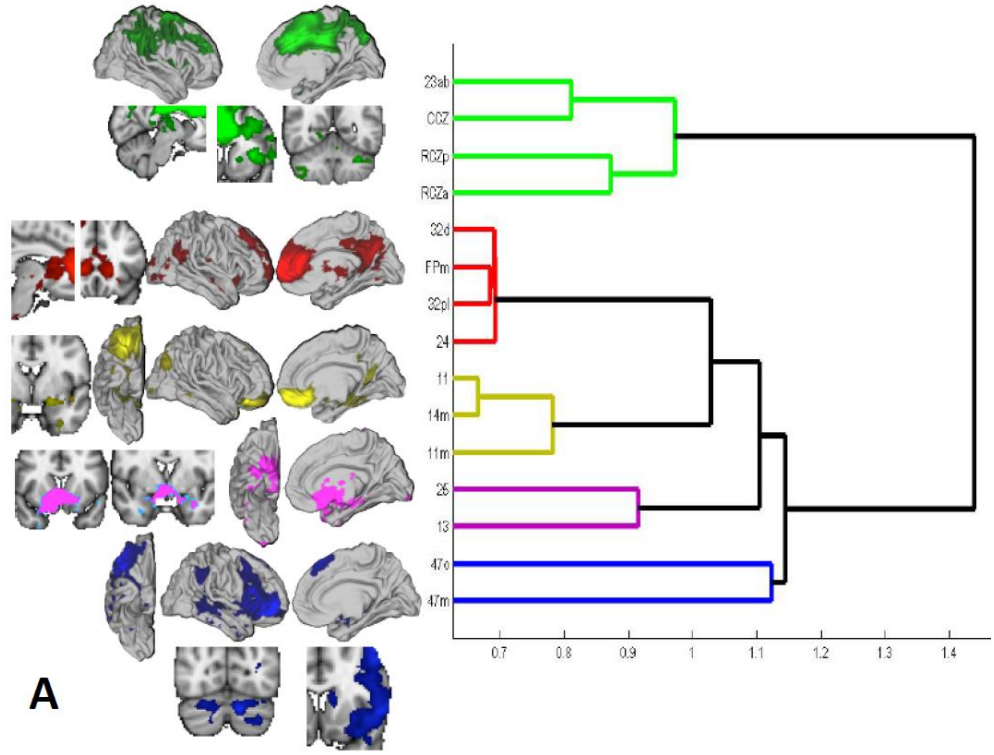
Second, we calculated two measures of the centrality of the medial and orbital frontal regions within the network of target areas. Centrality measures indicate how important a node is within a network. We note that these measures only provide an approximate measure of a

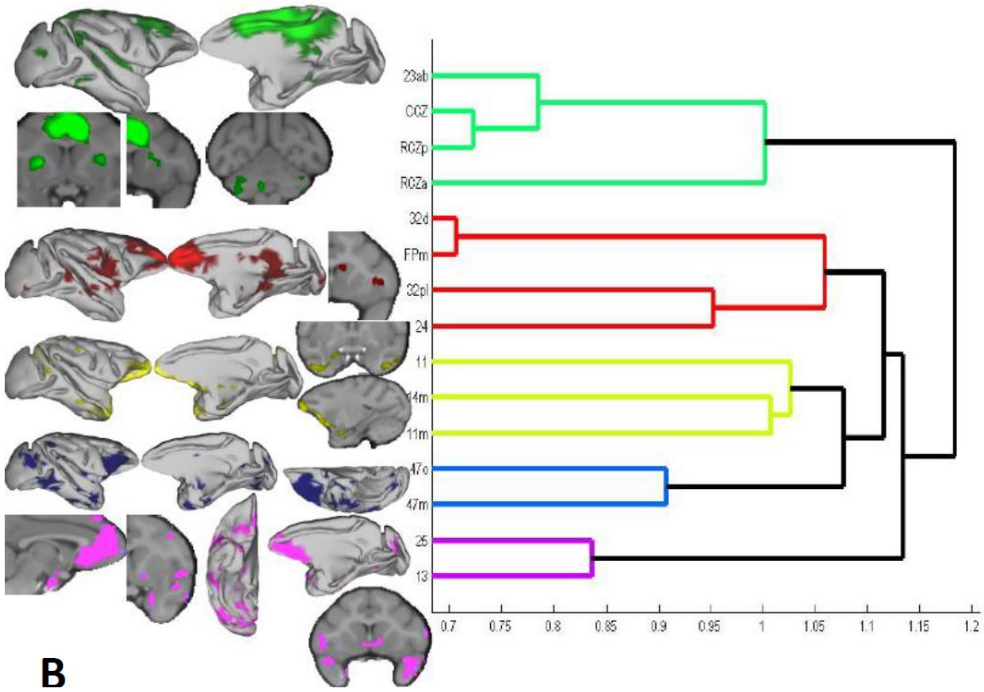
region's centrality, since by necessity we cannot include target areas in parts of the brain for which the homology between species has yet to be established (such as, for instance, large parts of temporal cortex (202)). We converted the Manhattan distance matrices into a similarity matrix for each species and calculated the degree centrality and eigenvector centrality (following (430)). The degree centrality of a region is a measure of the number of connections of that region or, in the context of the current study, the sum of the similarity scores of that region. The eigenvector centrality is a more elaborate measure that not only reflects whether a region has many strong connections, but also whether its neighbours are themselves strongly connected with the rest of the network. Again, the correlations between species in these measures were highly significant (degree centrality: $\rho = .57$, $p = .005$; eigenvector centrality: $\rho = .56$, $p = .005$)."

Finally we performed a hierarchical clustering analysis for both species on the whole-brain functional coupling of all areas derived from the tractography-based parcellation of the human medial and orbital cortex and their proposed monkey equivalents. As opposed to the between-species matching, this analysis was based on whole-brain coupling patterns because clustering was carried out in each species independently. The goal of this analysis was to provide another line of evidence regarding the "relational similarity" of the frontal areas under investigation.

SI Table 7-1 (following page): we show the results of the hierarchical clustering of the parcellation-derived human regions (A) and their proposed monkey homologues (B). The hierarchical clustering analysis did not reveal identical relationships between areas in the two species when they were examined at the finest level. However it can be seen that at the broader level clustering into families of regions is largely similar across species. Five broad groups of areas are identifiable in both species: (1) the cingulate motor areas, (2) the perigenual and medial frontopolar regions, (3) the vmPFC and anterior central OFC regions, (4) the posterior medial and central OFC regions, (5) lateral orbitofrontal cortex. This supports the idea that despite fine grained differences between species it is possible to identify broad similarities in the way that frontal areas interact in circuits in the two species. In order to understand which differences in functional coupling might drive these clustering results we tested for significant differences in coupling between the five families of areas using randomize permutation testing. The results are plotted next to the hierarchical clustering analysis for the respective species. It is apparent that each family showed a distinct coupling

pattern that was significantly different from those of the other areas in both cortical and subcortical regions.





7.1.1.1 Comparison of functional connectivity between areas identified in decision making studies and areas delineated by the tractography-based parcellation

In order to link the areas delineated by the connectivity based parcellation in the second part of the study back to the human decision making areas investigated in the first part of the study we used their respective group-derived functional coupling maps and performed a spatial correlation analysis on them. We here focused on ROIs and parcels in the left hemisphere including an equal number of subjects from both groups both with and without a paracingulate sulcus (noparacingulate_left; n=12; paracingulate_left; n=12). We used the FSL tool fslcc to calculate the spatial correlation of the unthresholded group–derived z-maps (Figure 3-5C).

Area of interest	Fig.	Coordinates (x,y,z) MNI152_standard & Frey et al., 2011			Study
Ventralmost aspect of an area in "Ventromedial PFC [which] encodes the relative chosen value during the decision-making phase"	2a	-6	32	-18	(56)
"A trinket [=noncomestible consumer items] category-dependent value representation was located in anterior mOFC"	2b	-15	57	-9	(254)
"Medial Prefrontal/ Rostral Anterior Cingulate" "tracks the revealed subjective value of delayed monetary rewards"	2c	-1	38	21	(32)
"subgenual anterior cingulate cortex […]may be a critical hub within distributed networks mediating depressive symptoms"	2d	-6	16	-10	(258)
"Dorsomedial prefrontal cortex [represents] modeled values (i.e., other values during self choice and self values during other choice)"	2e	6	42	36	(257)
"Single neuronal activity was recorded in the vmPFC […] from 7 to 15mmanterior to the anterior commissure (AC)"; neurons in anteroventral vmPFC "were persistently more active in the appetitive "reward" block"	3a	*-1*	*14.25*	*-4*	(242)
Neurons in posterodorsal vmPFC "were persistently more active in the aversive punishment block"	3b	*-1.25*	*10*	*0*	(242)
pregenual anterior cingulate cortex, "in the region of the ventral bank of the cingulate sulcus"	3c	*-1*	*16*	*6.25*	(72)
"dorsal anterior cingulate cortex/medial superior frontal cortex" "showed reliable start-cue and sustained activations across all or nearly all tasks [and] […] carried the most reliable error-related signals"	4a	-2	7	50	(186)
"Anterior CIngulate Cortex" "is part of a distributed neural system that is implicated in the representation and updating of decision values"	4b	-6	26	34	(82)
"213 ACC neurons from area 24c in the dorsal bank of the cingulate sulcus"	4c	*-2.75*	*12.25*	*12*	(271)
"Laterak OFC is concerned with the updating of precise associations between stimuli and outcomes during learning, regardless of whether outcomes are positive or negative"	5a	-32	42	-18	(431)
An "area in left OFC that exhibited higher activation levels for higher ambiguity levels as well as for lower winning probabilities"	5b	-43	52	1	(280)
"We tentatively identified the recording region as area 13m"	5c	*-7*	*12.25*	*0*	(273)
"The recordings came mainly from area 10"	5d	*-4.5*	*22.75*	*10.25*	(281)
An area in pregenual ACC is related to cost-benefit decision making. More specifically the costs of making foraging decisions and disinhibition of effortful choices	SI.1a	0	42	26	(54)

SI Table 7-2: Human and monkey areas under investigation. We drew ROIs (6mm isotropic for human and 3mm isotropic for macaque ROIs) centred on these coordinates. Coordinates refer

to MNI152_standard brain as provided by FSL (in the case of human ROIs, non- italicized coordinates) and the standardized monkey brain provided by Frey et al., 2011 (in the case of monkey ROIs, italicized coordinates).

Area	Abbreviation	Coordinates [x,y,z] in MNI152_standard & Frey et al., 2011	
Area 9/46 dorsal	9/46d	[42,36,38]	[-12.75,11.5,9.25]
Area 44v	44v	[52,20,8]	[-17,7.5,1.5]
Supplementary Motor Cortex	SMA	[4,-4,60]	[-2,-0.5,18]
Area 8 medial	8m	[-4,36,56]	[-1.75,16.75,13.25]
Primary Motor Cortex	M1	[-42,-18,48]	[-13.25,-6.5,16]
Primary Somatosensory Cortex	S1	[-44,-30,58]	[-12.25,-9.5,18.5]
Parietal Operculum	ParOp	[-52,-30,22]	[-18,-14.75,12.5]
Anterior inferior Parietal Sulcus	aIPS	[-56,-24,44]	[-13.5,-12,16.75]
Posterior inferior Parietal Sulcus	pIPS	[-34,-48,42]	[-9.75,-16.75,17.75]
Posterior inferior Parietal Lobule	pIPL	[-20,-76,44]	[-8,-24,16.25]
Area 23ab	23ab	[-2,-20,36]	[0.25,-8.5,12.75]
Retrosplenial Cortex	rsplC	[-8,-48,10]	[-2.5,-17,3.75]
Lateral occipital Cortex	latocc	[-56,-60,-10]	[-21.25,-22.5,1]
Perirhinal Cortex	perirhinal	[-46,0,-38]	[-16,-3.75,-13.25]
Temporal pole	temPol	[-44,12,-32]	[-17,4.25,-13.5]
Ventral Striatum	ventrStr	[-14,10,-10]	[-3.5,6.5,-3]
Head of caudate nucleus	Caud	[-14,10,16]	[-5,8.5,4.75]
Putamen	put	[28,6,-6]	[-10.5,-0.5,1.75]
Hippocampus	hippoc	[-26,-20,-16]	[11,-4,-11.25]
Amygdala	amygd	[-24,-4,-18]	[-9.5,0,-11.75]
Hypothalamus	hypoth	[-2,-2,-12]	[-1.25,-4,-7]
Ventral tegmental area	VTA	[0,-20,-16]	[-2.5,-9.75,-6.75]

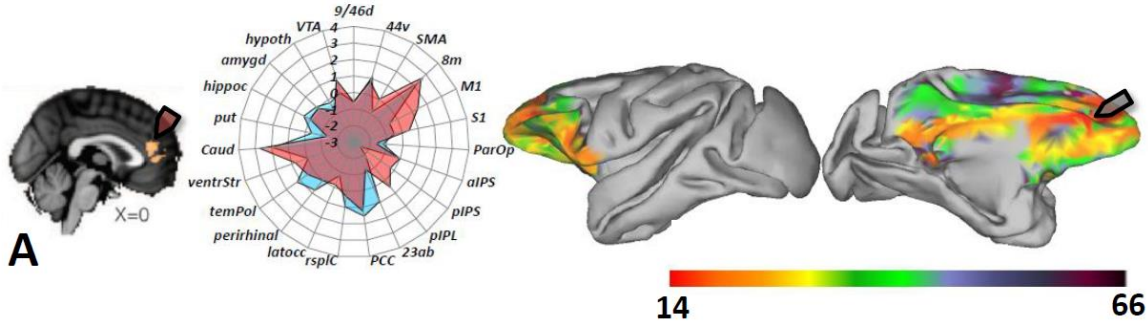
SI Table 7-3: Human and monkey target areas for spider plots. Coordinates refer to MNI152_standard brain as provided by FSL (3rd column) and the standardized monkey brain provided by Frey et al., 2011 (4th column).

Area	Abbrevi-ation	Comments	Supporting references
Area 9/46 dorsal	9/46d	In both comparative cytoarchitectonic and connectivity studies an area in human anterior middle frontal gyrus and superior frontal sulcus has been shown to resembe monkey 9/46d	(196, 235, 237, 238, 419)
Area 44v	44v	Based on a recent investigation of functional coupling patterns of ventrolateral frontal areas in humans and monkeys the ventral and opercular aspects of the pars opercularis gyri frontalis inferioris were linked with monkey area ProM	(113, 205, 237, 238, 350, 432, 433)
Supplementary Motor Cortex (MII, 6m, F3)	SMA	Electrophysiological micro-stimulation, cytoarchitectonic and connectivity studies in monkeys have established an area in the medial aspect of the superior frontal gyrus directly anterior to the primary motor cortex. This area has been linked with an area in the medial superior frontal gyrus posterior to the coronal plane of the anterior commissure in humans on the basis of cytoarchitecture, the activity patterns they exhibit and the connections that it is estimated to have with the motor cortex.	(196, 200, 237, 238, 293)
Area 8 medial	8m	In both comparative cytoarchitectonic and connectivity studies an area in the medial aspect of the human superior frontal gyrus extending down to the paracingulate sulcus has been shown to resemble monkey 8B / 8Bm	(196, 235, 237, 238, 419)
Primary Motor Cortex	M1	Correspondences between cytoarchitecture, activity patterns, estimated connectivity, and effects of stimulation in these areas in humans and monkeys is well established.	(103, 237, 238, 434)
Primary Somatosensory Cortex	S1	Correspondences between cytoarchitecture, activity patterns, and estimated connectivity, in these areas in humans and monkeys is well established.	(237, 293)
Parietal Operculum	ParOp	An area in the human anterior inferior parietal lobule and the parietal operculum, sometimes also referred to as parietal opercular region (PFop),and monkey area PFop/7op has been suggested on the basis of cytoarchitectonic, functional activity patterns and estimates of connectivity.	(192, 230, 237, 238, 318, 432, 435, 436)
Anterior inferior Parietal Sulcus	aIPS	Correspondence between an area in the human anterior inferior parietal sulculs - sometimes also referred to as hIP3, IPS4, DIPSA, aIPS or SPL ant – and the monkey anterior inferior parietal sulcus and anterior intraparietal area (AIP) have been proposed on cytoarchitectonic and functional grounds as well as on the basis of connectivity estimates.	(192, 230, 237, 238, 318, 432, 435, 436)
Posterior	pIPS	An area in the human posterior inferior parietal	(192, 230,

