Control and causal learning in the human brain and their links to mental health

Lisa Spiering

University College
University of Oxford

A dissertation submitted for the degree of
Doctor of Philosophy

Michaelmas term 2024

Acknowledgments

This journey would not have been possible without so many people supporting me. Thank you first of all to my supervisors Jacquie and Matthew. Jacquie is brilliant at computational modelling and data analysis, and there truly seems to be no methods problem she cannot tackle. Matthew is the kindest and most brilliant researcher I ever had the joy to meet, and I am so grateful for his guidance and support over the years.

It is odd to start a DPhil during a pandemic, and I want to thank all the people in Matthew's lab for providing such an ever-supportive environment; in particular I want to thank Caroline Harbison, Hailey Trier and Urs Schüffelgen. A big thank you to Miriam Klein-Flügge and her lab as well and especially to Maja Friedemann, Rong Bi and Lilian Weber for being such lovely and inspiring office roommates. I want to thank all my friends from outside the lab for making me feel at home in Oxford; thank you to Clémence, Verena, Ellie, Philly and Vale. I also wish to thank my friends from home, who have brought me so much joy and I truly appreciate their support and friendship over so many years: Gesine, Anne-Kathrin, Anna, Thi, Sophia, Olivia, Fredi, Carlotta, Josi and Anouk.

Thank you to Marco for so much – in particular for sharing my passion for morning coffees in bed, and for keeping me sane during this journey. Mein größter Dank geht an meine wunderbare Familie, meine Eltern, meine Schwester Anne und meine große Oma, die all dies überhaupt erst ermöglicht haben.

Table of contents

Abstract 7

CHAPTER 1 General Introduction 8

1.1 Cooperation in social interactions 8

1.2 Credit assignment 9

1.3 Behavioural flexibility to aid credit assignment 12

1.3.1 Information seeking 12

1.3.2 Causal interventions 14

1.4 Control 18

1.4.1 Controlling one's environment: Understanding control from the perspective of learned helplessness and depressive realism 18

1.4.2 Controlling one's actions: Understanding control from a motor perspective and links to schizophrenia 24

1.4.3 Comparison of the theoretical frameworks of control 27

1.5 Summary and relevance to the current thesis 30

1.6 Social versus non-social processing 31

1.7 Brain areas relevant for control and causal learning 33

1.7.1 Supramarginal gyrus 33

1.7.2 Striatum 37

1.8 Links to mental health and computational psychiatry 40

1.8.1 Depression 41

1.8.2 Schizotypy and schizophrenia 44

CHAPTER 2 Controllability and cause in human collaboration 48

2.1 Introduction 48

2.2 Methods 51

2.2.1 Participants 51

2.2.2 Experimental procedure 52

2.2.3 Experimental design 53

2.2.4 Bayesian optimal observer models 68

2.2.5 Behavioural analyses 72

2.3 Results 83

2.3.1 Estimates of controllability modulate social prediction errors 83

2.3.2 Active disambiguation as a behavioural strategy 86

2.3.3 Active disambiguation enables inference of controllability 89

2.3.4 Supplementary results 93

2.4 Discussion 96

CHAPTER 3 Supramarginal Gyrus mediates controllability learning and tracks social prediction errors 99

3.1 Introduction 99

3.2 Methods 101

3.2.1 Imaging data acquisition and preprocessing 101

3.2.2 fMRI whole brain analysis 101

3.2.3 fMRI time course analysis on ROIs 105

3.3 Results 111

3.3.1 Supramarginal Gyrus tracks uncertainty and active disambiguation at the action 111

3.3.2 Supramarginal Gyrus carries neural signals for learning controllability 113

3.3.3 Supramarginal Gyrus tracks prediction errors assigned to self and others 117

3.3.4 Nucleus accumbens tracks feedback and feedback change 120

3.4 Discussion 122

CHAPTER 4 Schizotypy modulates credit assignment and control inference 124

4.1 Introduction 124

4.2 Methods 127

4.2.1 Participants 127

4.2.2 Experimental procedure and design 128

4.2.3 Self-report psychiatric questionnaires 129

4.2.4 Behavioural analyses 132

4.3 Results 141

4.3.1 Schizotypy modulates participants' rating accuracy and prior beliefs 141

4.3.2 Schizotypy increases credit assignment to self but not others 142

4.3.3 No effects of Schizotypy and QIDS on active disambiguation 146

4.3.4 Schizotypy increases controllability inference from AD 149

4.4 Discussion 153

CHAPTER 5 Striatal and SMG responses to ambiguous outcome processing in schizotypy and depression 156

5.1 Introduction 156

5.2 Methods 159

5.2.1 fMRI whole brain analysis 159

5.2.2 fMRI time course analysis on ROIs 160

5.3 Results 167

5.3.1 Supramarginal Gyrus tracks uncertainty and AD while schizotypy-UE modulates putamen activity of AD 167

5.3.2 Effects of schizotypy-UE and QIDS on the neural correlates of control inference 169

5.3.3 Effects of schizotypy-UE and QIDS on the neural correlates of self- and other-assigned PEs 173

5.3.4 Effects of schizotypy-UE and QIDS on reward representations in the ventral striatum 177

5.4 Discussion 179

CHAPTER 6 General Discussion 182

6.1 Inferring control and learning about self and others 183

6.2 The neural correlates of control and causal learning 184

6.3 Interindividual differences in the behavioural and neural signatures of control and causal learning 185

6.4 Limitations, ongoing and future work 188

6.5 Summary 192

References 193

Control and causal learning in the human brain and their links to mental health

Abstract

Humans often collaborate towards joint outcomes and can navigate complex social situations. In this thesis, I investigated the computational and neural mechanisms underlying human learning in social contexts, and their links to schizotypy and depression.

Collaborating with others requires humans to ascertain their individual level of control over shared outcomes (reviewed in chapter 1). I demonstrated that in ambiguous social contexts, people engage in specific patterns of behaviour that I refer to as active disambiguation (chapter 2). This process helps individuals establish what they themselves, as opposed to others, control and what consequence they themselves cause or that another person causes. People identify when active disambiguation is needed and engage in it at that time. A pattern of activity in the supramarginal gyrus, emerging during and after active disambiguation, is linked to establishing controllability and tracking the outcomes from which control is inferred (chapter 3). Activity in this brain region also signals a second learning mechanism, by which individuals attribute outcomes to themselves versus others, in proportion to their perceived control.

Understanding control and agency estimation may inform our understanding of psychological illness. I found that individuals with schizotypal traits rely more on external cues to estimate their level of control and learn about their own abilities (chapter 4). Firstly, they take outcomes furnished from active disambiguation more into account to estimate their control – behaviour which is in fact closer to simulated behaviour from an optimal computational learning model. Secondly, they fail to assign outcomes to themselves when their perceived control is low. Neural signals of outcome processing from active disambiguation are reduced in schizotypy (chapter 5). This reveals a potentially distinct neural mechanism by which outcome attribution may be altered in schizotypal traits.

Lisa Spiering

University of Oxford

Approximate thesis word count: 48,000 words

Submitted for the degree of Doctor in Philosophy

CHAPTER 1 General Introduction

Human and nonhuman primates are inherently social species and typically work together with others to achieve their goals. Whenever we work together with others, we need to take into account the impact that others will have on the situation. We therefore need to consider the degree of control we have and they have over the potential outcome that we might produce together. How do people come to understand how much control they have in social interactions? How do people learn about themselves and others under different levels of controllability? Examining these questions is the focus of this thesis. In this first introductory chapter, I will outline previous studies that inform the approach that I have taken. In chapters 2 and 3, I examine the cognitive and neural mechanisms of how people learn about themselves, others and their control, and how they employ specific behavioural strategies to do so. In chapters 4 and 5, I investigate how these processes can go awry in depression and as a function of schizotypy. In the final chapter 6, I discuss the findings presented in this thesis.

1.1 Cooperation in social interactions

Working together and cooperating with others allows primates to achieve their goals more quickly and efficiently (Melis & Raihani, 2023). Sociality can provide an evolutionary advantage because larger group sizes reduce predator risk (Shultz et al., 2004). Among primates, relative brain size is associated with indicators of social complexity such as group size, which might be related to the additional cognitive demands that follow from living in groups (Dunbar & Shultz, 2007; Sallet et al., 2011). One such cognitive demand, when cooperating with others, is to monitor those others (Melis & Raihani, 2023). In long-term relationships, for example, it might be necessary to remember how a partner behaved in different situations, resulting in greater memory demands (Melis & Raihani, 2023).

Monitoring how well one does in relation to others is an important feature of cooperation behaviours. This may specifically be the case when humans cooperate in groups to achieve multi-step tasks that require more complex coordination of behaviour (Misyak et al., 2014; Tomasello et al., 2005). Wittmann et al. (2016b) examined how humans estimate their own and others' performances in small games from feedback. They found that even though participants received separate feedback reflecting how well they did and how well their partner did, participants found it difficult to estimate their own performance completely independent of the partner. In fact, they called this type of confusion between self and others "self-other mergence". Intriguingly, this self-other mergence depended on the social context: If participants were cooperating with others, they based their self-estimates not just on the feedback received for themselves, but also on how well the other person did. Interestingly, if they competed with the other player, they displayed the opposite behaviour and their own estimate of their own performance was diminished in proportion to how well the other player had done. This study illustrates how social environments influence our perception of ourselves in quite a powerful way. How we think about ourselves is often not just based on some intrinsic process based on evaluation of our own actions, but it also takes into account who we interact with and what social context we are in.

1.2 Credit assignment

Whenever something happens in our daily life, we need to figure out who or what caused it and to what extent. Causal reasoning is an ability that humans develop from childhood. Children as young as two to four years old are able to attribute physical phenomena, including sound, wind and light, to its causes; for example, they can correctly judge how sound is transmitted and caused by different tuning forks placed close to a resonator box (Shultz, 1982). People are astonishingly good at inferring the causal relationships in the world, and can rapidly infer causality after as little as just two observations (Kushnir et al., 2003). This poses the question

what exactly are the cognitive mechanisms and computations which humans use to infer causality?

In the field of machine learning and in the context of reinforcement learning, causal inference has been quantitatively described under the term of credit assignment (Minsky, 1961; Sutton & Barto, 2018). Credit assignment is about learning about the causes of events and is necessary whenever an event is ambiguous and has several potential causes.

Even though people can infer the causes behind events very well in some situations (as illustrated in the beginning of this section), correctly assigning credit is still not easy. For example, when a chess game is lost, it is difficult (or even impossible) to blame an individual choice out of the large number of decisions taken over the course of the game (Minsky, 1961). In fact, human and nonhuman primates have different biases when assigning credit to their past choices. Humans and other animals have been shown to misassign credit due to "spread of effect", the phenomenon whereby rewards reinforce not only the action that immediately preceded their delivery, but also other, independent actions that happened close in time to the reward (Jocham et al., 2016; Noonan et al., 2017; Thorndike, 1933; Walton et al., 2010; Wittmann et al., 2020). Scholl et al. (2015) investigated such biases using an experimental design in which participants had to learn about two options with different rewards and effort associated with them. A key manipulation in this experiment was that sometimes, the reward that informed the participant that their choice had been correct was real and contributed to their bonus payment, but every now and then, the reward, while equally informative about the correctness of the action that had been made, was only hypothetical. Notably, whether a reward was real or not was determined by chance and not by participants' behaviour. Therefore, participants should only have paid attention to whether any reward was indicated but they should not have taken into account whether the reward was real or hypothetical as a

guide as to whether they should repeat the choice in the future. Scholl et al. (2015) found that despite being aware of this, participants still showed an irrational choice bias and tended to repeat choices after which they, by chance, received a real reward. This study illustrates that credit assignment is difficult even in situations without any ambiguity as to which antecedents of an event can possibly cause an outcome.

In many situations in real life, it is unknown which actions or events could possibly have caused an outcome, making the credit assignment problem even more acute. Akaishi et al. (2016) examined this using a variant of the weather prediction task (Knowlton et al., 1996), and asked participants to choose cues to predict the weather. This design allowed them to delineate the credit assignment from the choice process, and investigate how people test different hypotheses for credit assignment. They found that when the weather feedback confirmed participants' predictions, they assigned credit almost exclusively to the cue they selected to guide their prediction. In contrast, when the weather feedback did not match the prediction, participants assigned credit to the alternative cue they did not select to guide their prediction. Even though the weather did not match their prediction, they still showed a confirmation bias and assigned some credit for the outcome to the cue which incorrectly predicted the weather. This suggests that people learn in a hypothesis-driven way. This has important consequences. On the one hand it can help credit assignment but, on the other hand, sometimes it also leads to confirmation biases if learning only happens about the hypothesized cues.

One feature that the studies described above have in common is that they examine how humans (or macaque monkeys) assign credit to their choices or to stimulus cues they experienced. However, in real life, we often work together with others in one way or another to achieve our goals. Therefore, collaborative social interactions give rise to another type of attribution problem: which agent's action was really responsible for causing an outcome. In this

thesis, I use credit assignment as a computational framework to explain how such attributions are made in social contexts.

1.3 Behavioural flexibility to aid credit assignment

1.3.1 Information seeking

As described in the previous section, credit assignment is often difficult because there are many potential causes to assign credit to. This poses the question of how people are able to assign credit and infer causality in their daily lives. One way to resolve this uncertainty and aid credit assignment is to flexibly adapt one's behaviour and actively seek information. Information seeking can help agents form and update their model of the environment which is important for survival. Indeed, biased information seeking can lead to a biased understanding of the environment (Charpentier et al., 2018; Garrett et al., 2018). In particular, in a simulated stock market, participants preferred to learn about future positive outcomes more than negative ones, and are also willing to pay more money to know about promising rather than ominous market news (Charpentier et al., 2018). This suggests that the valence of information influences and biases human information seeking, ultimately leading to a biased world view. Overall, this suggests that information seeking and credit assignment are strongly intertwined cognitive and behavioural processes.

Information seeking behaviours are broadly divided into two categories, instrumental or non-instrumental. Information seeking is instrumental if it helps people make better decisions in the future. In particular, it can help them obtain first- or second-order reinforcers, such as food or money (Monosov, 2024). Non-instrumental information seeking refers to exploratory behaviours aimed at novel items or uncertainty reduction, without necessarily resulting in a reward or aiding task completion (Bromberg-Martin & Monosov, 2020; Gottlieb et al., 2020). Balancing whether to engage in exploratory information seeking behaviours or whether to

exploit known opportunities for reward is difficult, and the trade-off between these two behavioural modes is also known as the explore-exploit dilemma (Sutton & Barto, 2018). In other words, the explore-exploit dilemma is the conundrum of having to choose between something one knows or an alternative one does not know but which might turn out to be better.

A classic example from the reinforcement learning literature is the k-armed bandit problem, in which an agent is in a room with a k number of gambling machines (each a one-armed bandit) (Berry & Fristedt, 1985; Kaelbling et al., 1996). The agent can then, for a limited number of times, pull the levers of the gambling machines. Pulling the lever of a machine results in a pay-off of either 1 or 0, with an underlying unknown probability. While pulling the levers does not incur a monetary cost, there is potentially an opportunity cost of wasting a pull on a suboptimal machine. Given that the agent only has a fixed number of pulls and its goal is to maximise the total pay-off, the agent is faced with the question whether to explore new machines, or exploit and stick to known machines. For example, if the agent has a longer game (a higher total number of pulls), then it should explore more because converging on (exploiting) a suboptimal machine early has worse consequences in the long run (Kaelbling et al., 1996). This means that the time horizon should affect how people explore.

Wilson et al. (2014) investigated this using the “horizon task” in which participants chose between two machines with different reward rates. The participants’ task was to win as many rewards from the machines as possible, and they were told that one was better than the other. To dissociate the effects of reward and information on exploration, Wilson et al. introduced a set of forced choices in the beginning of each game. These forced-choice trials meant that, sometimes, participants had the same information about each machine or sometimes they had unequal information about the machines. After the forced-choice trials, participants were then

free to choose either of the two machines for either just one or more free-choice trials (short and long time horizon). Overall, participants explored more if they had a longer time horizon. Specifically, participants exhibited two different kinds of exploration: Firstly, participants sampled the machine for which they had less information about if they had a longer time horizon. This suggests that when faced with unequal information and the opportunity of a longer time horizon, people demonstrated a type of exploration behaviour that has also been known as “directed” exploration. Directed exploration is a way to solve the explore-exploit dilemma by sampling options one knows less about, effectively increasing what is known about an initially unknown option (Wilson et al., 2014). The second strategy is “random” exploration. Here, irrespective of the information content, a longer time horizon made participants initially more random in their choice. This means that a longer time horizon increased decision noise, leading to more random exploration.

While initially it might seem that directed exploration should be favoured over random exploration, because it should lead to more strategic information acquisition, random exploration has the advantage of being a simpler heuristic with lower computational costs. It has, however, been confirmed that people explore uncertain options more early in the learning process, while later on they tend to avoid uncertainty and sample options with known higher reward (Trudel et al., 2021). Overall, these studies suggest that people solve the explore-exploit dilemma and seek information in an adaptive way, using different strategies that are calibrated to their current environment. The ability to actively seek information seems to be an effective way to resolve uncertainty and understand one’s own environment better.

1.3.2 Causal interventions

Information seeking is typically understood as a way to find out about the properties of individual cues such as the reward probabilities that are associated with them. However, in real

life, different events in our environment are often linked and not independent from one another. Therefore, information seeking can more effectively aid credit assignment if one takes into account the relationships between events. For example, a higher likelihood of clouds in the sky is often correlated with a higher likelihood of rain, in turn making it more likely for the pavement to be wet. Figuring out whether an event just coincided with another event or whether it actually caused it, means distinguishing correlation from causation. It is impossible to distinguish between the two without actively intervening and observing whether the second event changes if the first one changes. As early as in the 19th century, Mill (1843) suggested that to establish cause-effect relationships, it is necessary to remove a potential cause while ensuring that the other variables remain constant. This is not only relevant for causal reasoning in everyday life; Fisher (1934) argued that researchers need randomised experiments in which variables of interest are systematically manipulated to establish causal links.

In cognitive science, theories of causal reasoning have formulated how agents can learn the causal relationships in their environment (Pearl, 2009; Sloman, 2005). Pearl (2009) introduced causal Bayesian networks as a mathematical framework for how different variables (such as events) are causally and systematically linked in one network, and how interventions help establish causality between them. In such a network, each node refers to a variable of interest such as the slipperiness of a pavement, rain, a sprinkler, and the current season (Figure 1.1). These variables are linked by arrows representing their causal dependencies. These causal dependencies are formulated as probability distributions for each variable which is conditioned on its causes. Mathematically, it is then possible to compute the probability that a variable will hold a particular value once one is given the values of the other variables. Intuitively, for example, if the current season is autumn, rainy weather is more likely, in turn leading to slippery pavements, but sprinklers are unlikely to be switched on. If the season is summer, rain is more unlikely, however the pavements might still be slippery because sprinklers are more likely to

be switched on. In such a framework, causality can only be established if and only if changing the value of one variable will change the value of another variable. Carrying out such a test is referred to as carrying out a causal intervention. For example, if it does not rain and the pavement is slippery, then it can be inferred that the rain did not cause the pavement to be slippery. Instead, this can be attributed to the sprinkler if switching the sprinkler on leads to a slippery pavement. Such interventions (switching the sprinkler on to observe if this makes the pavement wet) allow agents to test the causal relationships governing their environment.

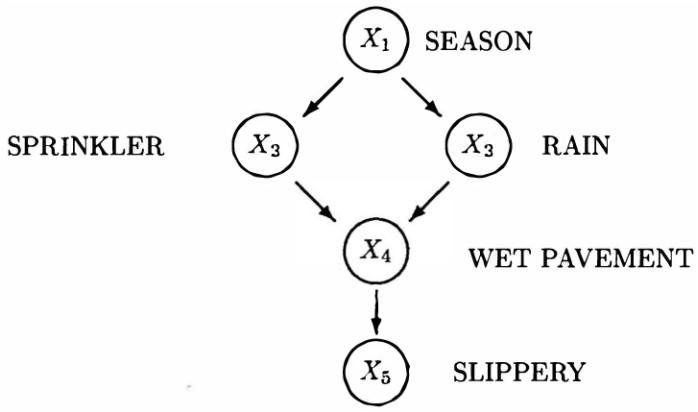


Figure 1.1. Example of a network of the causal relationships between five variables. Reprinted from Pearl (1999)

The theoretical predictions that can be derived from Bayesian causal networks have since been used to model human causal reasoning, and it has indeed been shown that people infer causality by making such interventions. Steyvers et al. (2003) investigated this in a series of experiments using a task in which participants had to learn, in a set of three alien mind readers, which alien can read the mind of which other alien. Participants were shown a series of nonsensical words that the aliens thought and could then observe how these words propagated from alien to alien. The authors found that in their first experiment, participants could, purely through passive observation of how the words propagated from alien to alien, already get a sense of the associations between the aliens. However, in this first experiment, participants

were not able to make any causal manipulations or interventions. In the second experiment, participants were given the opportunity to place a “mind zapper” into the alien group, thereby actively manipulating what word a particular alien thought. By employing such an intervention together with observing its resulting effects on the other aliens, participants were indeed considerably better at judging the causal links between the aliens.

It has since been shown that one mechanism by which interventions help people learn causal links is the temporal order of cause and effects. Any changes in the environment happening after an intervention are more likely to be caused by the intervention rather than by anything else (Lagnado & Sloman, 2004).

One feature of these studies is that they investigated how people use interventions to infer the causal links between observable variables. In the alien mind reading experiment by Steyvers et al. (2003), for example, the words thought by the aliens were displayed on the screen, and participants had to infer the unknown causal links between the aliens. However, people have also been shown to use interventions to learn about hidden or unobserved causes (Kushnir et al., 2003). Using a task in which participants had to judge whether two balls were controlled by a common mechanism hidden in a box, the authors found that participants indeed used interventions together with observations to infer the correct hidden common cause. This study extends the previous findings and demonstrates that interventions are a powerful way to learn about unobserved causal structures, suggesting a more sophisticated way of causal learning driven by interventions. Taken together, the reviewed findings suggest that behavioural flexibility, be it through information seeking or causal interventions, allows people to learn more about their environment.

1.4 Control

Behavioural flexibility is not only useful when developing a better understanding of the world around us but it can also help us better understand the effect of our actions on our environment. It is an important part of our daily lives to understand our own level of control and adapt our behaviour accordingly by, for example, investing no effort if we know our actions have no consequences. Indeed, the well-known serenity prayer starts with praying for "the serenity to accept the things I cannot change; courage to change the things I can; and wisdom to know the difference" (Niebuhr, 1943). How can we know how much control we have, that is whether we can change things or not?

1.4.1 Controlling one's environment: Understanding control from the perspective of learned helplessness and depressive realism

Misestimating one's own control can have severe consequences. In a classic Pavlovian conditioning study by Overmier and Seligman (1967), dogs were exposed to inescapable electric shocks and afterwards, did not escape electric shocks in a new situation even when it was possible to do so. Indeed, the result was very compelling and surprising when it was first reported because the dogs had to only jump over a small barrier to escape from the shock. This initial study was followed up by Seligman and Maier (1967), showing that it is indeed the experienced lack of control over the first shock, and not the mere exposure to the shock itself, that was the reason for the dogs not showing any escape behaviours later on, even when they had control over avoiding the shock. This lead to the formulation of the learned helplessness hypothesis (Maier & Seligman, 1976). According to this theory, if an animal experiences uncontrollable aversive events, it will learn that its actions have no consequences on their environment, an experience of learned helplessness which generalises to new situations. In this context, the authors formulised control such that an animal is in control if the conditional

probability p of an outcome (called RF for reinforcer) following an action (called R for response) is different to the probability of the outcome following no action ($\bar{R}$):

$$p(RF|R) \neq p(RF|\bar{R})$$

In other words, if an outcome occurs no matter whether one acts or not, then there is no contingency between action and outcome and one does not have any control over the outcome. Since then, it has been shown that the phenomenon of learned helplessness can be found in humans and rats (e.g. Hiroto & Seligman, 1975; Seligman & Beagley, 1975), and can predict symptoms of depression (Alloy et al., 1999). The learned helplessness model proposes that when people repeatedly experience stressful events that appear outside their control, this disrupts their learning, they come to expect future events to be uncontrollable as well and develop a sense of overall helplessness and persistent anhedonia (Seligman, 1974).

Cognitive theories of depression posit maladaptive attributional styles as a root cause of depression (Abramson et al., 1989; Beck, 2008). These negative attributional styles describe the phenomenon that depressed patients tend to believe that negative events are caused by themselves (internal cause) and that this persists over time (stable cause) and acts on many aspects of their lives (global cause). In other words, depressed patients believe they have more control over negative than over positive events, and attribute the causes for positive and negative events differently. In contrast, healthy participants show the opposite, positive attributional bias, also referred to as a "self-serving bias" (Miller & Ross, 1975); they tend to attribute positive events such as successes to themselves (internal cause) while negative events are attributed to external causes such as chance or other people. One interpretation of this finding is that such positive illusions about oneself may enhance and protect one's own mental health and self-esteem, especially when faced with adversities (Taylor & Brown, 1988).

More recently, such self-serving biases when identifying causes for positive and negative outcomes have also been found using novel tasks together with reinforcement learning models (Dorfman et al., 2019; Zamfir & Dayan, 2022). Dorfman et al. (2019) investigated how people attribute positive and negative outcomes to themselves or other agents. This could be formally described by a computational model that described how participants learned about the value of the options given their beliefs about whether the agent intervened or not. Participants' learning was best described with a lower learning rate when they believed an agent intervened, and interestingly, this learning rate was higher for wins than losses. This means that participants attributed positive outcomes more to themselves and their own choices, and negative outcomes to the hidden agents.

When faced with neutral uncontrollable events, people with mildly depressive symptoms are more accurate at detecting this while nondepressed people tend to overestimate their level of control (Alloy & Abramson, 1979; Benassi & Mahler, 1985; Martin et al., 1984; Vázquez, 1987). This means that healthy people display what is sometimes called an "illusion of control" or "optimistic bias", while depressed people demonstrate a "depressive realism". Alloy and Abramson (1979) found this pattern using an experimental paradigm, called a response-outcome contingency task, in which participants learned whether they could control a green light by pressing or not pressing a button. Studies using this or similar tasks often employ the Delta P rule (Allan, 1980) to quantify the contingency between events (Figure 1.2). This describes a simple 2-by-2 case where an outcome (O) can be either present or absent in the presence or absence of a precedent such as a stimulus or action (also referred to as signal S in this context). By recording the frequency of events in each of the cells of the resulting 2-by-2 contingency table, one can then compute the conditional probabilities of how likely an outcome is when an action occurred (p(O|S)) and when no action occurred (p(O|-S)). Delta P is the difference between these two probabilities, and therefore a measure of the contingency

between one's actions (e.g. pressing the button) and the observed outcome (e.g. the green light). If Delta P is 0, i.e. the outcome is just as likely with and without an action, this means no contingency or control. If Delta P is positive, then actions facilitate the outcome, while a negative Delta P means taking an action prevents an outcome. Note that this formulation, in its core idea, is very similar to the way Maier and Seligman (1976) formulated control (see equation above). However, the Delta P rule applies to any event-event contingencies and not just to action-outcome contingencies.

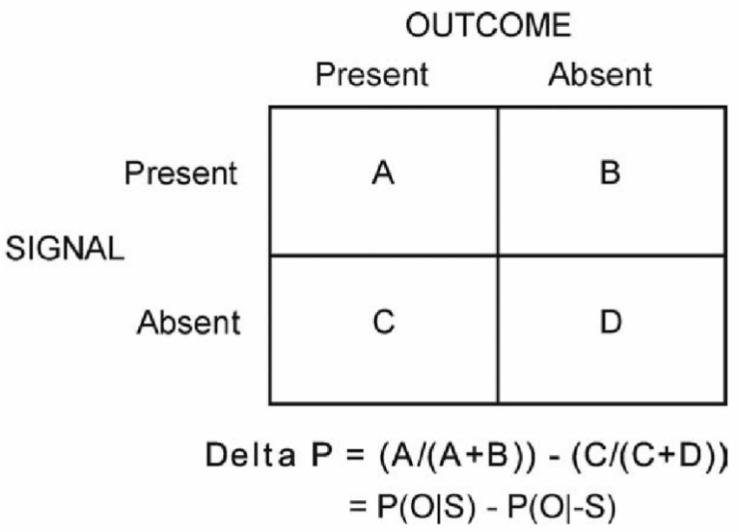


Figure 1.2. Delta P rule as a normative measure of contingencies between events. Reprinted from Murphy et al. (2005) with permission of Taylor & Francis Group.

Since the original studies describing depressive realism and illusion of control, other research has suggested that these effects are not quite as simple as the Delta P rule might suggest. It has been shown that the depressive realism effect is mediated by how often participants pressed the button, that is their activity level (Blanco et al., 2012). Specifically, people with depression respond less and for this reason, they estimate a lower level of control. In addition, studies found that the length of the waiting time in between trials (intertrial interval, ITI) impacts the contingency judgements of people with and without depressive symptoms differently (Msetfi et al., 2007; Msetfi et al., 2005). The reason for the ITI influencing people's continency

judgements is that people might perceive the ITI as an experimental condition in which there is neither action nor outcome. This means that people might see ITIs as evidence that if they do not make any action, then no outcome happens (i.e. cell D in Figure 1.2). This contrasts with the light being on only during the trial when one also presses the button to see if this controls the light. This means more exposure to a long ITI might lead to an overestimation of one's level of control. Indeed, if they were exposed to longer ITIs and had a low level of control, nondepressed participants overestimated their control, while depressed participants did not change or even decreased their control perception (Msetfi et al., 2005). Taken together, both the activity level and the overall context modulate the sense of control that people experience, but the degree to which this happens is modulated by depressive realism and individual variation in people's illusion of control.

More generally, there are several other drawbacks to the original formulation of how people estimate whether they are in control as a function of monitoring either $p(RF|R) \neq p(RF|\overline{R})$, from Maier and Seligman (1976), or the Delta P rule (Allan, 1980). One such drawback is that these formulations presuppose that only single actions are made and do not take into account the possibility of more realistic situations where people can take many different actions that might lead to the same outcome (Huys & Dayan, 2009); for example taking one out of two ways to get home. Another drawback is that this formulation can lead to illusions of control if outcomes are merely predictable but not controllable (Ligneul, 2021). For example, rain (outcome RF) is more likely to occur when the weather forecaster predicts rain (action R). However, even though this means that $p(RF|R) \neq p(RF|\overline{R})$, forecasters do not control the weather (Ligneul, 2021). Inspired by the engineering and information theoretic framework of control theory (Barnett et al., 2009; Touchette & Lloyd, 2004), Ligneul (2021) therefore suggested a reformulation of control in the following way:

$$p(RF|S) \neq p(RF|S,R)$$

Here, the main idea is that the probability of an outcome (RF) is not just contingent upon the agent taking an action (R) or not, but also on the state (S) of the system (or environment). This means that control can only be established if outcomes are contingent on actions over and above any contingency which already exists between outcomes and states of the environment. Coming back to the weather forecast example, if a) it rains (outcome, RF) if there are clouds in the sky (state, S), and b) if it also rains (RF) if there are clouds on the sky (S) and the forecaster announced rain (action, R), then it follows that p(RF|S)=p(RF|S,R) suggesting that the forecaster does, indeed, not control the weather. In other words, to estimate their control, people need to compare a passive "spectator" (p(RF|S) with an "actor" (p(RF|S,R)) model of their environment (Ligneul, 2021; Ligneul et al., 2022):

$$control = p(RF|S,R) - p(RF|S)$$

Note that the logic here is very similar to the Delta P model, but additionally takes into account the state-outcome transitions in the environment. Across three experiments, people have been shown to track their level of control by comparing such an actor with a spectator model (Ligneul et al., 2022). Interestingly, if people perform two tasks, the first one a more classic task exposing one group of participants to controllable electric shocks and the other group to uncontrollable shocks, then participants generalise their experience of being in control (or not) to a new second task in which they have to learn how much control they have over neutral stimuli (Ligneul et al., 2022). The group of participants exposed to uncontrollable stressors relied more on the spectator model, that is tended to ignore evidence that their actions influence their environment more, leading to a biased control perception. This effect was even more pronounced in participants with higher state anxiety scores.

Even though the formulation of control from the perspective of learned helplessness and depressive realism has been critiqued and revised since their original formulations, they still provide a useful framework, especially with more recent updates (Ligneul, 2021; Ligneul et al., 2022), to investigate how humans come to understand their level of control over their environment. One feature that all the proposed models have in common is that they encompass a comparison process between observing the environment in its default state and acting in order to infer control. This means that people need to strategically change their behaviour to estimate how this affects the environment.

1.4.2 Controlling one's actions: Understanding control from a motor perspective and links to schizophrenia

In the previous section, I described how in the learned helplessness and depressive realism literature, control is often conceptualised as, essentially, the difference it makes to the environment if an agent makes an action versus when it does not make an action. However, this requires the agent to have a precise understanding of its own behaviour and being able to control it in an intentional way. Being able to control one's own actions or movements leads to a perception or subjective experience that it is "I" who causes the action, which is also referred to as the sense of agency in this context (Gallagher, 2000; Haggard, 2017).

Perhaps the most prominent model of motor control or agency is the comparator model (Blakemore et al., 2000; Frith et al., 2000a, 2000b), illustrated in Figure 1.3. This model describes how a person comes to develop a sense of agency and that the central nervous system has a monitoring or internal "forward model" to do so. Specifically, when a person wants to make an action with a certain intention in mind, a motor command and a copy of the motor command, an efference copy, are issued. The motor command leads to the action which

interacts with the environment, leading to a sensory feedback. In contrast, an internal forward model uses the efference copy to predict the sensory feedback. Together, the experienced sensory feedback and the prediction are compared and result in a prediction error. Through this comparison, goal-directed movements are monitored and corrected. If the intended action matches the observed action then there is no prediction error, and then, perhaps at some higher level in the system, this leads to a sense of agency.

As predicted by this model, people report a diminished sense of agency if their actions produce unexpected tones (Ebert & Wegner, 2010; Majchrowicz & Wierzchoń, 2018), and they can experience an illusion of agency when events they did not cause happen to match their expectations (Sato & Yasuda, 2005; Wegner et al., 2004). This model was also used to explain why people cannot tickle themselves (Blakemore et al., 1999). Blakemore et al. (1999) tested this using a robotic arm setup in which participants either guided a robotic arm themselves to "tickle" their right hand (self-produced), or the same movement was externally produced or guided. Participants rated the externally produced sensations on their right hands as more tickly than if they were self-produced. Moreover, if a delay was introduced between how quickly the robotic arm executed the self-produced movement, then participants experienced this as more tickly as well. These findings were interpreted as evidence that there is an internal monitoring process that generates sensory predictions which are compared to the experienced sensory feedback, resulting in a prediction error which, in this case, leads to an experience of being tickled (Blakemore et al., 1999).

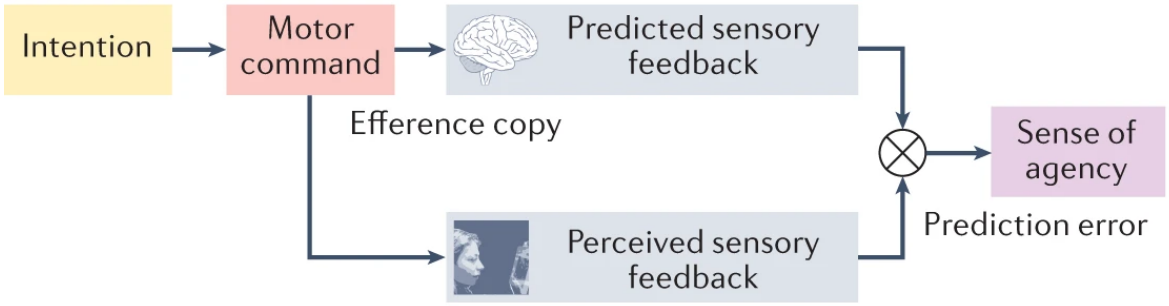


Figure 1.3. The comparator model of motor control. Adapted from Wen and Imamizu (2022) and reproduced with permission from Springer Nature.

This comparator model was also proposed to explain how people suffering from schizophrenia come to experience auditory hallucinations such as hearing voices, and passivity experiences such as delusions of control (Frith et al., 2000a, 2000b). Such control delusions or experiences of passivity include the patient feeling that actions, thoughts and even emotions are not the patients' but were made for them by some external agent, for example "The force moved my lips. I began to speak. The words were made for me." (Frith et al., 2000b; Mellor, 1970). Frith et al. (2000b) proposed that such delusions of control occur when the internal forward model fails to generate the predicted state of actions. This means that the comparison process between what is perceived and predicted is disrupted, ultimately leading to a distorted sense of agency and the patient feeling that they are not causing an action. In other words, while people with schizophrenia are aware that they intended to move, and that the movement happened, they are not aware of actually initiating the movement and therefore attribute it to external forces (Blakemore et al., 2000). It has indeed been shown that patients suffering from schizophrenia are worse at monitoring their own actions (Fourneret et al., 2002; Malenka et al., 1982), and fail to recognise their own movements as their own, instead attributing them to their environment (Farrer et al., 2003; Lindner et al., 2005). Past research found that patients with schizophrenia show deficits in predicting the consequences of their own actions (Shergill et al., 2005; Shergill et al., 2014), and these imprecise predictions result in relying more on

external cues of the environment (Synofzik et al., 2010; Voss et al., 2010). Such biases increase in social situations when patients are asked to infer the mental state of others, and it was suggested that this is caused by a faulty integration process of external sensory cues by others and patients' prior beliefs about others (Chambon et al., 2011).

1.4.3 Comparison of the theoretical frameworks of control

The studies described in the previous two sections attempt to understand control and its impacts on human experience and behaviour from different perspectives. Accounts focussing on attributional styles (Abramson et al., 1989; Beck, 2008) are a cognitive approach to understanding how people retrospectively think about a past event and its causes. Depressive realism highlights the importance of the accuracy of control beliefs, and that depression is linked to more accurate control estimates in certain contexts. In contrast, the theory of learned helplessness (Seligman, 1974) puts a stronger focus on the motivational and emotional consequences of experiences of low control and how such experiences affect behaviour in new situations. The comparator model (Frith et al., 2000a, 2000b) uses a sensorimotor perspective on how a sense of control arises from an internal comparison between the intentions that drive our actions and the sensory feedback that results from actual actions. In contrast, reinforcement learning provides a mechanistic understanding of how different control beliefs and attributional biases may unfold on a computational level and from a learning perspective (Dorfman et al., 2019; Zamfir & Dayan, 2022). In sum, these different frameworks suggest that control perception is a multifaceted construct which involves the internal monitoring of our actions (comparator model), learning action-outcome contingencies (reinforcement learning), attributions to oneself versus external causes (attributional styles) which can be more or less adaptive or accurate in psychiatric conditions (learned helplessness and depressive realism).

Some of the accounts of control complement each other. For instance, the comparator model might provide a motor explanation for how the phenomenon of learned helplessness arises. If an animal experiences electric shocks in a chamber it cannot escape from, initially the animal could try different actions to escape. However, since it cannot escape, the animal will experience large prediction errors between its intended action outcome (experiencing no electric shock) and the actual outcome (suffering from electric shocks). Ultimately, these prediction errors lead to a low sense of agency, according to the comparator model. Similarly, the theory of learned helplessness posits that an animal learns a low control level from repeated experience of having no control over its environment. In the learned helplessness framework, the acquired sense of low control generalises to passivity in new situations and a failure of updating of the low control belief. The comparator model suggests that this failure of control belief updating is caused by the animal not generating any actions which can result in prediction errors. This means that if the animal stays passive and does not try to escape the new electric shocks, there are no actions which lead to prediction errors and which can update the animal's sense of agency. In this way, passivity can lead to less learning and persistent maladaptive control beliefs.

In my thesis, I will draw on concepts from reinforcement learning and attributional biases to investigate control and causal learning. Reinforcement learning complements the attributional bias literature. This is because reinforcement learning algorithms are mathematical models that can be used to quantify psychological concepts such as attributional biases. In my thesis, I will use concepts from reinforcement learning to quantify attributional biases. Specifically, I will examine how humans attribute predictions errors to themselves and others in relation to their estimate of control (chapter 2), and as a function of depression and schizotypy (chapter 4).

While some control frameworks are complementary, others make different predictions about control experiences. For instance, accounts based on differences attributional styles suggest that people with depressive symptoms assign negative events to themselves, which means that they believe they have a lot of control over negative outcomes. The learned helplessness theory makes slightly different predictions. It posits that after experiencing low control over negative outcomes (such as electric shocks), animals will not realise they have any control over negative events (electric shocks) in new situations. In other words, the learned helplessness theory suggests that animals do not believe they have any control over negative outcomes.