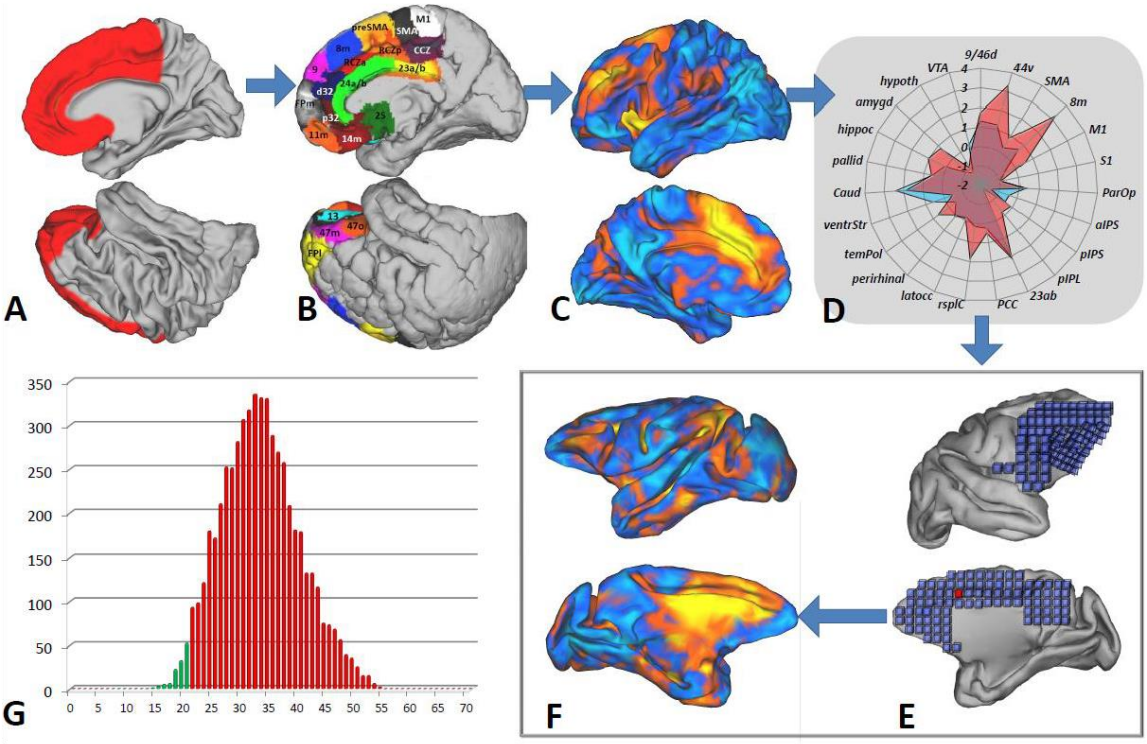
inferior Parietal Sulcus		sulcus, sometimes also referred to as POIPS, has been linked with monkey posterior inferior parietal sulcus on cytoarchitectonic and functional grounds and on the basis of connectivity estimates.	237, 238, 318, 432, 435, 436)
Posterior inferior Parietal Lobule	pIPL	An area in the human posterior inferior parietal lobule, sometimes also referred to as parietal area PG, and monkey posterior inferior parietal lobule area PG/7A has been suggested on cytoarchitectonic and functional grounds as well as on the basis of connectivity estimates.	(192, 230, 237, 238, 318, 432, 435, 436)
Area 23ab	23ab	This region in the posterior part of the cingulate gyrus of the human brain has based been linked with a monkey region in the posterior cingulate gyrus on the basis of cytoarchitecture and estimated connectivity patterns.	(201, 237, 238, 263, 264, 286, 293, 432)
Retrosplenial Cortex	rsplC	An area in the retrosplenial aspect of the human cingulate cortex falls into cytoarchitectonic area 23 and has been linked with monkey area 23 on the basis of cytoarchitecture and estimated connectivity patterns.	(201, 237, 238, 263, 264, 286, 293, 432)
Lateral occipital Cortex	latocc	The lateral occipital cortex in humans has on been linked on functional and anatomical grounds been linked with monkey area TE	(237, 238, 432, 437-440)
Perirhinal Cortex	perirhinal	Cytoarchitectonic, and functional and correspondences between areas 35 and 36 of the perirhinal cortex in humans and monkeys are well established in the literature	(195, 237, 238, 432, 441-443)
Temporal pole	temPol	The human temporopolar cortex (sometimes referred to as area 38, TG) has been linked with monkey area TG on the basis of similarities in cytoarchitecture and estimated connectivity.	(237, 238, 432, 437, 443, 444)
Ventral Striatum	ventrStr	Functional and anatomical structural correspondence between these areas in humans and monkeys is well established.	(237, 238, 445-448)
Head of caudate nucleus	Caud	Functional and anatomical structural correspondence between these areas in humans and monkeys is well established.	(237, 238, 445-448)
Putamen	put	Functional and anatomical structural correspondence between these areas in humans and monkeys is well established.	(237, 238, 445-448)
Hippocampus	hippoc	Functional and anatomical structural correspondence between these areas in humans and monkeys is well established.	(237, 238, 432, 449-452)
Amygdala	amygd	Functional and anatomical structural correspondence between these areas in humans and monkeys is well established.	(237, 238, 289, 432, 437, 445)
Hypothalamus	hypoth	Functional and anatomical structural correspondence between these areas in humans and monkeys is well established.	(237, 238, 289, 432, 445)

Ventral tegmental area	VTA	Functional and anatomical structural correspondence between these areas in humans and monkeys is well established.	(237, 238, 289, 432, 453)

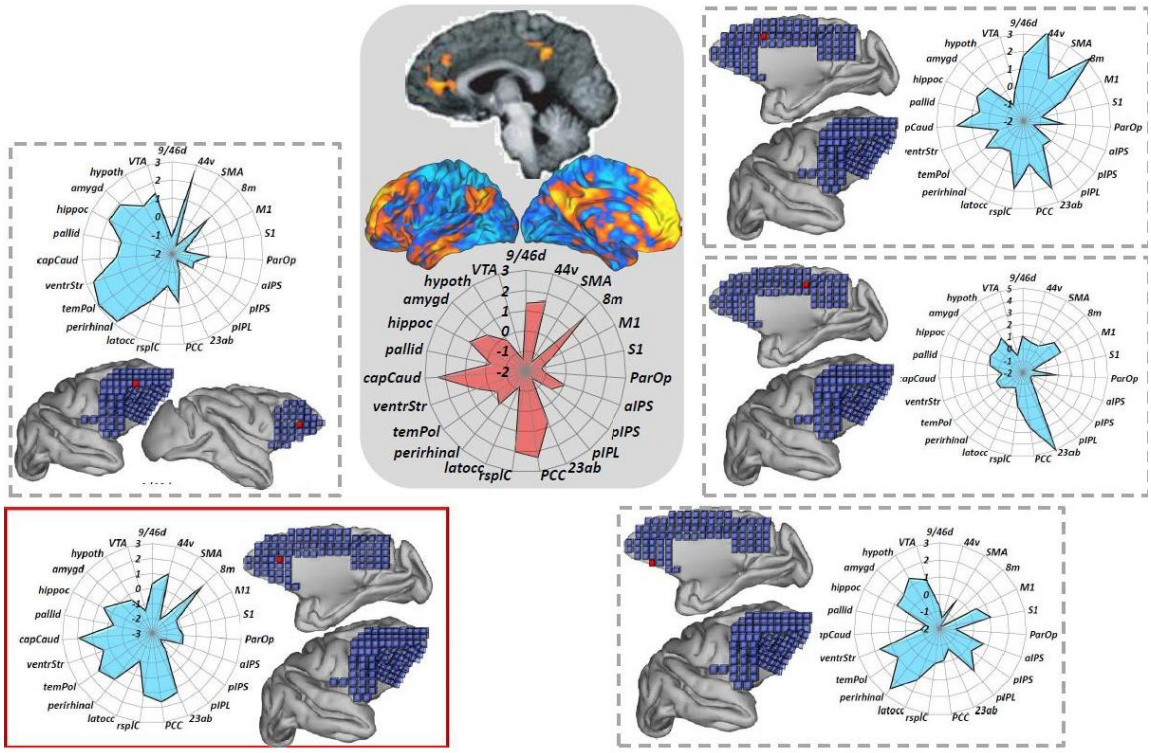
SI Table 7-4: Human and monkey target areas for spider plots. This table provides arguments and references for these region's cross-species homologies.



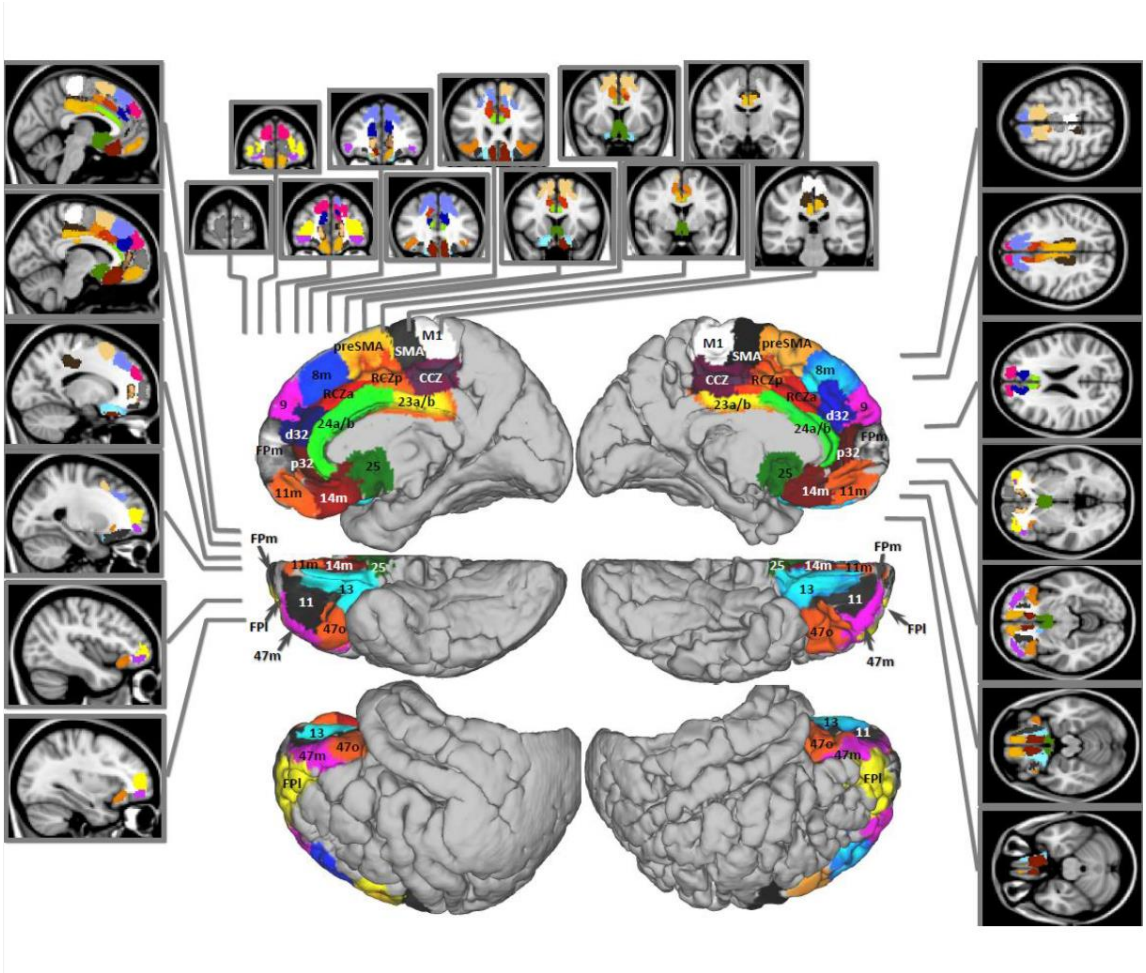
SI Figure 7-1: Human perigenual ACC region (left) linked with cost-benefit decisions (54, 60) could be linked to macaque brain regions (right) via similarities in their coupling patterns (center).



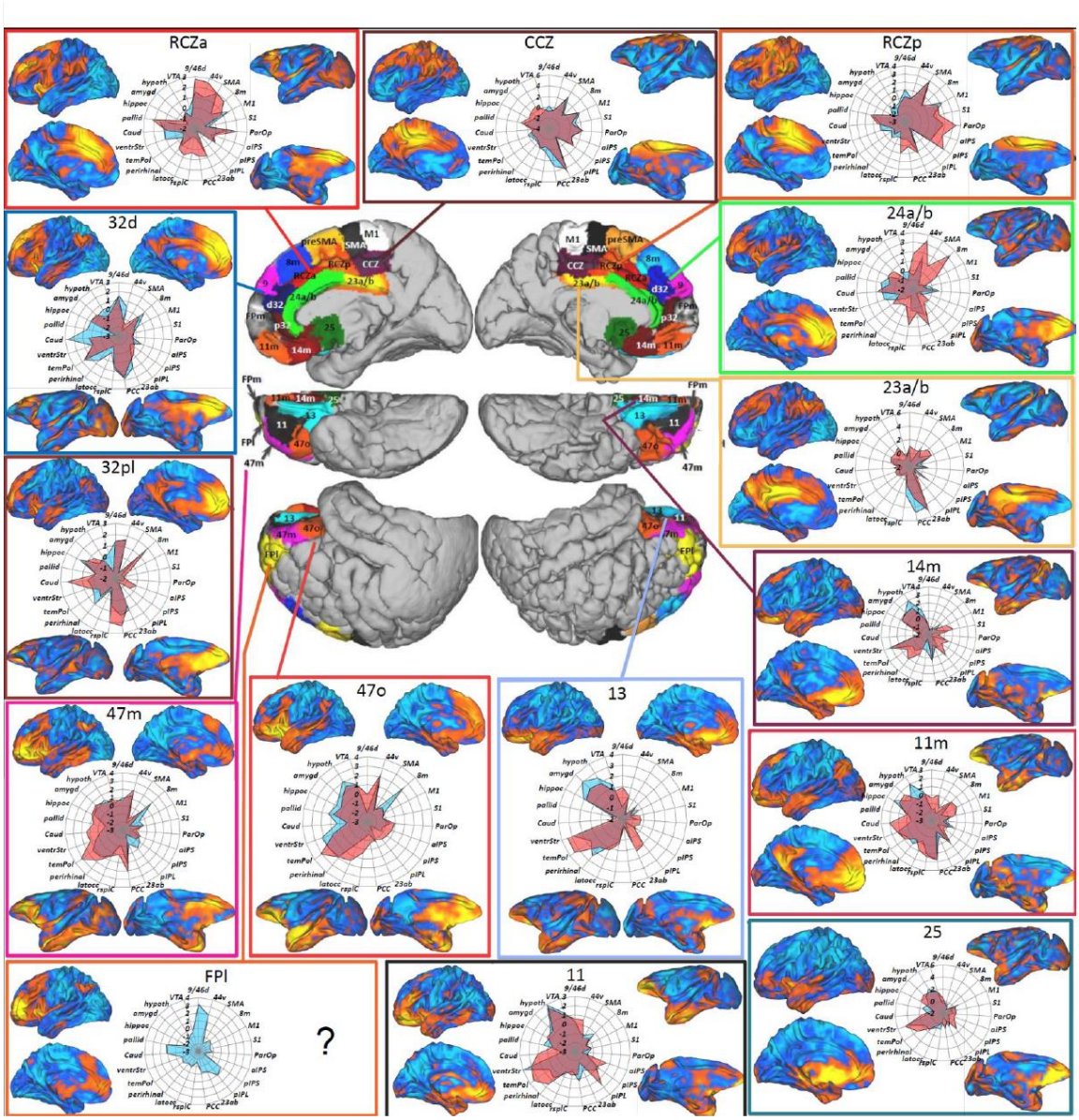
SI Figure 7-2: Overall approach of the second part of the study which used DWI-MRI tractography based parcellation of a medial and orbital frontal cortical ROI (A) to divide this bigger region of cortex into 21 subregions with consistent structural connectivity profiles (B). Resting-state fMRI in the same human participants was used to infer the functional coupling patterns of these subregions (C). Based on the functional coupling with 23 target areas (D), which can be considered to be homologous between humans and monkeys these regions were compared 448 monkey ROIs in cingulate, medial and orbital frontal cortex (E) based on the summed absolute differences in coupling profiles. For each of these human parcels and best match was established in the monkey brain (E-F). The distribution of the summed absolute differences of all 21 human areas with all 448 monkey areas (23 x 448) shares properties with a normal distribution. The distributions of the summed absolute differences (also called Manhattan distance) of all human to monkey matches shared properties with a normal distribution. A formal test confirmed the normality of all distributions (one-sample Kolmogorov-Smirnov test, all $p>0.05$) (G). We considered a summed absolute difference smaller than two standard deviations below the mean of this distribution as an acceptable human to monkey match (green bars in G).



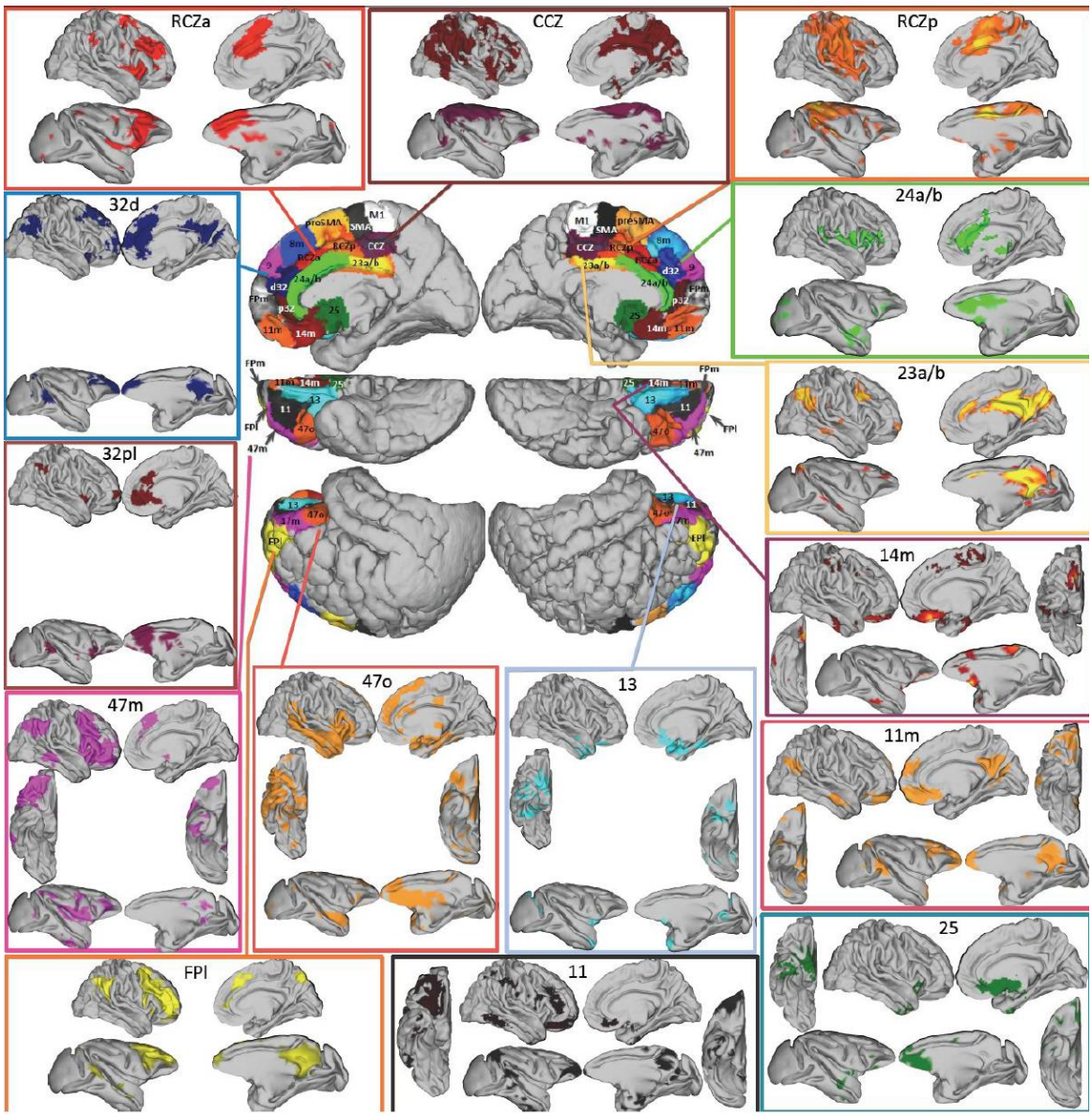
SI Figure 7-3: A further illustration of the "spider-matching approach". A region's coupling pattern (in this case pgACC as proposed by (32)) is compared to 448 monkey ROI's coupling patterns. ROIs in lOFC, posterior cingulate cortex, anterior cingulate cortex, vmPFC, pgACC all show different coupling "fingerprints". However monkey pgACC matches the "Kable-area" best.



SI Figure 7-4: A summary figure showing the results of the DWI-MRI tractography based parcellation with 21 distinct components plotted onto a group-average brain of the paracingulate_left and the noparacingulate_right group. Inserted pictures show slices of the 21 component parts of medial and orbital frontal cortex plotted onto the MNI152_T1_2mm standard brain as provided by FSL.

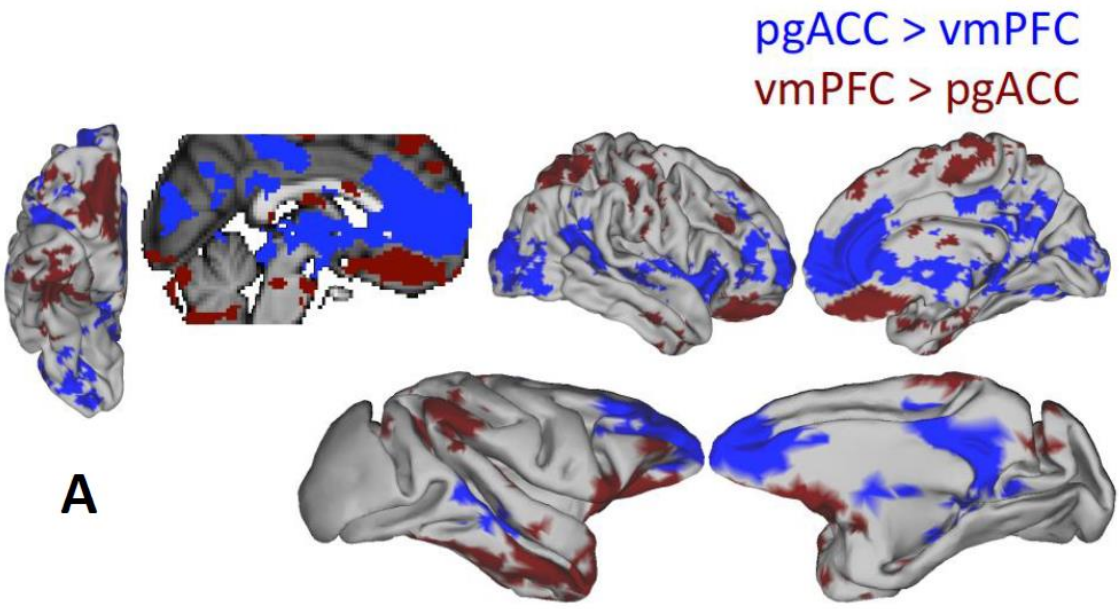


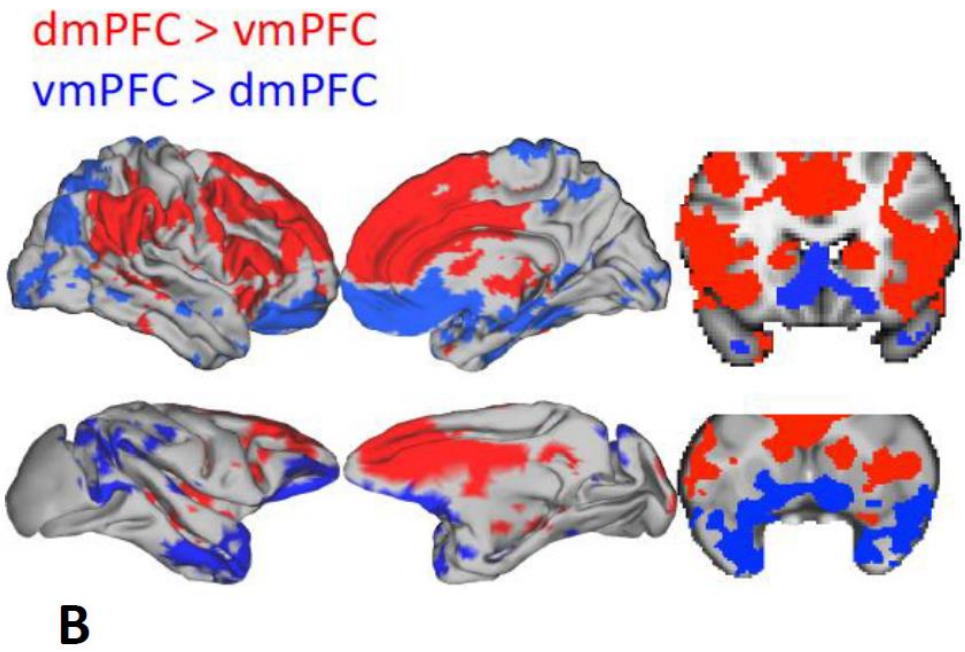
SI Figure 7-5: A summary figure showing the 21 distinct components of the medial and orbital frontal cortex parcellation together with their resting state fMRI based functional coupling profile of (except M1, SMA, pre-SMA, 8m, 9 and FPm, which are mentioned elsewhere). The spiderplots (red) summarize the coupling profiles and compare them to the best monkey area (blue), for which the coupling profile is shown on the right side on each panel.



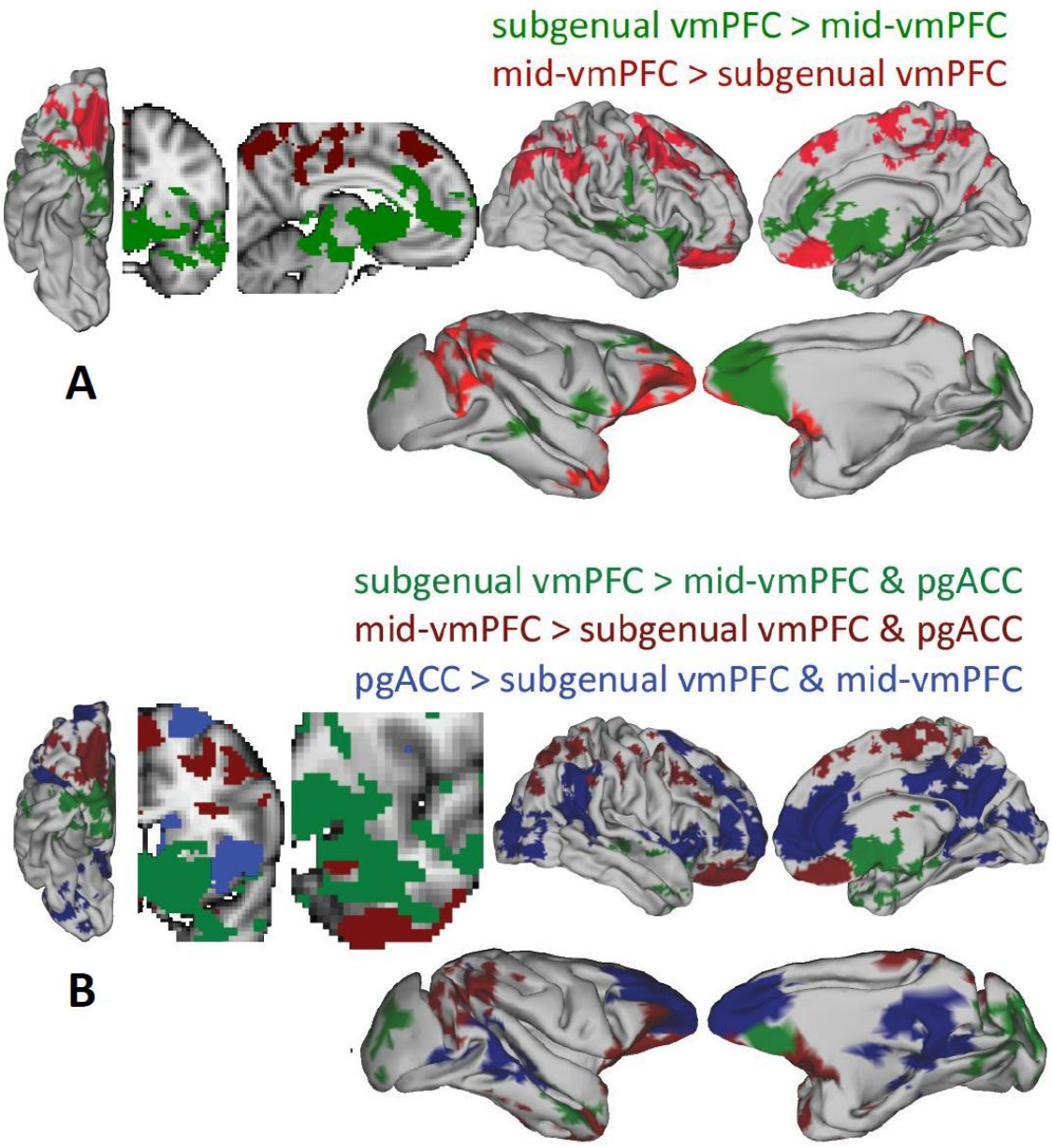
SI Figure 7-6: We also conducted a partial correlation analysis of all areas implicated by the structural-connectivity based parcellation that made up the second part of this study (figure 6). This supplementary analysis was conducted as follows: We extracted the BOLD-time courses of 17 areas derived from the connectivity-based parcellation (23ab, 24, CCZ, RCZp, RCZa, 32d, 32pl, 9, FPm, 11m, 14m, 25, 11, 13, FPl, 47o, 47m) and used them in a partial correlation (all these time-courses competed in one model in the first level analysis on the individual subject level) in order to infer the major differences in functional connectivity of the different sub-regions with the rest of the brain (controlling for confounding time series representing head movement and the Eigen time-series of white matter and corticospinal fluid). These partial correlation results on the individual subject level were subsequently subjected to a group-level analysis using 'nonparametric permutation testing' as implemented in FSL's Randomise tool (454).and cluster-mass-based thresholding ($p<0.05$) to look for significant partial correlations between each of the frontal areas and the rest of the brain. Note that areas FPm and 9 are not plotted because they were outside the focus of this study but their time-courses were included in the partial correlation because the areas, being on the border of the region investigated, appear in some of the figures. The same analysis was then

performed for the areas most likely to be homologous in the monkey. Surface-plots of the results are shown.

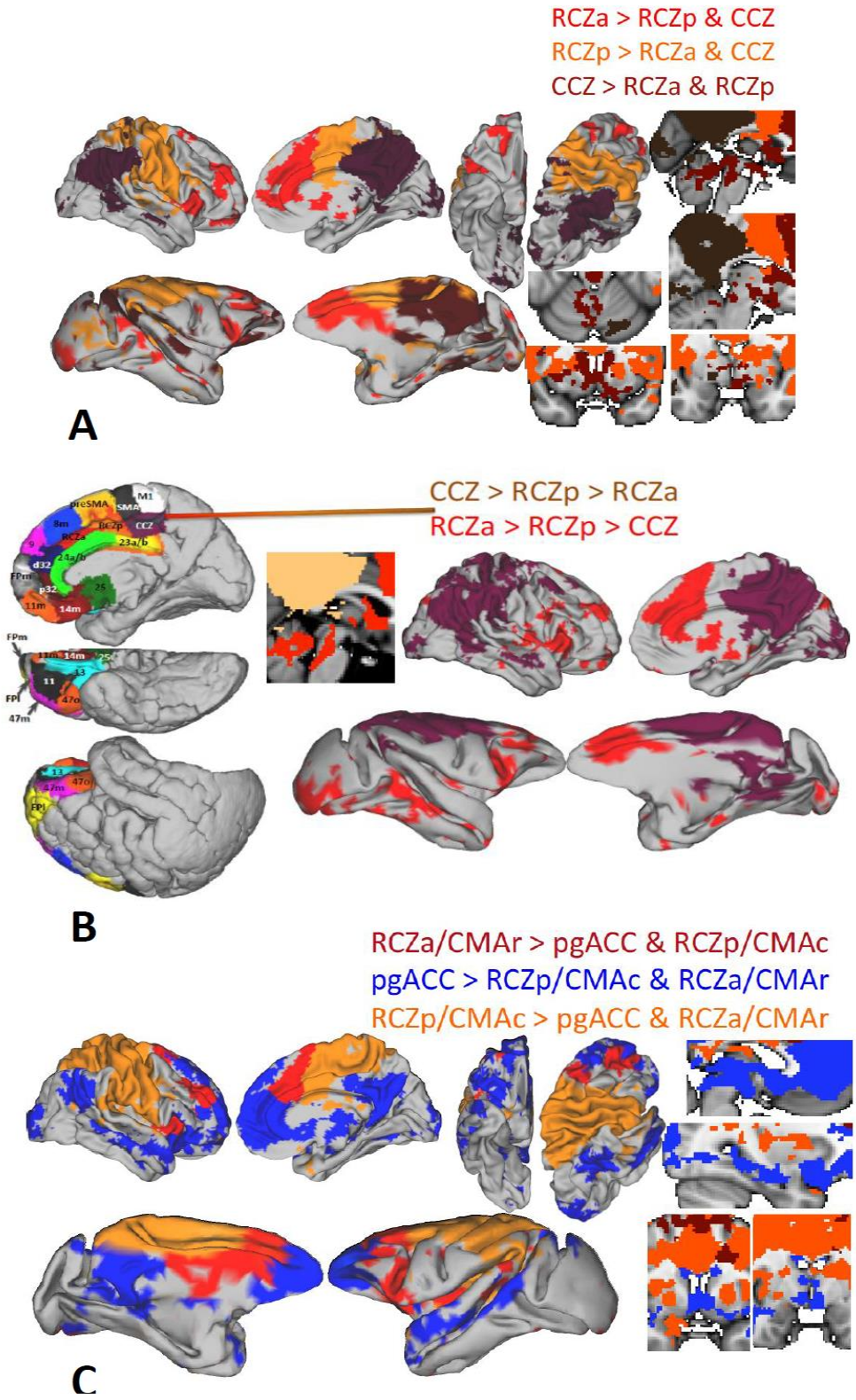




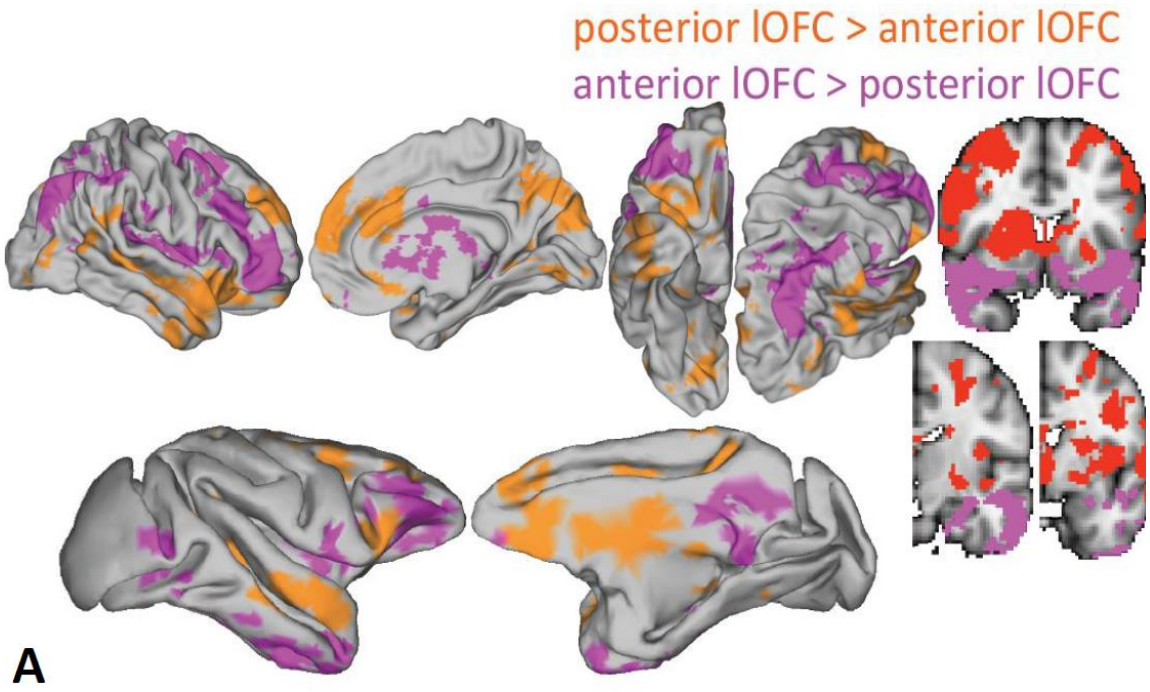
SI Figure 7-7: The coupling patterns between pgACC and vmPFC (A) were tested for significant differences using 'nonparametric permutation testing'. Significant differences between the coupling patterns of dorsomedial and ventromedial PFC (B).