Despite these differences, there are also some important similarities between the different frameworks for understanding control. Both reinforcement learning and the comparator models emphasize predictions and prediction errors. In reinforcement learning, the agent predicts future rewards, and updates its beliefs from the discrepancies (reward prediction errors) between the obtained outcome and its prediction. Similarly, the comparator model proposes that a sense of agency arises from the brain generating a prediction of the intended action outcome which is compared with the experienced action outcome (sensory feedback), generating a prediction error.

More generally, the frameworks of reinforcement learning, the comparator model, learned helplessness, the Delta P model and the model proposed by Ligneul et al. all involve some form of a comparison process to infer control. This means that to derive a sense of control, the agent needs to compare either observed and predicted action outcomes (comparator model, reinforcement learning) or the likelihood of outcomes caused by the environment and one's own actions (Ligneul et al., Delta P model, learned helplessness). In my thesis, I investigate how humans estimate their control in social environments when cooperating with others. I hypothesize that a similar comparison process might be at play when inferring control in social

contexts. However, an important difference might be that this requires taking into account not just the presence or absence of one's actions and states of the environment, but I also suggest that it is important for an agent seeking to learn what control it has, to make varied, exploratory actions and to observe the outcomes of those actions. In chapter 2, I will investigate this on a behavioural and cognitive level, before characterising the neural mechanisms of these processes in chapter 3. In chapters 4 and 5, I will investigate whether these mechanisms change as a function of depression and schizotypy.

1.5 Summary and relevance to the current thesis

In this thesis, I investigate the mechanisms involved in how people assign themselves and others the credit for joint outcomes and how they establish their control in such social interactions. Since I investigate this in social contexts, some of the behavioural and neural mechanisms involved in credit assignment and control learning might be different to the ones described in the previous sections. For example, from the perspective of learned helplessness and depressive realism, control is typically estimated in a simplified case of either performing an action or not, and an outcome being present or not. In this thesis, I present studies in which I investigate a more complicated but arguably an ecologically more frequent situation in which there are multiple agents present together.

In social situations, people often need to estimate their level of control as opposed to the level of control that another person has. If they are cooperating, the outcomes are continuously varying and not just binary. As a result, the degree of control experienced by a person might be similarly graded and varying rather than binary. I present a framework for how people can learn their control from ambiguous feedback in social contexts, which requires not just taking into account the presence of actions but also people's own skill levels. For this, I will draw on the frameworks of information seeking and causal interventions. I will, therefore, investigate

how humans engage in exploration and change their behaviour to actively disambiguate feedback in social situations. As a result, I show that they are able to estimate their level of control. In this thesis, I not only investigate the cognitive and behavioural mechanisms involved in social credit assignment and control learning (chapter 2), but also how these processes unfold on a neural level (chapter 3) and how they link to mental health on a behavioural (chapter 4) or neural level (chapter 5). In the next two sections, I will discuss social versus non-social processing, relevant brain areas and mental health constructs.

1.6 Social versus non-social processing

Social environments put unique demands on animals (including humans), as described at the beginning of this chapter (section 1.1). For instance, while working together with others might allow humans to achieve their goals more quickly, it requires monitoring other people (Melis & Raihani, 2023). Compared to fish and reptiles, primates have unusually large brains in relation to their body size. The social brain hypothesis argues that primates have more complex social lives than other species, and it has been found that social complexity correlates with relative brain size in primates (Barrett et al., 2002; Dunbar & Shultz, 2007). Living in larger groups leads to increases of grey matter of the mid-superior temporal sulcus and rostral prefrontal cortex in the primate brain (Sallet et al., 2011). However, since its formulation, the social brain hypothesis has been challenged and it was argued that primate brain size reflects general intelligence, including not only social, but also foraging and other skills (Reader et al., 2011).

An ongoing debate exists about whether certain neural, cognitive, and behavioural processes are uniquely social, and if so, which ones. Lockwood et al. (2020) argued that a process is socially specific if its algorithm or implementation is different from non-social processes, and if domain-general processes can be ruled out. Past research suggests that while social and non-social reward-based learning are mediated via neighbouring divisions of the anterior cingulate

cortex, information is ultimately combined in the ventromedial prefrontal cortex (Behrens et al., 2008). This work, together with other research (e.g. Boorman et al., 2013; Ruff & Fehr, 2014), suggests that at least in some brain regions, such as ventromedial prefrontal cortex, social information processing is not fully separate from non-social information processing. Trudel et al. (2023) investigated whether cognitive and neural mechanisms differ when people learn from social versus non-social sources. They found that even though the social and non-social task conditions were tightly matched, people were more confident when learning from social advisors, and activity in the dorsomedial prefrontal cortex and posterior temporoparietal junction tracked social confidence more strongly than in the non-social condition. Taken together, while some brain regions appear to be more specialised for social functioning, others support more domain-general processes.

It might be the case that some of the cognitive, behavioural and neural mechanisms of credit assignment and control learning shift in social and non-social contexts. In the learned helplessness framework, animals' behaviour was characterised when they were isolated from others and exposed to negative outcomes such as electric shocks (Hiroto & Seligman, 1975; Overmier & Seligman, 1967; Seligman & Beagley, 1975; Seligman & Maier, 1967). The classic depressive realism studies investigated how humans inferred their level of control over neutral stimuli such as a green light (Alloy & Abramson, 1979). Ligneul et al. (2022) investigated how humans infer their level of control over state transitions that could be either controlled by their own actions or a pre-determined schedule. In other words, in these experimental tasks, control was shared between the agent and a non-social entity – the general task context. However, humans are a very social species and therefore it is useful to understand how these processes unfold in social contexts. For example, Ligneul (2021) proposed that control can be inferred from comparing outcomes obtained via a passive "spectator" model (observing the environment in its default state) and an "actor" model (performing an action). A similar

mechanism might be at play when humans cooperate with others towards a shared outcome. In this thesis, I will investigate how humans learn their level of control and attribute outcomes to themselves and others on a behavioural (chapter 2) and neural level (chapter 3), and how these processes are altered as a function of depression and schizotypy (chapters 4 and 5).

1.7 Brain areas relevant for control and causal learning

The studies reviewed in sections 1.2, 1.3 and 1.4 focused on how human and non-human primates assign credit, seek information and estimate their control. They investigated how stimulus- or action-outcome associations are learnt in non-social settings and found regions in the frontal cortex such as the medial and orbitofrontal cortex. In this thesis, I investigate brain regions relevant for control and causal learning in social contexts. Since control and causal learning during cooperation involves different processes such as action monitoring (as reviewed before), a different set of brain substrates may be involved than in non-social contexts. In the studies which I conducted for this thesis, I focus on regions in the inferior parietal lobule and striatum and will give an overview over their anatomy and function in this section.

1.7.1 Supramarginal gyrus

The parietal cortex in the human brain is associated with a multitude of cognitive functions such as perception, action control and more higher-level cognition. One key area of interest for the studies I present in this thesis is the supramarginal gyrus (SMG, green and blue in Figure 1.4) which mainly comprises Brodmann's area 40 and is located in the more anterior part of the inferior parietal lobule. Often, the anatomical and functional properties of brain areas are better understood in the macaque than in humans, because invasive techniques have been used in the former but not the latter. In the macaque brain, Brodman area 7b is equivalent to the SMG and is positioned on the exposed surface of the anterior part of the posterior parietal lobe.

Using neuroanatomical tracers, it was found that area 7b has strong connections with somatosensory and motor areas such as the primary and secondary somatosensory cortex, representing head and forelimb positions, but also prefrontal regions involved in decision-making such as area 45 and the ventrolateral prefrontal / lateral orbitofrontal cortex (Cavada & Goldman-Rakic, 1989a, 1989b). Some information is available about the likely anatomical connections of human area 40 from studies employing diffusion-weighted magnet resonance imaging (MRI) and resting state functional MRI (fMRI) to estimate which brain areas are most likely to be connected (Mars et al., 2011). This approach yielded a set of five subclusters, two of which are located in the SMG (green and blue in Figure 1.4). Consistent with findings from macaque neural tracer studies, the SMG in the human brain is strongly interconnected with the dorsolateral prefrontal cortex (dlPFC) (Mars et al., 2011). The anterior part of the SMG is interconnected with the ventral premotor cortex (Mars et al., 2011), and indeed both are active during reaching and grasping movements (Tomassini et al., 2007). The posterior part of the SMG is connected to the anterior prefrontal cortex (Mars et al., 2011), and both activate when people decide to change their behaviour based on accumulating evidence (Boorman et al., 2009; Daw et al., 2006). Taken together, these results show that the inferior parietal lobule has similar anatomical connections across humans and macaques and that the SMG is a region at the intersection between lower-level motor regions and higher-level prefrontal regions.

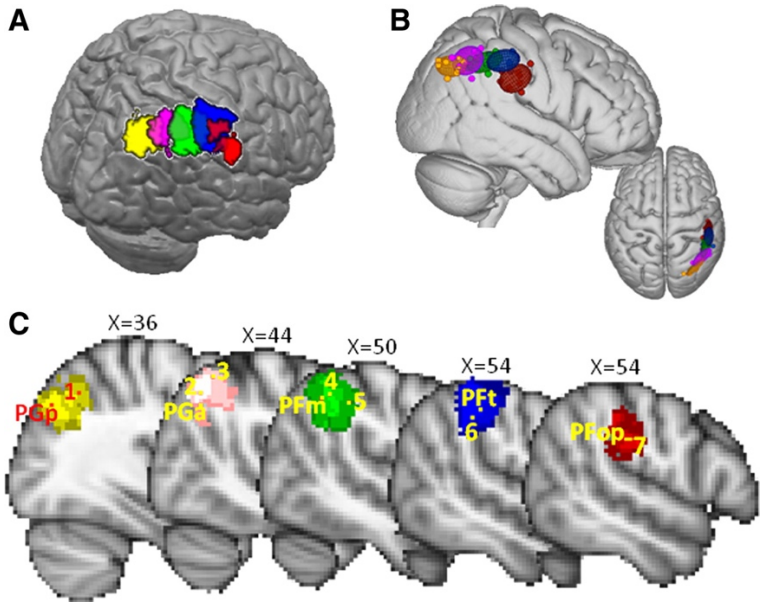


Figure 1.4. Tractography-based parcellation of the inferior parietal lobule. The SMG is highlighted in green (posterior) and blue (anterior part). Reprinted from Mars et al. (2011).

The SMG seems to be involved not just in movement execution but in more abstract motor observation, tool use and representing actions in terms of their purpose and meaning. In this paragraph, I will describe the evidence for this in more detail. The left SMG, in particular its anterior part, tracks actions not simply as hand movements but rather in a more abstract manner in terms of, for example, the kinematics of tool-related actions (Orban & Caruana, 2014; Peeters et al., 2009; Peeters et al., 2013), and shows stronger activations for pantomimed compared to real tool use actions (Chen et al., 2023) and activates during planning of pantomimed tool use (Króliczak & Frey, 2009). More generally, the SMG appears to track movement intentions; when the SMG is electrically stimulated, patients suddenly feel the intention to move and if the stimulation intensity is increased, believe that they did move even though no movement occurred (Desmurget et al., 2009). SMG activity is associated with motor attention in terms of preparing a movement (Rushworth et al., 2001b), and SMG lesions or disrupting SMG activity with TMS lead to deficits in motor attention (Rushworth et al., 2001a; Rushworth et al., 1997). When participants are asked to respond to cued target stimuli, the

SMG is involved in directing attention to stimuli that appear at cued locations (spatial attention) or at cued time points (temporal attention) (Coull & Nobre, 1998).

Relevant in particular for the topic of this thesis, the SMG tracks action outcomes in terms of action attributions and agency judgments. As described in the earlier section 1.4.2, theoretical models of motor control and agency posit that a person attributes action outcomes to themselves and feels a sense of agency if the internal forward model accurately predicts the sensory feedback that is experienced when an action is executed (Frith et al., 2000a, 2000b; Wolpert et al., 1995). In the healthy brain, SMG activity is associated with the degree of discrepancy (prediction error) between the experienced and predicted sensory feedback, leading to a lower sense of agency (Farrer et al., 2003). Sirigu et al. (1999) investigated whether the internal sensory prediction model is faulty in patients who suffer from apraxia, an inability to execute skilled motor actions in a voluntary manner and which has been linked to left parietal damage (Heilman et al., 1982). Participants had to produce hand gestures and observe them on a screen which showed either their own hand movement or one produced by an experimenter, while the hands were covered with identical gloves. Apraxic patients with lesions in the left SMG correctly attributed hand movements to themselves if they produced them, and could distinguish hand movements produced by the experimenter if the externally produced hand movement was different to their own. However, compared to controls, apraxic patients showed deficits in recognising a hand movement produced by the experimenter if it was congruent with the hand movement they executed at the same time. Interestingly, if they executed the instructed movement worse than the experimenter did on screen, they still attributed the movement on the screen to themselves. This was interpreted as evidence that apraxic patients with lesions in SMG fail to form an accurate prediction of their own movement and instead overly rely on external feedback, leading to biased action attribution. fMRI studies in healthy participants provided further evidence that SMG activity is associated with attributing

actions to oneself or others (Farrer & Frith, 2002; Ohata et al., 2020). Overall, this suggests that the SMG mediates agency estimation.

In section 1.4.2, I described that the comparator model of agency was proposed to explain delusions of control (passivity) and hallucinations in schizophrenia (Frith et al., 2000a, 2000b). In short, it proposed that these symptoms occur due to a lack of self-monitoring, resulting in a discrepancy between intention and execution and attributing experiences to external forces. Spence et al. (1997) found that during movement tasks, patients with schizophrenia and current passivity symptoms show increased activity in the SMG compared to patients with schizophrenia but without passivity symptoms, and healthy controls. Interestingly, this increased SMG activation reduced in a 4-6–week follow-up for patients whose passivity symptoms lessened in that time. This suggests that an internal action monitoring process in the SMG can go awry in patients suffering from schizophrenia and delusions of control.

1.7.2 Striatum

The striatum has been associated with reward processing, decision-making and voluntary motor control. Anatomically, it is typically divided into a ventral and dorsal part. The ventral striatum consists of the nucleus accumbens and the olfactory tubercle. The dorsal striatum can be further divided into its dorsomedial part, the caudate, and its dorsolateral part, the putamen (Figure 1.5A). The internal capsule, a white matter tract between cortex and brain stem, separates the caudate from the putamen. There are no cytoarchitectonic or histochemistry distinctions that can be used to separate the ventral from the dorsal striatum, and typically, their afferent projections are used to distinguish them (Haber & Knutson, 2010). The striatum receives anatomical connections from cortex, thalamus, midbrain, and brainstem (Haber, 2016; Haber & Knutson, 2010; Selemon & Goldman-Rakic, 1985; Selemon & Goldman-Rakic, 1988). The striatum receives dopaminergic inputs from the ventral tegmental area and substantia

nigra, located in the midbrain. Cortical projections are organised topographically (Figure 1.5A-B) (Haber, 2016). Motor and premotor cortex project mainly to the dorsolateral and central putamen but also slightly into the caudate (Figure 1.5A, blue). The dlPFC projects mostly to the caudate head and the medial parts of the putamen (Figure 1.5A, yellow). Projections from orbitofrontal cortex (OFC) and dorsal anterior cingulate cortex (dACC) terminate in the more rostral parts of the striatum, including the ventral striatum as well as more medial caudate and putamen. Ventromedial prefrontal cortex (vmPFC) projections are mainly found in the ventral striatum and on the medial border of the caudate, next to the ventricle.

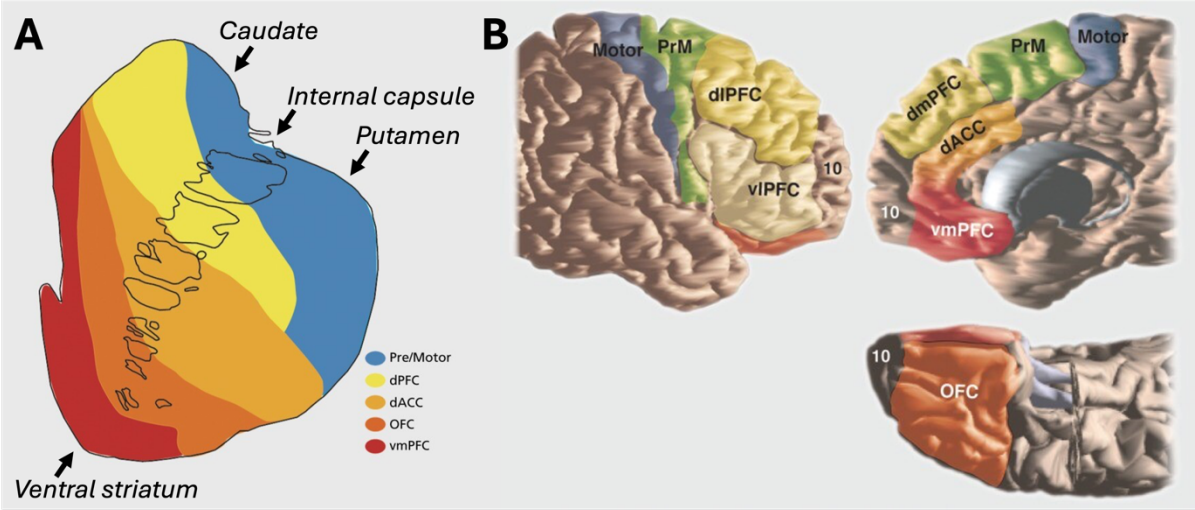


Figure 1.5. **Anatomical connections received by the striatum (A) from the cortex (B).** Reprinted from Haber (2016) by permission of Informa UK Limited, trading as Taylor & Francis Group.

Understanding these anatomical connections received by the striatum is useful to understand the variety of functions it is implicated in, from motor control functions (McFarland & Haber, 2000; Toni & Passingham, 1999) to reward processing (Cromwell & Schultz, 2003). The striatum is a key region in the reward circuit and receives dopaminergic input from the midbrain. Schultz et al. (1997) found that dopaminergic neurons in the midbrain encode reward prediction errors, and neural correlates of reward prediction errors are found in particular in the ventral but also the dorsal part of the striatum (McClure et al., 2003; O'Doherty et al., 2003;

Rutledge et al., 2010). O'Doherty et al. (2004) found that dorsal and ventral striatum encode prediction errors differently as a function of whether participants needed to act to receive the reward (instrumental conditioning) or not (Pavlovian conditioning). Prediction errors in the passive condition were represented in the ventral striatum, while both ventral and dorsal striatum tracked them in the active choice condition (O'Doherty et al., 2004). A similar result was found for rewards which participants were either able to actively chose or passively received them (Zink et al., 2004). Other work has found that striatum activity dissociates between when participants anticipate to make choices or not in order to receive rewards (Leotti & Delgado, 2011). Dorfman et al. (2021) investigated how the neural representations of reward prediction errors change when people believe they have control over them or not. For this, they used the same task as Dorfman et al. (2019) (described in section 1.4.1), in which participants chose between options which gave them either losses or wins. However, a key manipulation was that there were hidden agents that could interfere with the outcomes from the options. Using a reinforcement learning model to describe participants' behaviour, they found that participants adapted their learning based on their control beliefs; they attributed outcomes less to themselves if the hidden agent intervened, replicating Dorfman et al. (2019). The authors found that the ventral striatum carried a neural signal for reward prediction errors. In contrast, the dorsal striatum (putamen and caudate) tracked reward prediction errors modulated by participants' belief about their control over the prediction error. This, together with the studies described before, suggests that the ventral and dorsal striatum encode reward and learning signals as a function of whether participants believe they have any control in the current context. More generally, this suggests that causal beliefs modulate how the striatum encodes reward prediction errors. Indeed, it has recently been suggested that dopamine released in the ventral striatum, rather than supporting prospective learning via reward prediction errors (Schultz et al., 1997), encodes retrospective credit assignment instead (Jeong et al., 2022).

1.8 Links to mental health and computational psychiatry

In this thesis, I investigate the cognitive and neural mechanisms of control and causal learning in social contexts. Two mental health constructs often linked to biased causal attributions and control beliefs are depression and schizophrenia. In chapters 4 and 5 of my thesis, I investigate how people differ as a function of their depressive symptoms and schizotypy scores. Here, I will outline previous research on depression and schizotypy (and schizophrenia) that inform the approach that I have taken in my own studies.

Psychiatric disorders such as schizophrenia and major depressive disorder (MDD) are diagnosed with self-reported symptoms, using, for example, the Structured Clinical Interview for DSM-5 (SCID-5). This involves classifying individuals as healthy or, for example, depressed according to the number of symptoms experienced. Thresholds of sum scores are traditionally used in self-report rating scales such as the Beck Depression Inventory (Beck et al., 1988) or the Quick Inventory of Depressive Symptomatology (Rush et al., 2003). This approach has the advantage of having good face validity and psychiatrists can more easily diagnose, for example, MDD despite the heterogeneity of possible depressive symptoms. However, it has been suggested that to identify better biomarkers and develop more effective antidepressants, it is necessary to understand specific depression symptoms rather than reducing them to one diagnosis or sum score (Fried & Nesse, 2015). Another shortcoming of psychiatric classification systems such as the DSM-5 are that they lack any mechanistic explanations, and instead, clinicians diagnose based on symptoms rather than considering what the biological or psychological roots of a diagnosis are (but see the Research Domain Criteria [RDoC] project, e.g. Cuthbert and Insel (2013)). To improve our understanding of psychiatric disorders, their diagnosis and treatment, the goal of the computational psychiatry is to describe the

neurocognitive mechanisms of psychiatric disorders using computational models (Adams et al., 2016).

1.8.1 Depression

Depression is the most common mental illness and the leading cause of disability, according to the World Health Organization (2017). It is characterised by a variety of symptoms, including feelings of sadness, guilt or low self-worth, interest or pleasure, as well as bodily symptoms such as disturbed sleep and appetite. Depression has been linked to negative biases in memory, attention, perception and executive functions (Clark et al., 2009). Depressed patients tend to interpret ambiguous information in a negative way, remember negative information for longer and pay more attention to negative information (Bradley et al., 1995; Gur et al., 1992; Mogg et al., 1995).

As previously described in section 1.4.1, cognitive theories of depression propose that maladaptive attributional styles are a root cause of depression (Abramson et al., 1989; Beck, 2008). These attributional styles describe negative biases by which people with depressive symptoms attribute negative events internally (to themselves), and that this is stable over time and global. In other words, they believe that they cause negative events to happen, all the time and in many different situations of their lives. This suggests that people with depression believe they have more control over negative events than positive ones. In addition, compared to the healthy population, people with depression are more sensitive to having no or a low level of control over neutral events and this has been investigated under the term of “depressive realism” (Alloy & Abramson, 1979; Benassi & Mahler, 1985). Section 1.4.1 includes a more detailed discussion of the findings on depressive realism and negative attributional biases in depression. Despite long-standing findings on the importance of attributional biases in depression, how these biases arise and are maintained is less well understood. To address

this, in my thesis, I use credit assignment as a computational framework and quantitative description of how people attribute outcomes to potential causes.

People with depression typically view themselves negatively and withdraw socially, leading to social isolation and poorer mental health. Perceived social support and large, diverse social networks are protective factors against depression (Gariepy et al., 2016; Santini et al., 2015). Experiencing social exclusion can lead to higher sensitivity to rejection and low levels of social self-efficacy, that is believing that one is able to initiate social contact and build new friendships, which together increase the risk of depression (Niu et al., 2023; Sun et al., 2021). In tasks with facial emotional expressions, people with depression and those at risk of depression are worse at recognising positive faces (Chan et al., 2007; Harmer et al., 2009; Surguladze et al., 2004). Based on the literature on how depression is linked to social deficits, in my thesis, I investigate how credit assignment and outcome attribution evolve in social contexts.

In the field of computational psychiatry, a large number of studies have tried to characterise how reinforcement learning mechanisms are altered in depression. A meta-analysis across case-control studies found that depressed groups show impaired reward processing compared to nondepressed groups (Halahakoon et al., 2020). For example, people with depression show a weaker reward bias, that is a lower tendency to choose more frequently rewarded stimuli. In addition, depression is linked to slower reinforcement learning. However, the studies on reinforcement learning show a considerable heterogeneity in their results and use different parameters. To address the latter and estimate one common model space across studies, Pike and Robinson (2022) employed a simulation approach and found that depressed patients show higher punishment learning rates and lower reward learning rates. However, further research is needed to confirm these simulation results. One conclusion that might be drawn from such meta-analyses is that there is not a single learning-linked process that is uniquely and

consistently altered in depression but rather the existing evidence is consistent with the possibility of alteration in several different learning-related mechanisms. In this thesis I especially focus on the possibility that a process of credit assignment – the linking of an outcome to its specific cause – is altered in depression.

Work on learned helplessness in rats has demonstrated that uncontrollable stress leads to an activation of serotonin (5-HT) neurons in the dorsal raphe nucleus (Grahn et al., 1999; Maswood et al., 1998). The dorsal raphe nucleus has strong projections to the dorsal part of the striatum (Imai et al., 1986; Van der Kooy & Hattori, 1980). Additionally, and as discussed previously (section 1.7.2), the striatum receives dopaminergic projections from the midbrain and tracks reward prediction errors. Depression has been related to an irregular functioning of the dopaminergic system (see e.g. review by Papakostas, 2006). fMRI studies found that patients suffering from MDD show weaker reward prediction error signals in the striatum in reinforcement learning tasks (Dombrovski et al., 2013; Gradin et al., 2011; Kumar et al., 2008). If patients experienced a higher number of major depressive episodes, striatal signals of reward prediction errors become more blunted (Kumar et al., 2018). Indeed, meta-analyses across neuroimaging studies found that people with depression show reduced striatal activation related to rewarding outcomes (Ng et al., 2019; Zhang et al., 2013), and that this is even the case preceding the onset of depression in adolescence (Keren et al., 2018). In addition, MDD is associated with frontal brain regions in reward processing, such as increased activation in the orbitofrontal cortex and anterior cingulate cortex (Ng et al., 2019; Zhang et al., 2013). Interestingly, it has also been found that the inferior parietal lobe is increasingly activated during monetary reward processing in MDD (Zhang et al., 2013). The supramarginal gyrus has been found to track belief updating to bad news more strongly in patients with MDD compared to healthy individuals (Garrett et al., 2014). Taken together, this suggests that in depression,

the striatum and inferior parietal lobule are differentially involved in learning, potentially even in control and causal learning. I will investigate this in chapter 5.

1.8.2 Schizotypy and schizophrenia

Schizophrenia is a psychiatric disorder typically characterised by impairments in how individuals perceive reality and behavioural symptoms. People suffering from schizophrenia show symptoms such as persistent delusions and hallucinations, experiences of passivity (delusions of control) and disorganised thinking and behaviour. Symptoms of schizophrenia are typically classified as “positive” symptoms such as hallucinations and delusions, and “negative” symptoms such as social withdrawal and restricted experience of emotions. Schizotypy refers to personality traits and is often considered a subclinical expression of schizophrenia (Claridge, 1997; Lenzenweger, 2011; Raine, 2006). Akin to symptoms of schizophrenia, schizotypy can be characterised by several symptom groups; positive (subsyndromal hallucinations and delusions), negative (reduced emotional, social and physical functioning) and disorganised schizotypy (bizarre behaviours and disorganised thoughts) (Ettinger et al., 2015; Raine, 2006). Schizotypal traits are a risk factor for developing schizophrenia, although not everyone with schizotypal traits develops schizophrenia later on (Gooding et al., 2005; Meehl, 1962). Twin and family studies found that the relatives of people with schizophrenia are more likely to show schizotypal features (Kety et al., 1976; Mata et al., 2003).

Social cognition and cognitive biases play an important role for how well people with schizophrenia spectrum disorders can function in their daily lives (Halverson et al., 2019; Pinkham et al., 2014). Schizophrenia has been related to a self-serving attributional style, also referred to as externalizing bias in the psychosis literature (Kinderman & Bentall, 1996). This describes an attributional style in which people blame themselves less and others more for

negative events, and credit themselves for positive events (see also section 1.4.1 for self-serving biases in the healthy population). The “delusion of defence theory” proposed that such self-serving biases are critical for developing psychotic symptoms, and that they protect patients with persecutory delusions from low self-esteem (Bentall et al., 2001; Bentall et al., 1994). However, the cost of attributing negative events to other people can be a perception of hostility in one’s social environment, leading to a paranoid world view (Bentall et al., 2001; Bentall et al., 1994). Meta-analyses found that indeed, patients with schizophrenia spectrum disorders and persecutory delusions show a more pronounced self-serving bias than healthy controls, people with depression and patients with schizophrenia spectrum disorders but without persecutory delusions (Mueller et al., 2021; Murphy et al., 2018).

The comparator model of agency was developed to explain hallucinations and delusions of control in people suffering from schizophrenia (Frith et al., 2000a, 2000b), as described in more detail in section 1.4.2. In short, this model suggests that when an action is performed, the brain predicts the sensory feedback it expects and compares this with the actually experienced sensory feedback. If this comparison process results in no or a small discrepancy (or prediction error), then the individual experiences a sense of agency (“I” caused this action). Delusions of control and hallucinations in schizophrenia occur when the prediction process is disturbed which means that individuals are not aware of initiating a movement and therefore attribute it to alien forces (Blakemore et al., 2000; Frith et al., 2000b). Section 1.4.2 describes these findings and more recent studies on schizophrenia and the sense of agency in more detail. Interestingly, similar findings have been found relating schizotypy to the sense of agency. Asai (2016) investigated how healthy individuals process feedback which was morphed across movements participants made themselves and pre-recorded movements by others. Participants with higher positive schizotypy scores were worse at discriminating when the feedback was based more on their own movements and when it was more based on the other.

This means that they estimated their level of control or agency over the feedback differently than individuals with low schizotypy scores. Other research found that healthy participants with higher scores in schizotypy are more likely to confuse hand movements produced by an experimenter as their own, even if the participant does not move their hands but hears auditory instructions which describe the observed hand movement (Moore et al., 2021). This suggests that similar to findings in schizophrenia (Synofzik et al., 2010; Voss et al., 2010), people with higher schizotypy rely more on external cues to derive their sense of agency.

Neuroimaging work examining the sense of agency in schizophrenia has identified the inferior parietal lobe, in particular the supramarginal gyrus and adjacent angular gyrus. Spence et al. (1997) found that patients suffering from schizophrenia and current delusion of control symptoms show increased activations in the supramarginal gyrus, compared to healthy individuals and patients with schizophrenia but without current delusions of control (more details in section 1.7.1). This hyperactivation of the supramarginal gyrus was not present after a month when delusion of control symptoms remitted. Cued-choice versus free-choice trials change fronto-parietal connectivity in healthy individuals but not in patients with schizophrenia (Voss et al., 2017). Compared with healthy individuals, patients suffering from schizophrenia and schizotypal individuals show reduced grey matter volumes in regions of the inferior parietal lobe such as the supramarginal gyrus and adjacent angular gyrus (Venkatasubramanian et al., 2011; Zhou et al., 2007).

Another line of evidence found that dopamine plays a role in schizophrenia, originally based on the finding that dopamine receptor antagonists reduce auditory hallucinations and delusions (“dopamine hypothesis” of schizophrenia) (Carlsson, 1978). Schizophrenia has a heritability of 83% (Cannon et al., 1998), and dopamine receptor upregulation in the caudate has been related to the genetic risk of schizophrenia (Hirvonen et al., 2005). Dopamine is traditionally

linked to reward prediction errors in the context of reinforcement learning (Schultz et al., 1997) (but see also Jeong et al. (2022) for an interpretation of dopamine function in terms of credit assignment as opposed to prediction errors). During reward anticipation, unmedicated patients with schizophrenia show reduced activation of the ventral striatum (Esslinger et al., 2012; Juckel et al., 2006; Nielsen et al., 2012). A similar reduction of ventral striatal activity to anticipated rewards can be found in first-degree relatives of patients with schizophrenia (Grimm et al., 2014). When anticipating either rewards or losses compared to neural outcomes, healthy individuals show activations in the caudate and ventral striatum, while this activation is absent in participants with schizophrenia (Mucci et al., 2015). During outcome observation, Schlagenhauf et al. (2014) found blunted reward prediction error signals in the ventral striatum of patients with schizophrenia compared to healthy individuals.

Taken together, in this section, I outlined past findings on altered attributional styles and control beliefs in schizophrenia and schizotypy. In my thesis chapter 4, I will investigate how control and causal learning are altered as a function of schizotypy and depression. Finally, in chapter 5, I will examine the neural correlates of behavioural differences related to schizotypy and depression. Given the past literature, the supramarginal gyrus and striatum might be potential candidate regions for any interindividual differences related to schizotypy.

CHAPTER 2 Controllability and cause in human collaboration

2.1 Introduction

Humans are inherently social and collaborate towards shared goals. Achievements such as the moon landing were only made possible by people working together. To foster effective collaboration, it is therefore important to establish how much each individual contributes. In other words, it is essential to identify the causes behind shared outcomes. For instance, when a student receives mediocre feedback on a research project, she must determine the reason for it in order to improve. Whenever there are multiple antecedents to an event, credit assignment is needed to determine which factors are causally important and to what extent. It is especially important to determine the extent to which you, as an individual, as opposed to other people, caused a particular outcome to occur. For example, if the student took on a greater share of the project workload than her collaborators then the outcome reveals something important about her ability. However, the outcome reveals little about her ability if she had only limited control over the project's conduct. Therefore, credit assignment also requires assessing the level of responsibility, or control each agent has. Previous studies have examined how people attribute outcomes to one choice or another (Akaishi et al., 2016; Jocham et al., 2016; Noonan et al., 2017; Scholl et al., 2015). However, it remains an open question how people attribute ambiguous feedback to themselves or others under varying degrees of controllability, and how they infer their control over outcomes. Understanding how we ascertain control levels is important not just because it is a fundamental prerequisite for learning in humans and other animals but because alterations in perception of control have been identified in psychological illnesses such as depression (Alloy & Abramson, 1979; Msetfi et al., 2005) and schizophrenia (Moritz et al., 2014; Shergill et al., 2014).

It is clear that establishing the causes of outcomes is difficult; an outcome may be assigned not just to the action that led to it, but also to one's overall skill (Mancinelli et al., 2021; Rouault et al., 2019; Zamfir & Dayan, 2022), or to other choices (Jocham et al., 2016; Noonan et al., 2017; Thorndike, 1933; Walton et al., 2010; Wittmann et al., 2020) and stimuli (Akaishi et al., 2016; Boorman et al., 2016; Fischer et al., 2017; Scholl et al., 2015) that occurred close in time. Moreover, the problem is acute in social contexts in which multiple agents interact; in such situations people frequently confuse their own performance with that of others they cooperate with or compete against (Wittmann et al., 2016b; Wittmann et al., 2021). Theories of causal reasoning posit that causality can be established by actively intervening in the environment and observing the resulting changes (Pearl, 2009; Sloman, 2005). It might initially seem difficult to imagine how such a principle might operate in social contexts but one possibility is that people effectively remove their own contribution from the collaboration by performing badly on purpose to determine what, if any, are the consequences. This might appear self-defeating in typical goal-directed behaviour, and indeed it would be were such an approach employed all the time. But, employed strategically at the right time, it would allow people to disambiguate the level of control they exert over events from the level of control exerted by others. To return to the example of the student with the mediocre feedback, she can ascertain her control by observing how the feedback changes when she is an active member of the research team or when she does not bother to turn up.

In this chapter, I investigated how people attribute ambiguous feedback to their own and another's performance, and their control over the feedback (Spiering et al., 2024). For this, I present a novel task on social credit assignment in combination with behavioural analyses and computational modelling of behaviour. In this task, participants played games together with fictive other players. They had to infer from ambiguous feedback how well they did in the games, how well the other did and how much control they exerted over feedback. First, I tested

how participants assign feedback to themselves and others under varying levels of control. I show that participants adjusted their beliefs about themselves and others by assigning prediction errors in relation to their control. Second, I examined whether participants engage in active exploratory interventions to help them disambiguate their control level from that of others. I found that participants employed certain behavioural patterns during the games which we term ‘active disambiguation’ (AD). AD was driven by uncertainty, and in turn reduced uncertainty and led to better credit assignment. Thirdly, AD mediated inference about control.

2.2 Methods

2.2.1 Participants

36 healthy participants took part in the fMRI experiment (sample taken from a larger study, see below), and 69 participants took part in the online experiment. In subsequent analyses we refer to these as the fMRI and online samples. 2 MRI participants were excluded because they misunderstood how they should rate their own performance in the games (they rated their previous trial's performance rather than their overall best attempt, meaning that they rated their performance lower after an AD trial; one of these two participants reported having changed how they did their self-ratings midway through the experiment). These misunderstandings were prevented in the participants collected afterwards by an additional task-understanding question administered by the experimenter. 3 additional MRI participants were excluded because they performed fewer than 3 AD trials in the Control-Other phase. For the online sample, 3 participants were excluded because they took breaks longer than 5 min during the task blocks. The final samples included 31 MRI participants (22 female, aged 18–33 years), and 66 online participants (37 female, 2 diverse, aged 18–40 years). MRI participants were paid £50, plus a bonus based on task performance (range: £2–4). Online participants received £7.5 per hour plus a performance-dependent bonus (range: £0.5–2). The study was approved by the Medical Sciences Inter-Divisional Research Ethics Committee (MSD-IDREC, R40628/RE005 for MRI, R54722/RE008 for online) at the University of Oxford.

The analysed MRI data set originates from a larger sample with participants on a range of depression symptoms (findings reported in chapter 4). To be included in the MRI study, participants were right-handed, aged 18-45 years, with normal or corrected to normal vision, and fluent in English. Participants were excluded if they had any contraindications for MRI scanning, current alcohol or drug dependency, current or past history of bipolar or psychosis disorder, or current use of psychoactive medication or recreational drugs. Interested

individuals filled in a pre-screening online form hosted on Qualtrics. This online form included questions on demographics, MRI safety and a short depression questionnaire (Quick Inventory of Depressive Symptomatology, QIDS-SR) (Rush et al., 2003). Participants were then invited to a pre-screening video call during which MRI safety was checked again, and during which part of the Structured Clinical Interview for DSM-5 Disorders (SCID-5-CV) was administered (First et al., 2016). Participants were then invited to a single study visit. In this study, only participants with none or mild depressive symptoms (QIDS-SR total score < 11) were included.

The online study was advertised on the online platform Prolific (www.prolific.com). Participants were included into the study if they were aged 18-40 years old, had English as their first language, normal or corrected to normal vision, and at least 10 previous study submissions. Online participants could not move onto the main task if they failed the “system check” for the games. During this system check, participants practised the games and were instructed to do their best. For each of the games, participants had to produce five consecutive trials with sufficiently good game performance (defined as lower than 90% of largest absolute objective error from pilot data) to confirm that the games work correctly on their systems and that they paid attention. For each game, during task practice, participants had a maximum of 30 trials during which to produce those five good trials. Additionally, participants were excluded if after three attempts, they did not pass the multiple-choice test which checked their task understanding. After the first and second false attempt at this comprehension check, they were shown which questions they answered incorrectly and had to reread the instructions.

2.2.2 Experimental procedure

MRI participants took part in a single testing session. After providing written informed consent, they were instructed about the task and performed a task practice outside of the scanner. During the training, participants practised each of the games. They also practised one trial with

ratings, game and example feedback. Afterwards, they filled in a questionnaire assessing their task understanding. They then completed the main task inside the scanner. Afterwards, outside of the scanner, they filled in a task debrief questionnaire and psychiatric questionnaires (not analysed in this chapter, analysed in chapter 4). Online participants also took part in a single session, and followed a similar experimental procedure as MRI participants. The experimental task paradigm was created using java script. JATOS (Lange et al., 2015) was used for experiment hosting and data collection.

2.2.3 Experimental design

2.2.3.1 Overview

In this section, I give a summary over the experimental design while also referring to the later sections that contain more detailed information on the experiment. Participants were asked to assign a single, overall outcome (feedback) to three hidden latent causes: their own performance, another fictive player's performance, and their own control over the feedback (Figure 2.1A). The feedback varied parametrically over trials and was the sum of the participant's own and the other player's performance, weighted by their relative control levels over the outcome. To obtain estimates of each of the three beliefs, participants rated themselves, the other player, and their control at the start of each trial (Figure 2.1B left, see section 2.2.3.4 for more details). Then, they played one round of a simple game (Figure 2.1B middle). Participants were told that another player would play the same game at the same time hidden from the participants' view. The other player was described to participants as computer-generated but based on real people's previous performances in the game. We designed the games so that participants could not estimate their own performance from introspection with complete accuracy (see section 2.2.3.5 for more details) and instead had to rely on the outcome feedback they were shown (Figure 2.1B). To incentivise participants to assign credit as accurately as possible, they received a bonus payment for their rating accuracy.

As the credit assignment problem consisted of attributing a unitary outcome to three hidden latent causes (own/other performance and control), we systematically varied the information that participants were equipped with at different times in the block. This allowed us to precisely capture how they inferred different component aspects of the credit assignment problem. Participants completed four blocks, each with the same structure, totalling 144 trials. Each block was divided into distinct phases in which different latent causes had to be inferred (Figure 2.1C-D, see section 2.2.3.3 for more details): In the first phase, they started by playing a new game together with a new other player ('Self-Other phase', 16 trials, Figure 2.1Ci). Here, their true level of control was shown to them so that they only had to learn about performance levels for themselves and the other player. In the second phase, they played a few trials alone so that the feedback only reflected their own performance ('Self-only trials', four trials). Therefore, during this phase, they could confirm and if appropriate correct their beliefs about their own performance in the game. In the third phase, they were paired with a new other player, and told that their control had been reset and was unknown ('Control-Other phase', 16 trials, Figure 2.1Cii). This means that while knowing how good they themselves were at the game, they had to infer their control and the performance of the new other person. In summary, in the first phase, participants knew their control level but had to infer their own and the other player's performance levels while in the third phase, participants knew their own performance level but had to infer their control level and the other player's performance level. The brief second phase served as a manipulation to ensure the efficacy of the third phase and so all the key analyses were conducted on behavioural and neural data from the first and third task phases. Note that control was a single parameter reflecting the balance between participants' own and the other's influence over outcomes (other control=1-self control). In the following, we will refer to control from the perspective of the participant (e.g. 20% control means the participant and the other control 20% and 80% of the feedback respectively).

We used a pre-determined schedule of the true values of self, other and control that had to be learnt across the task (Figure 2.1D, see section 2.2.3.9 for more details). Each task block corresponded to the participants playing one game type and the games switched between each successive block. On each trial the participants' objective performances in a game were transformed with a logistic function into a self-performance score (Figure 2.3, see section 2.2.3.7 for more details). The latter modification was crucial, because it meant that self-performance was primarily based on the pre-determined schedule, and approximately similar across trials and participants. The only exception was if participants performed so badly that their performance was counted as zero. This becomes crucial in the definition and deployment of active disambiguation (AD) trials (see section 2.2.3.8 for more details).

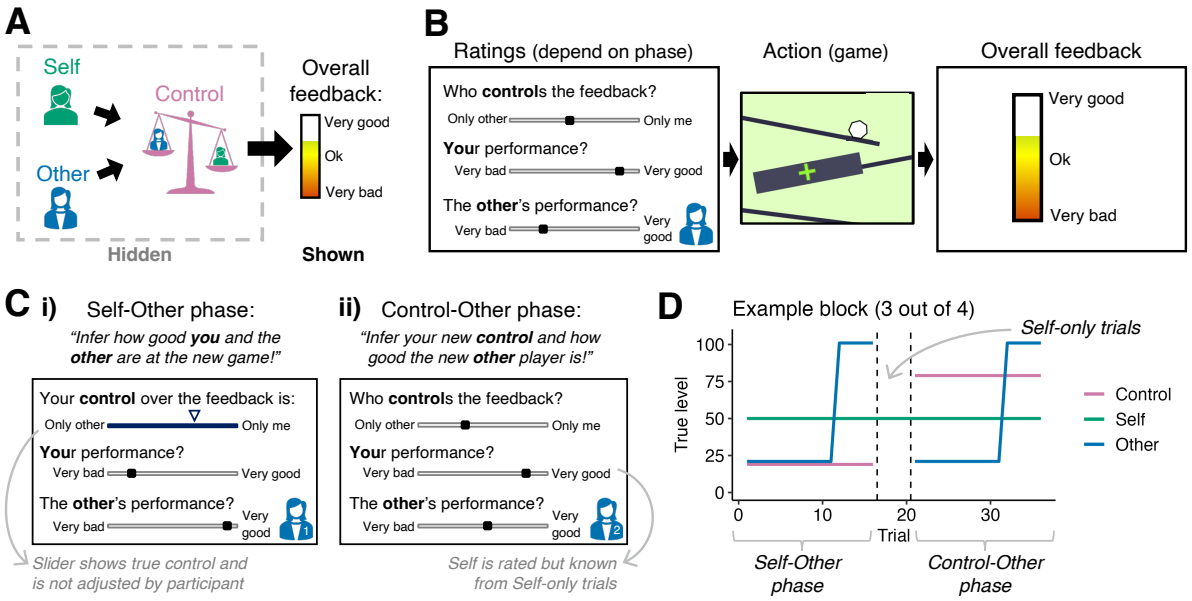


Figure 2.1. Structure of the credit assignment task. A) In our task, participants assigned outcomes (overall feedback) to three causes: their own performance, another fictive player's performance, and their respective control over the feedback. These causes were hidden and latent, and there was only one piece of feedback on each trial. The feedback parametrically varied over trials. **B)** Each trial started with a rating screen on which participants rated their beliefs about their own control, their own performance, and the other player's performance. Then, they played one round of a simple game (action phase): In the example here, a ball rolled down a slide and they had to press a button when they believed it had passed the green target cross. A large black box hid the ball's path which made it impossible to know exactly when the ball passed the cross. This meant that participants had to rely on the feedback

shown to them to infer their own performance, alongside the other player's performance and their own control. At the end of the trial, the overall feedback was shown which was the sum of the participants' own and the other player's performance, weighted by their respective control over the feedback. After participants pressed a button to continue, an ITI followed. **C)** Which ratings participants completed depended on the phase: **i)** In the beginning of each task block, participants entered the first, 'Self-Other phase'. Here, they were informed of the control level but they learned about their own and the other's performance level. This task phase allowed us to determine how participants assign credit to themselves and a new other player while knowing their control. Note that in the task, real faces were shown for the other players on each rating screen (here substituted with icons). **ii)** In the Control-Other phase, participants played with a new other player and their control was reset and unknown. In this phase, participants knew their own true performance level because of the preceding Self-only trials they had played between Self-Other and Control-Other phase. In the Self-only trials, participants played alone, and the feedback only reflected their own performance. This meant that they could double-check if they had accurately inferred their own performance level in the preceding Self-Other phase, and adjust their beliefs accordingly. Subsequently, therefore, in the following Control-Other phase, they inferred the new other player's performance and their new control over the feedback. **D)** Participants completed four blocks, each of which was divided into three phases, signalled to participants (example block is shown here). Each block started with the Self-Other phase, after which followed Self-only trials, ending with the Control-Other phase. Analyses are conducted on behavioural and neural data from the Self-Other and Control-Other phases when, respectively, participants were inferring their own and the other's performance levels from the feedback they received or inferring the other performance level and the control from the feedback.

2.2.3.2 Trial structure

In the task, participants had to infer, from feedback, their own performance level in a game, another player's performance level in the game, and their control over the feedback (Figure 2.1A). To measure participants' current estimates of these three variables, each trial started with a rating phase (reaction time, RT-paced, Figure 2.1B). Depending on the current task phase, participants rated different combinations of their own performance, the other's performance, and their control level on every trial (details below). After completing the ratings, participants played one round of a short and simple game (action phase, RT-paced). For example, in the rockslide game, they pressed a button to indicate when they thought a rolling ball would hit a target cross. After the game, they were shown a single piece of feedback. The feedback consisted of a parametrically varying score reflecting the aggregate of the

participant's own performance, the other's performance and their control over the feedback (outcome phase, RT-paced). Participants pressed a button to continue and move on to the next trial. After the outcome phase, a short interval was included (inter-trial interval, ITI), during which a fixation cross was shown. For the MRI sample, the duration of the ITI was drawn from a Poisson distribution with a range of 2 to 7s, in order to decorrelate variables of interest between trial phases. For the online sample, the ITI had a fixed duration of 500ms.

2.2.3.3 Task phases

Participants completed 144 trials, divided into four blocks. In each task block, participants played one game. Each block was further sub-divided into three distinct phases during which participants played the current game together with different players and different control level settings (Figure 2.1D). The beginning of each phase was explicitly signalled to participants with an instruction screen. The first 16 trials of each block belonged to the Self-Other phase. Here, participants were informed of the true control level and so only had to infer their own and the new other's performance. This allowed us to examine how participants assigned credit to themselves and others under varying levels of control. After completing the Self-Other phase, four trials followed which were the Self-only trials. In the Self-only trials, there was no other player and participants had full control over the feedback. This meant that the feedback directly reflected participant's self-performance on the given trial. Therefore, participants could double-check and correct their self-estimates during those trials. Afterwards, the third task phase followed, the Control-Other phase. During this phase, participants played with a new other player and their control level reset and was unknown. This meant that while they now knew their own performance level (self-performance) for the game, participants now had to infer from the feedback the new other's performance and their control level. Overall, participants were incentivised to provide ratings as accurately as possible, rather than playing the games as well as they could. For each task block, they received a bonus that depended only on how

accurate their ratings were, and not on their actual game performance. Additionally, the cover story of the task was that they were team captains and had to assess other players' performances in order to select teammates and assign control for each game. Therefore, at the end of each task block, they selected one of the two players they had played the current game with, and also, they chose how to distribute the control over the feedback from the two control levels they experienced.

2.2.3.4 Ratings

Which of the three ratings (self, partner, control) were completed on every trial depended on the current task phase (Figure 2.1C). In the Self-Other phase, participants only rated their own and the other's performance. Here, the control rating slider indicated the true control level and could not be moved by participants (see also Figure 2.1Ci). In the Self-only trials, participants only rated their own performance, and the control and other rating slider were not shown. In the Control-Other phase, participants rated their control level, self and other performance (Figure 2.1Cii). On the first trial of each phase, participants provided their ratings according to their 'gut feeling' which gave us a measure of their prior expectations about self, other, and control, before any feedback observations. For both self and other performance, participants were instructed to rate their best past attempt at the game, rather than the last trial (example rating questions: "On your best attempt, how good are you at the rockslide game?", "On their best attempt, how good is the other player at the rockslide game?"). This ensured that participants rated their overall abilities on normal trials as opposed to AD trials. Self and other was rated between 0 and 100, labelled as 'very bad' and 'very good' respectively. Control was rated (or, in the Self-Other phase, shown) on a scale between 0 and 1, ranging from 'only other' to 'only me'. To reduce motor confounds, the position of the rating pointer was initialised to a random position for the MRI sample. MRI participants then used the left and right buttons to move the pointer to the desired position, and pressed the middle button to submit their rating.

Online participants used their mouse cursor to rate and clicked on a ‘Continue’ button to submit their ratings (i.e. there was no initial rating pointer position).

2.2.3.5 Games

Each block corresponded to one game so that participants played four different games throughout the whole study. The exact nature of these games is less important than the fact that they were a vehicle to investigate how participants form beliefs about their own and others’ performances (for a previous example of this strategy, see Wittmann et al. (2016b); Wittmann et al. (2021)). Figure 2.2 shows all four games. For example, in the Rockslide game (Figure 2.1B and Figure 2.2A), a ball rolled down a set of slides and participants pressed a button when they thought that the ball passed a target cross. What made the game difficult to play was that the ball’s path was hidden behind a large black box that was centred around the target cross. This meant the participants had to guess when the ball passed by the target cross. On every trial, the position of the target cross on the middle slide changed slightly. The participant was told that the fictive other player would play the games at the same time but that they would be unable to see them play. This means that participants could not directly and accurately perceive how well they themselves had performed nor could they directly perceive how well the other player had performed in the current game. Therefore, they had to infer their own and the other’s performance from the feedback shown to them after the game phase. While they could not directly perceive their performance level accurately, they were able to perceive large, purposefully made errors that were made on AD trials.

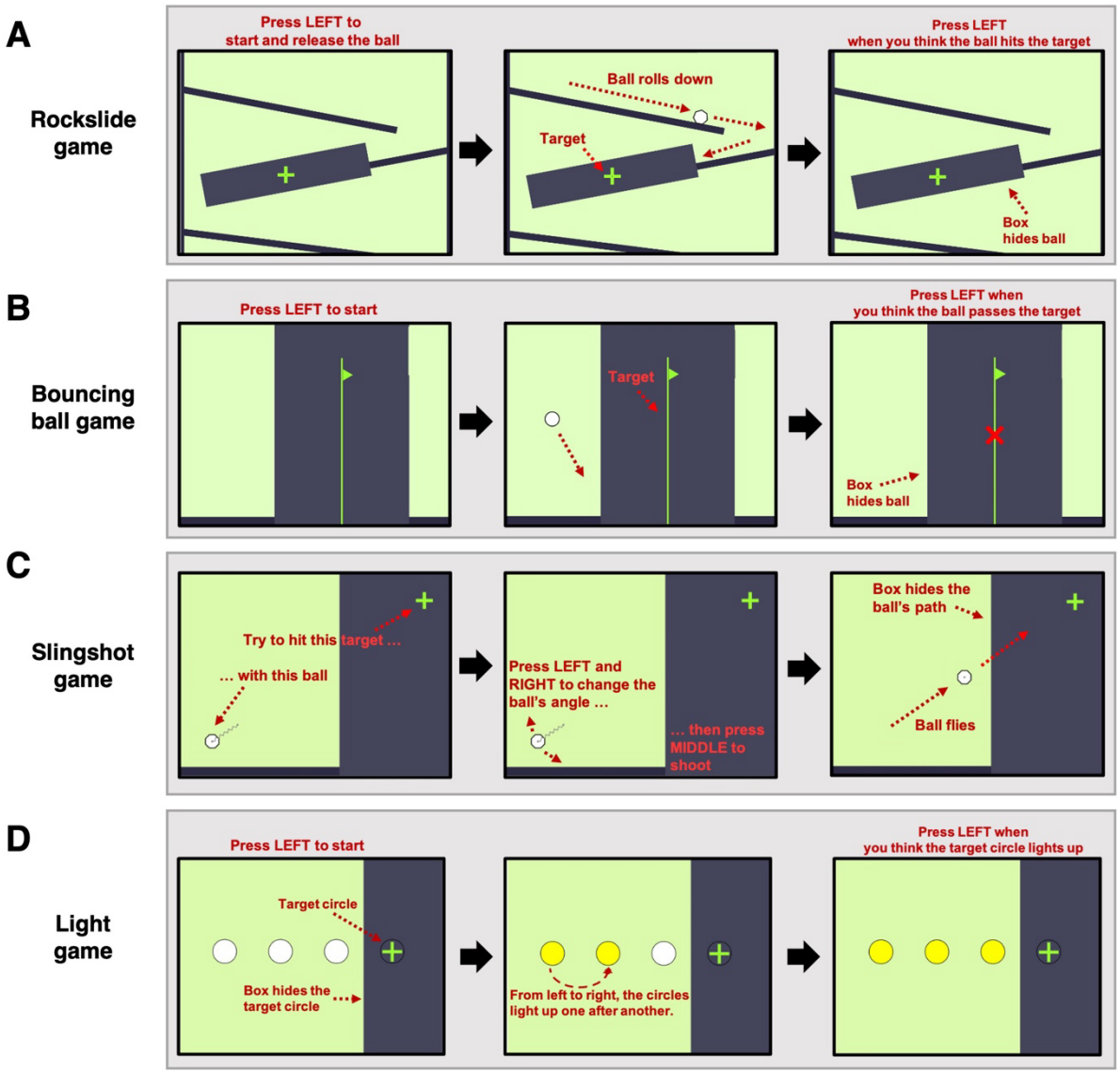


Figure 2.2. Games in the credit assignment task. A) In the Rockslide game, participants had to press the left button to release a ball that then rolled down a set of slides from the top. Participants then pressed the left button again to indicate when they thought that the ball passed the target cross. A black box hid the ball's path so that they could not see when the ball passes by the target exactly. The box was centred around the target cross. From trial to trial, the target cross changed its position slightly on the second slide. **B)** In the Bouncing ball game, participants pressed the left button to release a ball that would appear from the left and bounce towards the right. Participants pressed the left button again when they thought the ball passed the target flag. The black box hid the ball's path so that participants again were not able to tell when exactly the ball passed the target flag. On every trial, the target flag had a different horizontal position and the black box was centred around the target flag. **C)** In the Slingshot game, participants adjusted the angle of a ball (with left and right buttons) to hit a target cross. They pressed their middle button to shoot the ball. Again, a black box hid the ball's path so that participants were not able to see how accurately they shot the ball. On every trial, the target cross changed position vertically. **D)** In the Light game, the participants saw a set of white circles. When participants pressed the left button to start, the white circles lit up one after another from the left to the right. The fourth circle,

the target circle, was hidden behind a black box so that participants did not see when it lit up. They had to press the button when they thought that the target circle would light up. On every trial, the latency between start (when participants pressed the left button the first time) and when the target circle would light up changed.

2.2.3.6 Feedback

After playing the game, participants were shown an overall feedback (outcome phase, Figure 2.1B). This feedback reflected their own performance in the game on this trial, the other's performance on the current trial, weighted by their respective control levels. The following formula was used to compute the parametric feedback value on each trial:

$$feedback(t) = self(t) * control(t) + other(t) * (1 - control(t)) \quad (2.1)$$

where *feedback* is the parametric feedback value on the current trial *t*. *Self* is the self-performance on the current trial, and depends on how well participants played the game just before.

Importantly, note that we used veridical self-performance feedback, but nonetheless precisely controlled participants' self-performance in our experiment. We did this, essentially, by manipulating the mapping between the objective game performance to self-performance score. We used a logistic function to transform participants' objective game performance into a self-performance score that is manipulated according to our pre-determined schedule (see next section for details). *Control* is the control level the participant exerts over the feedback in the current phase, while *(1-control)* is the control level the other player exerts. *Other* is the other's performance level on the current trial, pre-determined according to the experimental schedule (see section on schedules). On Self-only trials, the feedback was also computed

according to this formula, but here *control = 1* so that the feedback(t) = self(t), meaning that the feedback directly reflects the self-performance.

2.2.3.7 Mapping of objective game performance to self-performance scores

We used a logistic function to transform participants' objective performance (absolute distance from target in pixels or msec, depending on game) on a given game round into a point score (*self-performance*) that then contributed to the overall feedback. This mapping of objective performance error to self-performance was precisely calibrated to participants' objective performance in the games. The goals of this mapping were 1) to manipulate the mean self-performance of each participant according to our pre-defined schedule, while 2) ensuring that, even after mapping, self-performance still reflected variations in the participant's true objective performance. This means that if on a given trial participants did worse than in the last trial, the self-performance value was lower (and *vice-versa* for a more accurate objective performance). Details on the exact mapping procedure are given in the following.

On every trial, participants' objective game performance was transformed into a self-performance, according to the following logistic function:

$$Self(t) = 100 * \left(1 - \left(1 + e^{Slope(t) * \left(Shift(t) - Error(t)\right)}\right)^{-1}\right) \quad (2.2)$$

with

$$Slope(t) = \frac{\ln\left((1 - 0.55)^{-1} - 1\right)}{Error_{mean} - Error_{min}} \quad (2.3)$$

and

$$Shift(t) = Error_{mean} + \ln\left(\left(1 - \frac{Self_{level}}{100}\right)^{-1} - 1\right) * Slope(t)^{-1} \quad (2.4)$$

where *Self(t)* is the transformed self-performance (point score between 0 and 100) on a given trial *t*. *Error(t)* is the absolute objective performance error on trial *t*. In the rockslide game, for example, this was the absolute Euclidian distance (in pixels) between the target cross and the ball when participants pressed the button. If the participant pressed the button much too early

or late, when the ball was not on the same slide as the target cross, their objective performance was always transformed to a *Self*=0 (AD trial, see next section). The parameter *Slope(t)* determines the steepness of the mapping curve and reflects how sensitive the self-performance is to the objective errors. It is determined based on the mean ($Error_{mean}$) and smallest error ($Error_{min}$) from the previous five trials, excluding AD trials. On the first five trials of a game block, this was based on the pre-task training trials of the respective game. This staircasing procedure was introduced because some participants' errors drifted over time. Calibrating the slope to the participant's previous error range ensured that the mapped self-performance has a constant variance across trials and is the same across participants. The *Shift(t)* parameter computes what error value is mapped onto the $Self_{level}$. $Self_{level}$ is the self-performance level pre-determined by our schedules (see section of schedules), and is the mean self-performance in the current task block. This is also the value that participants need to infer for their own self performance. The shift parameter depends on participants' mean errors across the previous five trials so that the $Self_{level}$ is given for participants' individual mean error. Calibrating the slope and shift parameters to the participants' recent error history ensured that the mapped self-performance remained constant and centred around the $Self_{level}$. Figure 2.3 illustrates the logistic function as well as the staircasing for an example participant.

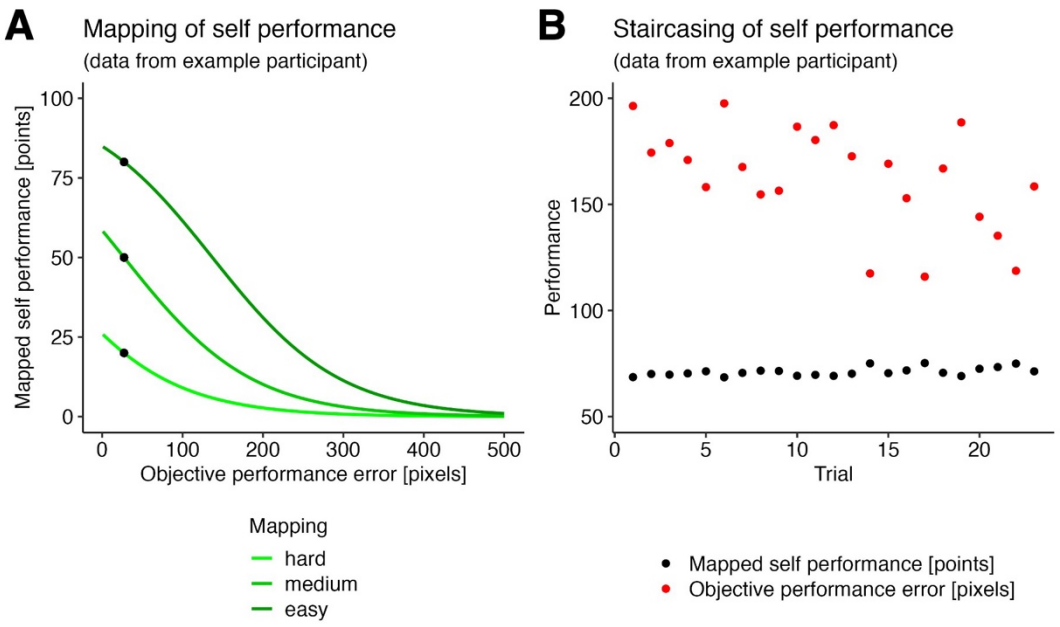


Figure 2.3. Transformation of objective game performance into self-performance. A) Mapping of objective performance error onto the self-performance. We used a logistic function to transform participants' objective performance errors into the self-performance. A lower objective performance error (i.e. higher accuracy) resulted in higher self-performance. Note that under different mapping conditions, the same objective performance error resulted in differently mapped self-performances (black dots). Values of an example participants are shown here. **B)** Staircasing of self-performance. Over time, this example participant drifted in their objective performance (red dots) and got better over time. By using a staircasing procedure, we transformed their drifting performance into a stable self-performance (black dots). This meant that even if participants objectively got better or worse at the games over time, the mapped self-performance remained the same.

2.2.3.8 Active disambiguation

Participants could change their behaviour in the games in order to observe its effect on the feedback. Specifically, they could on purpose play badly in the games. An objective performance error was detected as an AD trial, and therefore yielded a self-performance of 0 on that trial, if it fulfilled specific criteria. The detection criteria for AD were specific to the current game that participants played. Figure 2.4 illustrates the criteria for the rockslide game. In essence, trials were categorised as AD trials if the participant made an obvious and extreme mistake in the game. For example, in the rockslide game, this meant responding while the ball was still visible and not hidden behind the black box that hides the target cross. Panel C in

Figure 2.4 shows data from an example participant, illustrating that the ball position at response is distinctly different between normal and AD trials. For the other games, the criteria for AD trials were similar as for the rockslide game. To ensure that participants generated enough AD trials in the task that we could analyse neural correlates of AD, we took a series of measures in our experimental design. Firstly, participants were not rewarded with a bonus for how well they played the games, but for the accuracy of their ratings. Secondly, in the instructions participants received a hint that they can "try to change how they play the game to see how this affects the feedback". This ensured that participants were aware that they could alter their game performance. Note that we have several indications that AD trials are an intentional behavioural strategy rather than just accidental mistakes. Firstly, the proportion of AD trials strongly depends on the uncertainty of the feedback, that is participants employ AD when it can help them assign credit (Figure 2.7C). Secondly, participants reported using this strategy in the debrief questionnaire (Figure 2.7D).

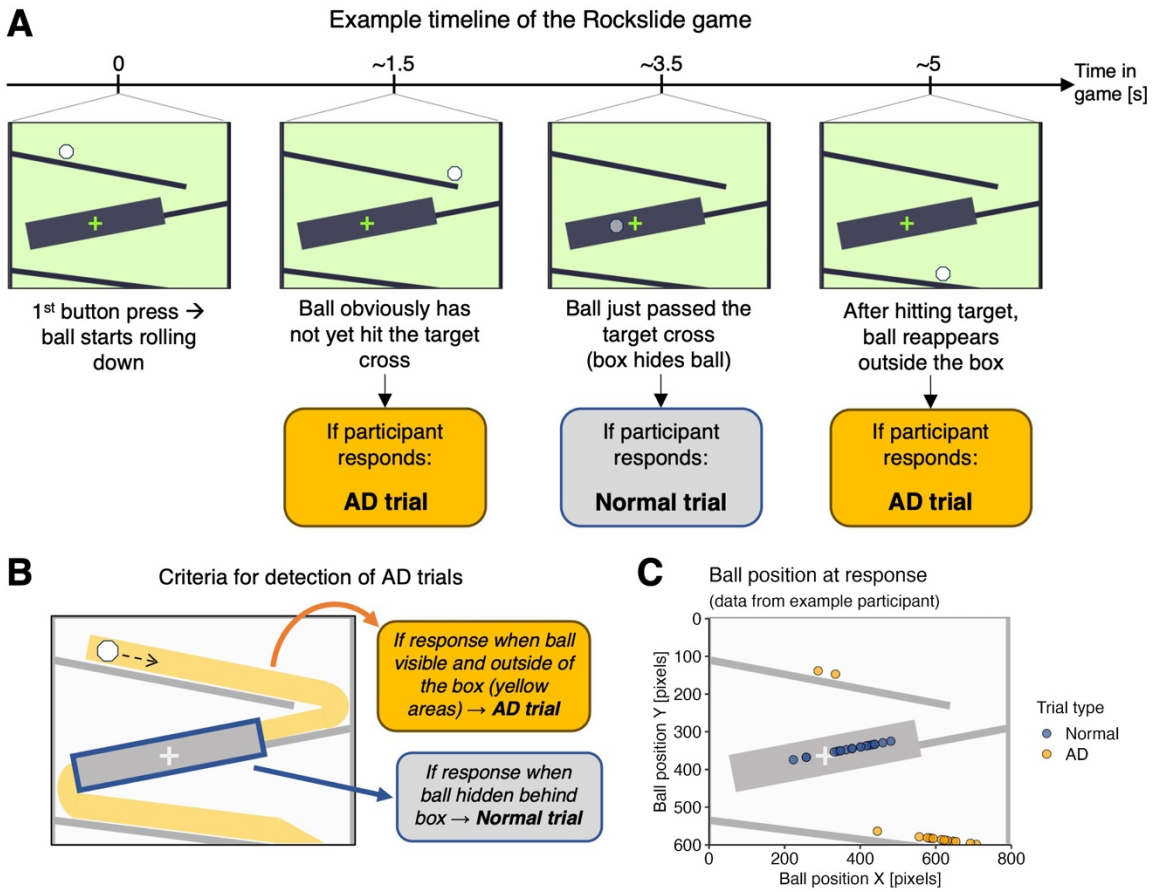


Figure 2.4. Detection of AD trials when participants play the games. A) Example timeline in the Rockslide game. The game starts when participants press the button, which releases a ball that rolls down a set of slides. Depending on when participants respond (by button press), the trial is detected as normal or AD. Trials are detected as AD trials if participants make 'obvious' mistakes in the games. **B)** Specifically, this is the case if participants respond when the ball is still visible and outside of the box that hides the target cross. The trial is detected as a normal trial if participants respond when the ball is close to the target cross and hidden behind the box. **C)** Data from an example participant shows the clear difference between ball positions in AD and normal trials. Data is overlaid with an example screenshot of the rockslide game. Note that the blue dots (ball position on normal trials) vary in their positions because between trials, the target cross changed its position on the 2nd slide and participants could not see the ball behind the black box (and therefore were sometimes too early or too late in pressing the button).

2.2.3.9 Schedule

We used a pre-determined schedule to manipulate the true self-performance level, other performance level and control levels per task block and phase. We used a within-participant design and each participant experienced the same schedule of self, other and control levels.

Figure 2.5A-B shows the schedule of true values participants had to learn in the four task blocks as well as their average ratings. Within each task block, the true self level stayed constant and only changed between blocks. On each trial within each task block, participants played one game. Each task block was divided into distinct task phases signalled to participants (see more information above). In both Self-Other and Control-Other phases, participants played 16 trials together with new other players and new control levels. To decorrelate the variables of interest, on the 12th trial of each phase, the other's performance level changed unannounced. The trial-wise performance of the other player was drawn from a normal distribution with a mean of the true other level and an SD of 6.5. To highlight the differences between the phases, the games had different background colours. The pictures that were shown for the other player were female neutral face stimuli taken from the Chicago face database (Ma et al., 2015). We created three schedule variants for the MRI sample to counterbalance the game order, game colours and face stimuli for the other players. For the online sample, four schedule variants were used.

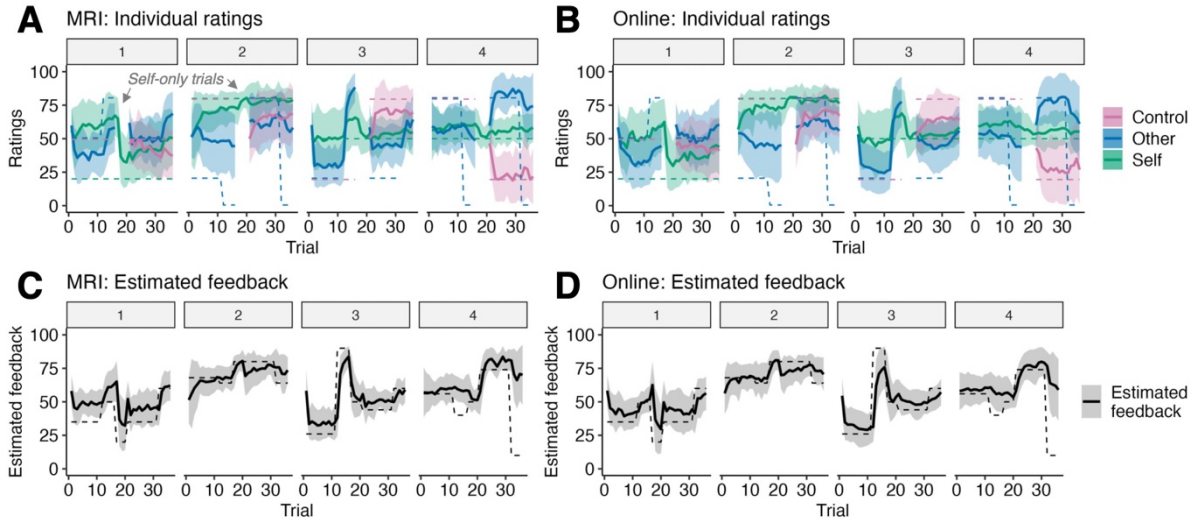


Figure 2.5. Participants' average ratings and estimated feedback in each of the four task blocks. Shaded intervals are SD. Individual ratings of the **A)** MRI sample and **B)** Online sample. During the Self-only trials, when the feedback reflects only the self and participants only rated their own performance, their self-estimates show a sudden adaptation to the true self level. Estimated feedback of the **C)** MRI sample and **D)** Online sample.

2.2.3.10 Post-task debrief questionnaires

After finishing the task, participants filled in a debrief questionnaire. One of the questions they answered was “In the study, how many times per game did you do badly on purpose?”, which was answered on a Likert scale from “never”, “less than once per game”, 1, 2, …, 10, “More than 10 times per game”. 5 acquired MRI participants (one excluded from analysis) did not fill in the debrief questionnaire because it was only added at a later stage. Online participants filled in a more extensive debrief questionnaire. One of the additional questions filled in only by the online sample was “If you did badly on purpose, how did it help you do each rating?”, which for each rating (self, other, control) could be answered on a Likert scale from “1 – Not at all”, “2”, “3”, “4”, “5”, “6 – Very much”.

2.2.4 Bayesian optimal observer models

We compared participants’ behaviour to Bayesian observer models. For this, we fit three different models to the sequence of feedback observed by participants. These models thus provide the beliefs participants could optimally hold at any point in the task (under different assumptions about ADs, see below).

2.2.4.1 Model types

We fit three different learners to participants’ data. Crucially, all three models were optimal Bayesian observer models that were applied to the feedback observed by participants. The models differed, however, in how they treated AD trials. Specifically, the “Active learner” model was given accurate information about when participants performed AD trials and when they performed normal trials. The “Ignorant learner” model was a similar, optimal Bayesian learner but did not know when an AD had occurred and instead assumed that each trial was normal, i.e. feedback on every trial (including the *de facto* AD trials) always reflected participants’ best possible performance on the game. The third model, the “Passive learner” was again an

optimal Bayesian learner, moreover it correctly knew when an AD as opposed to a normal trial had occurred. However, the Passive learner did not learn from the AD trials; it did not use them for updating its beliefs. In other words, the model knew when AD trials occurred and knew that they were different but simply discarded them from the learning process. Full details on how the three models were implemented are given in the next section (Modelling details). By comparing the three models, it was possible to show that learning and credit assignment depended on using the difference between the feedback on the AD and normal trials in the manner of the Active learner.

2.2.4.2 Modelling details

Models were fit using cmdstanr (v. 0.5.3) (Gabry, 2021). They were run with four chains and 10,000 iterations. Convergence was checked using the Rhat<1.1 criterion and the absence of divergent samples (Carpenter et al., 2017). Iterations were increased in the case of divergent samples and until the Rhat criterion was fulfilled.

The models were fit separately to each participant, each game block and each Self-Other and Control-Other phase. Each model comprised five fitted parameters: mean and standard deviation of the self-performance, mean and standard deviation of other performance, and mean of control level. For instance, the estimated mean self-performance corresponded to the average self-performance, while the standard deviation corresponded to the variability of self-performance. On every trial, the models updated these parameters based on the feedback observed by participants. To obtain beliefs for each trial, the models were fitted repeatedly (i.e. first fitted only to data from trial one, then data from trials one and two, then from one, two and three, etc.). While the three models differed in the assumptions they held about the AD and normal trial types (see end of this section for more details), in other respects they were identical; they updated their beliefs in the same manner and all followed Bayes' rule. Before

the very first feedback observation (i.e. phase trial 1), the models' priors were defined as follows:

$$self_{prior} \sim \mathcal{N}(self_{rating\ t1}, 30) \quad (2.5)$$

$$other_{prior} \sim \mathcal{N}(other_{rating\ t1}, 30) \quad (2.6)$$

$$control_{prior} \sim \mathcal{N}(control_{rating\ t1}, 0.3) \quad (2.7)$$

where $self_{prior}$, $other_{prior}$ and $control_{prior}$ are the priors for the mean parameters of self, other and control, respectively. These are modelled as Gaussians, centred around the participant's ratings on the first phase trial ($self_{rating\ t1}$, $other_{rating\ t1}$, $control_{rating\ t1}$) and with a wide standard deviation (30 for self and other, 0.3 for control). The latter led to a wide initial 90% credible interval of 98.70, spanning almost the whole range of possible beliefs if the prior mean is 50. The priors for the standard deviation parameters of self and other performances were defined as follows:

$$\sigma_{self} \sim \mathcal{N}(0,5) \quad (2.8)$$

$$\sigma_{other} \sim \mathcal{N}(0,5) \quad (2.9)$$

where σ_{self} and σ_{other} are the priors for the standard deviation parameters of self and other performance (i.e. the variabilities of these performances).

The models then updated the prior parameters of self, other and control (both means and standard deviations) based on the observed feedback. The way in which this happened depended on the model type (Active, Ignorant or Passive learner). Crucially, only the Active learner model was set up with the correct assumptions about when AD trials happened and

how the feedback is different on AD trials and normal trials. In the Active learner model, the priors were updated depending on whether the current trial was a normal trial or an AD trial:

$$feedback(normal) \sim \mathcal{N}\left(self * control + other * (1 - control), \sqrt{\sigma_{self}^2 * control^2 + \sigma_{other}^2 * (1 - control)^2}\right) \quad (2.10)$$

$$feedback(AD) \sim \mathcal{N}\left(other * (1 - control), \sqrt{\sigma_{other}^2 * (1 - control)^2}\right) \quad (2.11)$$

where on both normal and AD trials, the feedback is assumed to be drawn from a normal distribution. In the Self-Other phase of the task in which participants were given accurate information about the true control level, *control* simply corresponded to this true control level. In the Control-Other phase of the task in which participants had been given accurate information about the true self level, *control* was estimated but the models were given the true *self* level and $\sigma_{self} = 5$. In the last few trials of both Self-Other and Control-Other phases of the task, the other player changed performance in a manner that was unannounced to participants. On those trials, the models estimated the other as a new other player to allow for flexible learning so late in the task phase. All model estimates were constrained between 0 and 100 (1 for control), similar to the possible rating range for participants.

In the Ignorant learner model, all trials (including AD trials) were treated as normal trials. This means that on every trial, the estimates of self, other and control were updated as follows (similar to equation 2.10 in the Active learner model but applied to all trials):

$$feedback(all\ trials, normal\ and\ AD) \sim \mathcal{N}\left(self * control + other * (1 - control), \sqrt{\sigma_{self}^2 * control^2 + \sigma_{other}^2 * (1 - control)^2}\right) \quad (2.12)$$

In the Passive learner model, our goal was to mimic someone who gets confused by AD trials and discards them from learning. Therefore, the Passive learner model was fit to normal trials only and AD trials were removed from the data. Crucially, this meant that the parameters about self, other and control were only updated on the remaining normal trials (and not on AD trials), as follows (same as equation 2.10 in Active learner model):

$$feedback(normal) \sim \mathcal{N}\left(self * control + other * (1 - control), \sqrt{\sigma_{self}^2 * control^2 + \sigma_{other}^2 * (1 - control)^2}\right) \quad (2.13)$$

2.2.4.3 Model parameters

We extracted the mean parameters for self, other and control for all three computational models; these are referred to as estimates throughout the manuscript. Additionally, we used features of the Active learner model to extract a trial-wise measure of uncertainty in the task. For this, we first extracted the 90% credible intervals around the self, other, and control mean estimates on every trial. Trial-wise uncertainty was then defined as the average across these individual uncertainties on each trial, according to the task phase:

$$uncertainty = \begin{cases} \frac{uncertainty_{self} + uncertainty_{other}}{2}, & Self - Other\ phase \\ \frac{100 * uncertainty_{control} + uncertainty_{other}}{2}, & Control - Other\ phase \end{cases} \quad (2.14)$$

where the trial-wise *uncertainty* was the average across self and other uncertainties in the Self-Other phase, and across control and other uncertainties in the Control-Other phase.

2.2.5 Behavioural analyses

We applied a set of Bayesian regression models to the behavioural data, using the brms package (v. 2.17.0, Bürkner, 2017) built on Stan (Carpenter et al., 2017). For all non-

hierarchical regressions, we ran the GLMs per participant and normalised all regressors to a mean of 0 and SD of 1. Four chains were run with 8000 iterations and delta sampling parameter of 0.8. Model fit was checked with Rhat<1.1 and the absence of divergent samples (Carpenter et al., 2017). If these criteria were not fulfilled, the number of iterations and delta were increased. Weak priors were used (normal(0,5) for parameters). We tested the resulting beta estimates for significance using ANOVAs with sample (MRI or online sample) as a between-factor, and evaluated the p-value of the intercept to evaluate the significance of the beta estimates (irrespective of sample). In addition, we report effect size as Cohen's *d* (*d*) for t-tests and Eta-squared (η^2) for ANOVAs.

2.2.5.1 Rating accuracy

To test whether participants learnt from the feedback, we compared the accuracies of the estimated feedback to each of the individual ratings provided by the participants (i.e. the ratings for self, other, and control, Figure 2.6D). For this, we computed participants' feedback estimation on every trial of the Self-Other and Control-Other phase as follows:

$$estimated\ feedback = \\ self\ rating * control\ rating + other\ rating * (1 - control\ rating) \quad (2.15)$$

where the estimated feedback was the sum of participants' beliefs about themselves (*self rating*) and the other player (*other rating*), weighted by their belief about their respective control (*control rating*). In the Self-Other phase, *control rating* was substituted with the true control level shown to participants, as this was participants' control estimate in this task phase.