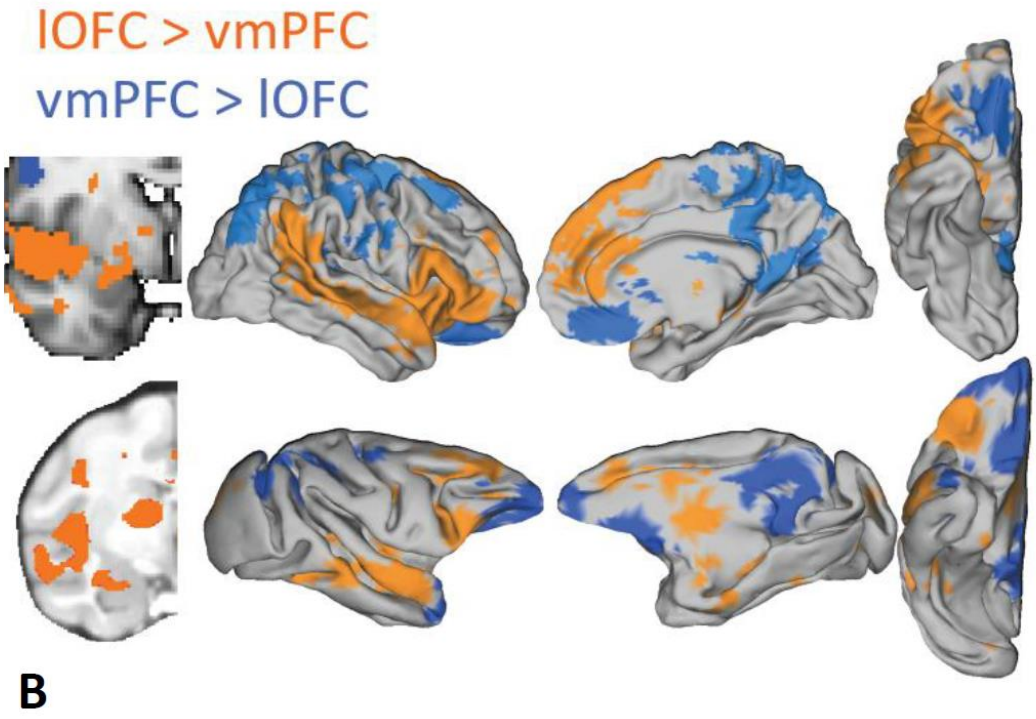


SI Figure 7-8: Significant differences in coupling patterns between subgenual and mid vmPFC (A), and between subgenaul vmPFC, mid-vmPFC, and pgACC (B).

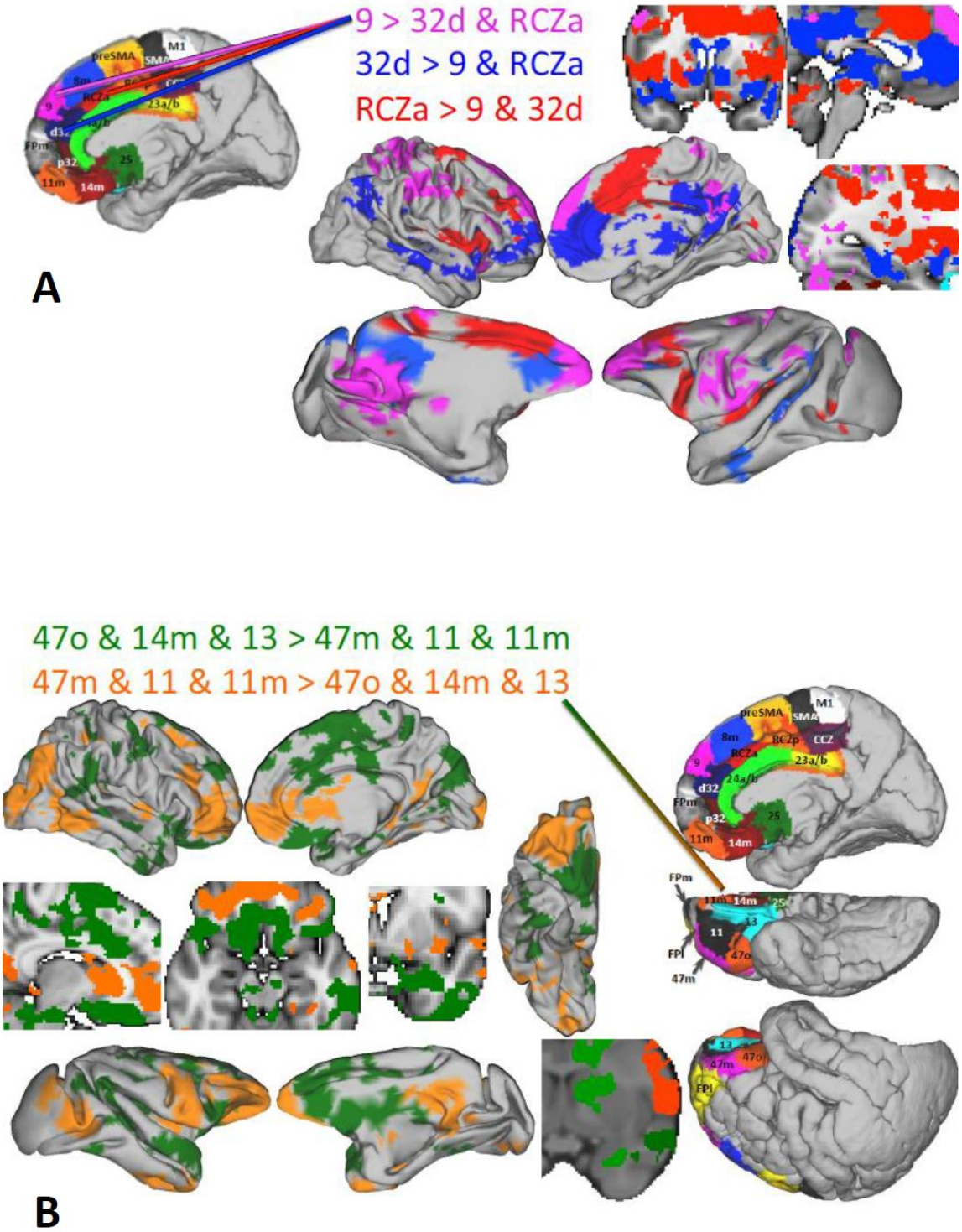


SI Figure 7-9: Significant differences in coupling patterns between RCZa/CMAr, RCZp/CMAv and CCZ/CMAd (A). Significant gradient change in functional coupling from anterior cingulate motor region RCZa/CMAr, via mid cingulate motor region RCZp/CMAv towards posterior cingulate motor region CCZ/CMAd (B). Significant differences in coupling patterns between RCZa/CMAr, RCZp/CMAv and pgACC.

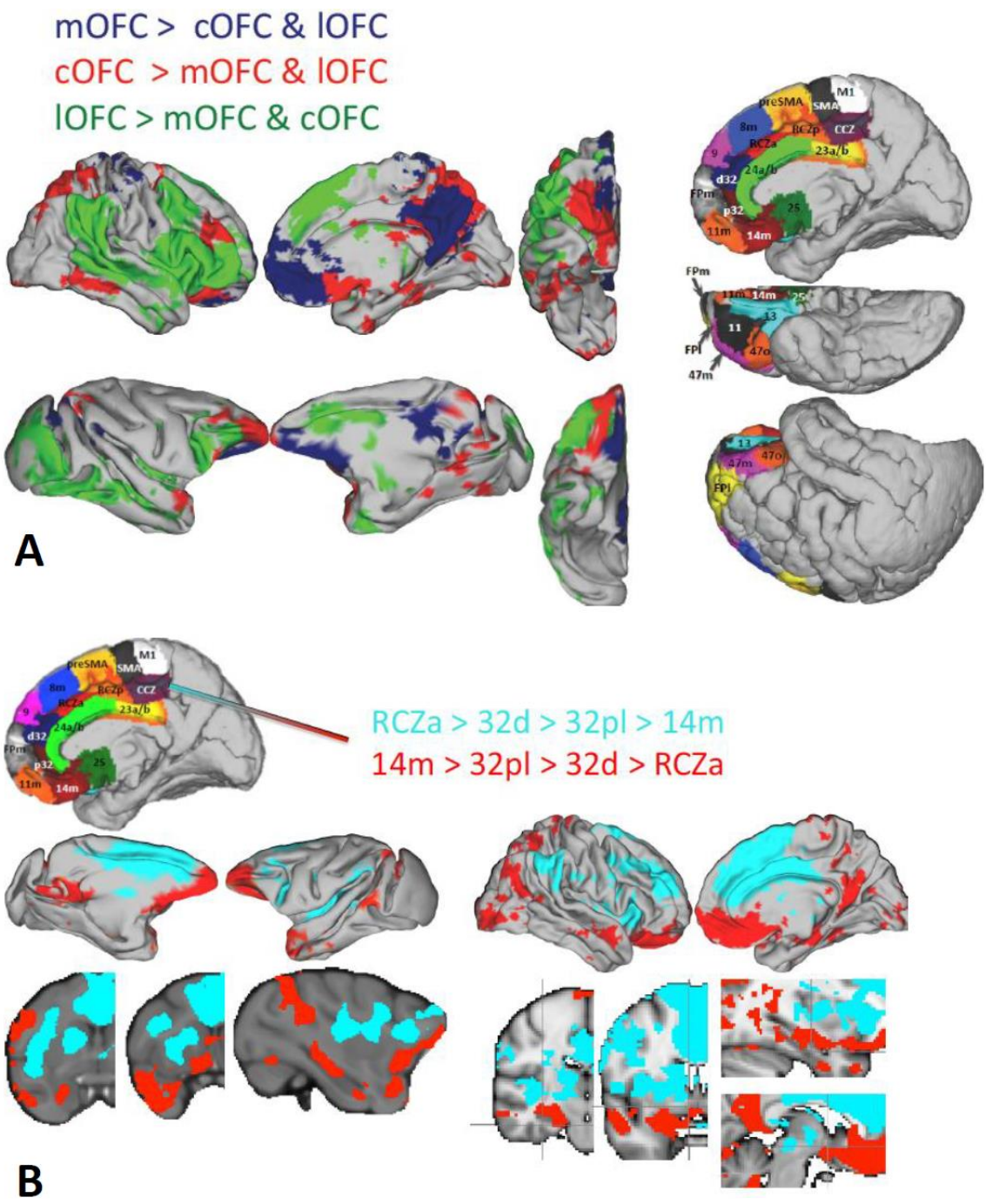




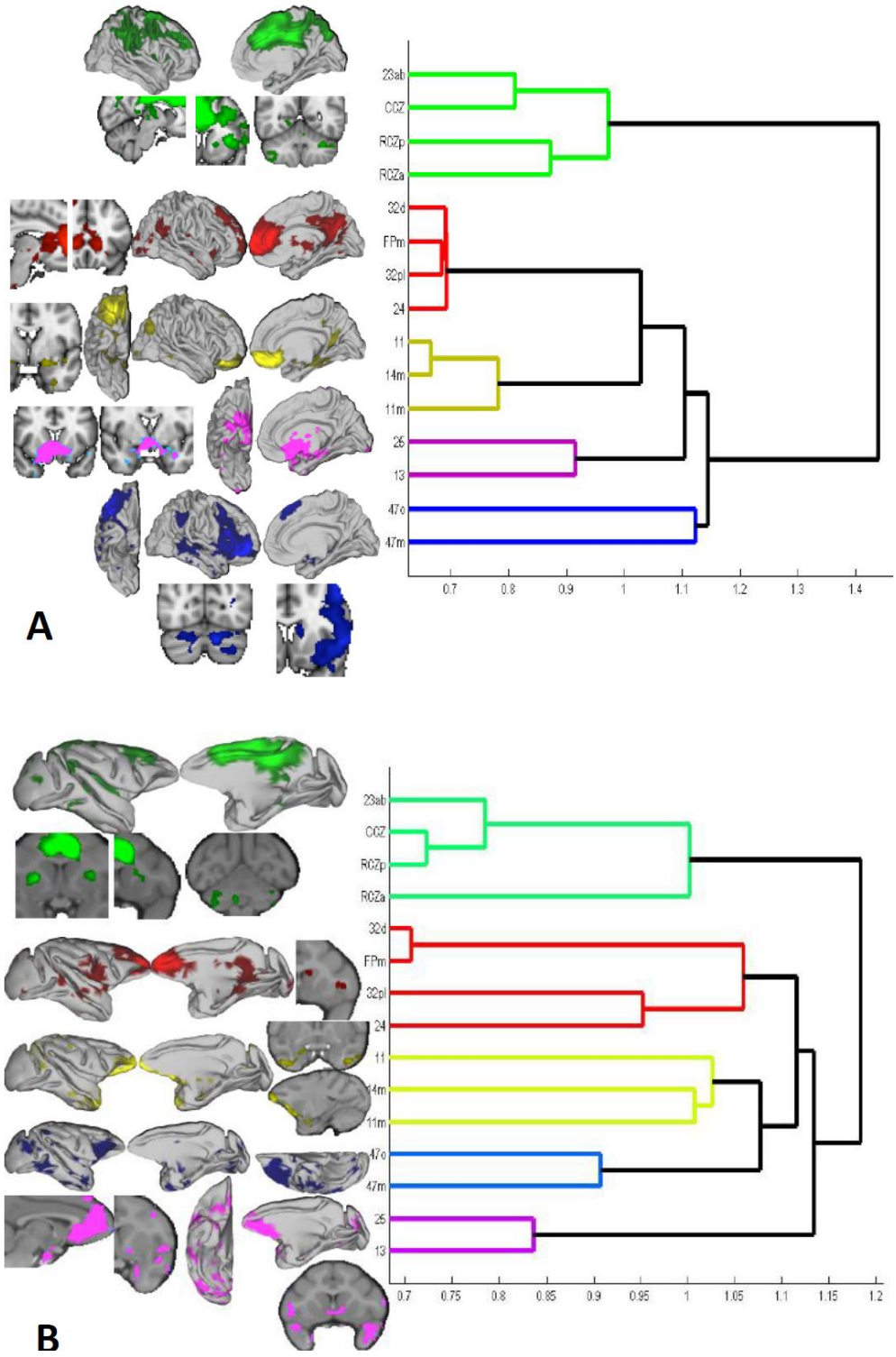
SI Figure 7-10: Significant differences in coupling patterns between anterior and posterior lOFC (A); and between lOFC and vmPFC / mOFCl (B).



SI Figure 7-11: Significant differences in coupling patterns between areas 9, 32d and RCZa/CMAr (A). Significant differences in coupling patterns between anterior (47m, 11, 11m) and posterior (470, 13, 14m) OFC (B).



SI Figure 7-12: Significant differences in coupling patterns between vmPFC / mOFC, cOFC and lOFC (A). Significant gradient effects from the from the most anterior of three cingulate motor areas RCZa, through the perigenual ACC regions 32d and 32pl towards the so-called ventromedial prefrontal cortex or area 14m.



SI Figure 7-13: Hierarchical clustering analysis for both species on the whole-brain functional coupling of all areas derived from the tractography-based parcellation of the human medial and orbital cortex and their proposed monkey equivalents. Clustering into families of regions is largely similar across species. Significant differences of these family's functional coupling patterns next to the hierarchical clustering analysis using the same colours (green: the cingulate motor areas, red: the perigenual and medial frontopolar regions, yellow: the vmPFC and anterior central OFC regions, pink the posterior medial and central OFC regions, blue: lateral orbitofrontal cortex).