Then, on every trial we extracted their rating accuracies for self, other and control ratings, and estimated feedback respectively:

$$rating\ accuracy = 1 - \frac{rating\ error}{maximum\ possible\ rating\ error} \quad (2.16)$$

with

$$rating\ error = |rating - true\ level| \quad (2.17)$$

$$maximum\ rating\ error = \max(100 - true\ level, true\ level) \quad (2.18)$$

For example, the accuracy of participants' self ratings was computed based on the difference between the participants' self ratings on a given trial and the participants' true self level on the same trial. For the estimated feedback, this variable was used instead of the *rating*, and the observed feedback, averaged across all non-AD trials of the current phase, was used as the true level.

Next, we tested whether the accuracies of the individual ratings and the estimated feedback improved over time. For this, we extracted the rating accuracies on the first ('Start') and last trial (11, the last trial before the other player changed performance unexpectedly, 'End') in each phase. For self and control, rating accuracies were only extracted from the Self-Other phase or Control-Other phase respectively because self and control were only learnt in one phase. For each participant, the rating accuracies were then averaged across the four task blocks. For each type of accuracy (self, other, control ratings or estimated feedback) as dependent variable, we then ran an ANOVA with *time in phase* ('Start' or 'End') as within, and *sample* (MRI or online) as between-participant factors. Here, significant intercept terms suggest that participants' rating accuracies improved over time.

Finally, we examined whether the accuracy of the estimated feedback improved more over time than the ratings individually. Here, we extracted the difference of the accuracies from start to the end of a phase, and averaged the accuracy difference of the individual ratings per participant. Then, we compared the accuracy change (from start to end) of the estimated

feedback and the average individual ratings using an ANOVA with *sample* as between- and *rating type* (binary, estimated feedback or individual ratings) as within-participant factor. Note that the accuracy of the estimated feedback can be higher than the accuracies of the individual ratings. For example, if participants are presented with very good feedback, they should believe that either they themselves or the other player did very well (or even both). However, there is an infinite number of possible combinations of self, other and control that can all lead to the same feedback. For example, a high feedback could result from one's own performance being high and a high control for oneself, or from the other's performance being high and the other having a high level of control. This means that while participants' individual ratings (or beliefs) might not be accurate, their ratings in combination should still closely match the ambiguous feedback.

2.2.5.2 Credit assignment in the Self-Other phase

To examine how participants assign credit in the Self-Other phase, we examined whether participants split up the signed total prediction error (tPE) into a self and other-assigned PE (sPE, oPE) according to their control (Figure 2.6E). For this, we computed the trial-wise tPE in this way:

$$tPE = feedback - estimated\ feedback \tag{2.19}$$

where the tPE of a current trial is the signed difference between the observed feedback and the estimated feedback from the current trial. Then, we fit these regression models:

$$rating\ update \sim tPE + control + tPE * control + trial\ number \tag{2.20}$$

where *rating update* was the signed difference between the current and previous trials' rating. This was fit separately to the updates of self and other ratings of the participant, as well as the respective estimates from the active Bayesian learning model. *tPE* is the total prediction error observed on the previous trial; for the active learning model it was computed based on its estimates. *Control* is the known control level. For the regression models predicting other rating updates, the standard control estimate was substituted with the control estimate for the other player (1-control). *Trial number* was included as a control regressor indexing time within a given task phase. Only normal trials during the Self-Other phase were included because only on those trials can the tPE can be split up according to the known control level.

2.2.5.3 Active disambiguation as a behavioural strategy

We hypothesized that participants used a specific behavioural strategy (AD) to assign credit more accurately and reduce uncertainty. For this, we first established that participants generated more AD trials in uncertain contexts than certain contexts (Figure 2.7C). We report the proportions of AD trials out of all trials per phase, averaged across participants.

Next, we validated the detection of AD trials in our task against participants' self-reports (Figure 2.7D). Our goal here was to show that our analysis methods detect more AD trials in the data from participants who reported having performed more AD trials. For this, we extracted participants' responses to the debrief question "In the study, how many times per game did you do badly on purpose?". Then, we computed each participant's average number of AD trials across game blocks, and correlated this number ('# AD trials in behaviour') with the debrief responses ('Reported # AD trials') using Spearman's correlation (ρ). In this correlation analysis, we included participants that were otherwise excluded because of the low number of detected or reported AD trials (see information on sample selection). Therefore, in the online sample, an additional 28 participants were included so that the total online sample was n=64 for this

analysis. For the MRI sample, 29 participants were included in this analysis because from the total sample of n=34, five participants did not fill in the debrief questionnaire.

Next, we tested whether AD reduces the trial-wise uncertainty (Figure 2.7E). The following regression model was fit:

$$uncertainty \sim AD_{t-1} + AD_{t-2} + trial\ number + phase \quad (2.21)$$

where *uncertainty* is the current trial's average uncertainty. AD_{t-1} and AD_{t-2} are binary indicator variables for whether there was an AD one or two trials before the current trial. *Trial number* is time in the current phase, and *phase* is a binary variable for whether the current trial is in the Self-Other phase or Control-Other phase.

To test the effect of AD on rating errors (Figure 2.7F), we computed the sum of all rating errors on every trial (i.e. combining errors across self, other and control) and fit a regression model as follows:

$$summed\ rating\ errors \quad (2.22)$$
$$= \begin{cases} self\ rating\ error + other\ rating\ error, Self - Other\ phase \\ self\ rating\ error + other\ rating\ error + control\ rating\ error, Control - Other\ phase \end{cases}$$

$$summed\ rating\ errors \sim AD_{t-1} + AD_{t-2} + trial\ number + phase \quad (2.23)$$

where *summed rating errors* are computed separately for participants' ratings and the Active learner's estimates. The regression model was fit separately to participants' and Active learner's rating errors, with the same regressors as for the regression model predicting uncertainty. The Active learner's summed rating errors were based on the sum of control and

other rating errors during the Control-Other phase. This is because in this task phase, participants had been informed of the veridical self-performance level, and accordingly, the Active learner assumed the true self-performance level in this task phase.

2.2.5.4 Inferring control from active disambiguation

We hypothesized that participants inferred their controllability over the feedback from comparing normal and AD feedback, taking into account their overall self-performance. Analytically, this follows from the equations of how the feedback is computed on normal and AD trials:

$$normal\ feedback = Self * Control + Other * (1 - Control) \quad (2.24)$$

$$AD\ feedback = Other * (1 - Control) \quad (2.25)$$

where the AD feedback is only based on the other's performance, weighted by their control. This is because on an AD trial, participants play the games so badly that on that trial *t*, their objective performance is transformed into Self(t)=0. From these equations, it then follows:

$$Control = \frac{feedback\ difference}{Self} \quad (2.26)$$

with

$$feedback\ difference = normal\ feedback - AD\ feedback \quad (2.27)$$

Intuitively, let's assume that overall, a participant performs at a medium level in a certain game. On a normal trial, when the participant does their best in the game, the overall feedback shown is around a medium or average level. If they engage in AD, that is they deliberately do badly in the game, then the feedback they receive will be very bad. This means that the participant must have substantial control over the feedback to account for the big feedback difference.

However, if they do badly on purpose and the feedback is not changed much (Figure 2.8B), then the participant can infer that they have only very little control over the feedback. In short, given that the participant knows their own performance, they can change their behaviour (or performance) in the game to see how this affects the feedback, and from the difference between the feedback on the two occasions they can infer how much control they have over the feedback. To translate this into a numerical example: The participant knows that their mean Self-performance level on a normal trial is *Self* = 50, and has observed a *normal feedback* = 44 and an *AD feedback* = 4 (all on a scale from 0 to 100). From this, a control of 0.8 can be inferred because *control = (44 – 4) / 50 = 0.8.*

We used regression models to test if participants used this intuition to learn about their control over the feedback (Figure 2.8C-D). We ran these regression models only on switch trials during the Control-Other phase because this is when participants can infer their control if they are comparing feedback on an AD and an adjacent (prior or subsequent) normal trial. Switch trials were defined as trials that are either a normal trial following an AD trial, or vice-versa. Since some participants had a low number of switch trials, hierarchical regression models were used and run across MRI and online samples. Models were run with weak priors (normal(0,5) for parameters), four chains, 6000 iterations and a delta sampling parameter of 0.9. Model convergence was checked as before. All regressors were normalised to a mean of 0 and SD of 1 across participants. The following hierarchical regression model was fit:

$$\begin{aligned} control \sim\ & feedback\ difference + inverse\ self + \\ & feedback\ difference * inverse\ self + \\ & previous\ control + (feedback\ difference + inverse\ self + \\ & feedback\ difference * inverse\ self + previous\ control \mid ID) \end{aligned} \quad (2.28)$$

where *control* is the next trial's control rating. *Feedback difference* is the difference between normal and AD feedback, as observed on the current and previous trial. *Inverse self* is ln((current trial's self rating)$^{-1}$), and *previous control* is the current trial's control rating. The hierarchical regression model was separately fit to participants' control ratings, and to the respective control estimates provided by the three Bayesian models (active, ignorant, and passive).

2.2.5.5 Supplementary analyses: Inferring self and other from active disambiguation

Just as control can be inferred in the Control-Other phase, self-performance levels ("self") can be inferred from those trials during the Self-Other phase (see Figure 2.10A left). Here, self can be inferred from the difference between normal and AD trials, taking into account the known control level. We tested whether participants used this logic (equation 2.29) with a regression model (equation 2.30):

$$Self = \frac{feedback\ difference}{Control} \tag{2.29}$$

$$\begin{aligned} self \sim feedback\ difference + inverse\ control + \\ feedback\ difference * inverse\ control + previous\ self + \\ (feedback\ difference + inverse\ control + \\ feedback\ difference * inverse\ control + previous\ self | ID) \end{aligned} \tag{2.30}$$

where *self* is the next trial's self rating. *Feedback difference* is the difference between normal and AD feedback, as observed on the current and previous trials. *Inverse control* is ln((current trial's control level)$^{-1}$), and *previous self* is the current trial's self rating. The hierarchical regression model was fit separately to participants' ratings, and to the respective control estimates of the active, ignorant and passive Bayesian learning models. Similar to the

equivalent control regressions, this regression was fit to switch trials only as this is when the feedback difference can be observed. We fit this regression to participants' self ratings as well as the self-estimates of the active, passive and ignorant Bayesian learning models. Results of this regression analysis are shown in Figure 2.10B.

From the feedback equations it follows that participants could infer the other from the AD feedback accounting for the other's control. This is because on AD trials, the feedback shown is only based on the other's performance, weighted by the other's control. We tested whether participants used this logic (equation 2.31) using regressions fit to their other ratings (equation 2.32):

$$Other = \frac{AD\ feedback}{Other\ control} \tag{2.31}$$

$$\begin{aligned} other \sim & AD\ feedback + inverse\ other\ control + \\ & AD\ feedback * inverse\ other\ control + previous\ other + \\ & (AD\ feedback + inverse\ other\ control + \\ & AD\ feedback * inverse\ other\ control + previous\ other | ID) \end{aligned} \tag{2.32}$$

where other is the next trial's other rating. *AD feedback* is the feedback observed on the current AD trial. *Inverse other control* is the inverse of the control that the other player exerts, i.e. ln((1-Control)$^{-1}$). This regression was fit only to AD trials as this is when the AD feedback can be observed, and fit separately to the Self-Other and the Control-Other phase. In the Self-Other phase, *inverse other control* is computed with the known control level, while it is computed based on the control ratings in the Control-Other phase. *Previous other* is the current trial's other rating, i.e. the rating just before observing the AD feedback. This regression was fit separately to participants' other ratings, and the other estimates of the active, passive and

ignorant learners. For the learning models, *inverse other control* was based on the control estimates from the respective model.

2.3 Results

2.3.1 Estimates of controllability modulate social prediction errors

A common finding across many social and non-social learning studies is that people change their beliefs based on prediction errors (PEs, differences between actual outcomes and the outcomes that had been expected) (Bang et al., 2022; Lockwood et al., 2018; Pisauro et al., 2022; Suzuki et al., 2012; Wittmann et al., 2016b; Wittmann et al., 2018). We hypothesized that in our task, participants also learn via a "total PE" (tPE) between the observed feedback and their feedback expectation (Figure 2.6A). However, crucially, this tPE should then be used to assign credit to the appropriate latent causes, and to derive PEs relating to either themselves (and concerning their own performance level), the other player (and concerning the other's performance level), or the control they have over the task.

We estimated participants' feedback expectations ('estimated feedback', see Methods) from the ratings that participants provided about their self, other and control beliefs at the start of each trial. We used a Bayesian optimal observer model ("Active learner") to extract trial-by-trial estimates of the optimal beliefs about self-performance, other performance, and control, and the uncertainty around these beliefs. The simulated data from this computational model allowed us 1) to assess whether participants' credit assignment behaviour was optimal, and 2) to derive trial-wise uncertainty estimates for participants' beliefs about self-performance, other-performance, and control.

We indeed found that participants' estimated feedback tracked the feedback shown (Figure 2.6B). In contrast, their ratings of self-performance, other-performance, and control followed the hidden true levels less closely (Figure 2.6C). While the accuracies of ratings and estimated feedback improved over time (ANOVA on rating accuracies with *sample [fMRI or online]* as between- and *Time in phase [start or end]* as within-participant factor; *Time in phase* effect:

estimated feedback, F(1,65)=150.67, p<0.001, η^2=0.54; control, F(1,65)=38.37, p<0.001, η^2=0.25; other, F(1,65)=18.55, p<0.001, η^2=0.16; self, F(1,65)=11.04, p=0.001, η^2=0.06; Figure 2.6D), the accuracies of estimated feedback improved more from the start to the end of a phase than the ratings did (ANOVA on rating accuracy change with *rating type* [estimated feedback or averaged individual ratings] as within- and *sample* with between-participant factor; *rating type* effect: F(1,65)=32.39, p<0.001, η^2=0.08; Figure 2.6D). This suggests that while participants could track the overall feedback, they inferred the latent causes less well. This is because the observed feedback could be attributed to a multitude of possible combinations of latent causes. In other words, different combinations of self, other and control can mathematically give the same overall estimated feedback. That is if one holds a wrong belief about self, there can be a corresponding wrong belief about other that together would make the observed feedback possible.

Next, we tested whether people assigned credit appropriately. A key feature of credit assignment in our task was that in the first, Self-Other phase, participants knew their control level (they were directly instructed) and so tPEs should be assigned to self vs. other in relation to how much control they had over the outcome (see red text in Figure 2.6A). Again, we found this to be true: overall, participants updated their beliefs about self and other more positively for positive tPEs (*tPE* effect on rating updates: self, F(1,65)=419.09, p<0.001, η^2=0.87; other, F(1,65)=310.16, p<0.001, η^2=0.83; Figure 2.6E), meaning they rated the performances higher if the overall feedback was better than expected. However, crucially, this effect was modulated by their level of control; they also assigned more of the tPE to themselves or the other player depending on who had more control (*tPE*Control*: self, F(1,65)=42.56, p<0.001, η^2=0.40; *tPE*(1-Control)*: other, F(1,65)=86.44, p<0.001, η^2=0.57). We found consistent effects in simulated data from the optimal active Bayesian learner (Self: *tPE*, F(1,65)=2203.97, p<0.001, η^2=0.97; *tPE*Control*, F(1,65)=150.66, p<0.001, η^2=0.70; Other: *tPE*, F(1,65)=1872.74, p<0.001,

η^2=0.97; *tPE*(1-Control),* F(1,65)=151.40, p<0.001, η^2=0.70). This suggests that participants, like a Bayesian optimal observer, attribute social PEs in accordance with their controllability.

In the regression results shown in Figure 2.6E, we also find that overall, participants update their beliefs about the other (but not themselves) depending on the control level (main effect *1-Control* on other rating update: F(1,65)=16.02, p<0.001, η^2=0.20; *Control* on self-rating update: F(1,65)=0.88, p=0.35, η^2=0.01). The active learner showed a main effect of control on both self and other (*Control* on self, F(1,65)=69.38, p<0.001, η^2=0.51; *1-Control* on other, F(1,65)=68.33, p<0.001, η^2=0.51). Since we did not have hypotheses about this effect of control and the overall effect size is very small, we will not interpret it further.

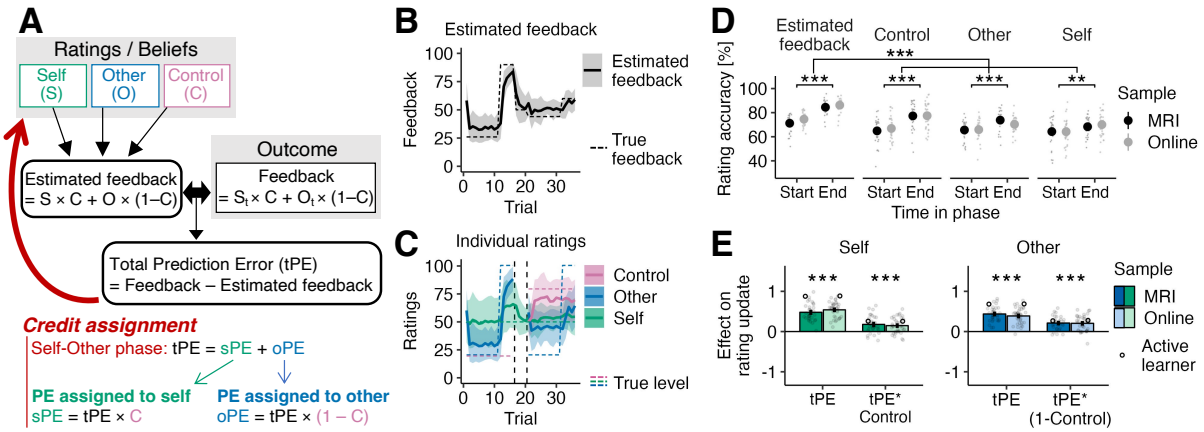


Figure 2.6. Credit assignment behaviour. A) Hypothesized cognitive process and behaviour. Participants' beliefs (measured as their ratings) together form an estimate of the feedback they expect to receive (estimated feedback). At the time of the outcome, the discrepancy between their estimated feedback and the actual feedback results in a total prediction error (tPE). Participants should use the tPE to update their beliefs about themselves (their own performance level), their own control, and the other player's performance level. In the, first, Self-Other phase, when participants know about their own control, they can use this knowledge to split up the tPE into a PE assigned to themselves (sPE) and the other (oPE). **B)** Participants' estimated feedback followed the observed mean feedback throughout the task (example block 3 out of 4, means plotted as lines, standard deviation (SD) plotted as shaded interval). **C)** Participants' mean ratings (lines) and their SD (shaded interval) for the same example block as in B. Figure 2.5 shows the averaged individual ratings and estimated feedback for all task blocks. **D)** Over time, participants improved the accuracy, both of the estimated feedback and the individual component ratings of self-performance, other performance, and control, from observing the feedback.

Overall, the estimated feedback improved more from the start of a phase to the end than did the individual component ratings individually. Accuracy is defined as the inverse of the absolute rating errors, normalised by the highest error possible. **E)** Participants updated their beliefs about self (left) and other (right) more positively after a high positive tPE on the previous trial (main effect of tPE on rating updates). As expected if participants assign credit appropriately, they here also take into account their control level over the feedback, and attribute more of the tPE to themselves (and less to the other) if they themselves have more control (interaction effects tPE*Control on self update, and tPE*(1-Control) on other update). Simulated data from an optimal Bayesian observer model ("Active learner") show the same result pattern as participants (open black dots). We used ANOVAs (with sample, fMRI versus online, as between-participant factor) and evaluated the p-value of the intercept term to assess the significance of participants' beta estimates (i.e. standardized regression coefficients) irrespective of sample. n=31 MRI, n=36 online; *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$.

2.3.2 Active disambiguation as a behavioural strategy

In the task, the feedback received on each trial was ambiguous because it could be attributed to three potential causes: one's own performance, the other player's performance, and one's own control. This made the credit assignment task inherently difficult to solve. However, we hypothesized that participants employed a behavioural strategy to help with this, which we call 'active disambiguation' (AD). We hypothesized that AD is a form of purposeful intervention (Pearl, 2009; Sloman, 2005) whereby the person aims to remove their own contribution from the outcome. In our task, this is possible via something that appears drastic and self-sabotaging in the context of standard goal-directed behaviour: to play the game (very) badly on purpose (e.g. pressing the button too late, Figure 2.7A). However, this allowed them to observe how much is due to internal causes (themselves) and how much is due to external causes (others), thereby simplifying the credit assignment task. If this reasoning is correct, then we should observe some critical features of AD trials. We hypothesized that participants make AD trials when the feedback is uncertain (when it is unclear how it reflected one or other player's performance levels, or both performance levels, and/or the control level), and that this helps them reduce uncertainty and their rating errors (Figure 2.7B).

As hypothesised, we found evidence of AD; participants sometimes made extreme errors in the games so that their self-performance on a trial was zero. Participants' behaviour exhibited high levels of AD under high uncertainty when feedback was caused by multiple sources, and AD was therefore informative (Self-Other phase: MRI, mean (*M)*=25.45%; online, *M*=20.16%; Control-Other phase: MRI, *M*=29.33%; online, *M*=25.40%; yellow bars in Figure 2.7C). Whenever the feedback was not ambiguous and only reflected one's own performance, the proportion of AD showed a striking drop (Self-only trials: MRI, *M*=3.63%; online, *M*=2.42%; blue bars in Figure 2.7C). We found that participants who objectively generated more AD trials also reported having done so after the experiment (MRI, Spearman's ρ=0.81, p<0.001, two-sided, 95% CI=(0.63,0.91); online, Spearman's ρ=0.85, p<0.001, two-sided, 95% CI=(0.76,0.90); Figure 2.7D), further suggesting that participants used AD trials in a deliberate way and that they were not unintentional slips.

Next, we tested whether in the trials immediately following AD, AD had reduced uncertainty and led to better credit assignment (measured as lower rating errors). Since participants did not self-report uncertainty in our task, we extracted from the Bayesian model ("Active learner", introduced before) a trial-wise measure of the overall uncertainty across self, other, and control. Previous work showed that uncertainty estimates derived from such a Bayesian model are both good estimates of people's uncertainty (confidence) reports and they drive exploration (Trudel et al., 2021). We found that uncertainty was indeed reduced in the trials directly after AD (Effect on uncertainty: *1 trial after AD*, F(1,65)=333.97, p<0.001, η^2=0.84; *2 trials after AD*, F(1,65)=385.22, p<0.001, η^2=0.86; Figure 2.7E), and AD led to a reduction in the model's rating errors (Effect on errors of Active learner: *1 trial after AD*, F(1,65)=152.08, p<0.001, η^2=0.70; *2 trials after AD*, F(1,65)=123.24, p<0.001, η^2=0.65; black open dots in Figure 2.7F). This shows that a Bayesian optimal learner benefits from AD trials. Crucially, participants' actual rating errors were similarly also reduced after AD (Effect on participants' rating errors: *1 trial after*

AD, F(1,65)=7.26, p=0.009, η^2=0.10; *2 trials after AD*, F(1,65)=12.56, p<0.001, η^2=0.16; Figure 2.7F).

Taken together, participants used AD when there was uncertainty to be resolved. After AD, both their uncertainty and their rating errors were decreased. This indicates that participants had a good intuition about when AD was useful for them, and that they used the additional information provided by AD in order to change their beliefs.

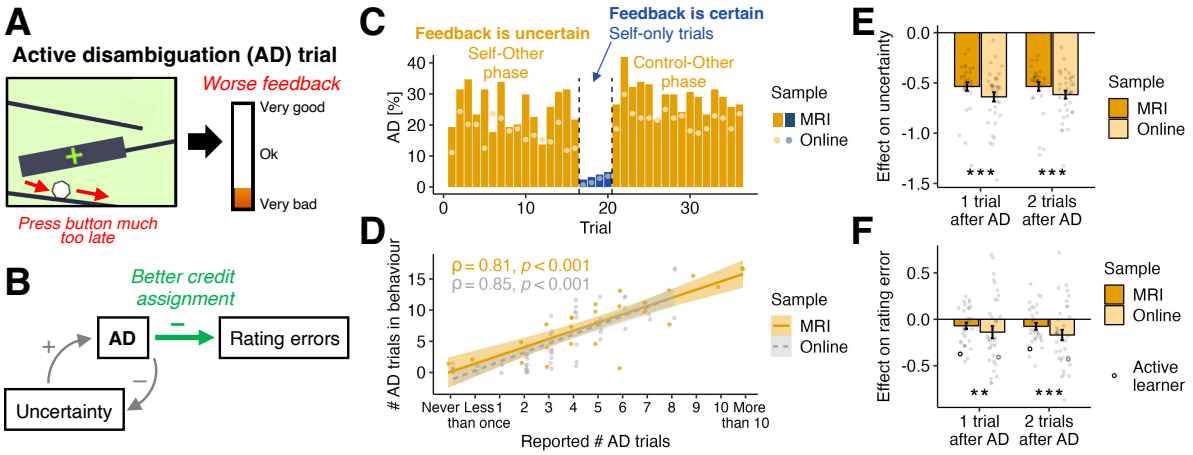


Figure 2.7. Active disambiguation (AD). A) We hypothesized that participants made deliberate mistakes in the games to assign credit better. In the game shown here, they could press the button too late to see how much of a difference this makes to the feedback. We refer to such a trial as an AD trial. During the task, AD trials were detected whenever participants made extreme errors during the games that were then transformed into a zero self-performance (see Methods for details). **B)** We hypothesized that participants generated more AD trials under high uncertainty, and that AD trials helped them reduce uncertainty and assign credit better (lower rating errors). **C)** We indeed found that participants generated AD in phases when the feedback was ambiguous (Self-Other and Control-Other phases). Whenever the feedback was certain (Self-only trials), AD dropped strikingly. **D)** After finishing the task, participants reported how many AD trials they generated per game. We found that their subjective reports of AD trials were significantly correlated with the objective frequency of AD trials identified in their behaviour. **E)** Trial-wise uncertainty, extracted from the active Bayesian learning model, is reduced if the previous trial or the one before was an AD trial. **F)** Crucially, participants' rating errors are reduced after AD, which shows that they can use the additional information afforded by AD (similar to the simulated data from the Active learner model, black open dots). In panels E and F, we tested the significance of the beta estimates (irrespective of sample) by evaluating the intercept terms from ANOVAs with sample (MRI or online) as between-participant factor. Note that in both E and F we did not find a significant

group difference between the samples (all p>0.05). Panel D includes n=29 MRI and n=64 online participants with otherwise excluded participants (that e.g. never did AD, see Methods for details); Panels C, E, F: n=31 MRI, n=36 online; n.s., not significant; **, p<0.01; ***, p<0.001.

2.3.3 Active disambiguation enables inference of controllability

We have shown that participants purposefully use ADs when they are informative, reduce uncertainty and enable better credit assignment. Next, we analysed the mechanism by which AD trials made this possible. We reasoned that AD trials are critical for inferring controllability in the following way. Crucially, they might compare feedback from normal and AD trials (Figure 2.8A). From this difference, integrated across trials, they could infer their control as long as they also took into account their own overall performance level (which was known in this task phase). Intuitively, if making a deliberate mistake (AD) makes hardly any difference to the feedback compared to when doing one's best (normal trial), then a low control can be inferred (Figure 2.8B). Conversely, if the feedback difference is high, then a higher control level can be inferred. While this is true even when high self-performance is assumed, it is especially the case when low self-performance is assumed; if low self-performance is assumed and there is a high feedback difference then a very high control level should be inferred. In summary, control estimates should be proportional to the feedback difference (normal – AD trials) divided by the self-performance level. This can also be rendered as the feedback difference multiplied by the inverse of the self-performance level (feedback difference * inverse self). We tested whether participants used this estimate in the Control-Other phase when control was unknown and had to be inferred. We found that participants indeed estimated their control by combining the information from normal and AD trials, taking into account their own performance (interaction *Feedback difference* * *Inverse Self*, β=0.14, 95% CI=(0.09,0.20); Figure 2.8C).

To test whether the behaviour observed is indeed the behaviour expected from an optimal learner that does perform AD trials and that knows that AD trials were performed (rather than

assumes that a trial reflects its best performance), we compared participants' behaviour to the optimal Active learner (introduced before) and two non-optimal learners we refer to as the "Ignorant" and the "Passive" learners. All learning models were given the same data from participants. However, the second and third models made wrong assumptions about the feedback: The Ignorant learner assumed that AD trials are not AD but normal trials. This mimics someone who does not realize that these trials are intentionally performed to achieve AD, and instead assumed the outcome reflects the best performance the agent was capable of, as was the case on all other trials. The Passive learner does not learn from AD trials at all, but only from normal trials. This is like someone who gets confused by the AD trials and discards that information from learning. Figure 2.9 shows the parameter/trial-to-trial estimates from the models. Participants' behaviour was consistent with the Active learner model, but importantly, different from the two alternative learners (interaction *Feedback difference* * *Inverse Self*: Active learner, β=0.22, 95% CI=(0.19,0.24); Passive learner, β=-0.01, 95% CI=(-0.04,0.03); Ignorant learner, β=-0.28, 95% CI=(-0.32,-0.24); Figure 2.8C). Both the Active learner and participants inferred high control when the feedback difference was high, and even more so if the self-performance was low, because then a large feedback difference must be explained by higher control (Figure 2.8D, top panel). In contrast, if the feedback difference is low, the Active learner and the participants inferred a low control level regardless of the level of self-performance. The ignorant and passive learner show a very different result pattern for the interaction effect (Figure 2.8D, bottom panel), suggesting that AD, and insight into when AD occurs, are essential to infer controllability.

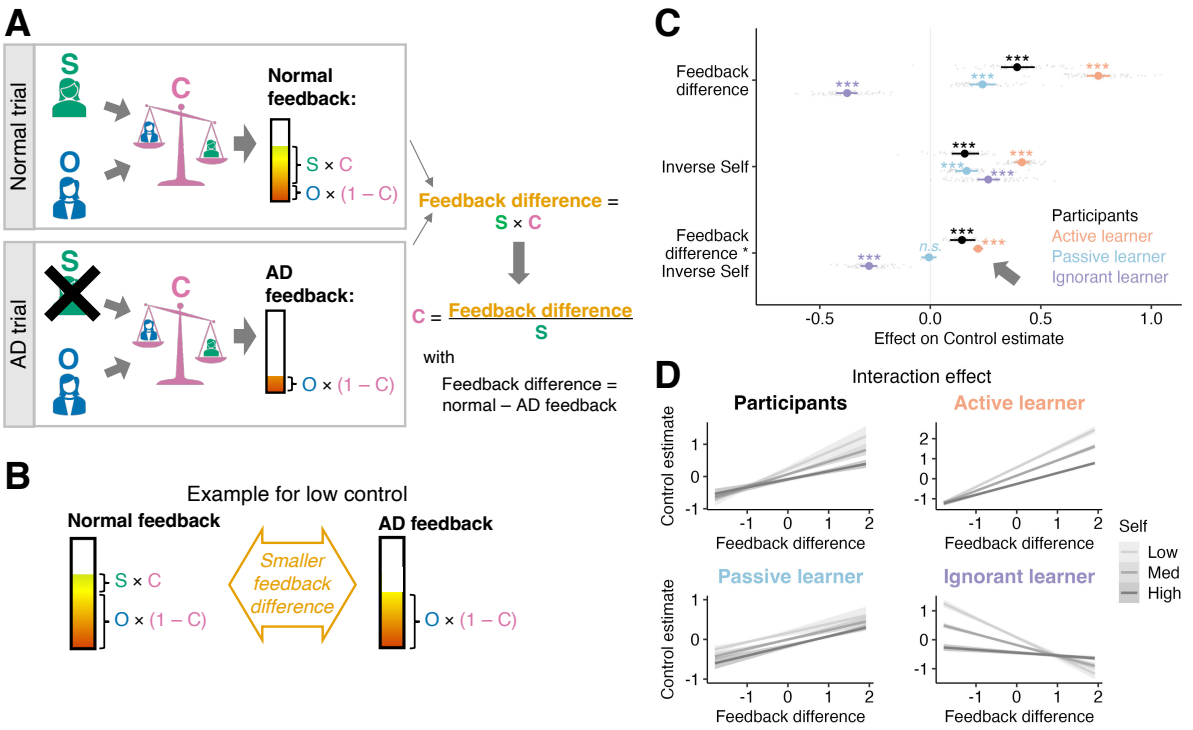


Figure 2.8. Inferring control. **A)** Optimally, one can infer one's control level from observing the difference between trials when doing one's best ('normal trials') vs. making a deliberate mistake ('AD trial'). On a normal trial, the observed feedback is: normal feedback = Self × Control + Other × (1-Control). On an AD trial, the participant plays so badly that Self = 0, which results in the AD feedback being: AD feedback = Other × (1-Control). This means that on an AD trial, the feedback only reflects the other player's performance, weighted by their control. Therefore, the difference between feedback on a normal versus an AD trial corresponds to the participant's own performance weighted by their control (Self × Control). This means that control can be inferred by observing the difference between a normal and an AD trial, taking into account one's own performance (Control = Feedback difference / Self). **B)** Intuitively, low control can be inferred if there is barely any difference between feedback on a normal versus an AD trial. **C)** To test whether participants used the strategy depicted in A, we analysed their control ratings with a hierarchical regression (MRI and online samples were pooled here). We examined whether, after 'switch' trials (either AD trials following normal trials, or *vice-versa*), the next trials' control ratings could be predicted by the observed feedback difference (normal feedback – AD feedback) and the inverse self-rating, and their interaction. As expected, participants showed effects similar to the optimal Active learner, meaning that they inferred their control from the feedback difference between AD and normal trials, taking into account their own overall game performance. As control analyses, we ran the same analyses on simulated data from two computational models that made wrong assumptions about AD trials. The passive learner ignored AD trials during learning, while the ignorant learner assumed that AD trials were in fact normal trials. Importantly, all agents (participants, as well as the active, passive and ignorant learners) received the same observations, which corresponded to the feedback participants actually received. We found that indeed the passive and ignorant learner showed different learning patterns: the interaction effect, *Feedback difference * Inverse Self*, differentiated

between, on the one hand, participants and the Active learner, and, on the other hand, the Passive and Ignorant learners. While participants and the Active learner had a significant positive interaction effect, the Passive learner had no significant effect, and the Ignorant learner had a significantly negative effect. **D)** Interaction effect *Feedback difference* * *Inverse Self*. Note that for easier interpretation, here, we relabelled the legend of 'inverse self' regressor to 'self' and flipped its levels accordingly (e.g. 'low self' is the effect of 'high inverse self'). Both participants and the Active learner infer a low control when there is a lower difference between normal and AD feedback. If there is a large difference, they infer a high control – but this depends on the level of their own performance. If low self-performance is assumed and there is still a high feedback difference, then a higher control is inferred to account for a high feedback difference. If, however, high self-performance is assumed and the feedback difference is high, then a slightly lower control can be inferred. The ignorant and passive learner show markedly different behaviour in this interaction effect. Overall, this means that, like the Active learner, participants were able to take advantage of the additional information AD trials offer in order to infer controllability. n=31 MRI, n=36 online; n.s., 95% CI includes 0; ***, 99.9% CI excludes 0.

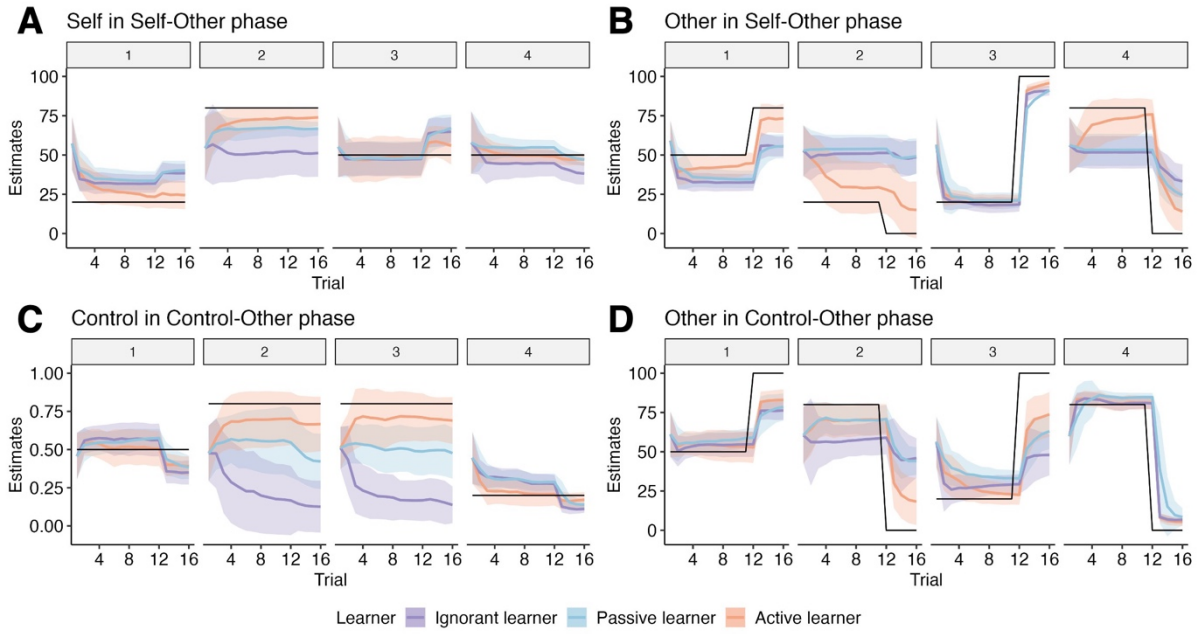


Figure 2.9. Average estimates per game block for the active, ignorant and passive Bayesian learner. Simulated data from MRI and online sample are pooled here. Lines are averaged estimates across participants, and shaded intervals are the standard deviation. Black lines indicate the true levels of each variable. **A)** Self estimates per game block in the Self-Other phase. **B)** Other estimates in the Self-Other phase. **C)** Control estimates in the Control-Other phase. **D)** Other estimates in the Control-Other phase. The active learner has lower rating errors than the suboptimal passive and ignorant learners (Main effect of learning model, $F(2,130)=447.24$, $p<0.001$, $\eta^2=0.83$; active vs. passive learner, $F(1,65)=461.54$, $p<0.001$, $\eta^2=0.72$; active vs. ignorant learner, $F(1,65)=485.83$, $p<0.001$, $\eta^2=0.86$; passive vs. ignorant learner, $F(1,65)=275.22$, $p<0.001$, $\eta^2=0.60$). Additionally, the active learner also has

the lowest uncertainty at the end of the phases compared to the other suboptimal learning models (Main effect of learning model, F(2,130)=325.07, p<0.001, η^2=0.79; active vs. passive learner, F(1,65)=350.98, p<0.001, η^2=0.79; active vs. ignorant learner, F(1,65)=351.19, p<0.001, η^2=0.72; passive vs. ignorant learner, F(1,65)=94.04, p<0.001, η^2=0.38). Crucially, these results suggest that an optimal learner, given the knowledge of AD trials and the ability to learn from it, performs better (in terms of errors and uncertainty) than the two suboptimal learners, which either discard AD trials for learning (passive learner) or assume that AD trials are normal trials (ignorant learner). Summed rating errors and average uncertainty were extracted from the last phase trial of each phase in each block, and averaged across the individual ratings.

2.3.4 Supplementary results

In Figure 2.10 we consider other ways in which participants might have used equivalent strategies to infer self and other, and show that participants did not use these strategies to the same degree to infer control. After finishing the task, participants also reported that they found AD trials most helpful for inferring their control (ANOVA main effect of *rating type [self, other, or control]*: F(2,70)=13.57, p<0.001, η^2=0.17; paired two-sided t-tests: self vs. other, t(35)=-3.29, p=0.002, d=0.55; self vs. control, t(35)=-4.50, p<0.001, d=0.75; other vs. control, t(35)=-2.22, p=0.03, d=0.37; Figure 2.10F). This suggests that participants used AD efficiently for learning about their control and that they had insight into this.