8 Bibliography

1. Schopenhauer A (1977) *Schopenhauer, Arthur: Zürcher Ausgabe ; 8 : Parerga und Paralipomena ; 1,2 . Aphorismen zur Lebensweisheit* (Diogenes, Zürich) pp S. 343 - 540.
2. Taleb NN (2007) *The black swan : the impact of the highly improbable* (Allen Lane, London) pp xxviii, 366 p.
3. Anonymous (May 2nd 2015) The new drug warriors. *The Economist* (17).
4. Koob GF (2013) Negative reinforcement in drug addiction: the darkness within. *Curr Opin Neurobiol* 23(4):559-563.
5. Koob GF & Le Moal M (2008) Addiction and the brain antireward system. *Annu Rev Psychol* 59:29-53.
6. Volkow ND & Li TK (2004) Drug addiction: the neurobiology of behaviour gone awry. *Nat Rev Neurosci* 5(12):963-970.
7. Schultz W, Dayan P, & Montague PR (1997) A neural substrate of prediction and reward. *Science* 275(5306):1593-1599.
8. Levy N (2013) Addiction is Not a Brain Disease (and it Matters). *Front Psychiatry* 4:24.
9. Peterson M (2009) *An introduction to decision theory* (Cambridge University Press, Cambridge) pp x, 317 p.
10. Hume D, Norton DF, & Norton MJ (2007) *A treatise of human nature : a critical edition* (Clarendon Press, Oxford).
11. Habermas Jr (1971) *Toward a rational society : student protest, science and politics* (Heinemann Educational, London) pp ix, 132 p.
12. Horkheimer M, Adorno TW, & Cumming J (1973) *Dialectic of enlightenment* (Allen Lane, London) pp xvii, 258 p.
13. Heidegger M (1977) *The question concerning technology, and other essays* (Harper & Row, New York) 1st Ed pp xxxix, 182 p.
14. Todd PM & Gigerenzer G (2012) *Ecological rationality : intelligence in the world* (Oxford University Press, New York) pp xviii, 590 p.
15. Gigerenzer G (2000) *Adaptive thinking : rationality in the real world* (Oxford University Press, New York) pp xi, 344 p.
16. Chau BK, Kolling N, Hunt LT, Walton ME, & Rushworth MF (2014) A neural mechanism underlying failure of optimal choice with multiple alternatives. *Nat Neurosci* 17(3):463-470.
17. Kahneman D & Tversky A (2000) *Choices, values, and frames* (Russell sage Foundation ; Cambridge University Press, New York Cambridge, UK) pp xx, 840 p.
18. Kahneman D (2011) *Thinking, fast and slow* (Farrar, Straus and Giroux, New York) 1st Ed p 499 p.
19. Anonymous (2009) *Kein neues Menschenbild zur Sprache der Hirnforschung* (Suhrkamp, Frankfurt am Main) Orig.-Ausg., 1. Aufl. Ed p 191 S.
20. Glimcher PW (2003) *Decisions, uncertainty, and the brain : the science of neuroeconomics* (MIT Press, Cambridge, Mass. ; London) pp xx, 375 p.
21. Knutson B, Adams CM, Fong GW, & Hommer D (2001) Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci* 21(16):RC159.
22. Platt ML & Glimcher PW (1999) Neural correlates of decision variables in parietal cortex. *Nature* 400(6741):233-238.
23. Caplin A, Dean M, Glimcher PW, & Rutledge RB (2010) Measuring Beliefs and Rewards: A Neuroeconomic Approach. *The quarterly journal of economics* 125(3):923-960.
24. Glimcher PW & Rustichini A (2004) Neuroeconomics: the consilience of brain and decision. *Science* 306(5695):447-452.
25. Louie K & Glimcher PW (2012) Efficient coding and the neural representation of value. *Ann N Y Acad Sci* 1251:13-32.
26. Wang XJ (2002) Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36(5):955-968.
27. Boorman ED, Behrens TE, & Rushworth MF (2011) Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS Biol* 9(6):e1001093.

28. Coricelli G & Nagel R (2009) Neural correlates of depth of strategic reasoning in medial prefrontal cortex. *Proc Natl Acad Sci U S A* 106(23):9163-9168.
29. Glimcher PW, Dorris MC, & Bayer HM (2005) Physiological utility theory and the neuroeconomics of choice. *Games and economic behavior* 52(2):213-256.
30. Dorris MC & Glimcher PW (2004) Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44(2):365-378.
31. Levy DJ & Glimcher PW (2011) Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *J Neurosci* 31(41):14693-14707.
32. Kable JW & Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10(12):1625-1633.
33. Anonymous (2015) More Kirk than Spock. *The Economist* (18).
34. Lewis M (2004) *Moneyball : the art of winning an unfair game* (W.W. Norton, New York ; London) pp xv, 316 p.
35. RP McAfee HM, SH Mialon (April 2010) Do sunk costs matter? *Econ. Inq.* Volume 48(Issue 2):pages 323–336.
36. Pearson JM, Watson KK, & Platt ML (2014) Decision making: the neuroethological turn. *Neuron* 82(5):950-965.
37. Wallis JD (2012) Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nat Neurosci* 15(1):13-19.
38. Watson KK & Platt ML (2012) Social signals in primate orbitofrontal cortex. *Curr Biol* 22(23):2268-2273.
39. Noonan MP, Kolling N, Walton ME, & Rushworth MF (2012) Re-evaluating the role of the orbitofrontal cortex in reward and reinforcement. *Eur J Neurosci* 35(7):997-1010.
40. Berridge KC (2007) The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl)* 191(3):391-431.
41. Levy DJ & Glimcher PW (2012) The root of all value: a neural common currency for choice. *Curr Opin Neurobiol* 22(6):1027-1038.
42. Padoa-Schioppa C (2011) Neurobiology of economic choice: a good-based model. *Annu Rev Neurosci* 34:333-359.
43. Padoa-Schioppa C (2007) Orbitofrontal cortex and the computation of economic value. *Ann N Y Acad Sci* 1121:232-253.
44. Hare TA, Schultz W, Camerer CF, O'Doherty JP, & Rangel A (2011) Transformation of stimulus value signals into motor commands during simple choice. *Proc Natl Acad Sci U S A* 108(44):18120-18125.
45. Gold JI & Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535-574.
46. Rushworth MF, Kolling N, Sallet J, & Mars RB (2012) Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Curr Opin Neurobiol* 22(6):946-955.
47. Rushworth MF, Noonan MP, Boorman ED, Walton ME, & Behrens TE (2011) Frontal cortex and reward-guided learning and decision-making. *Neuron* 70(6):1054-1069.
48. Noonan MP*, et al.* (2010) Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc Natl Acad Sci U S A* 107(47):20547-20552.
49. Rudebeck PH*, et al.* (2008) Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci* 28(51):13775-13785.
50. Camille N, Tsuchida A, & Fellows LK (2011) Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *J Neurosci* 31(42):15048-15052.
51. Behrens TE, Hunt LT, Woolrich MW, & Rushworth MF (2008) Associative learning of social value. *Nature* 456(7219):245-249.
52. Behrens TE, Hunt LT, & Rushworth MF (2009) The computation of social behavior. *Science* 324(5931):1160-1164.
53. Daw ND, Gershman SJ, Seymour B, Dayan P, & Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69(6):1204-1215.
54. Kolling N, Behrens TE, Mars RB, & Rushworth MF (2012) Neural mechanisms of foraging. *Science* 336(6077):95-98.
55. Hayden BY, Pearson JM, & Platt ML (2011) Neuronal basis of sequential foraging decisions in a patchy environment. *Nat Neurosci* 14(7):933-939.

56. Boorman ED, Behrens TE, Woolrich MW, & Rushworth MF (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.
57. Walton ME, Behrens TE, Buckley MJ, Rudebeck PH, & Rushworth MF (2010) Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron* 65(6):927-939.
58. Arrow KJ (1970) *Essays in the theory of risk-bearing* (North-Holland, Amsterdam ; Oxford) pp vii, 278 p.
59. Tervo DG*, et al.* (2014) Behavioral variability through stochastic choice and its gating by anterior cingulate cortex. *Cell* 159(1):21-32.
60. Kolling N, Wittmann M, & Rushworth MF (2014) Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron* 81(5):1190-1202.
61. Krebs JR & Davies NB (1997) *Behavioural ecology : an evolutionary approach* (Blackwell Science, Oxford) 4th Ed pp viii, 456 p.
62. Thomas Caraco WUB, Gina M Gregory, Jonathan A Newman, Gregg M Recer, Susan M Zwicker (1990) Risk-sensitivity: ambient temperature affects foraging choice. *Animal Behaviour* 39
(2):338-345.
63. Fisher I (1930) *The theory of interest : as determined by impatience to spend income and opportunity to invest it* (Macmillan, New York) pp xxvii, 566 p.
64. Sandmo A (2011) *Economics evolving : a history of economic thought* (Princeton University Press, Princeton, N.J. ; Oxford) pp viii, 489 p.
65. Green L, Myerson J, Lichtman D, Rosen S, & Fry A (1996) Temporal discounting in choice between delayed rewards: the role of age and income. *Psychol Aging* 11(1):79-84.
66. Lin H & Epstein LH (2014) Living in the moment: effects of time perspective and emotional valence of episodic thinking on delay discounting. *Behav Neurosci* 128(1):12-19.
67. Hare TA, Camerer CF, & Rangel A (2009) Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 324(5927):646-648.
68. Kouneiher F, Charron S, & Koechlin E (2009) Motivation and cognitive control in the human prefrontal cortex. *Nat Neurosci* 12(7):939-945.
69. Kerns JG*, et al.* (2004) Anterior cingulate conflict monitoring and adjustments in control. *Science* 303(5660):1023-1026.
70. McGuire JT & Kable JW (2015) Medial prefrontal cortical activity reflects dynamic re-evaluation during voluntary persistence. *Nat Neurosci* 18(5):760-766.
71. Mischel W, Ebbesen EB, & Zeiss AR (1972) Cognitive and attentional mechanisms in delay of gratification. *J Pers Soc Psychol* 21(2):204-218.
72. Amemori K & Graybiel AM (2012) Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making. *Nat Neurosci* 15(5):776-785.
73. Sharp C, Monterosso J, & Montague PR (2012) Neuroeconomics: a bridge for translational research. *Biol Psychiatry* 72(2):87-92.
74. Hasler G (2012) Can the neuroeconomics revolution revolutionize psychiatry? *Neurosci Biobehav Rev* 36(1):64-78.
75. Tabibnia G*, et al.* (2011) Different forms of self-control share a neurocognitive substrate. *J Neurosci* 31(13):4805-4810.
76. Leung M*, et al.* (2011) Gray matter in first-episode schizophrenia before and after antipsychotic drug treatment. Anatomical likelihood estimation meta-analyses with sample size weighting. *Schizophr Bull* 37(1):199-211.
77. Honea R, Crow TJ, Passingham D, & Mackay CE (2005) Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. *Am J Psychiatry* 162(12):2233-2245.
78. Yoshimura S*, et al.* (2014) Cognitive behavioral therapy for depression changes medial prefrontal and ventral anterior cingulate cortex activity associated with self-referential processing. *Soc Cogn Affect Neurosci* 9(4):487-493.
79. Houben K, Nederkoorn C, Wiers RW, & Jansen A (2011) Resisting temptation: decreasing alcohol-related affect and drinking behavior by training response inhibition. *Drug Alcohol Depend* 116(1-3):132-136.
80. Sweeney PD, Anderson K, & Bailey S (1986) Attributional style in depression: a meta-analytic review. *J Pers Soc Psychol* 50(5):974-991.

81. Yeh ZT & Liu SI (2007) Depressive realism: evidence from false interpersonal perception. *Psychiatry Clin Neurosci* 61(2):135-141.
82. Behrens TE, Woolrich MW, Walton ME, & Rushworth MF (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10(9):1214-1221.
83. Noonan MP*, et al.* (2014) A neural circuit covarying with social hierarchy in macaques. *PLoS Biol* 12(9):e1001940.
84. Greenier KD*, et al.* (1999) Individual differences in reactivity to daily events: examining the roles of stability and level of self-esteem. *J Pers* 67(1):185-208.
85. Andrews PW TJJ (2009) The bright side of being blue: depression as an adaptation for analyzing complex problems. *Psychol Rev.* 3(116):620-654.
86. Walkup J (1990) On the measurement of delusions. *Br J Med Psychol* 63 (Pt 4):365-368.
87. Garety P (1985) Delusions: problems in definition and measurement. *Br J Med Psychol* 58 (Pt 1):25-34.
88. Mirowsky J & Ross CE (1983) Paranoia and the structure of powerlessness. *Am Sociol Rev* 48(2):228-239.
89. LaRusso L (1978) Sensitivity of paranoid patients to nonverbal cues. *J Abnorm Psychol* 87(5):463-471.
90. Lyon HM, Kaney S, & Bentall RP (1994) The defensive function of persecutory delusions. Evidence from attribution tasks. *Br J Psychiatry* 164(5):637-646.
91. An SK*, et al.* (2010) Attribution bias in ultra-high risk for psychosis and first-episode schizophrenia. *Schizophr Res* 118(1-3):54-61.
92. Kinderman P & Bentall RP (2000) Self-discrepancies and causal attributions: studies of hypothesized relationships. *Br J Clin Psychol* 39 (Pt 3):255-273.
93. Young HF & Bentall RP (1997) Probabilistic reasoning in deluded, depressed and normal subjects: effects of task difficulty and meaningful versus non-meaningful material. *Psychol Med* 27(2):455-465.
94. Gilbert DT, Pelham BW, & Krull DS (1988) On cognitive busyness: When person perceivers meet persons perceived. *Journal of personality and social psychology* 54(5):733.
95. Davis PJ & Gibson MG (2000) Recognition of posed and genuine facial expressions of emotion in paranoid and nonparanoid schizophrenia. *J Abnorm Psychol* 109(3):445-450.
96. Kruglanski AW & Webster DM (1996) Motivated closing of the mind: "seizing" and "freezing". *Psychol Rev* 103(2):263-283.
97. Farris HH & Revlin R (1989) Sensible reasoning in two tasks: rule discovery and hypothesis evaluation. *Mem Cognit* 17(2):221-232.
98. Bentall RP & Slade PD (1985) Reality testing and auditory hallucinations: a signal detection analysis. *Br J Clin Psychol* 24 (Pt 3):159-169.
99. Rankin PM & O'Carroll PJ (1995) Reality discrimination, reality monitoring and disposition towards hallucination. *British Journal of Clinical Psychology* 34(4):517-528.
100. Crockett MJ, Clark L, & Robbins TW (2009) Reconciling the role of serotonin in behavioral inhibition and aversion: acute tryptophan depletion abolishes punishment-induced inhibition in humans. *J Neurosci* 29(38):11993-11999.
101. Zola-Morgan S (1995) Localization of brain function: the legacy of Franz Joseph Gall (1758-1828). *Annu Rev Neurosci* 18:359-383.
102. Catani M & ffytche DH (2005) The rises and falls of disconnection syndromes. *Brain* 128(Pt 10):2224-2239.
103. Dum RP & Strick PL (2002) Motor areas in the frontal lobe of the primate. *Physiol Behav* 77(4-5):677-682.
104. Catani M & Mesulam M (2008) What is a disconnection syndrome? *Cortex* 44(8):911-913.
105. James W (1890) The principles of psychology, Vol I. (Henry Holt and Co, Inc,, New York, NY), p p.
106. Marin O (2012) Interneuron dysfunction in psychiatric disorders. *Nat Rev Neurosci* 13(2):107-120.
107. Kwan KY*, et al.* (2012) Species-dependent posttranscriptional regulation of NOS1 by FMRP in the developing cerebral cortex. *Cell* 149(4):899-911.
108. Campbell AW (1905) *Histological studies on the localisation of cerebral function* (Cambridge University Press, Cambridge) pp xix, 360 p.
109. Brodmann K & Garey L (1999) *Brodmann's 'Localisation in the cerebral cortex'* (Imperial College Press, London) pp xi, 300 p.

110. Amunts K & Zilles K (2001) Advances in cytoarchitectonic mapping of the human cerebral cortex. *Neuroimaging Clin. N. Am.* 11(2):151-169, vii.
111. Mackey S & Petrides M (2010) Quantitative demonstration of comparable architectonic areas within the ventromedial and lateral orbital frontal cortex in the human and the macaque monkey brains. *Eur J Neurosci* 32(11):1940-1950.
112. Eickhoff SB, Schleicher A, Scheperjans F, Palomero-Gallagher N, & Zilles K (2007) Analysis of neurotransmitter receptor distribution patterns in the cerebral cortex. *Neuroimage* 34(4):1317-1330.
113. Amunts K, *et al.* (2010) Broca's region: novel organizational principles and multiple receptor mapping. *PLoS Biol* 8(9).
114. O'Shea J, Johansen-Berg H, Trief D, Gobel S, & Rushworth MF (2007) Functionally specific reorganization in human premotor cortex. *Neuron* 54(3):479-490.
115. Nicolelis MA, *et al.* (2003) Chronic, multisite, multielectrode recordings in macaque monkeys. *Proc Natl Acad Sci U S A* 100(19):11041-11046.
116. Scholtens LH, Schmidt R, de Reus MA, & van den Heuvel MP (2014) Linking macroscale graph analytical organization to microscale neuroarchitectonics in the macaque connectome. *J. Neurosci.* 34(36):12192-12205.
117. Markov NT, *et al.* (2013) The role of long-range connections on the specificity of the macaque interareal cortical network. *Proceedings of the National Academy of Sciences* 110(13):5187-5192.
118. Walsh V & Pascual-Leone A (2003) *Transcranial magnetic stimulation : a neurochronometrics of mind* (MIT Press, Cambridge, Mass. ; London) pp xx, 297 p., [292] p. of plates.
119. Barker AT, Jalinous R, & Freeston IL (1985) Non-invasive magnetic stimulation of human motor cortex. *Lancet* 1(8437):1106-1107.
120. Rossini PM & Rossi S (2007) Transcranial magnetic stimulation: diagnostic, therapeutic, and research potential. *Neurology* 68(7):484-488.
121. Hallett M (2000) Transcranial magnetic stimulation and the human brain. *Nature* 406(6792):147-150.
122. Walsh V & Cowey A (2000) Transcranial magnetic stimulation and cognitive neuroscience. *Nat Rev Neurosci* 1(1):73-79.
123. O'Shea J, Taylor PC, & Rushworth MF (2008) Imaging causal interactions during sensorimotor processing. *Cortex* 44(5):598-608.
124. Driver J, Blankenburg F, Bestmann S, Vanduffel W, & Ruff CC (2009) Concurrent brain-stimulation and neuroimaging for studies of cognition. *Trends Cogn Sci* 13(7):319-327.
125. Wasserman E, *et al.* (2008) *The Oxford handbook of transcranial stimulation* (Oxford University Press, Oxford) pp xiv, 747 p., [748] p. of plates.
126. Miniussi C, Ruzzoli M, & Walsh V (2009) The mechanism of transcranial magnetic stimulation in cognition. *Cortex*.
127. Huang YZ, Edwards MJ, Rounis E, Bhatia KP, & Rothwell JC (2005) Theta burst stimulation of the human motor cortex. *Neuron* 45(2):201-206.
128. Ruff CC, Driver J, & Bestmann S (2009) Combining TMS and fMRI: From 'virtual lesions' to functional-network accounts of cognition. *Cortex* 45(9):1043-1049.
129. Hartwigsen G, *et al.* (2012) Left dorsal premotor cortex and supramarginal gyrus complement each other during rapid action reprogramming. *J Neurosci* 32(46):16162-16171a.
130. Fox MD, Halko MA, Eldaief MC, & Pascual-Leone A (2012) Measuring and manipulating brain connectivity with resting state functional connectivity magnetic resonance imaging (fcMRI) and transcranial magnetic stimulation (TMS). *Neuroimage* 62(4):2232-2243.
131. Akaishi R, Morishima Y, Rajeswaren VP, Aoki S, & Sakai K (2010) Stimulation of the frontal eye field reveals persistent effective connectivity after controlled behavior. *J Neurosci* 30(12):4295-4305.
132. Buch ER, Johnen VM, Nelissen N, O'Shea J, & Rushworth MF (2011) Noninvasive associative plasticity induction in a corticocortical pathway of the human brain. *J Neurosci* 31(48):17669-17679.
133. Johnen VM, *et al.* (2015) Causal manipulation of functional connectivity in a specific neural pathway during behaviour and at rest. *Elife* 4.
134. Mosso A, Raichle ME, Shepherd GM, & Nockels Fabbri C (*Angelo Mosso's Circulation of blood in the human brain* pp xxxii, 202 pages.

135. Logothetis NK, Pauls J, Augath M, Trinath T, & Oeltermann A (2001) Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412(6843):150-157.
136. Logothetis NK (2003) The underpinnings of the BOLD functional magnetic resonance imaging signal. *J Neurosci* 23(10):3963-3971.
137. Kahn I, *et al.* (2011) Characterization of the functional MRI response temporal linearity via optical control of neocortical pyramidal neurons. *J Neurosci* 31(42):15086-15091.
138. Lee JH, *et al.* (2010) Global and local fMRI signals driven by neurons defined optogenetically by type and wiring. *Nature* 465(7299):788-792.
139. Sallet J, *et al.* (2011) Social network size affects neural circuits in macaques. *Science* 334(6056):697-700.
140. Bullmore E & Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10(3):186-198.
141. Coltheart M (2006) What has functional neuroimaging told us about the mind (so far)? *Cortex* 42(3):323-331.
142. Johansen-Berg H & Behrens TEJ (2009) *Diffusion MRI : from quantitative measurement to in-vivo neuroanatomy* (Elsevier/Academic Press, Amsterdam ; Boston) 1st Ed pp xi, 490 p.
143. Morecraft RJ, *et al.* (2002) Localization of arm representation in the corona radiata and internal capsule in the non-human primate. *Brain* 125(Pt 1):176-198.
144. Kuypers HG & Ugolini G (1990) Viruses as transneuronal tracers. *Trends Neurosci* 13(2):71-75.
145. Cai D, Cohen KB, Luo T, Lichtman JW, & Sanes JR (2013) Improved tools for the Brainbow toolbox. *Nat Methods* 10(6):540-547.
146. Livet J, *et al.* (2007) Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* 450(7166):56-62.
147. O'Reilly JX, *et al.* (2013) Causal effect of disconnection lesions on interhemispheric functional connectivity in rhesus monkeys. *Proc Natl Acad Sci U S A* 110(34):13982-13987.
148. Gonzalez RG, *et al.* (1999) Diffusion-weighted MR imaging: diagnostic accuracy in patients imaged within 6 hours of stroke symptom onset. *Radiology* 210(1):155-162.
149. Basser PJ & Jones DK (2002) Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review. *NMR Biomed* 15(7-8):456-467.
150. Westerhausen R, *et al.* (2004) Effects of handedness and gender on macro- and microstructure of the corpus callosum and its subregions: a combined high-resolution and diffusion-tensor MRI study. *Brain Res Cogn Brain Res* 21(3):418-426.
151. Westerhausen R, *et al.* (2003) The influence of handedness and gender on the microstructure of the human corpus callosum: a diffusion-tensor magnetic resonance imaging study. *Neurosci Lett* 351(2):99-102.
152. Gaser C & Schlaug G (2003) Brain structures differ between musicians and non-musicians. *J Neurosci* 23(27):9240-9245.
153. Bengtsson SL, *et al.* (2005) Extensive piano practicing has regionally specific effects on white matter development. *Nat Neurosci* 8(9):1148-1150.
154. Johansen-Berg H, Della-Maggiore V, Behrens TE, Smith SM, & Paus T (2007) Integrity of white matter in the corpus callosum correlates with bimanual co-ordination skills. *Neuroimage* 36 Suppl 2:T16-21.
155. Wolbers T, Schoell ED, & Buchel C (2006) The predictive value of white matter organization in posterior parietal cortex for spatial visualization ability. *Neuroimage* 32(3):1450-1455.
156. Boorman ED, O'Shea J, Sebastian C, Rushworth MF, & Johansen-Berg H (2007) Individual differences in white-matter microstructure reflect variation in functional connectivity during choice. *Curr Biol* 17(16):1426-1431.
157. Smith SM, *et al.* (2006) Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31(4):1487-1505.
158. Smith SM, *et al.* (2007) Acquisition and voxelwise analysis of multi-subject diffusion data with tract-based spatial statistics. *Nat Protoc* 2(3):499-503.
159. Douaud G, *et al.* (2011) DTI measures in crossing-fibre areas: increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimer's disease. *Neuroimage* 55(3):880-890.
160. Johansen-Berg H, *et al.* (2005) Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. *Cereb Cortex* 15(1):31-39.

161. Behrens TE, *et al.* (2003) Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn Reson Med* 50(5):1077-1088.
162. Johansen-Berg H & Rushworth MF (2009) Using diffusion imaging to study human connectional anatomy. *Annu Rev Neurosci* 32:75-94.
163. Shulman RG, Rothman DL, Behar KL, & Hyder F (2004) Energetic basis of brain activity: implications for neuroimaging. *Trends Neurosci* 27(8):489-495.
164. Fox MD & Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 8(9):700-711.
165. Ishai A, Haxby JV, & Ungerleider LG (2002) Visual imagery of famous faces: effects of memory and attention revealed by fMRI. *Neuroimage* 17(4):1729-1741.
166. Sadaghiani S, Hesselmann G, Friston KJ, & Kleinschmidt A (2010) The relation of ongoing brain activity, evoked neural responses, and cognition. *Front Syst Neurosci* 4:20.
167. Honey CJ, *et al.* (2009) Predicting human resting-state functional connectivity from structural connectivity. *Proc Natl Acad Sci U S A* 106(6):2035-2040.
168. Fox MD, Snyder AZ, Vincent JL, & Raichle ME (2007) Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. *Neuron* 56(1):171-184.
169. Smith SM, *et al.* (2009) Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A* 106(31):13040-13045.
170. Sporns O (2015) Cerebral cartography and connectomics. *Philos Trans R Soc Lond B Biol Sci* 370(1668).
171. Sporns O (2013) Structure and function of complex brain networks. *Dialogues Clin Neurosci* 15(3):247-262.
172. Pievani M, de Haan W, Wu T, Seeley WW, & Frisoni GB (2011) Functional network disruption in the degenerative dementias. *Lancet Neurol* 10(9):829-843.
173. Mars RB, *et al.* (2012) On the relationship between the "default mode network" and the "social brain". *Front Hum Neurosci* 6:189.
174. De Martino B, O'Doherty JP, Ray D, Bossaerts P, & Camerer C (2013) In the mind of the market: theory of mind biases value computation during financial bubbles. *Neuron* 79(6):1222-1231.
175. Burgess PW, Scott SK, & Frith CD (2003) The role of the rostral frontal cortex (area 10) in prospective memory: a lateral versus medial dissociation. *Neuropsychologia* 41(8):906-918.
176. Burgess PW, Quayle A, & Frith CD (2001) Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia* 39(6):545-555.
177. Kensinger EA & Schacter DL (2006) Neural processes underlying memory attribution on a reality-monitoring task. *Cereb Cortex* 16(8):1126-1133.
178. Wang XJ, Tegner J, Constantinidis C, & Goldman-Rakic PS (2004) Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory. *Proc Natl Acad Sci U S A* 101(5):1368-1373.
179. Hunt LT, *et al.* (2012) Mechanisms underlying cortical activity during value-guided choice. *Nat Neurosci* 15(3):470-476, S471-473.
180. Jocham G, Hunt LT, Near J, & Behrens TE (2012) A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. *Nat Neurosci* 15(7):960-961.
181. Friederici AD & Gierhan SM (2012) The language network. *Curr Opin Neurobiol*.
182. Aron AR (2007) The neural basis of inhibition in cognitive control. *Neuroscientist* 13(3):214-228.
183. Higo T, Mars RB, Boorman ED, Buch ER, & Rushworth MF (2011) Distributed and causal influence of frontal operculum in task control. *Proc Natl Acad Sci U S A* 108(10):4230-4235.
184. Neubert FX, Mars RB, Buch ER, Olivier E, & Rushworth MF (2010) Cortical and subcortical interactions during action reprogramming and their related white matter pathways. *Proc Natl Acad Sci U S A* 107(30):13240-13245.
185. Brass M, Derrfuss J, Forstmann B, & von Cramon DY (2005) The role of the inferior frontal junction area in cognitive control. *Trends Cogn Sci* 9(7):314-316.
186. Dosenbach NU, *et al.* (2006) A core system for the implementation of task sets. *Neuron* 50(5):799-812.
187. Petrides M & Pandya DN (2002) Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *Eur J Neurosci* 16(2):291-310.

188. Romanski LM (2012) Integration of faces and vocalizations in ventral prefrontal cortex: implications for the evolution of audiovisual speech. *Proc Natl Acad Sci U S A* 109 Suppl 1:10717-10724.
189. Petrides M, Cadoret G, & Mackey S (2005) Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. *Nature* 435(7046):1235-1238.
190. Rizzolatti G & Sinigaglia C (2010) The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat Rev Neurosci* 11(4):264-274.
191. Passingham RE & Wise SP (2012) *The neurobiology of the prefrontal cortex : anatomy, evolution, and the origin of insight* (Oxford University Press, Oxford) pp xxii, 399 p.
192. Mars RB*, et al.* (2011) Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. *J Neurosci* 31(11):4087-4100.
193. Power JD*, et al.* (2011) Functional network organization of the human brain. *Neuron* 72(4):665-678.
194. Vincent JL*, et al.* (2007) Intrinsic functional architecture in the anaesthetized monkey brain. *Nature* 447(7140):83-86.
195. Margulies DS*, et al.* (2009) Precuneus shares intrinsic functional architecture in humans and monkeys. *Proc Natl Acad Sci U S A* 106(47):20069-20074.
196. Sallet J*, et al.* (2013) The organization of dorsal frontal cortex in humans and macaques. *J Neurosci* 33(30):12255-12274.
197. Tomassini V*, et al.* (2007) Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. *J Neurosci* 27(38):10259-10269.
198. Zhang Y, Brady M, & Smith S (2001) Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 20(1):45-57.
199. Behrens TE, Berg HJ, Jbabdi S, Rushworth MF, & Woolrich MW (2007) Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* 34(1):144-155.
200. Johansen-Berg H*, et al.* (2004) Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proc Natl Acad Sci U S A* 101(36):13335-13340.
201. Beckmann M, Johansen-Berg H, & Rushworth MF (2009) Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J Neurosci* 29(4):1175-1190.
202. Mars RB, Sallet J, Neubert FX, & Rushworth MF (2013) Connectivity profiles reveal the relationship between brain areas for social cognition in human and monkey temporoparietal cortex. *Proc Natl Acad Sci U S A* 110(26):10806-10811.
203. Belmalih A*, et al.* (2009) Multimodal architectonic subdivision of the rostral part (area F5) of the macaque ventral premotor cortex. *J Comp Neurol* 512(2):183-217.
204. Anwander A, Tittgemeyer M, von Cramon D, Friederici A, & Knosche T (2007) Connectivity-Based Parcellation of Broca's Area. *Cereb Cortex* 17(4):816-825.
205. Amunts K*, et al.* (1999) Broca's region revisited: cytoarchitecture and intersubject variability. *J Comp Neurol* 412(2):319-341.
206. Petrides M & Pandya DN (2009) Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biol* 7(8):e1000170.
207. Rudebeck PH, Buckley MJ, Walton ME, & Rushworth MF (2006) A role for the macaque anterior cingulate gyrus in social valuation. *Science* 313(5791):1310-1312.
208. Chang SW, Gariepy JF, & Platt ML (2012) Neuronal reference frames for social decisions in primate frontal cortex. *Nat Neurosci*.
209. Petrides M & Pandya D (2007) Efferent Association Pathways from the Rostral Prefrontal Cortex in the Macaque Monkey. *Journal of Neuroscience*.
210. Schubotz RI, Anwander A, Knosche TR, von Cramon DY, & Tittgemeyer M (2010) Anatomical and functional parcellation of the human lateral premotor cortex. *Neuroimage* 50(2):396-408.
211. Luppino G, Murata A, Govoni P, & Matelli M (1999) Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the the ventral premotor cortex (areas F5 and F4). *Experimental Brain Research* 128:181-187.

212. Borra E, Gerbella M, Rozzi S, & Luppino G (2011) Anatomical evidence for the involvement of the macaque ventrolateral prefrontal area 12r in controlling goal-directed actions. *J Neurosci* 31(34):12351-12363.
213. Nelissen K, Luppino G, Vanduffel W, Rizzolatti G, & Orban GA (2005) Observing others: multiple action representation in the frontal lobe. *Science* 310(5746):332-336.
214. Nelissen K, *et al.* (2011) Action observation circuits in the macaque monkey cortex. *J Neurosci* 31(10):3743-3756.
215. Duncan J & Owen AM (2000) Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci* 23(10):475-483.
216. Verbruggen F, Aron AR, Stevens MA, & Chambers CD (2010) Theta burst stimulation dissociates attention and action updating in human inferior frontal cortex. *Proc Natl Acad Sci U S A* 107(31):13966-13971.
217. Chikazoe J, *et al.* (2009) Preparation to inhibit a response complements response inhibition during performance of a stop-signal task. *J Neurosci* 29(50):15870-15877.
218. Genovesio A, Tsujimoto S, & Wise SP (2012) Encoding goals but not abstract magnitude in the primate prefrontal cortex. *Neuron* 74(4):656-662.
219. Koechlin E & Jubault T (2006) Broca's area and the hierarchical organization of human behavior. *Neuron* 50(6):963-974.
220. Scott BH, Mishkin M, & Yin P (2012) Monkeys have a limited form of short-term memory in audition. *Proc Natl Acad Sci U S A* 109(30):12237-12241.
221. Gaffan D & Harrison S (1991) Auditory-visual associations, hemispheric specialization and temporal-frontal interaction in the rhesus monkey. *Brain* 114:2133-2144.
222. Koechlin E (2011) Frontal pole function: what is specifically human? *Trends Cogn Sci* 15(6):241; author reply 243.
223. Frith CD & Frith U (2007) Social cognition in humans. *Curr Biol* 17(16):R724-732.
224. Krubitzer L (2007) The magnificent compromise: cortical field evolution in mammals. *Neuron* 56(2):201-208.
225. Mantini D, Corbetta M, Romani GL, Orban GA, & Vanduffel W (2013) Evolutionarily novel functional networks in the human brain? *J Neurosci* 33(8):3259-3275.
226. Yeo BT, *et al.* (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106(3):1125-1165.
227. Klein JC, *et al.* (2007) Connectivity-based parcellation of human cortex using diffusion MRI: Establishing reproducibility, validity and observer independence in BA 44/45 and SMA/pre-SMA. *Neuroimage* 34(1):204-211.
228. Zhang Y, *et al.* (2012) Connectivity-Based Parcellation of the Human Posteromedial Cortex. *Cereb Cortex*.
229. Wang J, *et al.* (2012) Tractography-based parcellation of the human left inferior parietal lobule. *Neuroimage* 63(2):641-652.
230. Caspers S, *et al.* (2008) The human inferior parietal lobule in stereotaxic space. *Brain Struct Funct* 212(6):481-495.
231. Scheperjans F, *et al.* (2008) Probabilistic maps, morphometry, and variability of cytoarchitectonic areas in the human superior parietal cortex. *Cereb Cortex* 18(9):2141-2157.
232. Dyrby TB, *et al.* (2007) Validation of in vitro probabilistic tractography. *Neuroimage* 37(4):1267-1277.
233. Nelson SM, *et al.* (2010) A parcellation scheme for human left lateral parietal cortex. *Neuron* 67(1):156-170.
234. O'Reilly JX, Beckmann CF, Tomassini V, Ramnani N, & Johansen-Berg H (2010) Distinct and overlapping functional zones in the cerebellum defined by resting state functional connectivity. *Cereb Cortex* 20(4):953-965.
235. Petrides M & Pandya DN (1999) Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur J Neurosci* 11(3):1011-1036.
236. Carmichael ST & Price JL (1994) Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *The Journal of Comparative Neurology* 346:366-402.
237. Paxinos G, Huang XF, & Toga AW (2000) *The rhesus monkey brain in stereotaxic coordinates* (Academic, San Diego ; London) p 163 p.