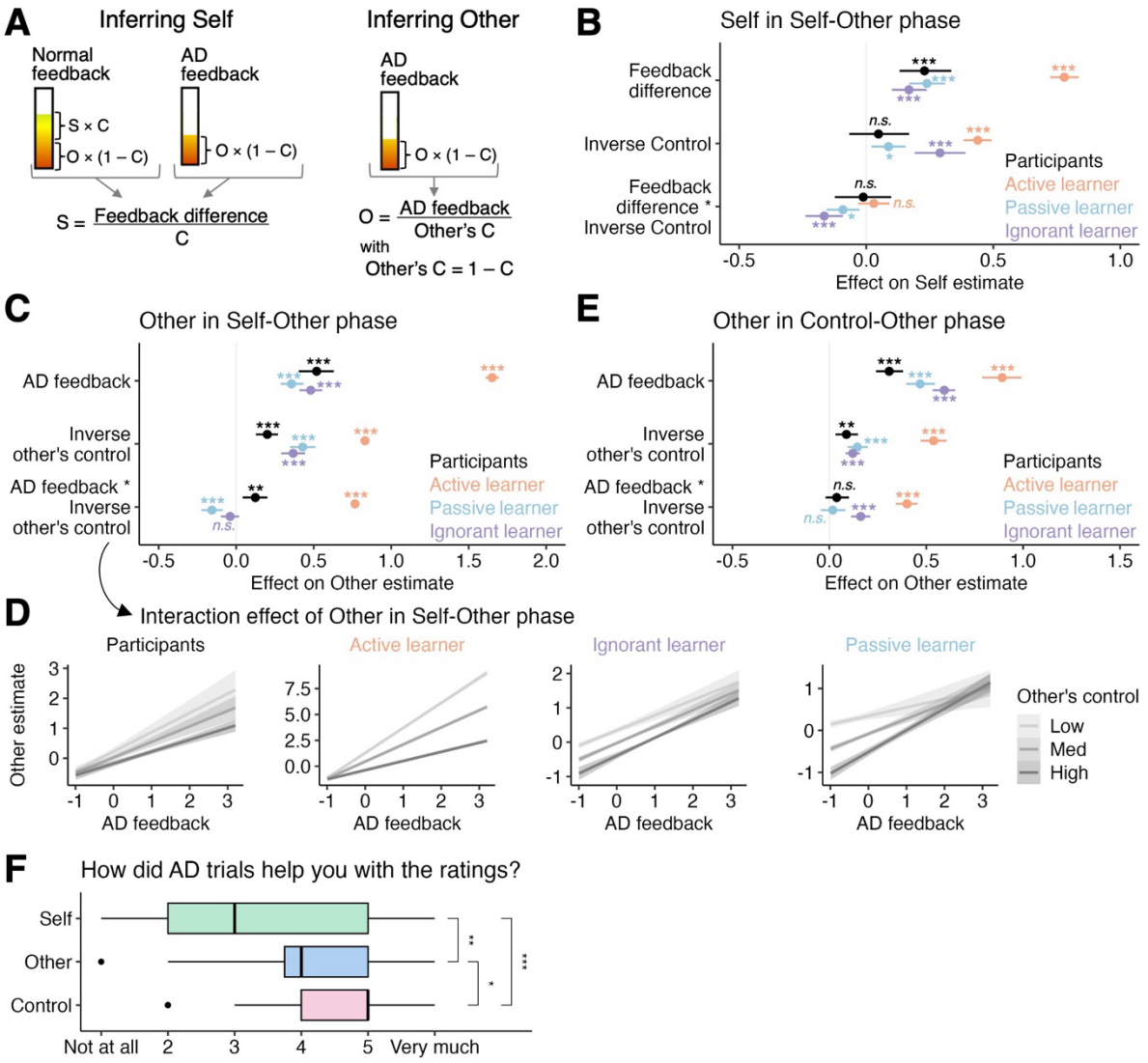


Figure 2.10. **Hierarchical regressions predicting how to assign credit from switch and AD trials to self and others**. **A)** Optimally, in the Self-Other phase, self can be inferred from the feedback difference between normal and AD trials accounting for one's control. Note that this strategy is symmetrical to how control is inferred in the Control-Other phase (Figure 2.8). In both Self-Other and Control-Other phases, the other's performance can be inferred by just observing the AD feedback and accounting for the other's control (1-control). Note that in the Self-Other phase, the control is known while in the Control-Other phase, this requires the control which needs to be inferred as well. **B)** For inferring self-performance in the Self-Other phase, we found that participant performance neither resembled the active Bayesian learner (main effect of *inverse control*), nor the passive or ignorant Bayesian learning models (*inverse control * feedback difference* interaction effect). This means that participants were not able to optimally extract the information afforded by AD and normal trials to infer their own performance. This is also reflected in their helpfulness ratings in the debrief questionnaire (panel F), where they reported AD trials as the least helpful for the self ratings. **C)** In contrast, participants inferred the other's performance similarly to the active Bayesian learner in the Self-Other phase. Here, the interaction effect *AD feedback * Inverse other's control* differentiated between, on one hand,

participants and the active learner, and on the other hand, the two suboptimal learning models. Note however that participants' beta weights are not as close to the active learner as they are for the respective control inference (Figure 2.8D). Participants also rated AD trials as more helpful for learning about their control than learning about the other (panel F). For these reasons, we focussed our analyses on the control inference process and did not follow up this effect in inferring the other. **D)** The interaction effect from C shows that both participants and the active Bayesian learner inferred that the other had a low performance if the AD feedback is low, irrespective of the other's control. If the AD feedback is high, however, they inferred a higher other performance level and even more so if the other had a low control because then the other's performance must account for the high AD feedback. The passive and ignorant Bayesian learners show markedly different behaviour in this interaction effect. Note that for easier interpretation, here, we relabelled the legend of "inverse other's control" regressor to "Other's control" and flipped its levels accordingly (e.g. "Low other's control" is the effect of "high inverse other's control"). **E)** In the Control-Other phase, we found that participants were not able to infer the other's performance like an optimal active Bayesian learner (see interaction effect). **F)** In the debrief questionnaire, participants reported that they found the AD trials most helpful for inferring their controllability. Only data from the online sample is shown here because the MRI sample did not receive this debrief question. n=31 MRI, n=36 online; panels B-E: n.s., 95% CI includes 0; *, 95% CI excludes 0; **, 99% CI excludes 0; ***, 99.9% CI excludes 0; panel F: *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$.

2.4 Discussion

A long-standing issue of interest in psychology and cognitive neuroscience, and one that has influenced the understanding of psychological illness, has been how humans and other animals learn which aspects of their environment are controllable (Maier & Seligman, 2016). Recently there has been renewed interest in this topic inspired by new computational approaches to learning (Dorfman et al., 2019; Ligneul, 2021; Ligneul et al., 2022; Moscarello & Hartley, 2017). Here we examined what might be considered an instance of this problem that is prominent for a social species such as humans: how we learn the extent to which we ourselves, as opposed to others, have responsibility or agency over events. Estimating control in this way goes hand-in-hand with estimating performance levels – how good at a task are we ourselves and how good is the other person we are working with.

We investigated how in collaborative two-person games, individuals assign credit to either themselves or others, while also estimating their control over the outcomes. Participants appropriately assigned credit by attributing prediction errors to either themselves or the other person after taking into account their control. The credit assignment problem in the games was inherently ambiguous and difficult to solve. To overcome this, participants used a behavioural strategy we refer to as ‘active disambiguation’ (AD). They employed this tactic especially when there was high uncertainty about who was responsible for what. AD enabled participants to more accurately estimate their control by comparing the outcomes observed during AD with those on non-AD trials. This comparison leads directly to an estimate of the participant’s control. Finally, identifying control from AD is intertwined with estimating one’s own performance level.

When participants needed to learn about their own ability on the task (self-performance) and the other person’s ability (other performance), feedback on any trial could be attributed to

either their own or the other players performance because it was inherently ambiguous. Participants, however, took this into account, splitting the information contained within each outcome (the prediction error furnished by the outcome) to either themselves or the other person according to their belief about how much control they or the other person had over the outcome (Figure 2.6). This pattern was rational, and resembled one exhibited by an optimal Bayesian learner. In a very general sense, using the feedback provided by outcomes in proportion to the degree to which it is informative is an important principle of learning (Akaishi et al., 2016; Behrens et al., 2007; Leong et al., 2017; Pearce & Hall, 1980). For example, in the context of credit assignment, it has been proposed that learning is modulated by attention (Akaishi et al., 2016; Leong et al., 2017), there is greater attribution of outcomes to stimuli that were attended to, which in turn leads to more attention to the stimuli because of their greater task relevance. Unequal attributions of consequences to causes have also been reported in other settings (Dorfman et al., 2019).

A key learning mechanism exhibited by participants was one of active information seeking. Specifically, people inferred their control by observing how outcomes change when they varied their behaviour, to the extent of even performing badly on purpose during AD. AD resembled other forms of exploratory behaviours that are guided not simply by the immediate prospect of reward, but also by the potential for information acquisition (Bromberg-Martin & Monosov, 2020; Hunt et al., 2021; Kaanders et al., 2021; Monosov, 2024; Monosov & Rushworth, 2022; Trudel et al., 2021; Wilson et al., 2014). Importantly, in our task, participants do not choose among options that simply differ in information content, but instead they actively participate in creating the information they seek by performing exploratory actions that test specific hypotheses about levels of agency. There is increasing interest in how information seeking is tied to active hypothesis testing (Akaishi et al., 2016; Chambon et al., 2020; Rouault et al., 2022; Song et al., 2022; Weiss et al., 2021). A similarly active process may be at work when inferences

about an environment's level of controllability are made (Ligneul, 2021; Ligneul et al., 2022). Crucially, in our task, the key question does not concern establishing the extent to which outcomes are contingent on our actions beyond existing contingencies between outcomes and states of the environment. Instead, the focus is on comparison of the degree of action-dependent contingency exhibited by each actor – the participant themselves and the other person with whom they collaborate in the task.

More generally, AD-guided learning can be considered a form of intervention-based causal learning (Lagnado & Sloman, 2004; Steyvers et al., 2003). Sometimes the contingencies between an agent and the environment can only be ascertained by actively intervening in their environment and observing the consequences. Learning in development may also depend on interventions (Gureckis & Markant, 2012). From childhood to adulthood, people get better at estimating controllability and change how they explore to assess control (Raab et al., 2022). AD is also, arguably, the target of some cognitive therapies for depression ("behavioural experiments") (Bennett-Levy, 2003; Bennett-Levy et al., 2004; Skilbeck et al., 2020). The current results suggest that a similar process of inference and information seeking is inherent in human collaboration.

CHAPTER 3 Supramarginal Gyrus mediates controllability learning and tracks social prediction errors

3.1 Introduction

The previous chapter examined the behavioural mechanisms of how humans assign credit and infer their control from ambiguous feedback in social contexts. In short, I showed that participants use their estimates of control to assign prediction errors to themselves and others. Participants also exhibit a second learning mechanism mediated by exploratory behaviour which I refer to as active disambiguation (AD). I found that AD guided control inference. This chapter focusses on the neural underpinnings mediating these behaviours.

A key aspect of establishing control is establishing whether outcomes are contingent on our actions over and above any contingency that already exists between outcomes and states of the environment (Ligneul, 2021; Ligneul et al., 2022). Analogously, establishing control in social situations requires the establishment of each individual's degree of agency relative to one another. The supramarginal gyrus (SMG) in the inferior parietal lobe has been linked to sense of agency. Patients with schizophrenia and who experience delusions of control show stronger activation in the SMG than healthy individuals or people suffering from schizophrenia but without control delusions (Spence et al., 1997). Apraxic patients with lesions in the left SMG confuse whether a hand movement is their own or someone else's (Sirigu et al., 1999) and SMG activity has been linked to realizing whether an action was one's own or not (Farrer & Frith, 2002; Ohata et al., 2020). SMG activity tracks the deviance between the expected and experienced sensory feedback to one's actions, leading to a diminished sense of agency (Farrer et al., 2003). More generally, the SMG has been linked to directing attention to movement (Rushworth et al., 2001a; Rushworth et al., 1997; Rushworth et al., 2001b) and

tracks movement intentions (Desmurget et al., 2009). The SMG tracks actions in an abstract manner such as the kinematics of tool use (Orban & Caruana, 2014; Peeters et al., 2009; Peeters et al., 2013) and shows increased activation during the planning of pantomimed tool use (Króliczak & Frey, 2009). The SMG is, therefore, a candidate brain area for interpreting the consequences of any active intervention strategy that a person might employ to disambiguate their own or another's level of control over a situation.

In this chapter, I investigated the neural mechanisms of how people assign ambiguous feedback to themselves, others, and their control over the feedback (Spiering et al., 2024). For this, I analysed the neural data from the MRI experiment presented in chapter 2. I found that the SMG represented uncertainty and AD at the time of action generation. During the outcome phase, the SMG also tracked the feedback difference that was observed from AD and normal behaviour, as well as the control estimate that was then inferred from the feedback difference. Finally, I show that the SMG was also involved in tracking prediction errors in an agent-based manner according to control.

3.2 Methods

3.2.1 Imaging data acquisition and preprocessing

Imaging data were acquired with a Siemens Prisma 3T MRI using a 32-channel head-coil. T1 weighted structural scans were obtained with repetition time (TR) = 1,900ms, echo time (TE) = 3.96ms and 1×1×1mm voxel size. Functional images were acquired using a multiband T2*-weighted echo planar imaging sequence with acceleration factor of two, and with an oblique angle of 30° to the posterior commissure (PC) – anterior commissure (AC) line to reduce signal dropout in orbitofrontal regions (Deichmann et al., 2003). Other acquisition parameters included 2.4×2.4×2.4 mm voxel size, TE = 30ms, TR = 1,360ms, 60° flip angle, a 240mm field of view and 66 slices per volume. For each session, a fieldmap (voxel size = 2.4×2.4×2.4 mm, TR = 482.0ms, TE1 = 4.92ms, TE2 = 7.38ms) was acquired to reduce spatial distortions. Bias correction was applied directly to the scan. We took physiological recordings during the task to measure pulse, breathing and skin conductance of participants.

Imaging data was analysed using FMRIB's Software Library (FSL) (Smith et al., 2004). Preprocessing stages included motion correction, correction for spatial distortion by applying the fieldmap, brain extraction, spatial smoothing (using a full-width at half-maximum of 3 mm), and high-pass temporal filtering. Functional images were first co-registered to an individuals' T1 structural image and then nonlinearly registered to the Montreal Neurological Institute (MNI) template using 12 degrees of freedom.

3.2.2 fMRI whole brain analysis

First-level analysis was carried out using FSL FEAT (version 6.00). First, data were pre-whitened with FSL FILM to account for temporal autocorrelations. Temporal derivatives, standard motion correction parameters and physiological noise regressors (using the FSL PNM toolbox, Brooks et al. (2008)) were included in the model. Results were obtained with automatic

outlier-deweighting, and using FSL FLAME1 with cluster-correction threshold of $z > 3.1$ and $p < 0.05$ (Worsley, 2001).

We used a single GLM for whole brain analysis, and included all three phases of a trial (rating, action, outcome) in the GLM. Constant and parametric regressors were time-locked to the onsets of the relevant trial event, and had the duration of the respective event. All parametric regressors were normalized to a mean of zero and standard deviation of 1. All regressors were convolved with the double-gamma HRF in FEAT. Correlations between regressors are shown in Figure 3.1.

Rating phase. The rating phase began at the time the rating screen appeared and lasted until the participant finished all ratings. To account for differences between Self-Other and Control-Other phases, the regressors related to control and self ratings were defined separately for each task phase. The regressors for the other rating were defined across both task phases because participants needed to learn about the other's performance level in a similar manner in both task phases.

- Control rating related regressors, time locked to rating screen onset:
    - constant for control rating
    - parametric control rating
- Self rating related regressors, time locked to self rating onset (in the Self-Other phase, when the true control was shown, this was time locked to the rating screen onset + 1.23sec, the mean RT for making a rating (computed from the Self-only trials)):
    - constant for self rating
    - parametric self rating
- Other rating related regressors, time locked to other rating onset
    - constant for other rating

- parametric other rating

Parametric regressors of no interest were included for the reaction times (RT) of each rating (separate by task phase, control rating RT only for Control-Other phase).

Action phase. The action phase began when participants played the games and lasted until the game finished. Regressors were defined across Self-Other and Control-Other phases because the action phase did not differ between task phases.

- game-type specific constants (four in total, one for each of the four games)
- AD (binary, prior to normalizing: one for AD trial, zero for normal trial)
- parametric uncertainty (extracted from active Bayesian learning model)
- switch trial (binary, prior to normalizing: one for switch trial, zero for stay trial)

In addition, a parametric regressor of no interest was included for the RT of the action phase to control for RT differences on AD trials. Another parametric regressor of no interest was included for the trial number across the game block to control for the effect of trial number (as a measure of time in block) on uncertainty.

Outcome phase. The outcome phase started when the feedback appeared on the screen and lasted until the participant pressed the 'Continue' button. Regressors were defined depending on the task phase because we hypothesized different learning mechanisms per task phase.

- Constant for outcome phase
- Parametric regressors in Self-Only phase:
  - sPE (on normal trials only)
  - oPE (on normal trials only)
  - signed uncertainty change (from the current to the next trial)
- Parametric regressors in the Control-Other phase:
  - Feedback difference (defined as normal – AD feedback, on switch trials only)

- signed uncertainty change (from the current to the next trial)

We included regressors of no interest in the outcome phase. To account for motor-related brain activity, we included a constant time-locked to when participants pressed the 'Continue' button and with a duration of one sec. For each task phase separately, we included binary regressors for switch trial and AD (both regressors were defined similarly to corresponding regressors in the action phase).

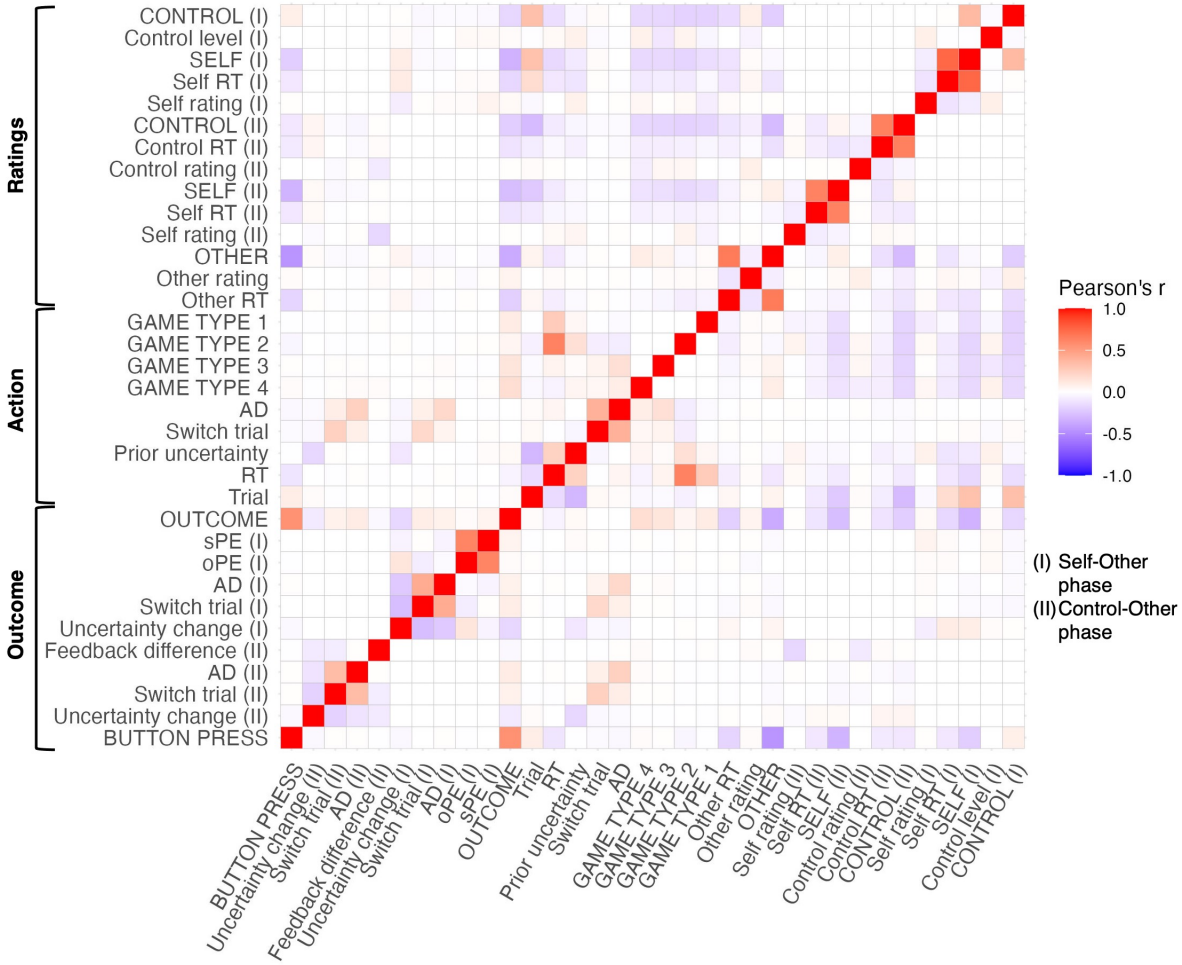


Figure 3.1. Correlations of regressors (convolved with hemodynamic response function) included in the fMRI whole brain analysis. Pearson's correlations were averaged across participants. Constant regressors are highlighted in uppercase. In the action phase when participants play the games, we included four constant regressors, one per game type (1=Light, 2=Slingshot, 3=Rockslide, 4=Bouncing ball).

3.2.3 fMRI time course analysis on ROIs

3.2.3.1 ROI selection and time series extraction

Spherical ROIs with a diameter of three voxels were constructed based on significant activation clusters in the whole-brain analysis. ROIs were centred on the peak voxels of the effect of uncertainty at the time of action, and where uncertainty-related clusters overlapped with the effect of AD at the time of action. This was the case for three brain regions, from which we discarded one for our analyses because it was in the occipital cortex. The two remaining ROIs were in the left SMG (MNI: x=-62, y=-24, z=36) and left area 7 in the superior parietal lobule (MNI: x=-12, y=-64, z=64). To ensure that only brain was included in the ROIs, we centred them two voxels more medially for the left SMG, and one voxel more inferiorly for the left area 7.

Spherical ROI masks were transformed from standard space to each participant's structural space using a standard-to-structural warp field, and from structural to functional space with a structural-to-functional affine matrix. The resulting transformed masks were thresholded and binarized. Then, motion and physiological noise were partialled out from the filtered functional data. The time-series of each ROI was extracted from the residual filtered functional data, and averaged across voxels. Timeseries data were normalised, and up-sampled 20 times using cubic spline interpolation. The resulting data were epoched into 17 sec windows of 2 sec before and 15 sec after each event of interest (described in GLM section below). We applied ordinary least squares GLMs to each time point, resulting in a time course of beta weights for each regressor. All parametric regressors were normalized (mean of 0 and a standard deviation of 1). In all of the following ROI-GLMs, we included constant regressors in addition to the described parametric regressors below.

3.2.3.2 Time course GLMs (ROI-GLMs)

We hypothesized that these regions which are involved in processing uncertainty and AD during action generation might also be involved in processing the feedback outcomes. Therefore, our time course analyses were independent of the whole-brain analyses based on which the ROIs were defined. Correlations between regressors for all ROI-GLMs are shown in Figure 3.2.

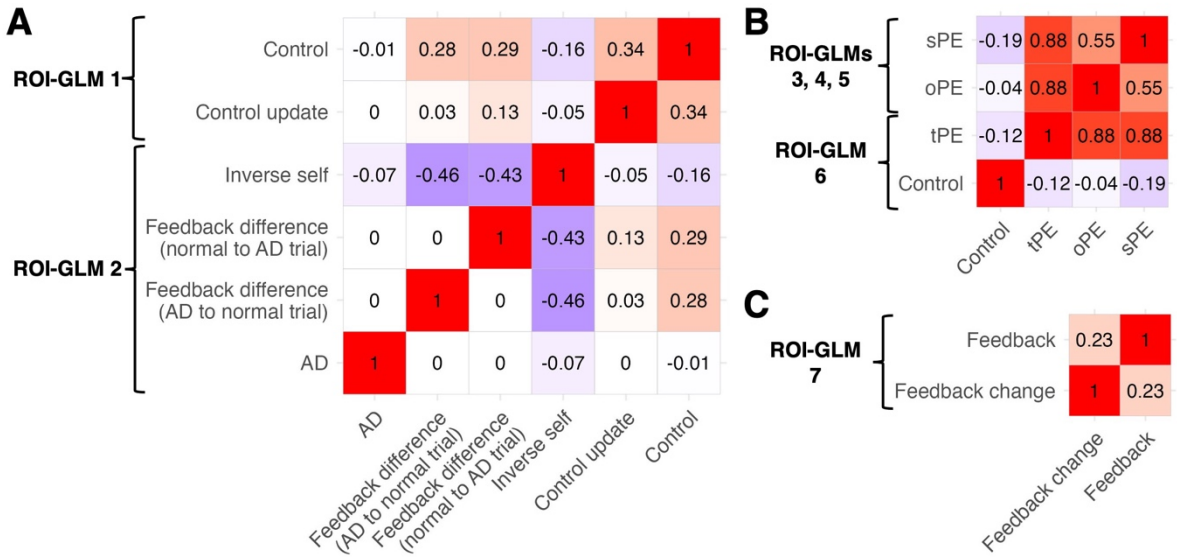


Figure 3.2. Correlation matrices for regressors included in time course analyses (ROI-GLMs). A) Correlation of regressors shown in Figure 3.4. Regressors for the ROI-GLMs 1 and 2 are shown in one correlation matrix to highlight that the regressors of control belief (ROI-GLM 1) and feedback differences (ROI-GLM 2) show low correlations ($r<0.3$). **B)** Correlations of regressors shown in Figure 3.6. Since sPE and oPE showed a correlation of $r=0.55$, we tested the time courses of these regressors first separately (ROI-GLMs 3 and 4), and then together (ROI-GLM5). We also tested the effects of tPE and control separately in ROI-GLM6. **C)** Regressor correlations for supplementary Figure S12 (ROI-GLM7).

Our overall hypothesis is that the brain represents the inferred control belief based on the currently observed feedback difference and the prior inverse self-estimate (Figure 3.4A). First, we tested whether SMG and area 7 represent the inferred control estimate and its update. Therefore, we ran the following ROI-GLM1, time-locked to the outcome of a current trial (Figure 3.4B):

$$BOLD \sim control + control\ update \quad (3.1)$$

where *control* is the control rating from the next trial and is therefore a measure of what control level participants infer from the currently observed feedback. *Control update* is the difference between the next trial's and the current trial's control rating, and is therefore a measure of how much participants change their control rating based on the currently observed outcomes. All trials in the Control-Other phase are included in this ROI-GLM1.

The goal of our next GLM was to examine if ROI activity correlates with the feedback difference and inverse self, which are the components necessary for control inference. Here, we examined two competing hypotheses (see Figure 3.4C-D) for how our ROIs might represent the feedback difference at the time of outcome. Firstly, we hypothesized that neural activity might represent the feedback difference as the comparison between normal and AD feedback to infer control (H1, Figure 3.4C). This means that neural activity does not depend on whether the currently observed outcome is from an AD trial and the previous was from a normal trial or *vice versa*. Alternatively, neural activity might track feedback differences between outcomes in a simpler comparison of current trial versus previous trial. In this case, neural activity depends only on the temporal order of current and previously seen feedback, and not on the identity of the trial as normal or AD. We tested these two hypotheses in a single GLM time-locked to the outcome phase (ROI-GLM2):

$$BOLD \sim feedback\ diff_{AD\ to\ normal} + feedback\ diff_{normal\ to\ AD} + inverse\ self + AD \quad (3.2)$$

with

$$feedback\ diff_{AD\ to\ normal} = normal\ feedback_{current\ trial} - AD\ feedback_{previous\ trial} \quad (3.3)$$

$$feedback\ diff_{normal\ to\ AD} = normal\ feedback_{previous\ trial} - AD\ feedback_{current\ trial} \quad (3.4)$$

Analogously to the behavioural analysis, only switch trials in the Control-Other phase are included in this analysis because these are the trials when participants can observe a feedback difference (between normal and AD trials), which is the critical variable to infer controllability. To test H1 and H2 in one GLM, we defined two regressors for feedback difference which are both computed as the difference of normal minus AD feedback. *Feedback diff*$_{AD\ to\ normal}$ is the feedback difference only on switch trials where the previous trial was AD and the current is normal. In contrast, *feedback diff*$_{normal\ to\ AD}$ is only defined on trials where the previous trial was normal and the current is AD. Defining the feedback difference separately by the type of switch trial allowed us to pit the two hypotheses H1 and H2 against each other (bar plot in Figure 3.4E). This is because the average across *feedback diff*$_{AD\ to\ normal}$ and *feedback diff*$_{normal\ to\ AD}$ is the overall effect in favour of H1, while the average across *feedback diff*$_{AD\ to\ normal}$ and the sign-flipped *feedback diff*$_{normal\ to\ AD}$ is the overall effect in favour of H2. By only including switch trials into this analysis, we also ensured that any feedback difference related effects cannot be driven by the feedback difference being higher on switch than stay trials. *Inverse self* is the inverse of the current trial's self rating (i.e. before observing the feedback, computed similarly as for the behavioural analysis) to measure what participants' prior for their own performance is. *AD* was included as a regressor of no interest to control for other features that might impact the feedback difference between outcomes.

Next, we investigated the neural correlates of how participants split up the tPE into sPE and oPE according to their known control level in the Self-Other phase. For this, we ran three GLMs. Analogously to the behavioural analysis, we analysed only normal trials during the Self-Other phase and computed sPE and oPE in the same way as for the behaviour. First, we tested whether SMG and area 7 represented the sPE and oPE at the time of outcome. Due to the high correlation between sPE and oPE (r=0.55), we tested their effects in two separate regressions

(ROI-GLMs 3 and 4), and jointly within a single regression model (ROI-GLM5) and ensured that similar results were found in all cases.

$$BOLD \sim sPE \tag{3.5}$$

$$BOLD \sim oPE \tag{3.6}$$

$$BOLD \sim sPE + oPE \tag{3.7}$$

We also tested an alternative approach for how neural activity might represent the splitting up of the tPE in the outcome phase. In ROI-GLM6, we tested whether there are representations of the tPE and control level, which are used to construct the sPE and oPE.

$$BOLD \sim tPE + control \tag{3.8}$$

where *control* is the true control level and *tPE* is the total prediction error. Analogously to our behavioural analysis, in ROI-GLMs 3 to 6, we analysed only normal trials in the Self-Other phase because this is when the feedback reflects both self and other, and therefore generates a tPE that can be split up into an sPE and oPE in relation to the known control level.

Finally, we tested whether neural activity correlates with feedback in our task, independent of any credit assignment or inference process. A key region that has been implicated in tracking reward is the ventral striatum. To test whether the ventral striatum signals feedback in our task, we extracted activity from an anatomically defined bilateral mask of the nucleus accumbens (Khalighinejad et al., 2021). We ran the following regression model time locked to the outcome phase (ROI-GLM7):

$$BOLD \sim feedback + feedback\ change \tag{3.9}$$

where *feedback* is the parametric feedback observed on the current trial, and *feedback change* is the signed difference between the feedback observed on the current compared to the previous trial. To ensure that any effects related to these two regressors are independent of AD, we tested this only on normal trials following normal trials, across Self-Other and Control-Other phases.

3.2.3.3 Leave-one-out (LOO) procedure

To test for significance and avoid temporal selection biases in our time course analyses, we used a similar approach to previous studies (Khalighinejad et al., 2021; Scholl et al., 2017; Trudel et al., 2021; Wittmann et al., 2016b). For each participant (n=31), we computed the average time course across the remaining participants (n=30). Next, we computed the time point between 4 and 10 sec (unless otherwise specified) when this average time course peaks, and extracted the beta value for the excluded participant at this peak time. This was repeated for all participants, resulting in a set of individual peak values that were determined in an unbiased way. These peak values were tested with a two-tailed one-sample t-test against zero.

3.3 Results

3.3.1 Supramarginal Gyrus tracks uncertainty and active disambiguation at the action

We analysed the neural data to identify the neural correlates of AD and AD-guided controllability estimation. We first examined neural activity when participants engaged in AD to resolve uncertainty. For this, we ran a whole-brain analysis to identify the neural correlates of AD and trial-wise uncertainty (from the Active learner model, as in the behavioural analyses in chapter 2 before) at the time when participants initiated action during the games.

The left anterior supramarginal gyrus (SMG) was sensitive to both uncertainty (Figure 3.3Ai) and AD (Figure 3.3Bi). For AD, this activation was bilateral (extending to and peaking in the temporoparietal junction [TPJ] in the right hemisphere). This means that activity in the SMG signals uncertainty as well as AD generation at the time of the action. While uncertainty-related activations were in the SMG, we also found some uncertainty-linked deactivations (Figure 3.3Aii). While activations for AD overlapped with uncertainty-related activity in SMG, they were more widespread (Figure 3.3Bii). The peak coordinates of significant clusters for uncertainty and AD at the time of action can be found in Table 3.1.

Our analysis incorporated three important controls. First, the activations emerged after controlling for the variety of game types participants played during the course of the experiment. Second, analyses controlled for variation in reaction time. Third, although ADs occurred intermittently amongst other normal trials, AD-associated activity reflect AD *per se* rather than the more general process of switching between AD and normal trials; a control regressor in the analysis indexed switches (a trial on which participants switched from a normal to an AD trial or vice-versa).

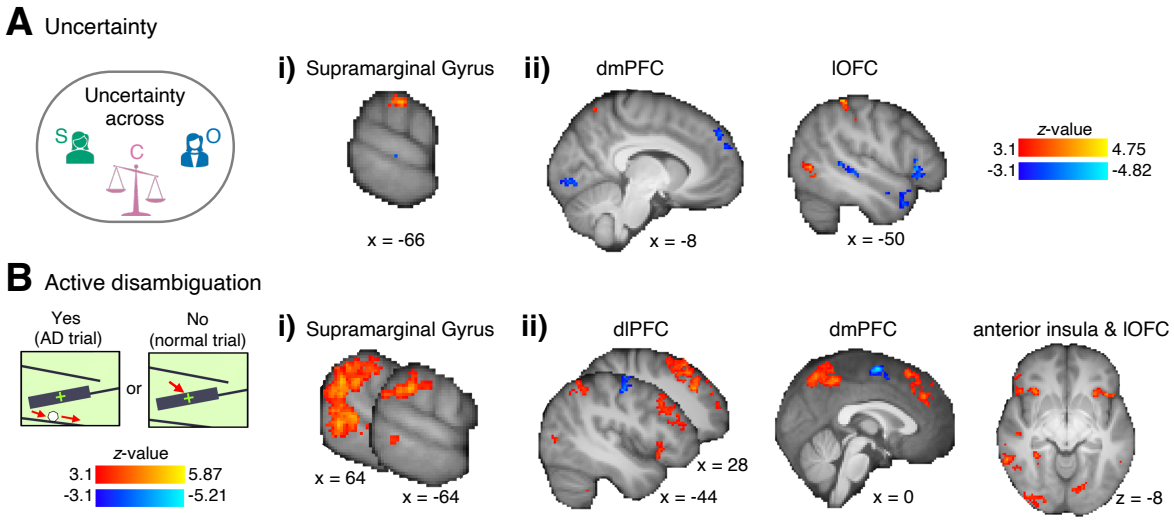


Figure 3.3. Whole-brain maps for average uncertainty and AD at the time of action. Since we found that participants generated AD trials under uncertainty in both task phases, we collapsed the data across both Self-Other and Control-Other phase in this neural analysis (see Methods for full details). **A)** i) Activity in the left SMG related to higher uncertainty. ii) Lower uncertainty (higher certainty) was reflected in dmPFC and left lOFC. **B)** i) We also found bilateral activation related to AD in SMG (with a peak in TPJ in the right hemisphere). ii) Additionally, we found AD-related activity in bilateral dorsolateral prefrontal cortex (dlPFC), extending to dorsomedial PFC (dmPFC), and adjacent dACC, bilateral anterior insula, and lOFC. The peak coordinates of significant clusters can be found in Table 3.1. n=31 MRI, whole-brain effects are cluster corrected with $z>3.1$ and $p<0.05$; effects are time-locked to the time of action, when participants played the games, and are across Self-Other and Control-Other phase.

Table 3.1. Peak coordinates of significant clusters in whole-brain fMRI analysis.

Contrast	Region	Peak coordinates x/y/z (in mm MNI space)	z value
Uncertainty	Lateral occipital cortex, superior division	-38, -80, 32	4.7
	Supramarginal gyrus (SMG)	-66, -24, 36	4.25
	Lateral occipital cortex, inferior division	-50, -66, 2	4.14
	Postcentral gyrus	-50, -38, 58	4.42
	Area 7	-12, -64, 66	3.94
	Superior parietal lobule	-38, -46, 64	3.78
	Intracalcarine Cortex	2, -84, 2	-4.77
	Superior temporal gyrus, posterior division	-56, -34, 4	-4.54
	Lateral orbitofrontal cortex (lOFC)	-50, 22, 0	-4.54
	Occipital pole	16, -90, 14	-4.17
	Dorsomedial prefrontal cortex (dmPFC)	-8, 52, 40	-3.92

	Temporal pole	-52, 12, -20	-3.9
AD	Temporoparietal junction (TPJ) - SMG subpeak (right):	52, -44, 34 SMG: 64, -34, 30	5.82 SMG: 4.70
	Dorsolateral prefrontal cortex (dlPFC) - dmPFC subpeak:	Right: 28, 32, 34 Left: -44, 38, 28 dmPFC: 0, 40, 34	Right: 4.93 Left: 3.76 dmPFC: 4.50
	Precuneus	2, -50, 46	4.98
	SMG (left)	-64, -36, 32	5.04
	Lateral frontal pole	40, 52, 0	4.44
	Inferior frontal gyrus, pars opercularis, stretching into lOFC (left)	52, 18, -2	4.93
	Lateral occipital cortex, inferior division	Right: 40, -86, -8 Left: -48, -74, 4	Right: 4.08 Left: 3.98
	Lingual gyrus	-6, -78, -6	4.65
	Cerebellum, uvula	-26, -78, -24	4.80
	2 clusters in right lOFC (area 47o)	Cluster 1: 32, 22, -8 Cluster 2: 48, 24, -8	Cluster 1: 4.57 Cluster 2: 4.23
	Parahippocampal gyrus	32, -44, -6	4.82
	Cerebellum, tonsil	-38, -44, -44	3.82
	Middle temporal gyrus	-64, -46, -2	3.70
	Anterior cingulate cortex	10, 50, 12	4.13
	Cerebellum, anterior lobe	-32, -60, -32	4.11
	Area 4	Left: -42, -14, 50 Right: 44, -8, 58	Left: -4.61 Right: -4.52
	Supplementary motor area	0, -4, 60	-5.16
Family-wise error cluster corrected, $z > 3.1$, $p < 0.05$			

3.3.2 Supramarginal Gyrus carries neural signals for learning controllability

Our behavioural results showed that participants inferred their control from the observed feedback difference between AD and normal trials. Next, we established the neural mechanisms mediating the inference process at the time of outcome when the feedback is shown (Figure 3.4A). We tested this in the SMG as the area sensitive to both uncertainty and AD at the time of action. As in the behavioural analysis, we focused on the Control-Other phase when participants learnt their control level and again focussed on how participants compared AD and normal trials. First, we tested if, at the time of outcome, the SMG represented the control estimate and the updating of control (ROI-GLM1). We indeed found that the SMG represented the inferred control level ($t(30)=2.34$, $p=0.03$, Figure 3.4B left). Additionally, SMG

activity was modulated by the control update from current to the next trial ($t(30)=-3.79$, $p<0.001$, Figure 3.4B right).

Next, we tested in a separate GLM whether neural activity reflected the components necessary for control inference. We contrasted two competing hypotheses for how neural activity might reflect the feedback difference between trials: First, neural activity might represent the feedback difference computed as the difference between normal minus AD trial (hypothesis one [H1], Figure 3.4C). This coding scheme encapsulates a comparison that directly leads to control inference. Alternatively, neural activity might track the feedback difference by comparing the current trial's feedback to the previous trial's feedback (hypothesis two [H2], Figure 3.4D). Such a contrast might be simpler to perform as each new piece of feedback is obtained and it resembles the types of contrasts at the heart of prediction error learning mechanisms when the current trial outcome is compared with weighted estimates of previous trial outcomes. However, despite its simplicity this contrast does not lead directly to control inference.

In summary, we found that at the time of outcome, the SMG activity reflected a comparison between normal and AD feedback in support of H1, rather than H2 (H1: $t(30)=3.77$, $p<0.001$; H2: $t(30)=0.06$, $p=0.95$; ROI-GLM2, Figure 3.4E). We ruled out that this signal reflected AD per se by including a categorical regressor for whether the current trial was AD. In the same GLM, we tested for the inverse of participants' prior estimate of self-performance (inverse self), the other component mediating control inference. We found that SMG activity was also sensitive to the inverse self (i.e. activation with high self-performance, $t(30)=-2.27$, $p=0.03$; ROI-GLM2, Figure 3.4F). We also observed a relatively early peak linked to the inverse self. While it lay outside our pre-defined time window of interest, a leave-one-out (LOO) procedure identified the effect as significant (-2 to 2sec time window: $t(30)=3.69$, $p<0.001$). This early peak could

result from the self being a prior estimate that is known ahead of the feedback. In Figure 3.5 we consider another brain region where uncertainty and AD activation clusters overlap.

In summary, SMG signalled all component variables needed for control inference, including the observed feedback difference between AD and normal trials, the self-performance estimate, and the updating of the control estimate at the time of the outcome. Crucially, it also represented the inferred control level at the same time. In addition, SMG activity reflected the initiation of the AD process and the uncertainty levels that, when high, preceded AD.

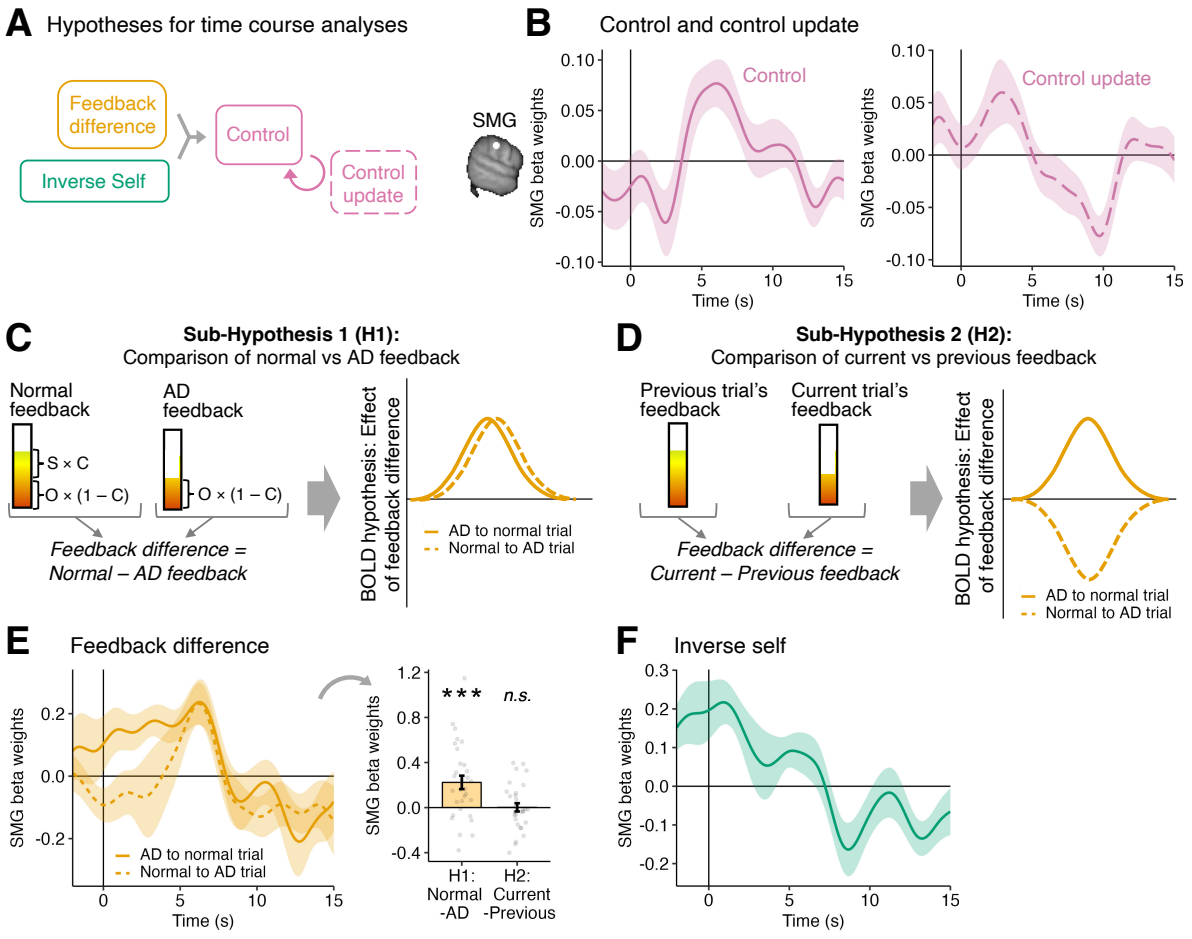


Figure 3.4. Representations of control, feedback difference and inverse self in the SMG at the time of outcome in the Control-Other phase. **A)** We used time-course analyses to identify the neural correlates of how participants infer their control from feedback difference and the inverse self. Since our whole-brain analyses showed that the SMG represents uncertainty and AD at the time of action, we selected this region as an ROI for the time-course analyses on the outcome phase. **B)** First, we tested

whether at the time of outcome (ROI-GLM1), the SMG represents participants' inferred control levels (next trial's control rating) and the updating of the control estimate (difference between current and next trial's control rating). We indeed found that SMG tracks the inferred control estimate (left panel) and the degree of control level updating (right panel) at the time of outcome. These activity patterns remained unchanged whether control and control update were examined in the same or separate GLMs. **C)** In a new GLM (ROI-GLM2), we next tested whether SMG also represents the components necessary for inferring control. We compared two neural hypotheses for how neural activity can track the difference in feedback between trials. Our first neural sub-hypothesis (H1) is that neural activity tracks feedback difference as the difference between normal and AD trials (left). This means that the neural activity does not depend on order: whether a normal trial followed an AD trial, or *vice-versa* (right). **D)** Alternatively (H2), neural activity might track the feedback difference in a simpler manner, by comparing the currently observed feedback to the feedback from the previous trial. We used a single GLM (ROI-GLM2) to test these two hypotheses. For this, we computed feedback difference as normal vs AD (according to H1) in all panels. However, we split trial transitions between AD to normal trial transitions and *vice versa*, because H1 and H2 make different predictions between those trial types. **E)** At the outcome time (ROI-GLM2), the SMG represents the feedback difference as normal vs AD feedback in accordance with H1 (left). When aggregating the participants' peak amplitudes using a LOO procedure, we found a significant effect of feedback difference in accordance with H1 rather than H2 (right). Note that the shown effects of feedback difference and control at the time of outcome (panels E and B left) have low regressor correlations ($r<0.3$, Figure 3.2A), and are run on different trial selections (control on all trials, feedback difference on switch trials only). **F)** In the same GLM (ROI-GLM2), we found that SMG also tracks participants' prior inverse self-estimate (current trial's inverse self rating) in the outcome phase. n=31 MRI; error bars and shaded intervals show s.e.m. across participants; n.s.; not significant; ***, $p<0.001$; effects are time locked to the outcome phase; in panel A all trials from the Control-Other phase are included; in panels E-F only switch trials in the Control-Other phase are included.

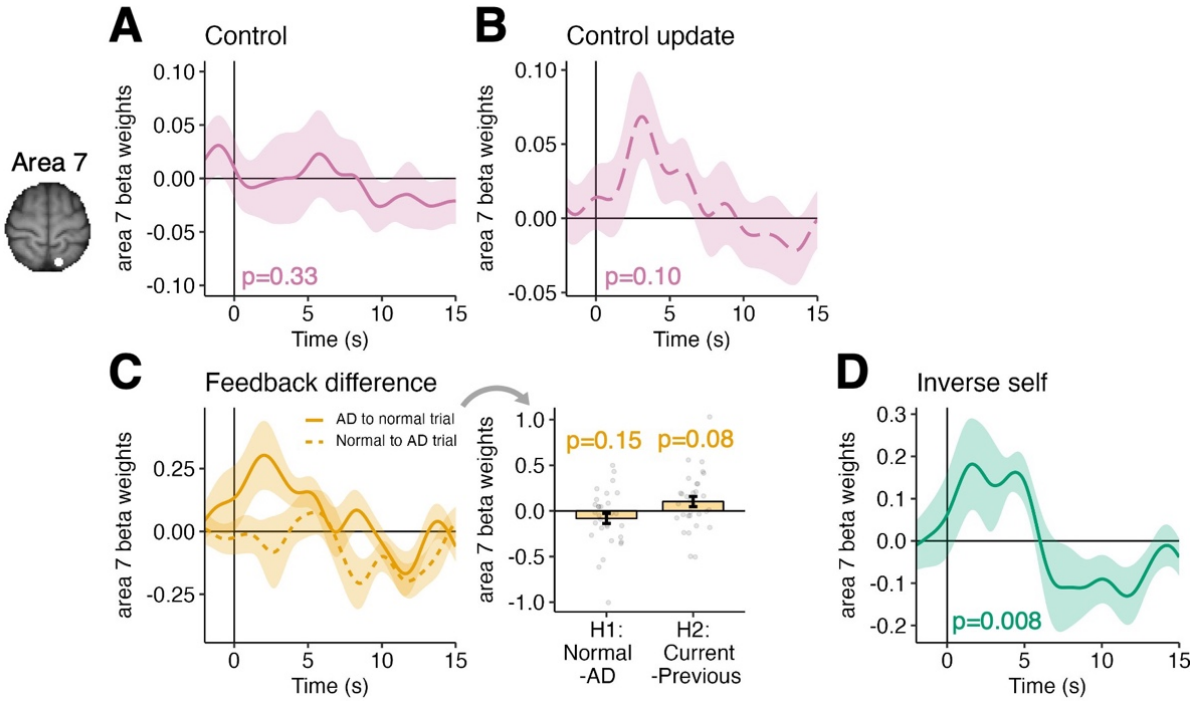


Figure 3.5. Time courses of feedback difference, control and self in left area 7, time-locked to the outcome phase. **A-B)** We found no evidence that area 7 activity signals the inferred control or the control update at the time of outcome (control, t(30)=-0.98, p=0.33; control update, t(30)=1.69, p=0.10; ROI-GLM1). **C)** While the waveforms of feedback differences look more alike H2 than H1, we found no evidence in favour of either hypothesis in area 7 (H1, t(30)=-1.48, p=0.15; H2, t(30)=1.83, p=0.08; ROI-GLM2). This is unlike the SMG result that we found (Figure 3.4). **D)** In the same GLM, we also tested for the inverse self. We indeed found that area 7 tracks the prior inverse self estimate at the time of outcome (inverse self, t(30)=2.82, p=0.008; ROI-GLM2). n=31 MRI, mean beta weights are plotted as lines, with s.e.m. as shaded intervals.

3.3.3 Supramarginal Gyrus tracks prediction errors assigned to self and others

Our behavioural analyses in chapter 2 showed that when participants know the level of their control over outcomes (Self-Other phase), they assign prediction errors to themselves (sPE) and others (oPE) by taking into account control. Next, we examined the neural correlates of this process time-locked to the time of credit assignment (i.e. outcome phase).

Given the activity patterns we had already identified in SMG, we again used the SMG as an ROI. We indeed found that SMG activity was modulated by PEs assigned to self and other (sPE,

t(30)=2.55, p=0.016, ROI-GLM3, Figure 3.6A; oPE, t(30)=2.34, p=0.026, ROI-GLM4, Figure 3.6B). Due to the correlation between oPE and sPE (r=0.55, Figure 3.2B), we tested their effects in separate GLMs (as reported before, Figure 3.6A-B), and also together in one GLM (sPE, t(30)=3.83, p<0.001; oPE, t(30)=-2.41, p=0.02; Figure 3.7A-B, ROI-GLM5). Each time, oPE and sPE effects remained significant in the SMG. Alternatively, we tested whether the SMG might track the tPE (the total prediction error or the sum of sPE and oPE). However, we found no evidence that the SMG tracked the overall tPE nor did it track control in this task phase, when participants had already been instructed about its level and it did not need to be inferred (control, t(30)=1.43, p=0.16, Figure 3.7C; tPE, t(30)=-0.36, p=0.72, Figure 3.6C; ROI-GLM6). We consider another brain region where uncertainty and AD activation clusters overlap in Figure 3.8. Together, our results suggest that in contexts when people already know their control, the SMG is involved in processing outcomes in an agent-based manner, attributed to either self or others.

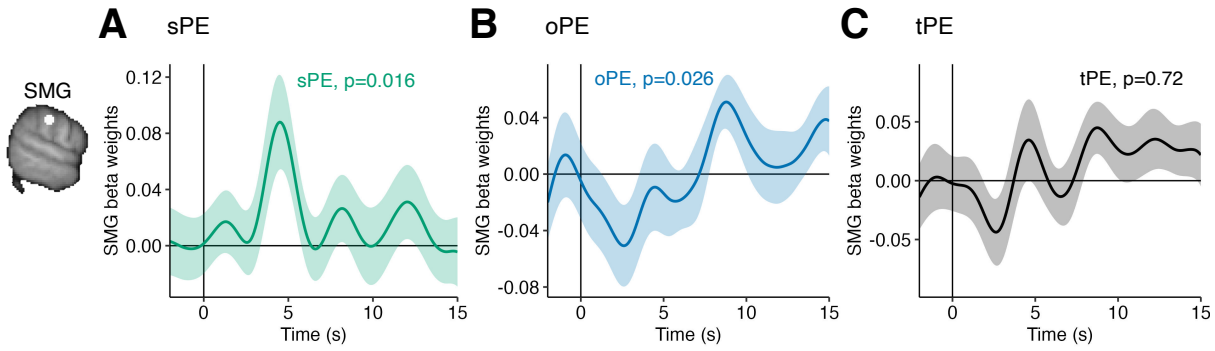


Figure 3.6. Representations of sPE and oPE in the SMG during the outcome phase in the Self-Other phase. In the Self-Other phase, the key credit assignment problem that participants face is to assign credit to themselves and other people according to their control. On normal trials, they split up the tPE into a sPE and oPE according to their known control. We therefore tested whether BOLD activity in the left SMG is modulated by sPE and oPE. Since sPE and oPE are correlated with r=0.55, we tested this both with two separate GLMs containing either sPE but not oPE and vice versa, and a GLM in which both variables were included. All the analyses, however, led to the same conclusion illustrated in the next two panels. **A)** We found that at the time of the outcome, activation of the left SMG is significantly related to the sPE (ROI-GLM3). **B)** In a separate GLM, we found that SMG BOLD activity is significantly modulated by the oPE (ROI-GLM4). **C)** In a separate analysis, we found that the SMG does not

represent the tPE which is the sum of sPE and oPE (ROI-GLM5). This is an interesting feature and together with panels A and B suggests that the SMG is involved in belief updating in an agent-based manner, rather than the total outcome per se. n=31 MRI; error bars and shaded intervals show s.e.m. across participants; effects are time locked to the outcome phase and analogously to our behavioural analysis, only normal trials in the Self-Other phase are included.

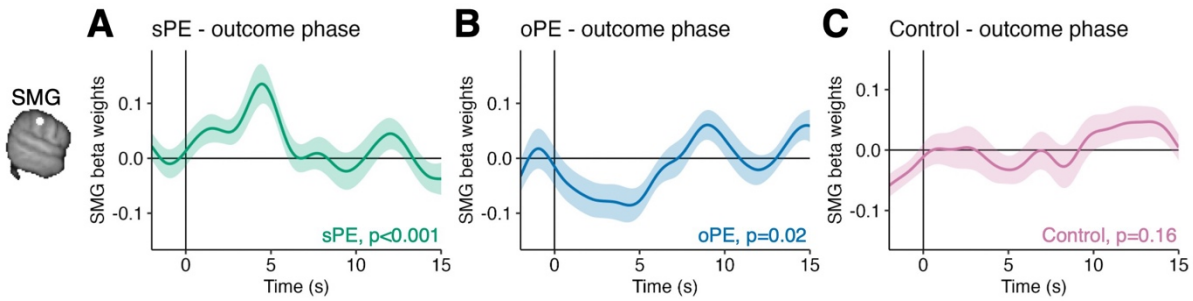


Figure 3.7. Time course analyses of sPE, oPE and Control in the SMG. A) Due to the correlation of sPE and oPE, we also ran a separate time course analysis including both sPE and oPE in one ROI-GLM (ROI-GLM5). Here, we found a significant effect of sPE ($t(30)=3.83$, $p<0.001$). **B)** oPE, which was included in the same GLM, also still showed a significant effect ($t(30)=-2.41$, $p=0.02$). **C)** In ROI-GLM6, we tested for effects of tPE (plotted in Figure 3.6C) and control in the outcome phase. We did not find that SMG represents the known control level ($t(30)=1.43$, $p=0.16$) in this phase of the task – the Self-Other phase – when participants had already been instructed about its level and did not need to infer it. n=31 MRI, mean beta weights are plotted as lines, with s.e.m. as shaded intervals.

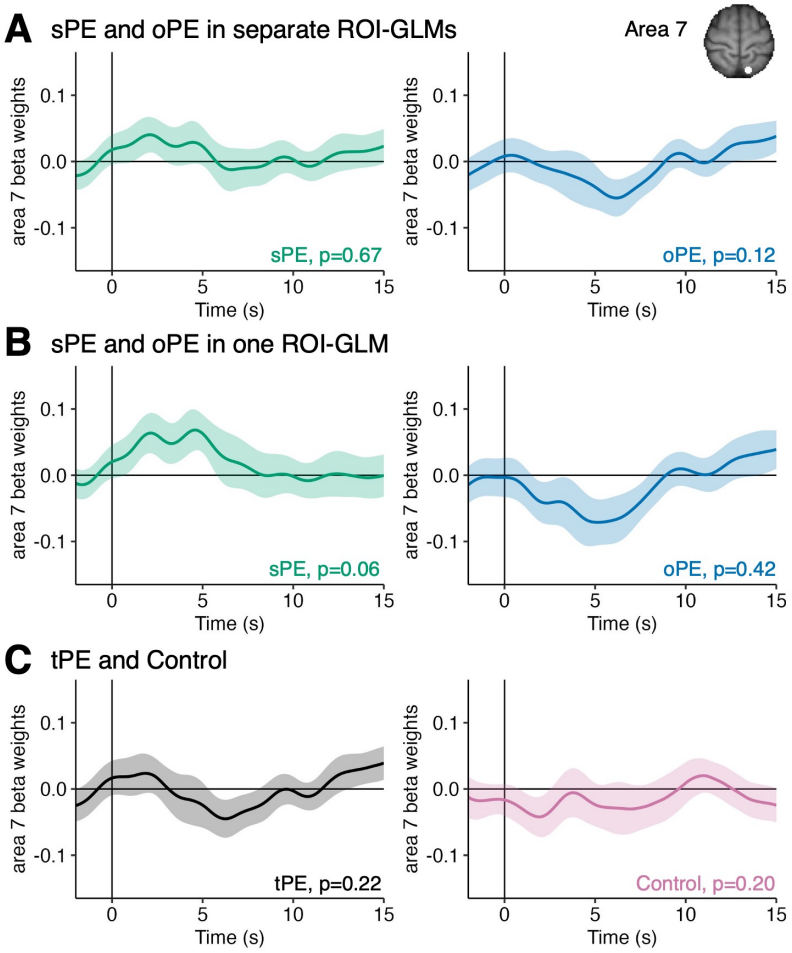


Figure 3.8. Time courses of sPE, oPE, tPE and control in area 7 at the time of outcome, during the Self-Other phase. A) We did not find that area 7 BOLD activity is significantly modulated by sPE nor oPE when tested in separate GLMs (sPE, t(30)=0.43, p=0.67, ROI-GLM3; oPE, t(30)=-1.58, p=0.12, ROI-GLM4). **B)** Tested in one GLM, sPE and oPE remained not significant in area 7 (sPE, t(30)=1.92, p=0.06; oPE, t(30)=-0.81, p=0.42; ROI-GLM5). **C)** BOLD activity in area 7 was not significantly modulated by tPE nor the known control either (tPE, t(30)=-1.26, p=0.22; control, t(30)=1.30, p=0.20; ROI-GLM6). n=31 MRI, mean beta weights are plotted as lines, with s.e.m. as shaded intervals.

3.3.4 Nucleus accumbens tracks feedback and feedback change

Finally, we tested whether neural activity tracked reward in our task, irrespective of any inference process. The ventral striatum has been implicated in representing reward in past studies. In an anatomically defined ROI of the ventral striatum, we found it indeed tracked the shown feedback, as well as trial-wise changes of the feedback at the time of the outcome (feedback, t(30)=2.23, p=0.03; feedback change, t(30)=2.48, p=0.02; ROI-GLM7, Figure 3.9).

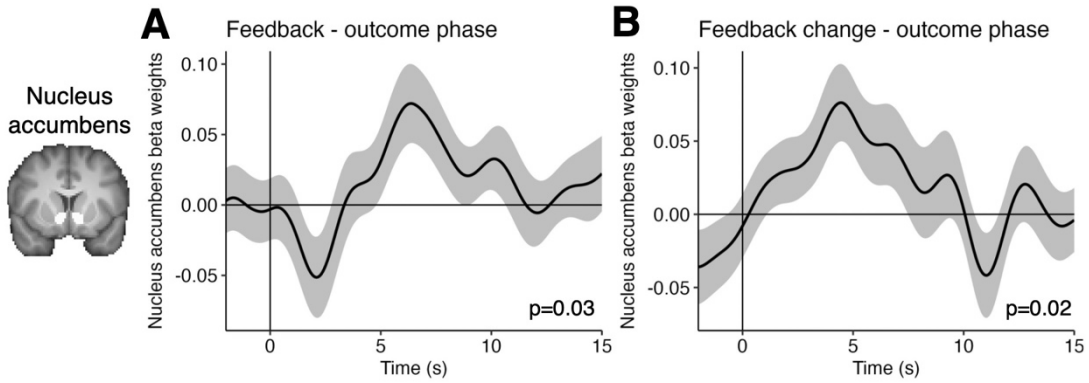


Figure 3.9. Time courses of feedback and feedback change in the ventral striatum in the outcome phase. Previous studies have shown that the ventral striatum is sensitive to reward (e.g. Wittmann et al., 2016a). To test whether the ventral striatum also signals reward in our task, we ran a time course analysis on an anatomically defined bilateral mask of the nucleus accumbens (ROI-GLM7). **A)** We found that activity in the nucleus accumbens tracks the feedback that is being observed at the time of outcome ($t(30)=2.23$, $p=0.03$, ROI-GLM7). **B)** In the same GLM, we also found that the nucleus accumbens signals the change of the feedback (difference between current and last trial's feedback) at the time of outcome ($t(30)=2.48$, $p=0.02$, ROI-GLM7). Analyses were run on normal trials that followed normal trials, across Self-Other and Control-Other phase. The regressors show low correlations ($r=0.23$, see Figure 3.2). n=31 MRI, mean beta weights are plotted as lines, with s.e.m. as shaded intervals.

3.4 Discussion

In this chapter, I presented neural findings related to control and causal learning from the novel task introduced in the previous chapter. I found that during action generation and when participants played the games, the SMG tracked uncertainty and AD. Importantly, these signals were controlled for lower level visual and motor differences and for whether participants switched their behaviour. At the time of outcome, the SMG signalled the control participants inferred and how much they changed their control beliefs based on the observed outcomes. In addition, I demonstrated that the SMG tracks the component variables necessary for control inference, including participants' self-performance estimates and the observed feedback difference between trials. Indeed, I found that the SMG tracks the difference between feedback observed on AD and non-AD trials, independent of their order. In addition, SMG tracked self- and other-assigned prediction errors in proportion to participants' level of control. This means that in our task, the SMG activity is associated with processing outcomes in an agent-based manner. This suggests that the SMG carries the neural correlates of the behavioural patterns on control inference and credit assignment which I presented in chapter 2. Finally, I found that activity in the nucleus accumbens signalled the observed feedback and how much it changed from trial to trial.

In the current task, SMG appeared to be central to credit assignment and identifying agency. Other brain regions such as area 47/12o on the border between ventrolateral and orbitofrontal cortex have been linked to credit assignment in non-social contexts (Folloni et al., 2021; Jocham et al., 2016; Noonan et al., 2017; Walton et al., 2010). However, monitoring the consequences of one's own and others' actions may depend on SMG, as has been found in previous work on action outcome attributions (Farrer & Frith, 2002; Ohata et al., 2020; Sirigu et al., 1999). Another difference between my and previous research is that previous work studied credit assignment of binary rewards to the agents' choices (e.g. Jocham et al., 2016;

Walton et al., 2010), while I examined credit assignment of continuous feedback to one's own and others' performances and degree of agency. This may make SMG important in establishing agents' relative levels of control – their degrees of agency – during social interactions.

I found widespread neural activation related to AD and to uncertainty about control and performance levels prior to AD at the time that participants generated actions. While activity in some components of the distributed network, such as SMG, may also be closely linked to credit assignment and agency identification processes, activity in dorsomedial prefrontal and adjacent anterior cingulate cortex and in dorsolateral prefrontal cortex may be more closely linked to the process of exploration itself (Kaanders et al., 2021; Stoll et al., 2016; Trudel et al., 2021).

Taken together, the results presented in this chapter suggest that the SMG is involved in control and performance estimation, and credit assignment during social interactions. The current study used fMRI to provide correlational evidence for this. In the future, to understand whether the SMG has a causal role in these processes, causal manipulations such as disrupting SMG activity with transcranial magnetic stimulation would be necessary.

CHAPTER 4 Schizotypy modulates credit assignment and control inference

4.1 Introduction

In chapters 2 and 3, I presented results on the behavioural and neural mechanisms of how healthy participants estimate their control and attribute outcomes to themselves and others. Understanding control and agency estimation may inform our understanding of psychological illness. In this chapter, I present findings on how these processes are altered as a function of participants' depression and schizotypy scores. I focus on depression and schizotypy because they have been linked to attributional biases and misestimations of control in the past.

Individuals with schizophrenia or schizophrenia-like traits show delusions of control (Moritz et al., 2014; Na et al., 2022; Shergill et al., 2014). Psychosis has been linked to difficulties in attributing experiences to oneself or external sources (Lavallé et al., 2021; McGuire et al., 1995; Shergill et al., 2005). In a 15-year longitudinal study, it was found that if patients with schizophrenia believed life events to be under their control (rather than due to external causes such as chance or powerful others), they were less likely to experience depressive symptoms and experienced more periods of recovery (Harrow et al., 2009). It has been proposed that people with schizophrenia attribute action outcomes more to alien forces than to themselves because of a lack of awareness and internal action monitoring (Frith et al., 2000a, 2000b). People with schizophrenia spectrum disorders and persecutory delusions tend to display stronger self-serving biases than healthy individuals (Mueller et al., 2021; Murphy et al., 2018). It has been suggested that such a bias to attribute negative events to other people can, in fact, lead to paranoia and a perception of hostility in one's social environment (Bentall et al., 2001; Bentall et al., 1994). Understanding better how people come to understand their own levels of

agency in a social context may, therefore, shed light on how this estimation of ability, perhaps even self-worth, is altered in schizophrenia and schizotypy.

People with depressive symptoms tend to attribute negative events to themselves (negative attributional biases) (Abramson et al., 1989; Beck, 2008), while nondepressed individuals show the inverse attributional bias (self-serving attributions, attributing positive events to self) (Dorfman et al., 2019; Miller & Ross, 1975; Zamfir & Dayan, 2022). Cognitive theories of depression proposed that aberrant attributional styles are a root cause of depression (Abramson et al., 1989; Beck, 2008), and cognitive therapies of depression target such biases (Msetfi et al., 2016; Msetfi et al., 2007; Norbury et al., 2024; Seligman, 1972). Depressed people estimate their level of control over uncontrollable outcomes more accurately than nondepressed individuals ("depressive realism") (Alloy & Abramson, 1979; Benassi & Mahler, 1985; Martin et al., 1984; Msetfi et al., 2005). It has been found that the activity level of depressed individuals mediates the depressive realism effect, meaning that people with depression tend to be more passive and therefore believe they have little control over outcomes (Blanco et al., 2012). Depression also affects people's social abilities, for example, people with depression are worse at recognising positive faces (Chan et al., 2007; Harmer et al., 2009; Surguladze et al., 2004). Therefore, understanding how individuals on a range of depression symptoms assign credit in social contexts and use (AD) to seek information about ambiguous outcomes may provide a mechanistic understanding of attributional biases and delusions of control.

In this chapter, I investigate how these attributional biases and misestimations of control arise and are maintained in depression and high schizotypy. Participants performed the credit assignment task described in the previous chapter while they were inside the MRI scanner. In the previous two chapters, I focussed on participants who had few depressive symptoms and

here again I analysed data from the same participants but additionally I also included participants with higher depression scores so that the total group investigated was larger (n=47). In addition to using questionnaires to measure depression symptoms I also measured schizotypy scores. In this chapter, I applied the previously developed analysis methods to the behavioural data of this larger sample, which allows me, firstly, to reproduce the behavioural findings from before in this larger sample, and secondly and crucially, to show how schizotypy and depression modulate these behaviours. I compare participants' behaviour to an optimal Bayesian observer model, developed in chapter 2, and examine if low or high scorers in depression and schizotypy behave similarly to an optimal learner. In summary, I found that participants with higher schizotypy scores demonstrated a lower prior belief about their control and overestimated their own performance levels. Individuals with higher schizotypy scores overadjusted for their control level when assigning themselves (but not others) portions of the prediction errors that occurred during joint behaviours. Finally, higher schizotypy was associated with a more optimal control inference based on the participants' self-estimates and the observed difference between AD and normal feedback. No significant relationships between depression and task behaviour were found.

4.2 Methods

4.2.1 Participants

53 participants took part in the study. Two participants were excluded because they misunderstood how to rate their own performance in the games (see chapter 2 methods for details). Two participants were excluded because they performed less than two switches from AD to normal trials or vice-versa across the experiment. One additional participant was excluded due to an incidental finding and one participant was excluded because they could not complete the task within the maximum scanning time. The final sample included 47 MRI participants (n=33 female, aged 18-43 years).

As described in chapter 2, participants were included in the study if they were MRI safe, right-handed, aged 18-45 years, fluent in English and had normal or corrected-to-normal vision. They were excluded if they had current alcohol or drug dependency, current or past history of bipolar or psychosis disorder, or current use of psychoactive medication or recreational drugs. These exclusion criteria were checked both during the initial pre-screening online form and again during the SCID-5-CV in the video call prior to the study visit. Participants were recruited into the study based on their depression scores in the Quick Inventory of Depressive Symptomatology (Self-Report, QIDS-SR) (Rush et al., 2003) which they completed in the pre-screening online form. Specifically, we aimed to include similar participant numbers in three symptom severity groups (QIDS scores: 0-5 none; 6-10 mild, >=11 moderate/severe depressive symptoms), while ensuring that the three groups did not differ in age, gender and education levels. The SCID-5-CV was administered to double-check the psychiatric exclusion criteria. For this, I administered the sections A (mood episodes), B (psychotic and associated symptoms), E (substance use disorders), F (anxiety disorders) and G (obsessive-compulsive disorder and posttraumatic stress disorder). Sample characteristics can be found in Table 4.1.

Table 4.1. MRI sample characteristics

Measure	Count (% of total n=47), or *M*±SD
Age	23.53 ± 4.84
Education level	
1: GCSE or middle school graduation (education up to age 15/16)	1 (2.13%)
2: A-Levels or high school diploma	20 (42.55%)
3: Bachelor's degree	12 (25.53%)
4: Master's degree	13 (27.66%)
5: Doctorate or similarly advanced qualification	1 (2.13%)
Gender	
Female	33 (70.21%)
Male	14 (29.79%)
Other/diverse	0 (0%)
QIDS-SR	7.68 ± 4.91
Schizotypy (unusual experiences subscale)	1.53 ± 2.09

4.2.2 Experimental procedure and design

The experimental procedure for this larger MRI data set analysed here is identical to the smaller data set presented in chapters 2 and 3. In short, MRI participants took part in one study testing session, during which they first provided written informed consent, they were instructed and practiced the task outside the scanner. After a task understanding questionnaire, they completed the main task inside the scanner. The task participants completed is the same as presented in chapter 2. Finally, outside of the scanner, they filled in a task debrief questionnaire and psychiatric questionnaires (see section 4.2.3).

4.2.3 Self-report psychiatric questionnaires

At the end of the study visit, participants completed self-report psychiatric questionnaires, assessing depression and schizotypy using the following:

- Quick Inventory of Depressive Symptomatology (Self-Report, QIDS-SR) (Rush et al., 2003)
- "Unusual Experiences" subscale from the Short Scales for Measuring Schizotypy (schizotypy-UE) (Mason et al., 2005).

In addition, participants filled the following questionnaires or questionnaire subscales which were not analysed in this thesis:

- Symptoms of Depression Questionnaire (SDQ, anger-specific items SDQ-24 and SDQ-25 were included) (Pedrelli et al., 2014)
- Cognitive Styles Questionnaire Very Short Form (CSQ-VSF) (Huys et al., 2016)
- Brief Locus of Control scale (Sapp & Harrod, 1993)
- Learned Helplessness Scale (LHS, Ability-Inability subscale was included) (Quinless & Nelson, 1988)
- Five-Dimensional Curiosity Scale Revised (5DCR, subscales Joyous Exploration, Deprivation Sensitivity and Stress Tolerance were included) (Kashdan et al., 2020)
- Single-Item Self-Esteem Scale (SISE) (Robins et al., 2001)
- Apathy Motivation Index (AMI) (Ang et al., 2017)
- Snaith-Hamilton Pleasure Scale (SHAPS) (Snaith et al., 1995)
- Liebowitz Social Anxiety Scale, self-report version (LSAS-SR, fear/anxiety of interactions subscale was included) (Liebowitz, 1987)
- State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA, only Trait Anxiety subscale was included) (Ree et al., 2008)

- Obsessive-Compulsive Inventory – Revised (OCI-R, subscales Hoarding, Checking, Ordering, Neutralizing, Washing were included, i.e. the Obsessing subscale was excluded) (Foa et al., 2002)

Our main hypotheses for this sample were about symptoms of depression and schizotypy and their effects on credit assignment and control inference. For this, we first computed participants' total scores on the QIDS and schizotypy-UE questionnaires, by summing their individual item responses according to the scoring instructions for both questionnaires. Figure 4.1 shows the distribution of total scores on the questionnaires. Since participants were recruited based on their QIDS scores in the pre-screening online form, the QIDS total scores (from the study visit) were uniformly distributed with equal participant numbers in each depression symptom severity group (n=16 with no depressive symptoms [QIDS 0-5], n=17 with mild depressive symptoms [QIDS 6-10], n=14 with moderate or severe depressive symptoms [QIDS >=11], Figure 4.1A).

Since participants were not recruited based on their schizotypy scores, we found that the distribution of schizotypy-UE scores was right-skewed, i.e. most participants had low scores (Figure 4.1B). To ensure that any effects of schizotypy-UE on behaviour or fMRI were not driven by outliers but related to meaningful differences in schizotypy-UE, we square-root transformed participants' raw schizotypy-UE scores as follows:

$$schizotypy_{UE} = \sqrt{schizotypy_{UE,raw} + 1} \quad (4.1)$$

where the $schizotypy_{UE}$ is the participant's transformed total score on the unusual experiences subscale of the schizotypy questionnaire. $Schizotypy_{UE,raw}$ is the raw total score. As shown in Figure 4.1C, this square-root transformation leads to a more compressed range of schizotypy

values. In the following, we use this square-root transformed schizotypy score (schizotypy-UE) for all analyses.

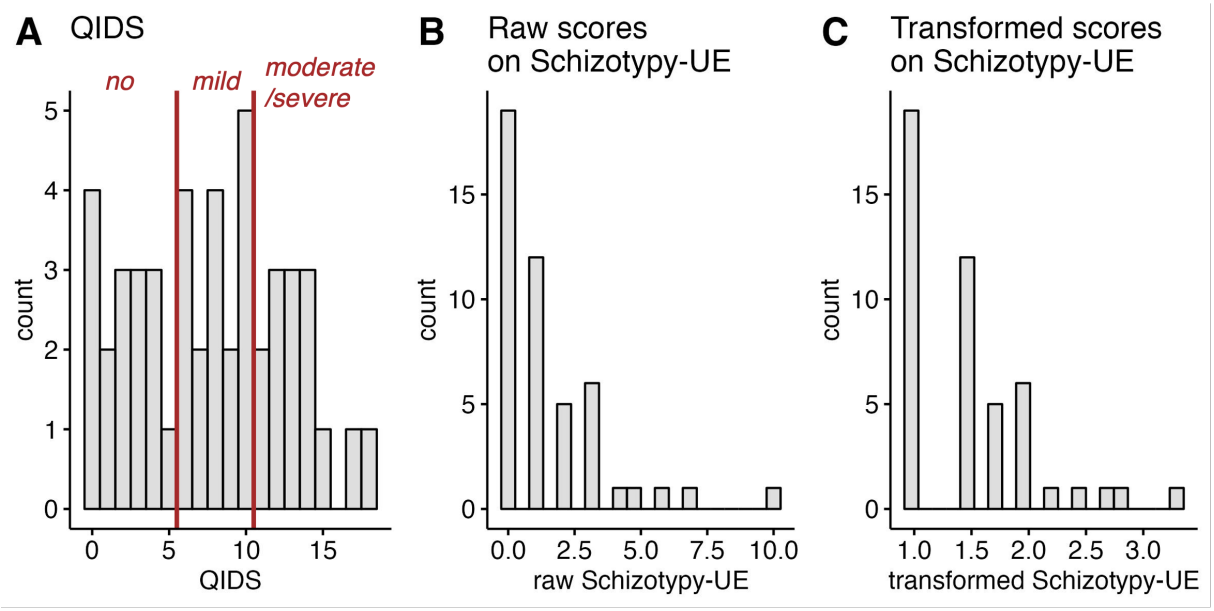


Figure 4.1. Distribution of depression and schizotypy scores in the sample. A) QIDS (depression) scores. Highlighted in red are the three QIDS score groups that were used for recruitment (QIDS total score 0-5: no, 6-10: mild, >=11: moderate/severe depressive symptoms). We aimed to recruit similar sample sizes in each QIDS group (no depressive symptoms: n=16, mild: n=17, moderate/severe: n=14). **B)** Participants' total raw scores on the unusual experiences subscale on the schizotypy questionnaire. Note that because we did not recruit based on schizotypy scores, the raw total schizotypy scores have a skewed distribution with outliers. **C)** Transformed schizotypy-UE scores. Due to the skewed distribution of raw scores (panel B), we square-root transformed them and used the transformed scores for our analyses.

For our analyses, our clinical variables of interest were participants' depression (QIDS) and schizotypy-UE scores. We first assessed how correlated QIDS and schizotypy-UE scores were with each other and with the demographic variables (age, education and gender). Figure 4.2 shows the relationships between these variables. Since education was scored as an ordinal variable (see Table 4.1), we report spearman's rank correlation in panels C and F. Since gender is a binary variable, we report point biserial correlation in panels D and G. We found that QIDS and schizotypy-UE were not significantly correlated with each other (Figure 4.2A), and neither

were they significantly correlated with any of the demographic variables (Figure 4.2B-G). Therefore, in all our subsequent analyses relating behaviour to depression and schizotypy-UE, we only include the two clinical variables of interest and do not control for demographic factors.

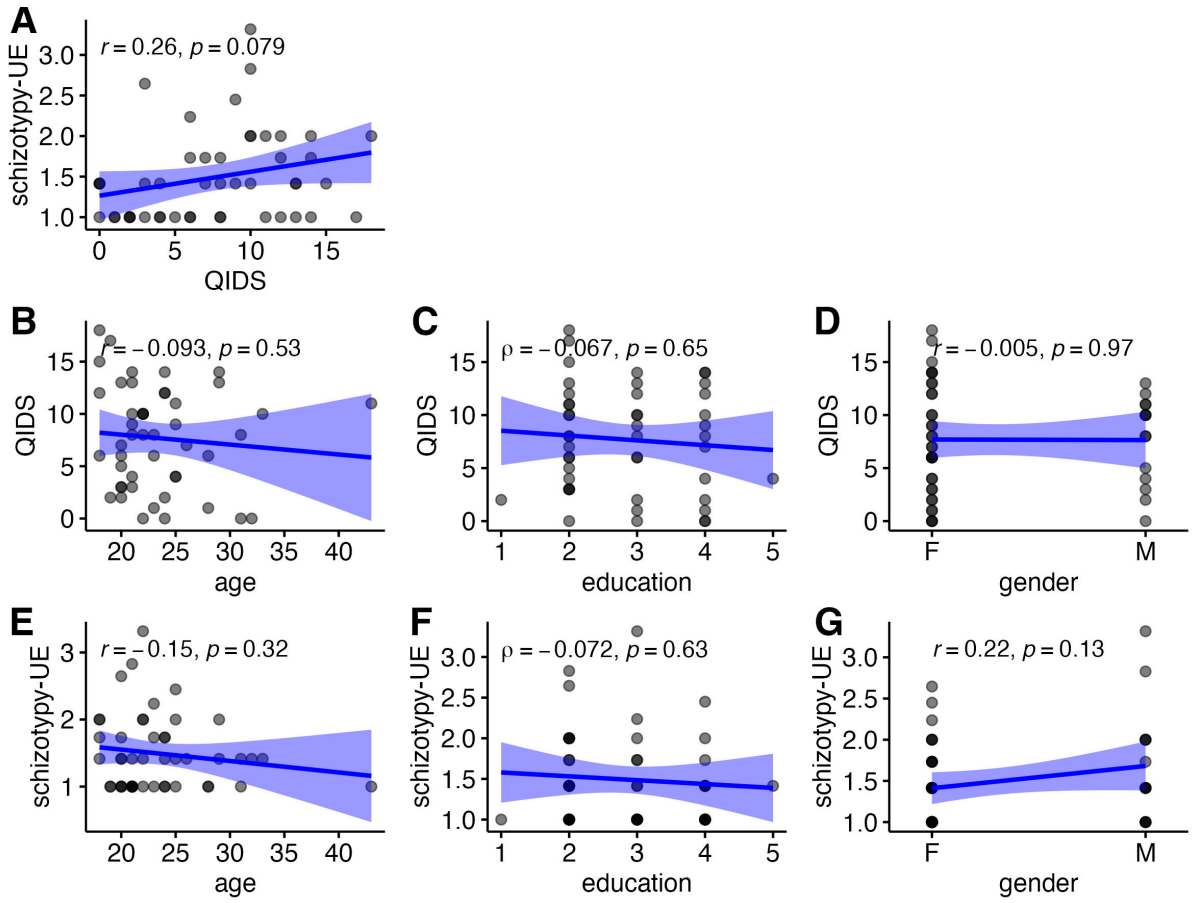


Figure 4.2. Correlation between clinical variables of interest and demographic scores. Relationships between **A)** depression (QIDS) and schizotypy-UE scores, **B-D)** depression and age, education and gender, and **E-G)** schizotypy-UE and age, education and gender. Overall, we do not find significant relationships between our clinical variables of interest (depression and schizotypy-UE) and demographics. Reported p-values were not corrected for multiple comparisons.