238. Saleem KS & Logothetis NK (2007) *A combined MRI and histology atlas of the rhesus monkey brain in stereotaxic coordinates* (Academic, London ; Burlington, MA) pp ix, 326 p.
239. Rangel A & Hare T (2010) Neural computations associated with goal-directed choice. *Curr Opin Neurobiol* 20(2):262-270.
240. O'Doherty JP (2011) Contributions of the ventromedial prefrontal cortex to goal-directed action selection. *Ann N Y Acad Sci* 1239:118-129.
241. Bouret S & Richmond BJ (2010) Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. *J Neurosci* 30(25):8591-8601.
242. Monosov IE & Hikosaka O (2012) Regionally distinct processing of rewards and punishments by the primate ventromedial prefrontal cortex. *J Neurosci* 32(30):10318-10330.
243. Padoa-Schioppa C & Cai X (2011) The orbitofrontal cortex and the computation of subjective value: consolidated concepts and new perspectives. *Ann N Y Acad Sci* 1239:130-137.
244. Cole MW, Yeung N, Freiwald WA, & Botvinick M (2009) Cingulate cortex: diverging data from humans and monkeys. *Trends in neurosciences* 32(11):566-574.
245. Cai X & Padoa-Schioppa C (2012) Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. *J Neurosci* 32(11):3791-3808.
246. Ridderinkhof KR, Ullsperger M, Crone EA, & Nieuwenhuis S (2004) The role of the medial frontal cortex in cognitive control. *Science* 306(5695):443-447.
247. Passingham RE, Stephan KE, & Kotter R (2002) The anatomical basis of functional localization in the cortex. *Nat Rev Neurosci* 3(8):606-616.
248. Neubert FX, Mars RB, Thomas AG, Sallet J, & Rushworth MF (2014) Comparison of Human Ventral Frontal Cortex Areas for Cognitive Control and Language with Areas in Monkey Frontal Cortex. *Neuron* 81:700-713.
249. Plassmann H, O'Doherty J, & Rangel A (2007) Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J Neurosci* 27(37):9984-9988.
250. Glascher J, Hampton AN, & O'Doherty JP (2009) Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb Cortex* 19(2):483-495.
251. Lim SL, O'Doherty JP, & Rangel A (2011) The Decision Value Computations in the vmPFC and Striatum Use a Relative Value Code That is Guided by Visual Attention. *J Neurosci* 31(37):13214-13223.
252. Lebreton M, Jorge S, Michel V, Thirion B, & Pessiglione M (2009) An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* 64(3):431-439.
253. Cooper JC, Kreps TA, Wiebe T, Pirkl T, & Knutson B (2010) When giving is good: ventromedial prefrontal cortex activation for others' intentions. *Neuron* 67(3):511-521.
254. McNamee D, Rangel A, & O'Doherty JP (2013) Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nat Neurosci* 16(4):479-485.
255. De Martino B, Fleming SM, Garrett N, & Dolan RJ (2013) Confidence in value-based choice. *Nat Neurosci* 16(1):105-110.
256. Barron HC, Dolan RJ, & Behrens TE (2013) Online evaluation of novel choices by simultaneous representation of multiple memories. *Nat Neurosci* 16(10):1492-1498.
257. Nicolle A*, et al.* (2012) An agent independent axis for executed and modeled choice in medial prefrontal cortex. *Neuron* 75(6):1114-1121.
258. Johansen-Berg H*, et al.* (2008) Anatomical connectivity of the subgenual cingulate region targeted with deep brain stimulation for treatment-resistant depression. *Cereb Cortex* 18(6):1374-1383.
259. Grabenhorst F & Rolls ET (2011) Value, pleasure and choice in the ventral prefrontal cortex. *Trends in cognitive sciences* 15(2):56-67.
260. Rolls ET, Grabenhorst F, & Parris BA (2010) Neural systems underlying decisions about affective odors. *Journal of cognitive neuroscience* 22(5):1069-1082.
261. Rolls ET, Grabenhorst F, & Deco G (2010) Decision-making, errors, and confidence in the brain. *J Neurophysiol* 104(5):2359-2374.
262. Rolls ET, Grabenhorst F, & Deco G (2010) Choice, difficulty, and confidence in the brain. *Neuroimage* 53(2):694-706.

263. Vogt BA, Vogt L, Farber NB, & Bush G (2005) Architecture and neurocytology of monkey cingulate gyrus. *J Comp Neurol* 485(3):218-239.
264. Morecraft RJ, Cipolloni PB, Stilwell-Morecraft KS, Gedney MT, & Pandya DN (2004) Cytoarchitecture and cortical connections of the posterior cingulate and adjacent somatosensory fields in the rhesus monkey. *J Comp Neurol* 469(1):37-69.
265. Murray EA, Wise SP, & Drevets WC (2011) Localization of dysfunction in major depressive disorder: prefrontal cortex and amygdala. *Biol Psychiatry* 69(12):e43-54.
266. Braver TS, Barch DM, Gray JR, Molfese DL, & Snyder A (2001) Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cereb Cortex* 11(9):825-836.
267. Frank MJ, Samanta J, Moustafa AA, & Sherman SJ (2007) Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. *Science* 318(5854):1309-1312.
268. Picard N & Strick PL (1996) Motor areas of the medial wall: a review of their location and functional activation. *Cerebral Cortex* 6:342-353.
269. Walton ME, Devlin JT, & Rushworth MFS (2004) Interactions between decision making and performance monitoring within prefrontal cortex. *Nat Neurosci* 7(11):1259-1265.
270. O'Reilly JX*, et al.* (2013) Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proc Natl Acad Sci U S A* 110(38):E3660-3669.
271. Kennerley SW, Behrens TE, & Wallis JD (2011) Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat Neurosci* 14(12):1581-1589.
272. Matsumoto M, Matsumoto K, Abe H, & Tanaka K (2007) Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci* 10(5):647-656.
273. Padoa-Schioppa C & Assad JA (2008) The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat Neurosci* 11(1):95-102.
274. Noonan MP, Mars RB, & Rushworth MF (2011) Distinct Roles of Three Frontal Cortical Areas in Reward-Guided Behavior. *Journal of Neuroscience* 31(40):14399-14412.
275. Kringelbach ML & Rolls ET (2004) The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog Neurobiol* 72(5):341-372.
276. Rolls E (2008) Functions of the orbitofrontal and pregenual cingulate cortex in taste, olfaction, appetite and emotion. . *Acta Physiologica Hungarica* 95:131-164.
277. Morrison SE & Salzman CD (2009) The convergence of information about rewarding and aversive stimuli in single neurons. *J Neurosci* 29(37):11471-11483.
278. Rich EL & Wallis JD (2014) Medial-lateral organization of the orbitofrontal cortex. *Journal of cognitive neuroscience* 26(7):1347-1362.
279. Klein-Flugge MC, Barron HC, Brodersen KH, Dolan RJ, & Behrens TE (2013) Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. *J Neurosci* 33(7):3202-3211.
280. Levy I, Snell J, Nelson AJ, Rustichini A, & Glimcher PW (2010) Neural representation of subjective value under risk and ambiguity. *J Neurophysiol* 103(2):1036-1047.
281. Tsujimoto S, Genovesio A, & Wise SP (2010) Evaluating self-generated decisions in frontal pole cortex of monkeys. *Nat Neurosci* 13(1):120-126.
282. Johansen-Berg H*, et al.* (2004) Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proc Natl Acad Sci U S A* 101(36):13335-13340.
283. Bludau S*, et al.* (2013) Cytoarchitecture, probability maps and functions of the human frontal pole. *Neuroimage*.
284. Paus T*, et al.* (1996) Human cingulate and paracingulate sulci: pattern, variability, asymetry, and probabilistic map. *Cerebral Cortex* 6:207-214.
285. Sallet J, *et al.* (2013) The Organization of Dorsal Frontal Cortex in Humans and Macaques. *The Journal of Neuroscience* 33(30):12255-12274.
286. Vogt BA, Nimchinsky EA, Vogt LJ, & Hof PR (1995) Human cingulate cortex: surface features, flat maps, and cytoarchitecture. *J Comp Neurol* 359(3):490-506.
287. Vogt BA*, et al.* (2013) Cingulate area 32 homologies in mouse, rat, macaque and human: cytoarchitecture and receptor architecture. *J Comp Neurol* 521(18):4189-4204.
288. Kahnt T, Chang LJ, Park SQ, Heinzle J, & Haynes JD (2012) Connectivity-based parcellation of the human orbitofrontal cortex. *J Neurosci* 32(18):6240-6250.
289. Ongur D & Price JL (2000) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex* 10(3):206-219.
290. Noonan MP, Sallet J, Rudebeck PH, Buckley MJ, & Rushworth MF (2010) Does the medial orbitofrontal cortex have a role in social valuation? *Eur J Neurosci* 31(12):2341-2351.

291. Rudebeck PH & Murray EA (2011) Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. *J Neurosci* 31(29):10569-10578.
292. Rudebeck PH & Murray EA (2011) Balkanizing the primate orbitofrontal cortex: distinct subregions for comparing and contrasting values. *Ann N Y Acad Sci* 1239(1):1-13.
293. Morecraft RJ*, et al.* (2012) Cytoarchitecture and cortical connections of the anterior cingulate and adjacent somatomotor fields in the rhesus monkey. *Brain Res Bull* 87(4-5):457-497.
294. Walton ME, Bannerman DM, Alterescu K, & Rushworth MFS (2003) Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *J Neurosci* 23:6475-6479.
295. Kennerley SW, Walton ME, Behrens TE, Buckley MJ, & Rushworth MF (2006) Optimal decision making and the anterior cingulate cortex. *Nature Neurosci* 9:940-947.
296. Nelissen N, Stokes M, Nobre AC, & Rushworth MF (2013) Frontal and parietal cortical interactions with distributed visual representations during selective attention and action selection. *J Neurosci* 33(42):16443-16458.
297. Yoshida W & Ishii S (2006) Resolution of uncertainty in prefrontal cortex. *Neuron* 50(5):781-789.
298. Koechlin E & Hyafil A (2007) Anterior prefrontal function and the limits of human decision-making. *Science* 318(5850):594-598.
299. Volman I, Roelofs K, Koch S, Verhagen L, & Toni I (2011) Anterior prefrontal cortex inhibition impairs control over social emotional actions. *Curr Biol* 21(20):1766-1770.
300. Fleming SM, Weil RS, Nagy Z, Dolan RJ, & Rees G (2010) Relating introspective accuracy to individual differences in brain structure. *Science* 329(5998):1541-1543.
301. Amiez C & Petrides M (2012) Neuroimaging Evidence of the Anatomo-Functional Organization of the Human Cingulate Motor Areas. *Cereb Cortex*.
302. Frey S*, et al.* (2011) An MRI based average macaque monkey stereotaxic atlas and space (MNI monkey space). *Neuroimage* 55(4):1435-1442.
303. Keller SS, Crow T, Foundas A, Amunts K, & Roberts N (2009) Broca's area: nomenclature, anatomy, typology and asymmetry. *Brain Lang* 109(1):29-48.
304. Brodmann K (1909) *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues* (Barth, Leipzig) pp x, 324 p.
305. Thiebaut de Schotten M, Dell'Acqua F, Valabregue R, & Catani M (2012) Monkey to human comparative anatomy of the frontal lobe association tracts. *Cortex* 48(1):82-96.
306. Yeterian EH, Pandya DN, Tomaiuolo F, & Petrides M (2011) The cortical connectivity of the prefrontal cortex in the monkey brain. *Cortex*.
307. Heilman KM & Watson RT (2008) The disconnection apraxias. *Cortex* 44(8):975-982.
308. Epelbaum S*, et al.* (2008) Pure alexia as a disconnection syndrome: new diffusion imaging evidence for an old concept. *Cortex* 44(8):962-974.
309. Geschwind N (1965) Disconnexion syndromes in animals and man. II. *Brain* 88(3):585-644.
310. Geschwind N (1965) Disconnexion syndromes in animals and man. I. *Brain* 88(2):237-294.
311. Sporns O (2011) *Networks of the brain* (MIT Press, Cambridge, Mass. ; London) pp xi, 412 p., 418 p. of plates.
312. Doricchi F, Thiebaut de Schotten M, Tomaiuolo F, & Bartolomeo P (2008) White matter (dis)connections and gray matter (dys)functions in visual neglect: gaining insights into the brain networks of spatial awareness. *Cortex* 44(8):983-995.
313. Buckner RL*, et al.* (2009) Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *J Neurosci* 29(6):1860-1873.
314. Sporns O, Honey CJ, & Kotter R (2007) Identification and classification of hubs in brain networks. *PLoS One* 2(10):e1049.
315. van den Heuvel MP & Sporns O (2013) Network hubs in the human brain. *Trends in cognitive sciences* 17(12):683-696.
316. Bava S*, et al.* (2010) Longitudinal characterization of white matter maturation during adolescence. *Brain research* 1327:38-46.
317. Sullivan EV, Rohlfing T, & Pfefferbaum A (2010) Quantitative fiber tracking of lateral and interhemispheric white matter systems in normal aging: relations to timed performance. *Neurobiology of aging* 31(3):464-481.

318. Caspers S, *et al.* (2013) Organization of the human inferior parietal lobule based on receptor architectonics. *Cereb Cortex* 23(3):615-628.
319. Margulies DS & Petrides M (2013) Distinct parietal and temporal connectivity profiles of ventrolateral frontal areas involved in language production. *J Neurosci* 33(42):16846-16852.
320. Yeterian EH, Pandya DN, Tomaiuolo F, & Petrides M (2012) The cortical connectivity of the prefrontal cortex in the monkey brain. *Cortex* 48(1):58-81.
321. Morecraft R, Geula C, & Mesulam MM (1992) Cytoarchitecture and neural afferents of orbitofrontal cortex in the brain of the monkey. *Journal of Comparative Neurology* 323(3):341-358.
322. Pievani M, de Haan W, Wu T, Seeley WW, & Frisoni GB (2011) Functional network disruption in the degenerative dementias. *The Lancet Neurology* 10(9):829-843.
323. Supekar K, *et al.* (2013) Brain hyperconnectivity in children with autism and its links to social deficits. *Cell reports* 5(3):738-747.
324. Geschwind DH & Levitt P (2007) Autism spectrum disorders: developmental disconnection syndromes. *Current opinion in neurobiology* 17(1):103-111.
325. Filippini N, *et al.* (2009) Distinct patterns of brain activity in young carriers of the APOE-epsilon4 allele. *Proc Natl Acad Sci U S A* 106(17):7209-7214.
326. Petrella J, Sheldon F, Prince S, Calhoun V, & Doraiswamy P (2011) Default mode network connectivity in stable vs progressive mild cognitive impairment. *Neurology* 76(6):511-517.
327. Seeley WW, Crawford RK, Zhou J, Miller BL, & Greicius MD (2009) Neurodegenerative diseases target large-scale human brain networks. *Neuron* 62(1):42-52.
328. Zhou J, *et al.* (2010) Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease. *Brain* 133(Pt 5):1352-1367.
329. Braak H & Braak E (1995) Staging of Alzheimer's disease-related neurofibrillary changes. *Neurobiol Aging* 16(3):271-278; discussion 278-284.
330. Minoshima S, *et al.* (1997) Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann Neurol* 42(1):85-94.
331. Richiardi J, *et al.* (2015) BRAIN NETWORKS. Correlated gene expression supports synchronous activity in brain networks. *Science* 348(6240):1241-1244.
332. Frost B & Diamond MI (2010) Prion-like mechanisms in neurodegenerative diseases. *Nat Rev Neurosci* 11(3):155-159.
333. Crawford JR & Parker DM (1992) *A handbook of neuropsychological assessment* (Psychology Press).
334. Corbetta M & Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience* 3(3):201-215.
335. Bartolomeo P, De Schotten MT, & Doricchi F (2007) Left unilateral neglect as a disconnection syndrome. *Cerebral cortex* 17(11):2479-2490.
336. Doricchi F & Tomaiuolo F (2003) The anatomy of neglect without hemianopia: a key role for parietal–frontal disconnection? *Neuroreport* 14(17):2239-2243.
337. Van Essen DC, *et al.* (2013) The WU-Minn Human Connectome Project: An overview. *Neuroimage* 80:62-79.
338. Uğurbil K, *et al.* (2013) Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *Neuroimage* 80:80-104.
339. Sotiropoulos SN, *et al.* (2013) Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *Neuroimage* 80:125-143.
340. Glasser MF, *et al.* (2013) The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* 80:105-124.
341. Andersson JLR, Skare S, & Ashburner J (2003) How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *Neuroimage* 20:870-888.
342. Jbabdi S, Sotiropoulos SN, Savio AM, Grana M, & Behrens TEJ (2012) Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems. *Magn Res Med* 68:1846-1855.
343. Smith SM, *et al.* (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23 Suppl 1:S208-219.
344. Marcus DS, *et al.* (2011) Informatics and data mining tools and strategies for the Human Connectome Project. *Front Neuroinformatics* 5:4.