4.2.4 Behavioural analyses

Our aim in this study was to test whether there are any behavioural differences between people with different QIDS or schizotypy-UE scores. For this, we used the analyses methods developed in chapter 2 and included interaction terms of the behavioural variables of interest

(e.g. rating accuracy) with our clinical variables of interest (schizotypy and QIDS). We applied Bayesian regression models using the *brms* package with the same parameters and fitting procedure as in chapter 2. Since we did not find significant correlations between our clinical variables of interest and the demographic variables, we did not control for demographics in the following analyses. To evaluate whether the resulting beta estimates were significant, we used the 95% confidence intervals (95% CI) and interpreted an effect as significant if the 95% CI did not include 0. Our analyses linking schizotypy-UE and QIDS to behaviour were exploratory and therefore, we did not correct for multiple comparisons (see discussion at the end of this chapter for further considerations on this point).

We compared participants' behaviour to Bayesian optimal observer models, as in chapter 2. Specifically, we examined how participants with higher depression or schizotypy scores compare to low-scoring participants and to the "Active learner" model. In this chapter, we only used the Active (and not the Passive or Ignorant) learning model because the Active learner made correct assumptions about AD trials and was therefore the most optimal in terms of credit assignment behaviour. Therefore, this allowed us to assess to what extent which participants' behaviour was optimal. Additionally, like in chapter 2, we used the confidence intervals of this model to derive a trial-wise measure of overall uncertainty. As in chapter 2, we simulated data from the Active learning model based on the feedback participants observed. We used the same model fitting procedure as before and full details on the model can be found in chapter 2.

4.2.4.1 General effects of schizotypy and QIDS on participants' ratings

First, we examined any general effects of schizotypy-UE and QIDS on participants' ratings irrespective of any specific hypotheses on credit assignment. For this, we tested whether schizotypy-UE and QIDS had any relationship with how well participants did the task (final trials'

unsigned rating accuracy and signed rating errors) and their prior beliefs (first trial's ratings). The results of these analyses are shown in Figure 4.3. First, we computed participants' accuracies for self, other and control ratings, and estimated feedback respectively in the same manner as in chapter 2, as follows:

$$rating\ accuracy = 1 - \frac{rating\ error}{maximum\ possible\ rating\ error} \quad (4.2)$$

with

$$rating\ error = |rating - true\ level| \quad (4.3)$$

$$maximum\ rating\ error = \max(100 - true\ level, true\ level) \quad (4.4)$$

Note that this gives us a measure of rating accuracy based on the absolute difference between their ratings and the true levels, normalised by the highest rating error possible. We extracted the accuracies of self, other, control and estimated feedback for the final trials of each task phase (self only from Self-Other phases, control only from Control-Other phases). These final rating accuracies were then averaged across the four task blocks, so that we obtained one final rating accuracy per person for each self, other, control and estimated feedback. We tested for interindividual differences in participants' rating accuracies with the following regression models:

$$self\ rating\ accuracy \sim QIDS + Schizotypy_{UE} \quad (4.5)$$

$$other\ rating\ accuracy \sim QIDS + Schizotypy_{UE} \quad (4.6)$$

$$control\ rating\ accuracy \sim QIDS + Schizotypy_{UE} \quad (4.7)$$

$$estimated\ feedback\ accuracy \sim QIDS + Schizotypy_{UE} \quad (4.8)$$

where *QIDS* is the individual score on the QIDS, and $schizotypy_{UE}$ is the individual score on the Schizotypy-UE.

Next, we tested whether schizotypy-UE and QIDS relate to how participants over- or underestimate themselves, others, their control or the feedback at the end of the task phases. For this, we extracted their signed rating errors of self, other, control and estimated feedback as follows:

$$rating\ error_{signed} = rating - true\ level \tag{4.9}$$

where *rating error*$_{signed}$ was computed separately for self, other, control and estimated feedback. *True level* refers to the true level of these variables (as in chapter 2, the observed feedback was used for estimated feedback). As above, these signed rating errors were extracted for the last trial of each task phase (for self and control only Self-Other and Control-Other phase, respectively), and averaged across all task blocks. This again lead to one rating error value per participant for each self, other, control and estimated feedback. We then used the following regression models to test for interindividual differences:

$$self\ rating\ error_{signed} \sim QIDS + Schizotypy_{UE} \tag{4.10}$$

$$other\ rating\ error_{signed} \sim QIDS + Schizotypy_{UE} \tag{4.11}$$

$$control\ rating\ error_{signed} \sim QIDS + Schizotypy_{UE} \tag{4.12}$$

$$estimated\ feedback\ error_{signed} \sim QIDS + Schizotypy_{UE} \tag{4.13}$$

Next, we tested whether participants differed in their initial beliefs about self, other, control and the estimated feedback, that is whether they differed in terms of their priors. For this, we extracted participants ratings on the first trial of each task phase (self only Self-Other phase, control only for Control-Other phase), and averaged those across task blocks. Note that the ratings on the very first trial from a task phase were prior to seeing the feedback (see also task scheme in chapter 2) and the rating questions were therefore rephrased to probe participants'

expectations (e.g. “Going on your gut feeling, how good do you expect to be at this game?”). We fitted the following regression models to examine whether participants differed in their initial expectations:

$$self\ prior \sim QIDS + Schizotypy_{UE} \quad (4.14)$$

$$other\ prior \sim QIDS + Schizotypy_{UE} \quad (4.15)$$

$$control\ prior \sim QIDS + Schizotypy_{UE} \quad (4.16)$$

$$estimated\ feedback\ prior \sim QIDS + Schizotypy_{UE} \quad (4.17)$$

where *self prior*, *other prior, control prior* and *estimated feedback prior* are the participants’ initial ratings on the first phase trials.

4.2.4.2 Effects of schizotypy und QIDS on credit assignment in the Self-Other phase

We hypothesized that participants with different schizotypy-UE and QIDS scores differ in how they assign outcomes to themselves and others. In particular, this could be the case during the Self-Other phase when participants should split up the total prediction error (tPE) into a self- and other-assigned PE (see also chapter 2, Figure 2.6A and E). To test for interindividual differences in this behaviour, we adapted the regression models probing this behaviour (see chapter 2) by including interaction terms of schizotypy-UE and QIDS with our behavioural variables of interest:

$$\begin{gathered} rating\ update \sim \\ (tPE + control + tPE * control) * (QIDS + schizotypy_{UE}) + trial\ number + \\ (tPE + control + tPE * control + trial\ number \mid ID) \end{gathered} \quad (4.18)$$

where *rating update* was the signed difference between participants’ current and previous trial’s ratings. The model was fit separately to participants’ self and other ratings. It was also fit

to the self and other estimates from the Active learning model, however here without the QIDS and schizotypy-UE scores (i.e. same regression model as in chapter 2). *tPE* is the signed total PE experienced on the previous trial and computed as before (see chapter 2, section 2.2.5.2 for details). For the regression models fit to the other ratings from participants and the Active learner, *control* was substituted with the control from the perspective of the other player (1-control). Only normal trials in the Self-Other phase were included. The results of this analysis are shown in Figure 4.4.

4.2.4.3 Effects of schizotypy and QIDS on active disambiguation

We hypothesized that participants with higher scores on schizotypy-UE and QIDS demonstrated a different use of AD in the task. Therefore, in this section, we included two participants that were otherwise excluded because of their low number of trials switching from normal to AD or vice-versa. This means that the data set analysed in this section comprised n=49 participants.

Specifically, we first hypothesized that they might differ in terms of their total number of AD trials that were detected in their behaviour or the number of AD trials they reported in the post-task debrief questionnaire. To test this, the following two regression models were fit to the number of AD trials from their behaviour and self-reports:

$$number\ of\ AD\ trials \sim QIDS + Schizotypy_{UE} \quad (4.19)$$

where *number of AD trials* is either the number of AD trials participants reported to have done in the post-task debrief questionnaire, or the total number of AD trials detected in the behaviour. For the behavioural number of AD trials, Self-Only trials were excluded because on those trials, feedback was not ambiguous and therefore, AD trials did not help disambiguate

the feedback. The regression model was fit separately to the behavioural number and the self-reported number of AD trials. For the regression on the self-reported number of AD trials, only n=44 participants were included into the analysis because five participants did not fill in the debrief questionnaire because it was added at a later stage during data collection.

Secondly, we examined whether participants with higher schizotypy-UE and QIDS scores employed AD differently depending on the current trial's uncertainty. In chapter 2, I presented results in the smaller MRI sample showing that participants use AD under high uncertainty and when useful for credit assignment (Figure 2.7C). For this, I compared the proportion of AD during uncertain phases (Self-Other and Control-Other phases) with task section with certain, unambiguous feedback (Self-Only trials). In this chapter and the present analysis, I tested whether during uncertain phases (Self-Other and Control-Other phases), AD was driven by trial-wise uncertainty and whether this relationship was modulated by schizotypy-UE or QIDS. The following regression model was fit:

$$AD \sim uncertainty * (QIDS + Schizotypy_{UE}) + phase + (uncertainty + phase \mid ID) \quad (4.20)$$

where *AD* denotes whether the current trial was AD (1) or normal (0). Therefore, AD was modelled binomially with the Bernoulli logit link function in brms. *Uncertainty* is the trial-wise overall uncertainty, extracted from the Active learning model. *Phase* denotes whether the current trial is in the Self-Other or Control-Other phase. Only trials in the Self-Other and Control-Other phase are included.

Finally, we hypothesized that participants with higher schizotypy-UE and QIDS scores differed in how AD reduced uncertainty and rating errors. For this, the regression analyses from chapter

2 (Figure 2.7E-F) were adapted to include interaction terms between our behavioural and clinical variables of interest. The following two regression models were fit:

$$uncertainty \sim (AD_{t-1} + AD_{t-2}) * (QIDS + schizotypy_{UE}) + trial\ number + phase + (AD_{t-1} + AD_{t-2} + trial\ number + phase \mid ID) \quad (4.21)$$

$$summed\ rating\ errors \sim (AD_{t-1} + AD_{t-2}) * (QIDS + schizotypy_{UE}) + trial\ number + phase + (AD_{t-1} + AD_{t-2} + trial\ number + phase \mid ID) \quad (4.22)$$

where *uncertainty* is the overall uncertainty (averaged across self, other and control uncertainties) from the Active learning model. *Summed rating errors* are the participants' absolute rating errors, summed across self, other and control ratings. AD_{t-1} and AD_{t-2} denote whether the current trial was one or two trials after an AD trial. *Trial number* is time in task phase. The results of these analyses are shown in Figure 4.5.

4.2.4.4 Effects of schizotypy and QIDS on controllability inference from AD trials

We hypothesized that participants with higher schizotypy-UE and QIDS scores infer their level of control differently from the feedback difference between normal and AD trials, while accounting for their overall self-performance. To test this, we used our analysis approach presented in chapter 2 (Figure 2.8) and adapted the regression model to include interaction terms between our behavioural variables of interest (feedback difference, inverse self, feedback difference * inverse self) and our clinical variables of interest (schizotypy-UE and QIDS). The following regression model was fit:

$$control \sim (feedback\ difference * inverse\ self) * (QIDS + schizotypy_{UE}) + previous\ control + (feedback\ difference * inverse\ self + previous\ control \mid ID) \quad (4.23)$$

where *control* is the next trial's control rating of participants, or the control estimate of the Active learning model. The regression model was separately fit to participants' and the Active learners' control estimates to compare participants' behaviour to that of an optimal learner. This comparison of participants' behaviour to the optimal Active learner model was useful because it could reveal whether participants, or a subgroup with low or high schizotypy scores for example, performed better or worse than an optimal learner would. *Feedback difference* is the difference between normal and AD feedback, as observed on the current and last trials. *Inverse self* is the inverse of participants' or the Active learner's self estimate of the current trial. *Previous control* is the current trial's control estimate of participants or the Active learner. Only switch trials (from AD to normal trials or vice-versa) during the Control-Other phase were included (similar to chapter 2).

4.3 Results

4.3.1 Schizotypy modulates participants' rating accuracy and prior beliefs

We first tested whether there are any interindividual differences in terms of how participants generally did their ratings in the task. We found that participants with higher schizotypy-UE scores rated their own performance less accurately at the end of task phases (self accuracy: β=-0.74, 95% CI=(-0.74,-0.17); Figure 4.3A), but did not differ in their accuracies of control, other or the total estimated feedback (control accuracy: β=0.09, 95% CI=(-0.23,0.4); other accuracy: β=0.07, 95% CI=(-0.24,0.38); estimated feedback accuracy: β=-0.19, 95% CI=(-0.50,0.11); Figure 4.3A). An analysis of the signed rating errors revealed that the lower self-rating accuracy in schizotypy-UE was driven by participants overestimating their own performance (signed rating errors of self: β=0.36, 95% CI=(0.06,0.66); Figure 4.3B), with no significant differences in the other signed rating errors (signed rating errors: control, β=0.01, 95% CI=(-0.3,0.32); other, β=-0.08, 95% CI=(-0.39,0.23); estimated feedback, β=0.15, 95% CI=(-0.17,0.46); Figure 4.3B). These effects probed participants' beliefs at the end of the task phases, after observing feedback. Next, we examined whether participants showed interindividual differences in their initial beliefs, or prior expectations, about themselves, others, their control or the feedback they expect. We found that before observing the feedback, participants with higher schizotypy-UE scores showed a lower prior for their own control (prior control rating: β=-0.41, 95% CI=(-0.70,-0.13); Figure 4.3C). In other words, prior to observing the feedback, they expected to have less control than participants scoring higher on schizotypy-UE. We also found that participants with higher schizotypy-UE scores expected to see higher feedback than participants with lower schizotypy-UE scores (prior estimated feedback: β=0.37, 95% CI=(0.08,0.67); Figure 4.3C). We found no effects of depression (QIDS) score on participants' final rating accuracies, signed rating errors or initial prior beliefs.



Figure 4.3. General effects of QIDS and schizotypy-UE on participants' ratings. A) While depression score (QIDS) did not modulate how accurately participants did their ratings, a higher schizotypy-UE score was associated with lower self-rating accuracy at the end of the task phases. **B)** We found that this lower self-rating accuracy in people with higher schizotypy-UE scores was driven by them overestimating their own performance (more positive signed self-rating errors). **C)** At the beginning of task phases, participants with higher schizotypy-UE scores had lower priors about their control, i.e. rated their control lower than participants with lower schizotypy-UE scores. We also found that they expected higher feedback, that is a higher schizotypy-UE score was associated with a higher prior estimated feedback.

4.3.2 Schizotypy increases credit assignment to self but not others

We hypothesized that schizotypy and depression modulate how people assign credit to themselves and others. In particular, we hypothesized that people with higher schizotypy or QIDS scores split up tPEs differently to themselves and others depending on their control. We tested this by adapting the analysis methods developed in chapter 2 (Figure 2.6). This allowed us firstly, to test how participants assigned tPEs in general (on a group level), and secondly, whether there are any interindividual differences related to schizotypy-UE or QIDS.

Firstly, we investigated how participants assigned tPEs to themselves and others in general. In the smaller MRI sample presented in chapter 2, we found that participants assigned themselves and others prediction errors proportional to their level of control. Here, in this larger MRI sample, we found that participants showed the same behavioural pattern on a group level: They updated their beliefs about themselves and others depending on the overall tPE (tPE effect on

self updates: β=0.48, 95% CI=(0.41,0.56), Figure 4.4A; tPE effect on other updates: β=0.38, 95% CI=(0.33,0.43), Figure 4.4B). Additionally, they assigned even more of the tPE to themselves or the other according to who had more control (interaction effect tPE * control on self-rating update: β=0.18, 95% CI=(0.13,0.24), Figure 4.4A; interaction effect tPE * other-control on other rating update: β=0.23, 95% CI=(0.03,0.29), Figure 4.4B). These behavioural effects are again consistent with the simulations of the optimal Active learning model (black dots in Figure 4.4A-B), suggesting that participants, like an optimal observer, attribute social PEs according to their control estimates.

Next, we examined the relationship between assigning tPEs and our clinical variables of interest. Strikingly, we found that compared to low schizotypy-UE scorers, participants with higher schizotypy-UE scores assigned themselves, but not others, more of the tPE according to their control. This is demonstrated by a significant 3-way interaction effect of tPE, control and schizotypy-UE on self but not other rating updates (interaction effect tPE * control * schizotypy-UE on self-rating updates: β=0.07, 95% CI=(0.02,0.12), Figure 4.4A; interaction effect tPE * other-control * schizotypy-UE on other rating updates: β=0.05, 95% CI=(-0.01,0.11), Figure 4.4B). Figure 4.4C shows the significant interaction effect on self-rating updates. Assuming a low control, participants with high schizotypy-UE scores did not assign the tPE to themselves (flat light green line in Figure 4.4C left). In contrast, low scoring schizotypy-UE participants assigned themselves at least a small amount of the tPE (positively sloped dark green line in Figure 4.4C left). Note that low control here refers to task phases when participants had 20% control over the feedback. Therefore, it is still appropriate to assign some of the tPE (even a small amount) to self, even if the control is low. This is also evident from the optimal Active learning model which assigned the tPE to self even for a low control (dashed black line in Figure 4.4C left), similar to the low schizotypy-UE participants. For medium control (when control was equally shared between self and other, i.e. 50%), participants did not show

any differences in credit assignment according to schizotypy-UE scores (green lines overlap in Figure 4.4C middle). If participants had a high level of control (80%), they should attribute most of the tPE to themselves, similar to the Active learning model (positively sloped black line in Figure 4.4C right). We found that participants did so (both light and dark green lines in Figure 4.4C right are positively sloped). However, compared to the optimal Active learning model, low schizotypy-UE participants assigned slightly too little of the tPE to themselves (dark green line is less positively sloped than the dashed black line), and high schizotypy-UE participants more accurately assigned themselves portions of the tPE (light green line is closer to the dashed black line than the dark green line, Figure 4.4C right). We found no significant effects of QIDS on how participants assigned themselves or others credit in terms of tPEs and their control (interaction tPE * control * QIDS on self updates: β=0.01, 95% CI=(-0.10,0.08), Figure 4.4A; interaction tPE * other-control * QIDS on other updates: β=0.00, 95% CI=(-0.10,0.10), Figure 4.4B).

Taken together, participants, in general, in this larger MRI sample assigned credit to themselves and others according to the tPE while taking their controllability estimate into account. However, there were individual differences in exactly how they did so; strikingly, we found interindividual differences in this behaviour which were related to schizotypy-UE and not QIDS. Participants with higher schizotypy-UE scores updated their beliefs about themselves, and not others, more based on the tPE and their controllability estimate. They appear to over-adjust for the control level they believe they have when learning about themselves. As a consequence, they assigned themselves too little credit, or even no credit for outcomes when they believed having low control. In contrast, when their perceived control was high, participants with higher schizotypy-UE scores more accurately assigned themselves credit according to their control than low-scoring individuals did.

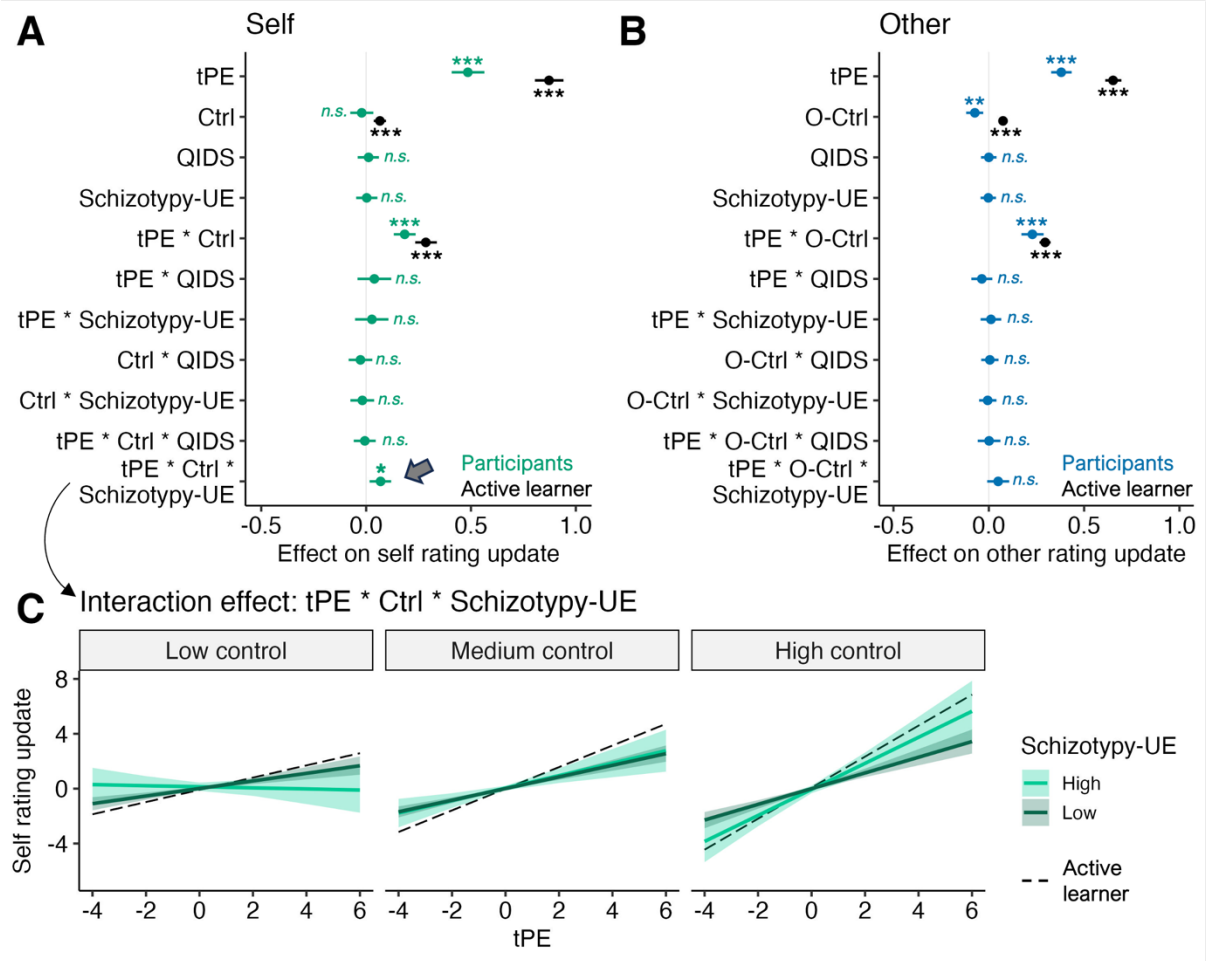


Figure 4.4. Effects of schizotypy and depression on credit assignment behaviour. When they know their control level in the Self-Other phase, participants should assign credit (tPEs) to themselves and others according to the control level they have (see also Figure 2.6 in chapter 2). **A)** We indeed found that in this larger MRI sample, participants updated their ratings about themselves more positively if they encountered a positive total PE (effect of tPE on self-rating updates), and assigned themselves more of the tPE if they had a high control over it (interaction tPE * Ctrl). Participants with higher schizotypy-UE scores, however, in contrast to those with low schizotypy-UE scores interpreted their own performance levels in a manner that reflected an increased influence of control (interaction tPE * Ctrl * Schizotypy-UE). **B)** As expected, we found that in this larger MRI sample, participants changed their ratings about the other based on the tPE (effect of tPE on other rating updates), and assigned more of the tPE to the other if the other had more control (interaction tPE * O-Ctrl). The other player's level of control is denoted as O-Ctrl, which is 1-Ctrl. We found no interindividual differences related to either QIDS or schizotypy-UE in how participants learnt about the other player. **C)** Interaction effect between tPE, control and schizotypy-UE on self-rating updates. Schizotypy-UE was related to the degree to which participants assigned themselves portions of the tPE according to their control. Compared with low-scoring participants, participants with high schizotypy-UE scores assigned themselves either less credit (when they had low control, left panel) or more credit (when they have a high level of control, right panel). Compared to the optimal Active learner, high schizotypy-UE participants assigned themselves too little credit when their control was low (light green line is flatter than the dashed black line, left panel).

However, when their control was high, high schizotypy-UE participants assigned themselves credit more optimally than low scoring individuals (light green line is closer to dashed black line than dark green line, right panel).

4.3.3 No effects of Schizotypy and QIDS on active disambiguation

In chapter 2, we found that in the smaller MRI sample, participants made deliberate mistakes (AD) to assign credit better. Here, in this larger MRI sample, we hypothesized that participants with higher schizotypy-UE and QIDS scores use AD differently to low-scoring participants. We tested whether participants differ in their use of AD in our task using a set of four analyses. First, we tested whether the total number of AD trials that participants did in the task or the number they self-report in the post-task debrief questionnaire depends on schizotypy-UE or QIDS scores. For example, people with higher depression scores might do less AD trials because they might not want to do even worse than they already think they are doing, and let the other player down. Contrary to our hypotheses, we found that this is not the case and that participants did not differ in terms of their total number of AD trials in their behaviour or their self-report (behavioural number of AD: schizotypy-UE, β=0.07, 95% CI=(-0.24,0.39); QIDS, β=0.01, 95% CI=(-0.30,0.33); self-reported number of AD: schizotypy-UE, β=0.01, 95% CI=(-0.32,0.36); QIDS, β=0.05, 95% CI=(-0.29,0.38); Figure 4.5A).

Next, we examined whether uncertainty drives AD differently in people with higher scores in schizotypy-UE or the QIDS. For example, people with higher scores on our clinical variables of interest might be less sensitive to the current uncertainty and do fewer uncertainty-driven AD trials. It could also be that people scoring higher on our clinical variables of interest are more sensitive to uncertainty and do more uncertainty-driven AD trials. We tested whether this is the case using a regression analysis predicting how trial-wise uncertainty drives AD within the uncertain phases (Self-Other and Control-Other phases). We found that QIDS and schizotypy-UE did not change how participants did AD depending on the current uncertainty (interaction

uncertainty * QIDS, β=0.04, 95% CI=(-0.04,0.12); interaction uncertainty * schizotypy-UE, β=0.04, 95% CI=(-0.03,0.12); Figure 4.5B). Somewhat surprisingly, we do not find that trial-wise uncertainty drives AD on a group-level (uncertainty, β=-0.01, 95% CI=(-0.10,0.07); Figure 4.5B). In chapter 2, we presented results showing that AD is driven by high uncertainty, by comparing the proportion of AD between task phases when feedback was uncertain (Self-Only trials) to phases when feedback was certain (Self-Other and Control-Other phases). In the present analysis, however, we took a different approach and analysed how within those uncertain phases, trial-wise uncertainty (derived from the Active learning model) drives AD. Therefore, we tested for a more subtle effect of trial-wise uncertainty on AD. This could explain why we do not find the expected uncertainty effects on a group-level in this analysis.

We hypothesized that people show interindividual differences in how they take advantage of AD trials. For this, we firstly tested whether uncertainty is reduced differently after AD depending on the individual schizotypy-UE or QIDS score. We found that generally, uncertainty is reduced after AD (similar finding as reported in chapter 2, Figure 2.7E; *1 trial after AD* effect on uncertainty, β=-0.50, 95% CI=(-0.58,-0.44); *2 trials after AD* effect on uncertainty, β=-0.50, 95% CI=(-0.57,-0.43); Figure 4.5C). This uncertainty effect was not modulated by QIDS or schizotypy-UE (interaction effects on uncertainty: 1 trial after AD * QIDS, β=0.02, 95% CI=(-0.05,0.08); 1 trial after AD * schizotypy-UE, β=0.00, 95% CI=(-0.07,0.07); 2 trials after AD * QIDS, β=0.01, 95% CI=(-0.05,0.08); 2 trials after AD * schizotypy-UE, β=-0.01, 95% CI=(-0.08,0.05); Figure 4.5C). Note that uncertainty was not self-reported by participants but estimated from the Active learning model. Since the dependent variable is not from participants' behaviour, this could explain why we did not find interindividual differences in how AD modulates uncertainty.

Finally, we tested whether participants with higher schizotypy-UE or QIDS scores assigned credit better or worse after AD, in terms of their overall rating errors. We did not find that this is the case (interaction effects on rating error: 1 trial after AD * QIDS, β=0.01, 95% CI=(-0.06,0.07); 1 trial after AD * schizotypy-UE, β=-0.03, 95% CI=(-0.10,0.04); 2 trials after AD * QIDS, β=-0.02, 95% CI=(-0.09,0.04); 2 trials after AD * schizotypy-UE, β=0.00, 95% CI=(-0.07,0.08); Figure 4.5D). In chapter 2, we found that participants assigned credit more accurately in the trials immediately following AD (Figure 2.7F). In this larger MRI sample, we found that this group-level effect held only on the second trial and not the first trial following AD (effects on rating error: 1 trial after AD, β=-0.05, 95% CI=(-0.11,0.02); 2 trials after AD, β=-0.08, 95% CI=(-0.15,-0.02); Figure 4.5D). Note that while not reaching significance, the effect on the trial immediately after AD (1 trial after AD) is in the expected direction (reduced rating error after AD).

Taken together, schizotypy-UE and QIDS scores did not modulate how participants used AD. Participants did not differ in the number of AD trials in their behaviour or post-task self-reports, nor did they differ in their sensitivity to uncertainty, or how AD leads reduced uncertainty and lead to better credit assignment.

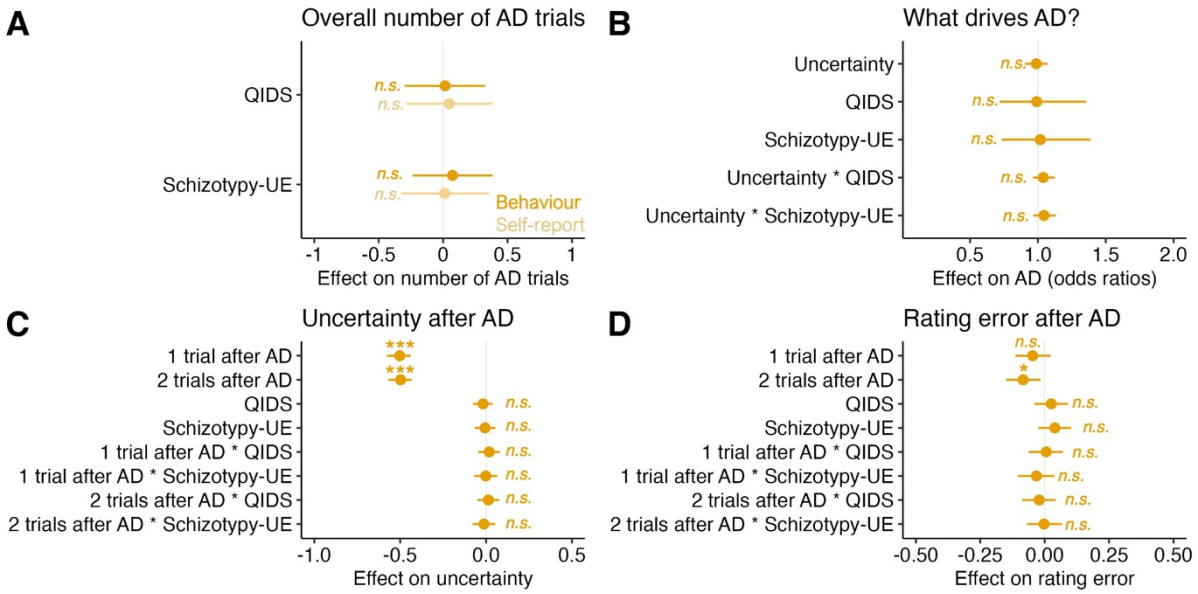


Figure 4.5. No effect of schizotypy-UE and QIDS on the use of AD. We found no relationships between QIDS or schizotypy-UE and **A)** how many AD trials participants did in the task or how many they reported in the post-task debrief questionnaire, or **B)** how trial-wise uncertainty drove AD. Schizotypy-UE and QIDS did not modulate **C)** how uncertainty was reduced after AD, or **D)** how rating errors were reduced after AD.

4.3.4 Schizotypy increases controllability inference from AD

In chapter 2, we showed that participants infer their control over feedback by purposefully varying their behaviour, i.e. doing their best (normal trial) and making deliberate mistakes (AD trial). They then estimated their control from the difference in feedback between normal and AD trials, while taking into account their overall self-performance (Figure 2.8). We hypothesized that this behavioural effect might be modulated by participants' schizotypy-UE or QIDS scores. Past research has shown that misestimations of control can be found in depression (Alloy & Abramson, 1979; Benassi & Mahler, 1985), schizotypy (Asai, 2016; Moore et al., 2021) and schizophrenia (Frith et al., 2000b; Synofzik et al., 2010; Voss et al., 2010). Therefore, here, we hypothesized that participants with higher schizotypy-UE and QIDS scores over- or underestimate their control after varying their behaviour.

To test this, we used our analysis approach from chapter 2 and included interaction terms between our behavioural variables of interest (feedback difference and inverse self) and our clinical variables of interest (QIDS and schizotypy-UE). We found that as in chapter 2, on a group level, participants estimated their control based on the feedback difference between normal and AD trials, while accounting for their own performance (interaction Feedback difference * Inverse Self on participants' control ratings: β=0.19, 95% CI=(0.12,0.27); Figure 4.6A). Again, this behavioural was consistent with simulated data from the optimal Active learning model (interaction Feedback difference * Inverse Self on the Active learner's control estimates: β=0.22, 95% CI=(0.20,0.25); Figure 4.6A), suggesting this is how participants should infer their control. Strikingly, we found that schizotypy-UE increases how high of a control participants infer based on the feedback difference and their own performance (3-way interaction effect Feedback difference * Inverse Self * Schizotypy-UE on participants' control ratings: β=0.10, 95% CI=(0.03,0.17); Figure 4.6A). Figure 4.6B illustrates this 3-way interaction effect together with the estimates from the optimal Active learning model. Comparing low and high schizotypy-scoring participants to the Active learning model revealed that high schizotypy-UE participants inferred their control more optimally (i.e. more similarly to the Active learner) than low schizotypy-UE scorers. This was especially the case when participants overall believed they had a low or medium self-performance in the games. In those cases, participants with higher schizotypy-UE scores overall took the feedback difference more into account than low scorers (light pink line [high schizotypy-UE] is steeper and closer to the dashed black line [Active learner] than the dark pink line [low schizotypy-UE], Figure 4.6B left and middle panels). This suggests that participants with higher schizotypy-UE scores were more sensitive to the difference making AD did to the feedback, and inferred their control appropriately. In scenarios when participants assumed a high self-performance in the games, they did not differ in terms of their schizotypy-UE scores and took the feedback difference less into account, as is appropriate (Figure 4.6B right). We also observed that the interaction between feedback

difference and schizotypy-UE, and the main effect of schizotypy-UE on participants' control ratings almost reached significance (interaction Feedback difference * schizotypy-UE, β=0.09, 95% CI=(0.00,0.19); schizotypy-UE, β=0.06, 95% CI=(-0.00,0.12); Figure 4.6A). However, since these effects were not significant, we will not interpret them further. We found no effects of QIDS on how participants estimate their control (Figure 4.6A).

Taken together, in this larger MRI sample, participants inferred their control based on the difference varying their behaviour made to the feedback, taking into account their normal overall self-performance. We found that this overall behavioural effect was modulated by schizotypy-UE. In particular, participants with higher scores in schizotypy-UE were more sensitive to the feedback difference and their own performance when inferring their control. This behaviour was more similar to the optimal Active learning model than the behaviour of the low schizotypy-UE scoring participants was.

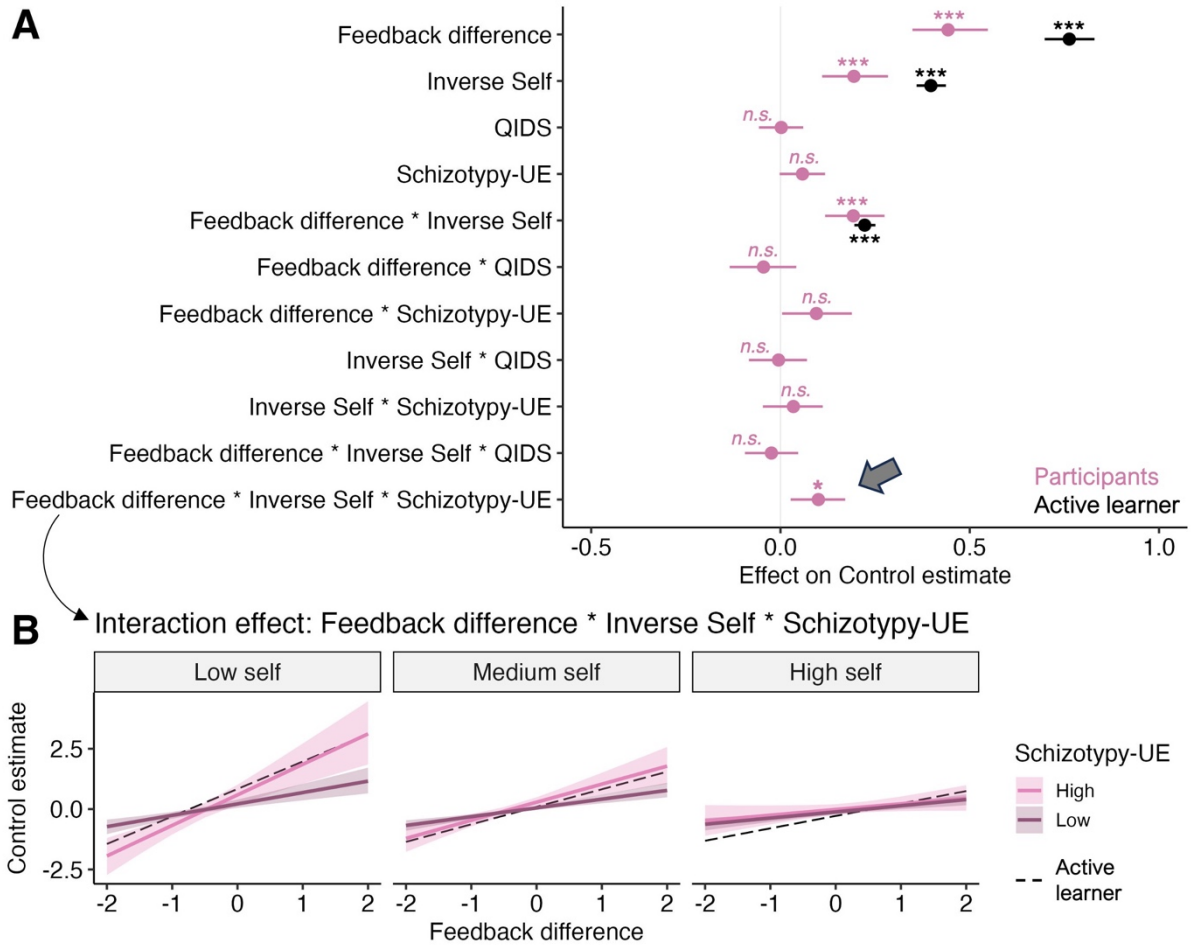


Figure 4.6. Schizotypy-UE modulates controllability inference from AD and normal trials. A) Participants adjusted their control beliefs based on the feedback difference observed between AD and normal trials, taking into account their own performance (pink group level effects of Feedback difference, Inverse Self, Feedback difference * Inverse Self). This behaviour is consistent with simulated data from the Active learner (black). These effects are similar to the behavioural effects we found in the smaller MRI sample in chapter 2. Strikingly, we found that people with higher schizotypy-UE scores base their control beliefs more on the Feedback difference and Inverse self (3-way interaction effect Feedback difference * Inverse Self * Schizotypy-UE, see grey arrow). We did not find any significant effects related to QIDS. **B)** 3-way interaction between Feedback difference, Inverse Self and Schizotypy-UE. Assuming a low or medium own performance (self), participants with higher schizotypy-UE scores infer a higher control from the feedback difference than participants with low schizotypy-UE scores. If participants assume a high own performance (high self), participants do not differ depending on their schizotypy-UE scores. Simulated data from the Active learner model is plotted in dashed black lines. Interestingly, participants with high schizotypy-UE scores show behaviour more similar to the Active learner model than participants with low schizotypy scores (i.e. the light pink line is closer than the dark pink line to the black dashed line). For easier interpretation, the effect of Inverse Self in B has been flipped and relabelled to Self, i.e. for example, 'low self' is the relabelled 'high inverse self'.