345. Rousseeuw PJ (1987) Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 20:53-65.
346. Schmahmann JD, Pandya DN, & Oxford University Press. (2006) Fiber pathways of the brain. (Oxford University Press,, Oxford), pp 1 online resource (xviii, 654 p.) ill. (some col.).
347. Catani M & Thiebaut de Schotten M (2012) *Atlas of human brain connections* (Oxford University Press, Oxford) pp xii, 519 p.
348. Winkler AM, Ridgway GR, Webster MA, Smith SM, & Nichols TE (2014) Permutation inference for the general linear model. *NeuroImage* 92:381-397.
349. Passingham RE, Stephan KE, & Kotter R (2002) The anatomical basis of functional localization in the cortex. *Nat Rev Neurosci* 3(8):606-616.
350. Neubert FX, Mars RB, Thomas AG, Sallet J, & Rushworth MF (2014) Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. *Neuron* 81(3):700-713.
351. Sallet J, *et al.* (2011) Social network size affects neural circuits in macaques. *Science* 334:697-700.
352. Noonan MP, *et al.* (2014) A neural circuit covarying with social hierarchy in macaques. *Plos Biol* 12:e1001940.
353. Reese TG, Heid O, Weisskoff RM, & Wedeen VJ (2003) Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. *Magnetic Resonance in Medicine* 49:177-182.
354. Behrens TE, Johansen-Berg H, Jbabdi S, Rushworth MF, & Woolrich MW (2007) Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* 34(1):144-155.
355. Stejskal EO & Tanner JE (1965) Spin diffusion measurements: Spin echos in the presence of a time-dependent field gradient. *J Chem Phys* 42:288-292.
356. D'Arceuil H & De Crespigny A (2007) The effects of brain tissue decomposition on diffusion tensor imaging and tractography. *Neuroimage* 36:64-68.
357. Dyrby TB, *et al.* (2011) An ex vivo imaging pipeline for producing high-quality and high-resolution diffusion-weighted imaging datasets. *Human Brain Mapping* 32:544-563.
358. Miller KL, *et al.* (2011) Diffusion imaging of whole, post-mortem human brains on a clinical MRI scanner. *Neuroimage* 57:167-181.
359. Eickhoff SB, *et al.* (2009) Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping* 30:2907-2926.
360. Boets B, *et al.* (2013) Intact but less accessible phonetic representations in adults with dyslexia. *Science* 342(6163):1251-1254.
361. Silk TJ, Vance A, Rinehart N, Bradshaw JL, & Cunnington R (2009) White‐matter abnormalities in attention deficit hyperactivity disorder: A diffusion tensor imaging study. *Human brain mapping* 30(9):2757-2765.
362. Connally EL, Ward D, Howell P, & Watkins KE (2014) Disrupted white matter in language and motor tracts in developmental stuttering. *Brain and language* 131:25-35.
363. Friston KJ & Frith CD (1995) Schizophrenia: a disconnection syndrome. *Clin Neurosci* 3(2):89-97.
364. Petrides M & Pandya DN (1984) Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. *J Comp Neurol* 228(1):105-116.
365. Schmahmann JD & Pandya DN (2006) *Fiber pathways of the brain* (Oxford University Press, New York) pp xviii, 654 p.
366. Catani M & Thiebaut de Schotten M (*Atlas of human brain connections* pp xii, 519 pages.
367. Buckner RL, *et al.* (2009) Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *The Journal of Neuroscience* 29(6):1860-1873.
368. Noonan M, Mars, R.B., Sallet, J., Dunbar, R.I.M., & Fellows, LK. (under review) The structural and functional brain networks that support human social networks.
369. Waiter GD, *et al.* (2005) Structural white matter deficits in high-functioning individuals with autistic spectrum disorder: a voxel-based investigation. *Neuroimage* 24(2):455-461.
370. Barnea-Goraly N, Lotspeich LJ, & Reiss AL (2010) Similar white matter aberrations in children with autism and their unaffected siblings: a diffusion tensor imaging study using tract-based spatial statistics. *Archives of general psychiatry* 67(10):1052-1060.

371. Riva-Posse P, *et al.* (2014) Defining critical white matter pathways mediating successful subcallosal cingulate deep brain stimulation for treatment-resistant depression. *Biological psychiatry* 76(12):963-969.
372. Kennedy SH, *et al.* (2001) Changes in regional brain glucose metabolism measured with positron emission tomography after paroxetine treatment of major depression. *American Journal of Psychiatry*.
373. Rudebeck PH, Saunders RC, Prescott AT, Chau LS, & Murray EA (2013) Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nature neuroscience* 16(8):1140-1145.
374. Huettel SA, *et al.* (2004) Linking hemodynamic and electrophysiological measures of brain activity: evidence from functional MRI and intracranial field potentials. *Cereb Cortex* 14(2):165-173.
375. Huettel SA, Song AW, & McCarthy G (*Functional magnetic resonance imaging* Third edition. Ed p 621 pages (in various pagings).
376. Gannon PJ, Holloway RL, Broadfield DC, & Braun AR (1998) Asymmetry of chimpanzee planum temporale: humanlike pattern of Wernicke's brain language area homolog. *Science* 279(5348):220-222.
377. Powell HR, *et al.* (2006) Hemispheric asymmetries in language-related pathways: a combined functional MRI and tractography study. *Neuroimage* 32(1):388-399.
378. Toga AW & Thompson PM (2003) Mapping brain asymmetry. *Nature Reviews Neuroscience* 4(1):37-48.
379. Association AP (2003) *Diagnostic and Statistical Manual of Mental Disorders:: DSM-5* (ManMag).
380. Heilman MKM & Valenstein E (2011) *Clinical neuropsychology* (Oxford University Press).
381. Mesulam M-M (2015) Fifty years of disconnexion syndromes and the Geschwind legacy. *Brain*:awv198.
382. Bauer RM (1982) Visual hypoemotionality as a symptom of visual-limbic disconnection in man. *Archives of Neurology* 39(11):702-708.
383. Catani M & Mesulam M (2008) The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex* 44(8):953-961.
384. Masdeu JC & Biller J (2011) *Localization in clinical neurology* (Lippincott Williams & Wilkins).
385. Johansen-Berg H, *et al.* (2002) The role of ipsilateral premotor cortex in hand movement after stroke. *Proc Natl Acad Sci U S A* 99(22):14518-14523.
386. Buch ER, Johnen VM, Nelissen N, O'Shea J, & Rushworth MF (2011) Noninvasive associative plasticity induction in a corticocortical pathway of the human brain. *The Journal of Neuroscience* 31(48):17669-17679.
387. Johnen VM, *et al.* (2015) Causal manipulation of functional connectivity in a specific neural pathway during behaviour and at rest. *eLife* 4:e04585.
388. Heilbronner SR & Haber SN (2014) Frontal cortical and subcortical projections provide a basis for segmenting the cingulum bundle: Implications for neuroimaging and psychiatric disorders. *Journal of Neuroscience* 34:100041-110054.
389. Stagg CJ, *et al.* (2014) Local GABA concentration is related to network-level resting functional connectivity. *eLife* 3:e01465.
390. Greicius MD, Srivastava G, Reiss AL, & Menon V (2004) Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proceedings of the National Academy of Sciences of the United States of America* 101(13):4637-4642.
391. Broyd SJ, *et al.* (2009) Default-mode brain dysfunction in mental disorders: a systematic review. *Neuroscience & biobehavioral reviews* 33(3):279-296.
392. Zilles K & Amunts K (2009) Receptor mapping: architecture of the human cerebral cortex. *Curr Opin Neurol* 22(4):331-339.
393. Fodor JA (1983) *The modularity of mind: An essay on faculty psychology* (MIT press).
394. Fodor JA (2001) *The mind doesn't work that way: The scope and limits of computational psychology* (MIT press).
395. Barrett HC & Kurzban R (2006) Modularity in cognition: framing the debate. *Psychological review* 113(3):628.
396. Barrett HC (2005) Enzymatic computation and cognitive modularity. *Mind & language* 20(3):259-287.

397. Baron-Cohen S (1997) *Mindblindness: An essay on autism and theory of mind* (MIT press).
398. Hermer L & Spelke E (1996) Modularity and development: The case of spatial reorientation. *Cognition* 61(3):195-232.
399. Dehaene S & Cohen L (1995) Towards an anatomical and functional model of number processing. *Mathematical cognition* 1(1):83-120.
400. Fleming SM, Huijgen J, & Dolan RJ (2012) Prefrontal contributions to metacognition in perceptual decision making. *The Journal of Neuroscience* 32(18):6117-6125.
401. Lim S-L, O'Doherty JP, & Rangel A (2013) Stimulus value signals in ventromedial PFC reflect the integration of attribute value signals computed in fusiform gyrus and posterior superior temporal gyrus. *The Journal of Neuroscience* 33(20):8729-8741.
402. Pinker S (2005) A reply to Jerry Fodor on how the mind works. *Mind & Language* 20(1):33-38.
403. Zuo X-N, *et al.* (2012) Network centrality in the human functional connectome. *Cerebral cortex* 22(8):1862-1875.
404. Tarr MJ & Gauthier I (2000) FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature neuroscience* 3:764-770.
405. Chiappe D & MacDonald K (2005) The evolution of domain-general mechanisms in intelligence and learning. *The Journal of general psychology* 132(1):5-40.
406. Corbetta M, Patel G, & Shulman GL (2008) The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58(3):306-324.
407. Derrfuss J, Brass M, Neumann J, & von Cramon DY (2005) Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. *Hum Brain Mapp* 25(1):22-34.
408. Aron AR & Poldrack RA (2006) Cortical and subcortical contributions to Stop signal response inhibition: role of the subthalamic nucleus. *J Neurosci* 26(9):2424-2433.
409. Friederici AD (2006) Broca's area and the ventral premotor cortex in language: functional differentiation and specificity. *Cortex* 42(4):472-475.
410. Fedorenko E, Behr MK, & Kanwisher N (2011) Functional specificity for high-level linguistic processing in the human brain. *Proc Natl Acad Sci U S A* 108(39):16428-16433.
411. Chikazoe J, *et al.* (2009) Functional dissociation in right inferior frontal cortex during performance of go/no-go task. *Cereb Cortex* 19(1):146-152.
412. Vossel S, Weidner R, & Fink GR (2011) Dynamic coding of events within the inferior frontal gyrus in a probabilistic selective attention task. *J Cogn Neurosci* 23(2):414-424.
413. Hirose S, *et al.* (2009) Sub-centimeter scale functional organization in human inferior frontal gyrus. *Neuroimage* 47(2):442-450.
414. Friederici AD, Bahlmann J, Heim S, Schubotz RI, & Anwander A (2006) The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proc Natl Acad Sci U S A* 103(7):2458-2463.
415. Fedorenko E, Duncan J, & Kanwisher N (2012) Language-Selective and Domain-General Regions Lie Side by Side within Broca's Area. *Curr Biol* 22(21):2059-2062.
416. Eickhoff SB, *et al.* (2009) Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. *Hum Brain Mapp* 30(9):2907-2926.
417. Laird AR, Lancaster JL, & Fox PT (2009) Lost in localization? The focus is meta-analysis. *Neuroimage* 48(1):18-20.
418. Laird AR, Lancaster JL, & Fox PT (2005) BrainMap: the social evolution of a human brain mapping database. *Neuroinformatics* 3(1):65-78.
419. Petrides M (2005) Lateral prefrontal cortex: architectonic and functional organization. *Philos Trans R Soc Lond B Biol Sci* 360(1456):781-795.
420. Mayka MA, Corcos DM, Leurgans SE, & Vaillancourt DE (2006) Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: a meta-analysis. *Neuroimage* 31(4):1453-1474.
421. Belin P, Zatorre RJ, Lafaille P, Ahad P, & Pike B (2000) Voice-selective areas in human auditory cortex. *Nature* 403(6767):309-312.
422. Mai JrK, Voss T, & Paxinos G (2008) *Atlas of the human brain* (Elsevier/Academic Press, London) 3rd Ed pp viii, 271 p.
423. Ongur D, Ferry AT, & Price JL (2003) Architectonic subdivision of the human orbital and medial prefrontal cortex. *J Comp Neurol* 460(3):425-449.

424. McLaren DG, *et al.* (2009) A population-average MRI-based atlas collection of the rhesus macaque. *Neuroimage* 45(1):52-59.
425. Vogt BA, Finch DM, & Olson CR (1992) Functional heterogeneity in cingulate cortex: the anterior executive and posterior evaluative regions. *Cereb Cortex* 2(6):435-443.
426. Bell AH, *et al.* (2011) Relationship between functional magnetic resonance imaging-identified regions and neuronal category selectivity. *J Neurosci* 31(34):12229-12240.
427. Petkov CI, *et al.* (2008) A voice region in the monkey brain. *Nat Neurosci* 11(3):367-374.
428. Fox MD, Zhang D, Snyder AZ, & Raichle ME (2009) The global signal and observed anticorrelated resting state brain networks. *J Neurophysiol* 101(6):3270-3283.
429. Neubert FX, Mars RB, Thomas AG, Sallet J, & Rushworth MF (2014) Comparison of Human Ventral Frontal Cortex Areas for Cognitive Control and Language with Areas in Monkey Frontal Cortex. *Neuron*.
430. Newman MEJ (2010) *Networks : an introduction* (Oxford University Press, Oxford) pp xi, 772 p., [774] p. of plates.
431. Noonan MP, Mars RB, & Rushworth MF (2011) Distinct roles of three frontal cortical areas in reward-guided behavior. *J Neurosci* 31(40):14399-14412.
432. Paxinos G & Mai JK (2004) *The human nervous system* (Elsevier Academic Press, Amsterdam ; Boston) 2nd Ed pp xvii, 1366 p.
433. Amunts K & Zilles K (2012) Architecture and organizational principles of Broca's region. *Trends Cogn Sci* 16(8):418-426.
434. Grol MJ, *et al.* (2007) Parieto-frontal connectivity during visually guided grasping. *J Neurosci* 27(44):11877-11887.
435. Eickhoff SB, *et al.* (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25(4):1325-1335.
436. Caspers S, *et al.* (2006) The human inferior parietal cortex: cytoarchitectonic parcellation and interindividual variability. *Neuroimage* 33(2):430-448.
437. Catani M, Jones DK, Donato R, & Ffytche DH (2003) Occipito-temporal connections in the human brain. *Brain* 126(Pt 9):2093-2107.
438. Moss MB & Rosene DL (1984) A perfusion-fixation procedure for the concurrent demonstration of Timm's, horseradish peroxidase (HRP), and acethycholinesterase (AChE) histochemistry. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* 32(10):1113-1116.
439. Karnath HO, *et al.* (2005) Normalized perfusion MRI to identify common areas of dysfunction: patients with basal ganglia neglect. *Brain* 128(Pt 10):2462-2469.
440. Van Essen DC, *et al.* (2001) Mapping visual cortex in monkeys and humans using surface-based atlases. *Vision Res* 41(10-11):1359-1378.
441. Squire LR, Stark CE, & Clark RE (2004) The medial temporal lobe. *Annu Rev Neurosci* 27:279-306.
442. Suzuki WA (1996) The anatomy, physiology and functions of the perirhinal cortex. *Curr Opin Neurobiol* 6(2):179-186.
443. Kondo H, Saleem KS, & Price JL (2003) Differential connections of the temporal pole with the orbital and medial prefrontal networks in macaque monkeys. *J Comp Neurol* 465(4):499-523.
444. Blaizot X, *et al.* (2010) The human parahippocampal region: I. Temporal pole cytoarchitectonic and MRI correlation. *Cereb Cortex* 20(9):2198-2212.
445. Haber SN & Knutson B (2010) The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35(1):4-26.
446. Draganski B, *et al.* (2008) Evidence for segregated and integrative connectivity patterns in the human Basal Ganglia. *J Neurosci* 28(28):7143-7152.
447. DeLong M & Wichmann T (2009) Update on models of basal ganglia function and dysfunction. *Parkinsonism & related disorders* 15 Suppl 3:S237-240.
448. DeLong MR (1990) Primate models of movement disorders of basal ganglia origin. *Trends Neurosci* 13(7):281-285.
449. Squire LR (1992) Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol Rev* 99(2):195-231.
450. Saunders RC, Rosene DL, & Van Hoesen GW (1988) Comparison of the efferents of the amygdala and the hippocampal formation in the rhesus monkey: II. Reciprocal and non-reciprocal connections. *The Journal of comparative neurology* 271(2):185-207.

451. Saunders RC & Rosene DL (1988) A comparison of the efferents of the amygdala and the hippocampal formation in the rhesus monkey: I. Convergence in the entorhinal, prorhinal, and perirhinal cortices. *The Journal of comparative neurology* 271(2):153-184.
452. Rosene DL & Van Hoesen GW (1977) Hippocampal efferents reach widespread areas of cerebral cortex and amygdala in the rhesus monkey. *Science* 198(4314):315-317.
453. Haber SN & Fudge JL (1997) The primate substantia nigra and VTA: integrative circuitry and function. *Critical reviews in neurobiology* 11(4):323-342.
454. Winkler AM, Ridgway GR, Webster MA, Smith SM, & Nichols TE (2014) Permutation inference for the general linear model. *Neuroimage* 92:381-397.