4.4 Discussion

In this chapter, I presented behavioural results on how depression and schizotypy relate to interindividual differences in how people estimate their own and others' performances and their level of control in my task. I found that participants with higher schizotypy scores had lower initial beliefs about their performance and control (but not others) than participants with lower schizotypy scores. Participants with higher schizotypy assigned themselves portions of the prediction errors but they overadjusted for their level of control. They gave themselves too little credit when they had a low level of control, but credited themselves more accurately when they had a high level of control. Participants with higher schizotypy scores took the observed feedback difference between AD and normal trials together with their own performance estimates more into account to infer their control. Interestingly, this resembled behaviour from an optimal Bayesian model more closely than was the case for participants with lower schizotypy scores.

Overall, we found effects of schizotypy only on learning about the self and about control, and not about the other players. This fits with previous work on control delusions (Moritz et al., 2014; Na et al., 2022; Shergill et al., 2014), and self-serving biases in schizophrenia which might protect patients from low self-esteem (Bentall et al., 2001; Bentall et al., 1994). These self-serving biases describe tendencies to make unequal attributions of positive events to oneself and negative events to external causes. While it is promising that schizotypy modulates self- and control-related learning in my task independent of the outcome valence, it would be interesting to examine whether schizotypy relates to biased credit assignment after positive versus negative outcomes in future data analyses.

Moreover, we found that higher scores in schizotypy are associated with participants relying more on external cues in terms of a) the prediction errors and their known control for learning

about their own performance, and b) the feedback difference and their learned self-estimate for learning about their level of control. This fits with previous studies demonstrating that healthy people with high schizotypy scores and patients suffering from schizophrenia rely more on external cues to derive their sense of agency (Moore et al., 2021; Synofzik et al., 2010; Voss et al., 2017).

We did not find that either depression or schizotypy modulates how participants employ AD in our task. Previous work has, however, suggested that information seeking and exploration behaviours are altered in schizophrenia (Waltz et al., 2020) and depression (Smith et al., 2022). In cognitive behavioural therapies for depression, "behavioural experiments" are activities aimed at challenging the patients' beliefs and constructing new views about themselves, others, and the world (Bennett-Levy et al., 2004). In the studies presented in this thesis, participants received a hint during the task instructions that they could change their behaviour in the games. We made this suggestion in order to ensure that there was a reasonable baseline level of AD to make data analysis possible. While this means that AD was still a spontaneous behaviour that participants chose to perform as when they found it useful, we found that every MRI participant performed AD. In the future, it would be of benefit to collect a task version without this hint in a large online sample. I hypothesize that under such conditions participants will exhibit greater variation in the degree to which they use AD and it will be possible to examine whether this is associated with depression or schizotypy.

While we designed the task in order to examine how attributional biases in depression emerge on a mechanistic level, we did not find that depression modulates credit assignment or control learning in our task. There are multiple possibilities for why this might be the case. Firstly, it might be that our task is not sensitive to the cognitive biases in depression. In clinical research, Public and Patient Involvement (PPI) describes the process of engaging patients or the public

into the research process, with the aim to improve the quality, relevance and impact of research (National Institute for Health Research, 2014). While I did not take this approach when designing my study, in the future, it would be interesting to discuss the experiences of patients with depression with my task in order to develop new and possibly more sensitive analyses. Accordingly, a second potential reason might be that my current analyses are not sensitive to the cognitive biases in depression. Negative attributional bias in depression describe a tendency to attribute negative events more to oneself, while positive events are attributed to external causes such as chance or other people (Abramson et al., 1989). In the analyses presented in chapter 4, I did not compare credit assignment after positive or negative outcomes. In the future, it would be interesting to assess whether, even in the current task, depression modulates how people assign themselves and others positive compared to negative prediction errors, or how depression might be related to different control perceptions after good or bad feedback.

The analyses presented in this chapter are exploratory and did not control for multiple comparisons. A test of whether these results are replicable will be possible in an independent sample such as a large online study with the same task and questionnaires. This is a currently ongoing project (see also general discussion, chapter 6).

In sum, in this chapter I explored how depression and schizotypy influence how individuals estimate their own and others' performances and control. Participants with higher schizotypy scores took their perceived control more into account when assigning themselves portions of prediction errors. Participants with higher schizotypy benefited more from doing AD for inferring their control; they based their control beliefs more on the difference between AD and normal feedback and their self-estimates, resembling an optimal Bayesian model more closely than participants with lower schizotypy scores.

CHAPTER 5 Striatal and SMG responses to ambiguous outcome processing in schizotypy and depression

5.1 Introduction

In the previous chapter, I presented behavioural results from a large MRI sample and examined how people with different depression and schizotypy scores attributed outcomes differently to themselves, others and their level of control. I found that schizotypy was related to different outcome attributions to self and control, but not other people. In this chapter, I draw on the neural analysis methods developed in chapter 3 to examine the neural correlates of schizotypy effects on behaviour.

Delusions of control in schizophrenia have been linked to the supramarginal gyrus (SMG). Compared with healthy participants and patients with schizophrenia but without control delusions, increased activation of the SMG has been found in patients with schizophrenia who experience delusions of control, and this increased activation vanishes as delusion of control symptoms subside (Spence et al., 1997). It has been proposed that delusions of control and hallucinations in schizophrenia arise due to a faulty internal action monitoring process, which results in a larger prediction error between expected and experienced action outcomes, ultimately leading to a distorted sense of agency (Frith et al., 2000b). SMG activity signals the prediction error between predicted and experienced sensory feedback in the healthy brain (Farrer et al., 2003). The SMG is involved in attributing actions to oneself or others (Farrer & Frith, 2002; Ohata et al., 2020; Sirigu et al., 1999). More generally, the SMG has been related to attention to movements (Rushworth et al., 2001a; Rushworth et al., 1997; Rushworth et al., 2001b), tool use (Orban & Caruana, 2014; Peeters et al., 2009; Peeters et al., 2013), and

representing actions in terms of their intentions and meaning (Desmurget et al., 2009; Króliczak & Frey, 2009).

Other lines of research have highlighted the roles of dopamine and the striatum in schizophrenia. During reward anticipation, ventral striatal activity is reduced in patients with schizophrenia and their first-degree relatives (Esslinger et al., 2012; Grimm et al., 2014; Juckel et al., 2006; Nielsen et al., 2012). Reward prediction error signals are reduced in the ventral striatum in patients with schizophrenia (Schlagenhauf et al., 2014). In sum, both the SMG and striatum are candidate regions that might be sensitive to different outcome attributions and control learning in schizotypy.

Depression has been linked to weaker signalling of reward prediction errors in the striatum (Dombrovski et al., 2013; Gradin et al., 2011; Kumar et al., 2008), and this effect becomes stronger in MDD patients with a higher number of depressive episodes (Kumar et al., 2018). Meta-analyses across neuroimaging studies have found that depression is associated with reduced striatal signalling of rewarding outcomes and increased orbitofrontal and anterior cingulate activation during reward processing (Ng et al., 2019; Zhang et al., 2013). In addition, the inferior parietal lobe shows stronger activation during the processing of monetary rewards in MDD (Zhang et al., 2013), and the SMG in patients with MDD signals belief updating to bad news more strongly than in healthy individuals (Garrett et al., 2014). Overall, this indicates that the striatum and inferior parietal lobe play distinct roles in learning process, which might possibly extend to control and causal learning.

In this chapter, I investigate the neural correlates of the interindividual differences identified in the previous chapter. For this, I analysed the neural data acquired from the large MRI sample (n=47) presented in the previous chapter. On a group level, I found overlapping whole-brain

effects of uncertainty and active disambiguation (AD) in the SMG, similar to the results presented in chapter 3. Since in chapter 4, I found that schizotypy-UE (but not depression) affected credit assignment and control learning, I expected stronger neural effects relating to schizotypy-UE than to depression. Indeed, whole-brain corrected activity in the putamen was sensitive to AD in schizotypy-UE. Putamen activity appeared to encode the estimated control at the time of outcome more strongly in participants with higher schizotypy scores. Activity in the ventral striatum (nucleus accumbens) related to trial-wise changes of the observed feedback seemed to be increased in schizotypy.

5.2 Methods

5.2.1 fMRI whole brain analysis

As in chapter 3, we carried out first-level analysis using FSL FEAT (version 6.00). Our aim in this chapter was to find the interindividual differences related to QIDS and schizotypy-UE scores in the fMRI data. Therefore, in this chapter, we analysed the fMRI data of the same participants as in chapter 4 (n=47). First, we set up the same first level analysis with the same regressors for the whole brain analysis of this large MRI sample as in chapter 3 (full details in section 3.2.2). Correlations between regressors are shown in Figure 5.1. To test for interindividual differences, in the higher level analysis we included the demeaned QIDS and schizotypy-UE scores as covariates in addition to the standard group average regressor. Results were obtained with automatic outlier-deweighting, and using FSL FLAME1 with cluster-correction threshold of $z > 3.1$ and $p < 0.05$ (Worsley, 2001).

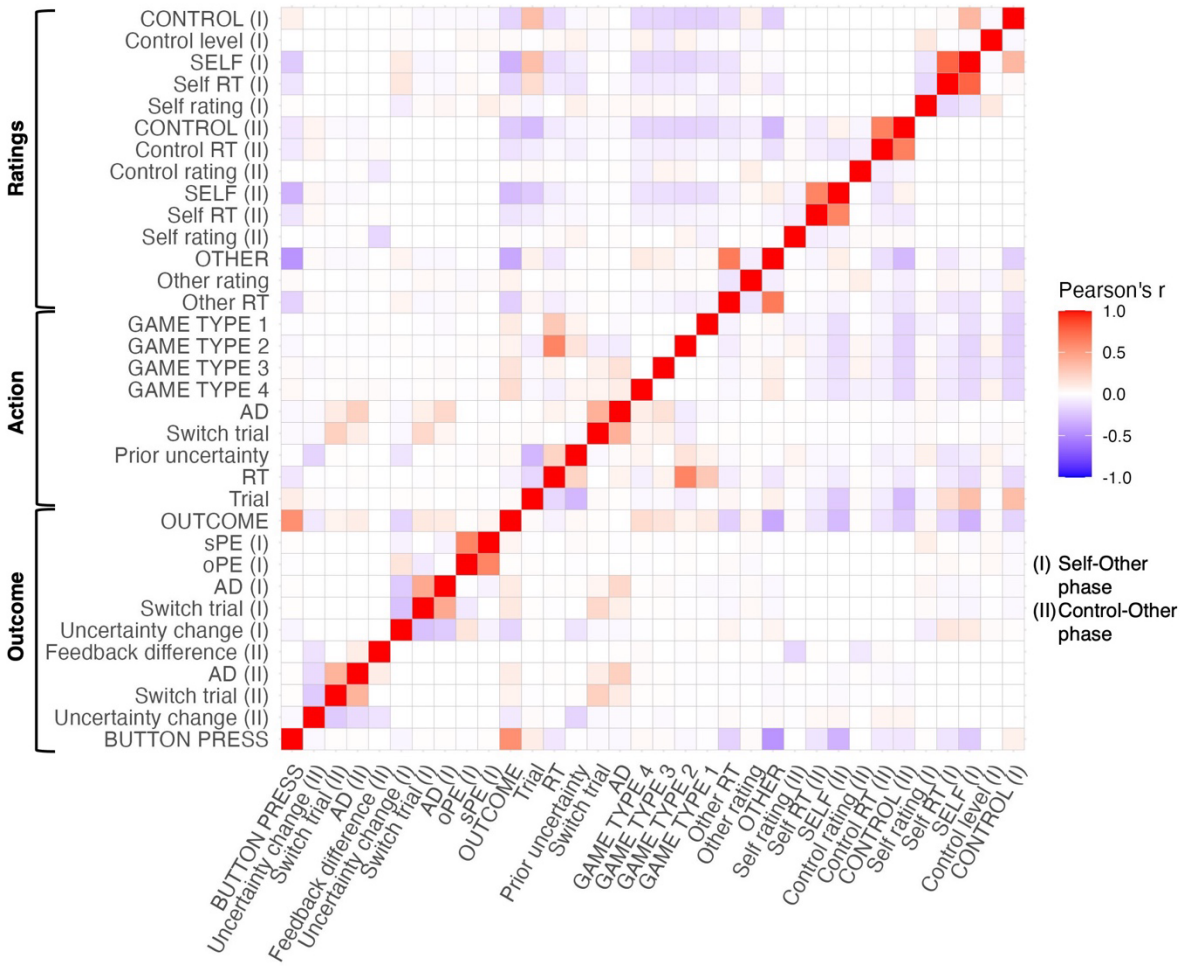


Figure 5.1. Correlations of convolved regressors included in the fMRI whole brain analysis. Here, we averaged Pearson's correlations across participants. Uppercase indicates constant regressors. See section 3.2.2 in chapter 3 for full details.

5.2.2 fMRI time course analysis on ROIs

5.2.2.1 ROI selection and time series extraction

We used the same method for constructing ROIs and extracting the time series functional data as in chapter 3. Spherical ROIs with a diameter of three voxels were centred on the peak voxels of significant activation clusters from the whole brain analysis. As in chapter 3, we selected an ROI in the left SMG (MNI: x=-60, y=-30, z=44) based on where, in the large MRI sample, the group average clusters of uncertainty overlapped with AD at the time of action. In the whole-brain analysis, we found a deactivation cluster in the left putamen for how schizotypy-UE

modulated AD-related activity at the time of outcome in the Control-Other phase. To ensure that the spherical ROI in the left putamen included mostly grey matter, we centred it two voxels more medially (MNI: x=-30, y=-4, z=4).

5.2.2.2 Time course GLMs (ROI-GLMs)

In chapter 3, we found that the SMG was involved in processing uncertainty and AD during action generation, as well as credit assignment and control inference during outcome phase. In this chapter, we hypothesized that interindividual differences related to schizotypy-UE and QIDS might be present in the BOLD activity of the SMG. Furthermore, we hypothesized that the left putamen, which showed schizotypy-UE differences in AD processing at the time of outcome, might also show schizotypy-UE or QIDS effects related to credit assignment and control inference. Note that the time course analyses were independent of the whole brain analyses which defined the ROIs. This is because firstly, the SMG was identified in an analysis of activity that was time-locked to action but, by contrast, the time course analyses focused on the later outcome phase. Secondly, the left putamen was identified in an analysis of schizotypy-UE effects on AD at the outcome during the Control-Other phase. All regressors of interest in the time course analyses were orthogonal to AD (see Figure 5.2 for regressor correlations with AD), and therefore the following time course analyses were independent of the manner in which the putamen ROI was selected.

Our overall hypothesis was that schizotypy-UE and QIDS modulate how the putamen and SMG track credit assignment and control inference during the outcome phase. Since we found behavioural differences in credit assignment and control inferences related to schizotypy-UE but not to QIDS, we expected to observe stronger effects of schizotypy-UE in the neural data. However, for completeness, we included both schizotypy-UE and QIDS into our analyses. To test for interindividual differences, we first applied the ROI-GLMs presented in chapter 3 to the

large MRI sample (n=47). As in chapter 3, all of the following ROI-GLMs were run on the outcome phase because this is when participants learn and can assign credit for an outcome to one actor or the other. The ROI-GLMs are summarised in the following and full details on the ROI-GLMs can be found in chapter 3. In the second step, we then tested for interindividual differences (see next section 5.2.2.3).

First, we tested whether the ROIs represent what control beliefs participants held and how much they updated their control beliefs at the time of outcome. The following ROI-GLM 1 was fit (identical to ROI-GLM 1 in chapter 3), time-locked to the outcome of the current trial:

$$BOLD \sim control + control\ update \quad (5.1)$$

where *control* is the participant's control rating from the next trial and *control update* is the signed difference between the next and current trials' control ratings. As in chapter 3, all trials in the Control-Other phase were included.

Next, we tested whether the ROI activity correlates with the components necessary for control inference, the feedback difference and the inverse of participants' ratings of their own performance (self ratings). The following ROI-GLM 2 was fit (identical to ROI-GLM 2 in chapter 3), time-locked to the outcome phase:

$$BOLD \sim feedback\ diff_{AD\ to\ normal} + feedback\ diff_{normal\ to\ AD} + inverse\ self + AD \quad (5.2)$$

where *feedback diff*$_{AD\ to\ normal}$ and *feedback diff*$_{normal\ to\ AD}$ are the feedback difference between normal and AD trials, either on switch trials when the current trial is AD and the previous is normal or vice-versa. In chapter 3, the motivation for defining the feedback difference as

separate regressors for these two different trial types was to investigate two competing hypotheses for coding of feedback difference using a single GLM. In chapter 3, we found that feedback difference is represented as normal vs AD feedback, as is optimal for control inference. Therefore, in this chapter, we averaged across the resulting beta estimates from the two feedback difference regressors (*feedback diff*$_{AD\ to\ normal}$ and *feedback diff*$_{normal\ to\ AD}$), in order to obtain the overall effect of feedback difference, coded as normal – AD. *Inverse self* is the inverse of participants' self ratings of the current trial (before observing the feedback), as a measure for participants' self prior. As in chapter 3, *AD* was included as a regressor of no interest. Again, only switch trials in the Control-Other phase were included. Switch trials were trials when participants performed AD after a normal trial or vice-versa, and are therefore the trials when the feedback difference between these two trial types could be observed.

Next, we hypothesized that ROI activity tracks how participants learn about themselves and others by splitting up the tPE by their known control. This splitting up of the tPE resulted in a self-assigned PE (sPE) and other-assigned PE (oPE). To test this, we ran two time course analyses (ROI-GLMs 3 and 4, identical to ROI-GLMs 3 and 4 in chapter 3). As in chapter 3, we used two ROI-GLMs to test for sPE and oPE separately due to their high correlation (r=0.54):

$$BOLD \sim sPE \tag{5.3}$$

$$BOLD \sim oPE \tag{5.4}$$

In this chapter, we did not run the additional analysis with both sPE and oPE in one model (ROI-GLM 5 in chapter 3) because this was a control analysis in chapter 3. In the present chapter, we focussed on the main neural analysis of chapter 3 and related them to interindividual differences in the large MRI sample. We tested for an alternative way of how the brain might

represent splitting up tPEs by the control at the time of outcome, using the following ROI-GLM 5 (identical to ROI-GLM 6 in chapter 3):

$$BOLD \sim tPE + control \quad (5.5)$$

where *tPE* is the total prediction error and *control* is the true control level (as displayed to participants on the rating screen). As in chapter 3, only normal trials during the Self-Other phase were included in ROI-GLMs 3, 4 and 5.

Finally, we examined whether activity in the ventral striatum correlates with feedback and how much the feedback changed in our task. For this, as in chapter 3, we used an anatomically defined bilateral mask of the nucleus accumbens (Khalighinejad et al., 2021). The following time course analysis was fit (ROI-GLM 6, identical to ROI-GLM 7 in chapter 3):

$$BOLD \sim feedback + feedback\ change \quad (5.6)$$

where *feedback* is the feedback observed on the current trial, and *feedback change* is how much the current feedback changed compared to the previous trial. As in chapter 3, to test for AD-independent effects, we included only normal trials following normal trials from both Self-Other and Control-Other phases.

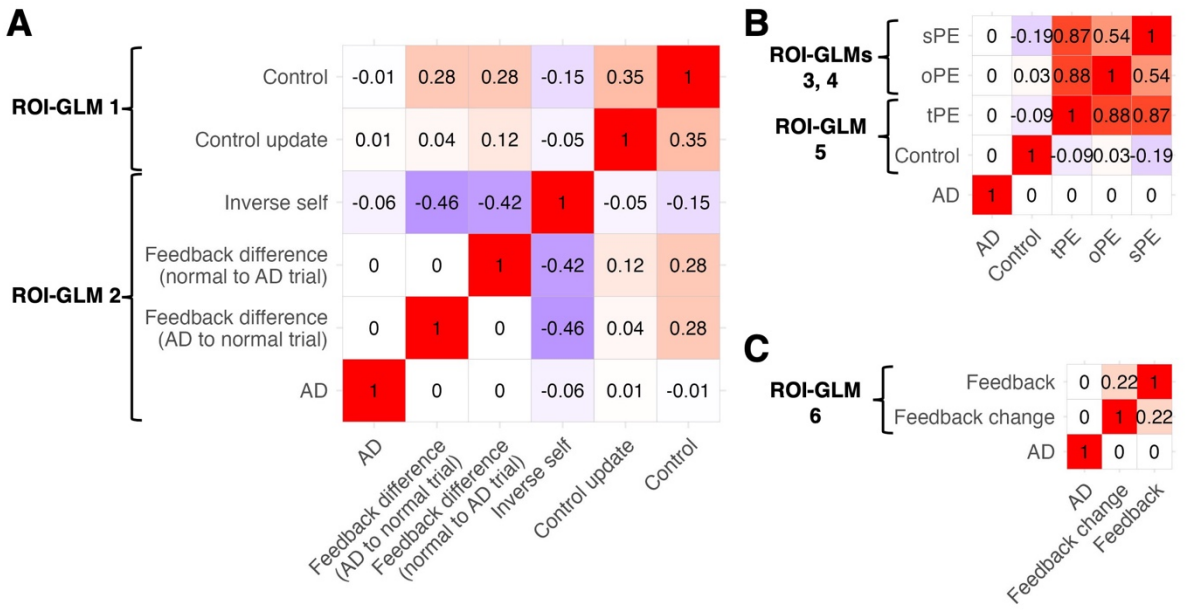


Figure 5.2. Correlation matrices for the regressors included in the time course analyses (ROI-GLMs) of the large MRI sample (n=47). A) As in the smaller MRI sample presented in chapter 3, in the large MRI sample analysed in this chapter, we found low correlations (r<0.3) between the regressors of control belief (ROI-GLM 1) and feedback differences (ROI-GLM 2). **B)** Since sPE and oPE were correlated with r=0.54, we tested their time courses separately (ROI-GLM 3 and 4). The effects of control and tPE were separately examined in ROI-GLM 5. AD was not included as a regressor into these ROI-GLMs but is shown to illustrate that the regressors of interest are orthogonal to AD in the Self-Other phase. Note that our regressors of interest (sPE, oPE, tPE and control) were only defined on normal trials and were therefore, by definition, orthogonal to AD. **C)** Correlations of regressors included in ROI-GLM 6. AD again is orthogonal to the feedback and feedback change. Note that in this panel C, feedback and feedback change are only defined on normal trials following normal trials, which is why AD is unrelated to these two regressors.

5.2.2.3 LOO procedure and interindividual differences analysis

The ROI-GLMs described in the previous section resulted in beta estimates of our regressors of interest (e.g. feedback difference) for each participant and time point in the outcome phase. In a second step, we used two statistical procedures to a) test for the overall effects of the regressors across participants, and to b) test the resulting beta estimates for interindividual differences related to schizotypy-UE and QIDS. To test for group-level effects across participants, we used the same Leave-one-out (LOO) procedure as in chapter 3.

Our main goal in this chapter was to examine interindividual differences related to schizotypy-UE and QIDS in SMG and putamen activity. We tested the effect of schizotypy-UE and QIDS on the resulting beta estimates for each regressor of interest of each ROI-GLM (described before, in section 5.2.2.2). The regressors included in the following regression analyses were normalised to a mean of 0 and standard deviation of 1. Using the *lm* function from the *stats* package (v. 4.2.1) in R, we applied the following regression model to the beta estimates resulting from the ROI-GLMs:

$$\beta_{ROI-GLM,t} \sim QIDS + schizotypy_{UE} \quad (5.7)$$

where *QIDS* and $schizotypy_{UE}$ are the participants' individual QIDS and schizotypy-UE scores. $\beta_{ROI\text{-}GLM,t}$ is the beta estimate of a regressor of interest from a ROI-GLM at time point *t*. For instance, this was the beta estimate of *control* on SMG activity at 5 sec after outcome onset, from ROI-GLM 1. This regression model was fit to the beta estimates of each time point (from -2 to 15sec around outcome onset) for all ROI-GLMs across participants. Therefore, this regression analysis resulted in a time course of how much, for example, schizotypy-UE modulated the representation of control in the SMG. The advantage of this analysis approach was that it allowed us to test effects of schizotypy-UE and QIDS on the time courses across the entire time window around outcome onset (from -2 sec to 15 sec). However, it also resulted in a high number of statistical tests and did not control for multiple comparisons or autocorrelations in the BOLD time courses. This is an important drawback which is further discussed in the discussion.

5.3 Results

5.3.1 Supramarginal Gyrus tracks uncertainty and AD while schizotypy-UE modulates putamen activity of AD

We analysed the neural data of participants to identify interindividual differences in how the brain represents AD and AD-guided control estimation. For this, we employed the same whole-brain and time course analyses as presented in chapter 3, but applied them to the large MRI sample (n=47) while including terms for schizotypy-UE and QIDS scores.

On a group level, we found a set of regions processing uncertainty and AD at the time of action, similar to our results in the smaller MRI sample presented in chapter 3. In this larger MRI sample, we again found that the left SMG was sensitive to both uncertainty and AD (Figure 5.3A-B left). Activity in the dmPFC and left lOFC tracked higher certainty (lower uncertainty, Figure 5.3B). We again found widespread activation patterns related to AD, including, bilateral SMG, bilateral dlPFC, extending into dACC and dmPFC, and bilateral lOFC extending into anterior insula (Figure 5.3B). In the outcome phase when participants observe the feedback, we found that bilateral putamen is sensitive to AD (Figure 5.3Ci). These effects were present across participants and not related to any interindividual differences. We found no effects of schizotypy-UE and QIDS on the neural processing of uncertainty and AD at the time of action. Instead, we found that at the time of outcome during the Control-Other phase, schizotypy-UE modulated neural activity in the left putamen related to AD (Figure 5.3C).

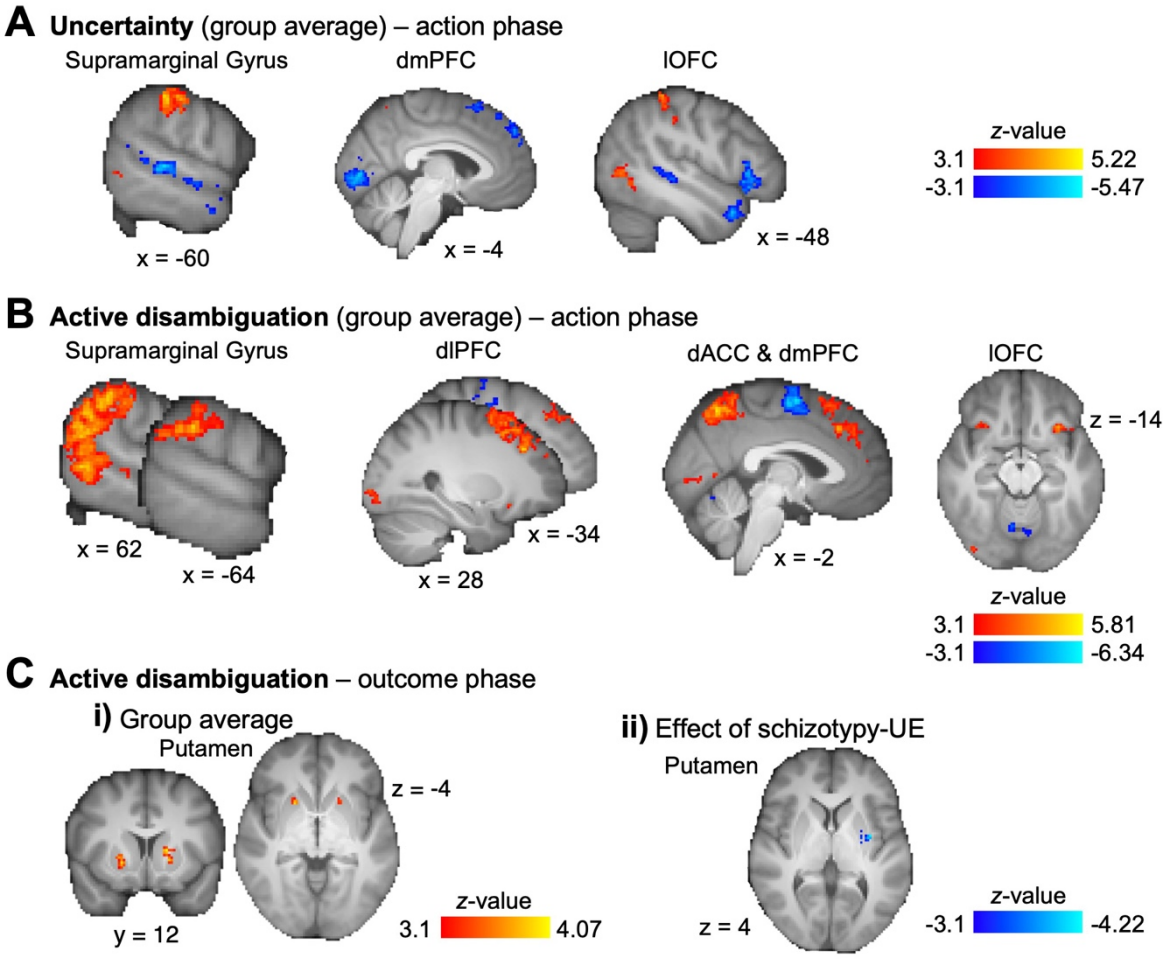


Figure 5.3. Whole-brain results of uncertainty and AD in the large MRI sample. A) In this larger MRI sample, we identified the same regions as in chapter 3 that carry signals of uncertainty at the time of action. Activity in the left SMG was associated with higher uncertainty, while dmPFC and lOFC activity related to lower uncertainty (higher certainty). **B)** We found similar regions as in chapter 3 that reflect AD at the time of action generation. Bilateral activation of SMG reflected AD. In addition, AD-related activity was found in bilateral dlPFC, extending into the dACC and adjacent dmPFC, and bilateral lOFC (extending into anterior insula). The results shown in panels A and B reproduce the uncertainty- and AD-related neural findings from the smaller MRI sample (n=31, chapter 3, Figure 3.3) in the large MRI sample (n=47). **Ci)** At the time of outcome, we found that AD is represented bilaterally in the putamen. **ii)** While we did not find that at the whole-brain level, the action-locked effects shown in panels A and B vary with schizotypy-UE or QIDS, we found effects of schizotypy-UE on neural activity related to AD at the time of outcome. We identified a deactivation cluster in the left putamen related to how participants' individual schizotypy-UE score modulates AD-related neural activity. We found no effect of QIDS on AD-related neural activity at the time of outcome. Panels A, B and Ci show group average effects across Self-Other and Control-Other phase. Panel Cii shows schizotypy-UE effects on neural activity which were present only in the Control-Other phase. Whole-brain effects are cluster corrected with $z>3.1$ and $p<0.05$.

5.3.2 Effects of schizotypy-UE and QIDS on the neural correlates of control inference

Our behavioural analyses in chapter 4 showed that participants with higher schizotypy-UE scores estimate their level of control more based on the difference between normal and AD feedback while taking into account their self-performance (Figure 4.6 in chapter 4). Here, we examined the neural correlates of this inference process during the outcome phase, and tested for schizotypy-UE effects. For completeness and consistency with the behavioural analysis, we also tested for QIDS effects on neural activity, although we did not find effects of QIDS on behaviour. We focussed our analyses on two ROIs (Figure 5.4A); the SMG which tracked uncertainty and AD during action generation, and the putamen which demonstrated schizotypy-UE modulated AD-activity at the outcome. We applied a set of time course analyses established in chapter 3 to the neural data of the large MRI sample, and then in a second step, examined the effects of schizotypy-UE and QIDS on the time course activity in SMG and putamen. Note that all of the following analyses are independent of the analyses that lead to the definition of SMG and putamen. This is because a) the following analyses focussed on the outcome phase and the SMG was defined based on action-related effects, and b) putamen was identified based on an AD-effect at the outcome and only regressors orthogonal to AD were examined in the following time course analyses.

First, we tested if on a group level, putamen and SMG activity represent the inferred control and control update at the time of outcome. As in the previous chapters, we focussed our analysis on the Control-Other phase when participants inferred their control from the feedback. We found that in this large MRI sample, neither putamen nor SMG were sensitive to the inferred control or control update on a group level (*control*: putamen, $t(46)=1.28$, $p=0.21$; SMG, $t(46)=0.48$, $p=0.64$, Figure 5.4Bi; *control update*: putamen, $t(46)=0.68$, $p=0.50$; SMG, $t(46)=0.74$, $p=0.46$, Figure 5.4Ci). This is somewhat surprising given that in the neural results

in chapter 3, we found that the SMG represented control and control update at the time of outcome. We discuss possible reasons for the non-significant group level result in this larger MRI sample in the discussion.

Next, we investigated if these control- and control update-related time courses were modulated by either schizotypy-UE or QIDS. We found that schizotypy-UE increased the neural representation of control in the putamen (peak of schizotypy-UE effect on control representation in the putamen: β=0.49, p<0.001, 8.10 sec after outcome onset, Figure 5.4Bii left). In other words, putamen activity appeared to be positively associated with control in participants with higher schizotypy-UE scores. This effect was present from 7 to 10 sec after outcome onset, corresponding with the time at which outcome-related hemodynamic changes would be expected to peak. We also found that schizotypy-UE had a negative correlation with control update representations in the putamen (peak of schizotypy-UE effect on control update in putamen: β=-0.35, p=0.02, 10.22 sec after outcome onset, Figure 5.4Cii left). However, this effect was less temporally extended and weaker than the effect of schizotypy-UE on control. We also found effects of schizotypy-UE and QIDS on control representations in the SMG (Figure 5.4Bii and iii right). However, they were relatively late (more than 10 sec after outcome onset) and may therefore not be outcome-related anymore.

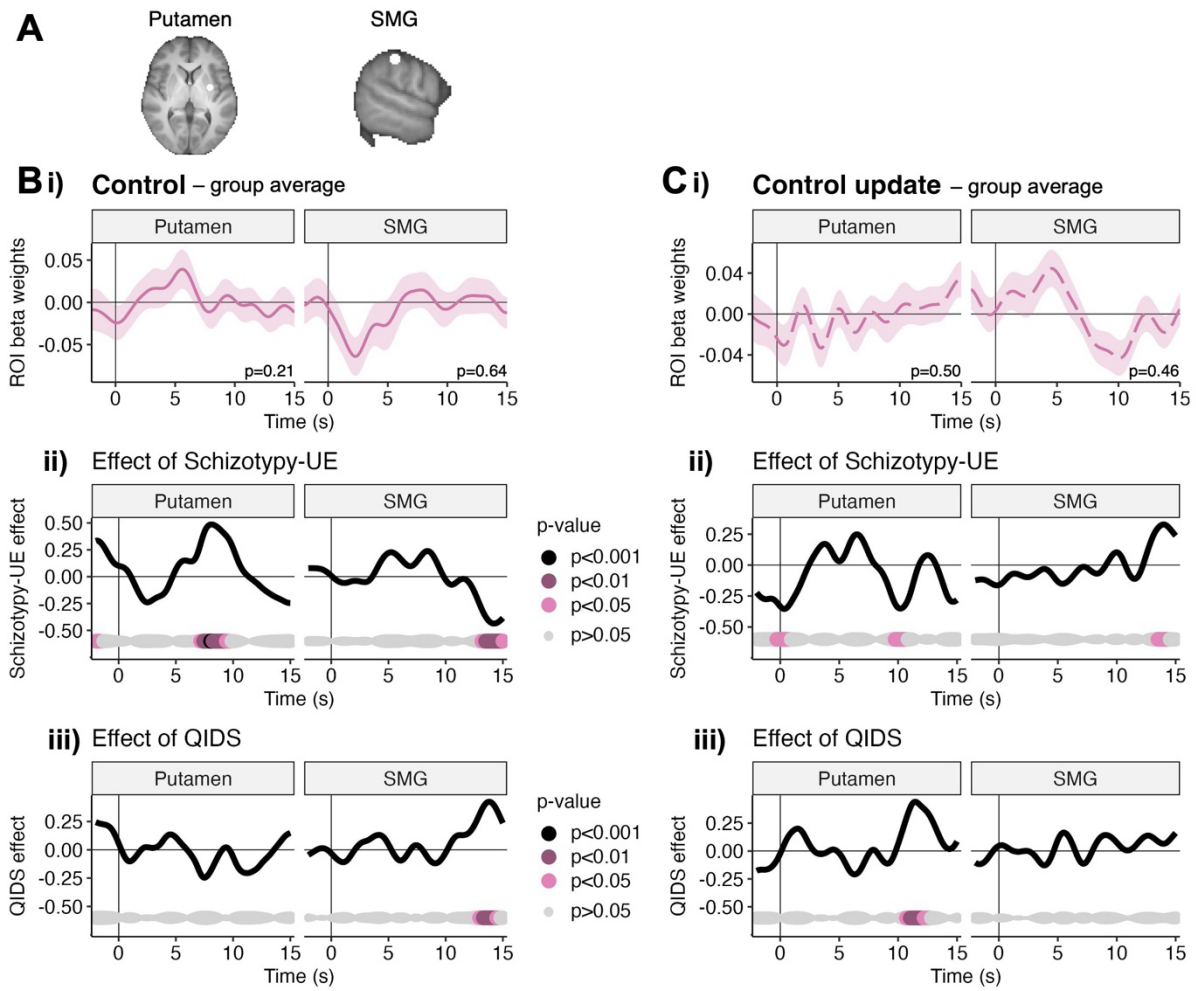


Figure 5.4. Time courses of control and control update in the SMG and putamen. A) Based on our whole-brain results, we selected two ROIs: the SMG, which we had identified in chapter 3, and the putamen which was identified in the whole brain analysis reported earlier in the current chapter. **Bi)** On a group level, putamen and SMG were not sensitive to the inferred control at the time of outcome. **ii)** We found that higher schizotypy-UE lead to a stronger control representation in the putamen (peak effect at 8sec). We found that schizotypy-UE modulated how sensitive the SMG was to control. However, this effect requires further investigation because it is relatively late (more than 10sec after outcome onset), and may therefore be related to the rating phase of the next trial. The solid black line indicated the effect of schizotypy-UE on the control time courses of putamen and SMG. The horizontal grey/violet line indicates the p-values of the schizotypy-UE betas on every time point, with bigger size and darker colours signalling lower p-values. **iii)** We found no effect of QIDS on control-related putamen activity (left), but a very late effect on SMG activity (right). Again, we will not interpret this effect further because it is very late (more than 10sec after outcome onset). **Ci)** We found no representations of the control update in putamen and SMG. While the LOO procedure identified neither of the ROI time courses as significant, we observed that the overall waveform of the control update in the SMG resembles the significant SMG time course of control update in chapter 3 (Figure 3.4). **ii)** We found that schizotypy-UE appeared to modulate how putamen and SMG activity was related to the control update. However, these effects appear less convincing because they were relatively short. **iii)** We found that QIDS increased

control update-related activity in the putamen. This effect was relatively late (more than 10sec after outcome onset), and requires further investigation. Time courses are time locked to the outcome onset and include trials in the Control-Other phase.

Next, we tested whether putamen and SMG represented the components necessary for control inference, and whether schizotypy-UE affected these representations. On a group level, we found that only the SMG represented the difference between normal and AD feedback (feedback difference: SMG, $t(46)=2.78$, $p=0.008$; putamen, $t(46)=-1.59$, $p=0.12$; Figure 5.5Ai), reproducing the SMG effect we found in chapter 3. We also found group level representations of the other component necessary to infer controllability, participants' self-performance ('inverse self'). SMG again tracked the inverse self at baseline, as in chapter 3, but in this large MRI sample, the later negative peak was not significant (inverse self in SMG: -2 to 2 sec time window: $t(46)=2.22$, $p=0.03$; 4 to 10 sec time window: $t(46)=-0.19$, $p=0.85$; Figure 5.5Bi). Putamen also represented the inverse self on a group level ($t(46)=-2.73$, $p=0.009$; Figure 5.5Bi). While we found effects of schizotypy-UE and QIDS on SMG and putamen representations of the inverse self, they were relatively late and short (Figure 5.5Bii-iii).

Taken together, we found that in this large MRI sample, neither putamen nor SMG represented participants' control estimate or how much they updated their control belief in the outcome phase. This SMG result is somewhat surprising given our results in the smaller MRI sample in chapter 3. Interestingly, we found that schizotypy-UE modulated whether the putamen represented participants' inferred control. SMG tracked the components necessary for control inference, the feedback difference and self-performance. These effects did not show strong modulations by either schizotypy-UE or QIDS.

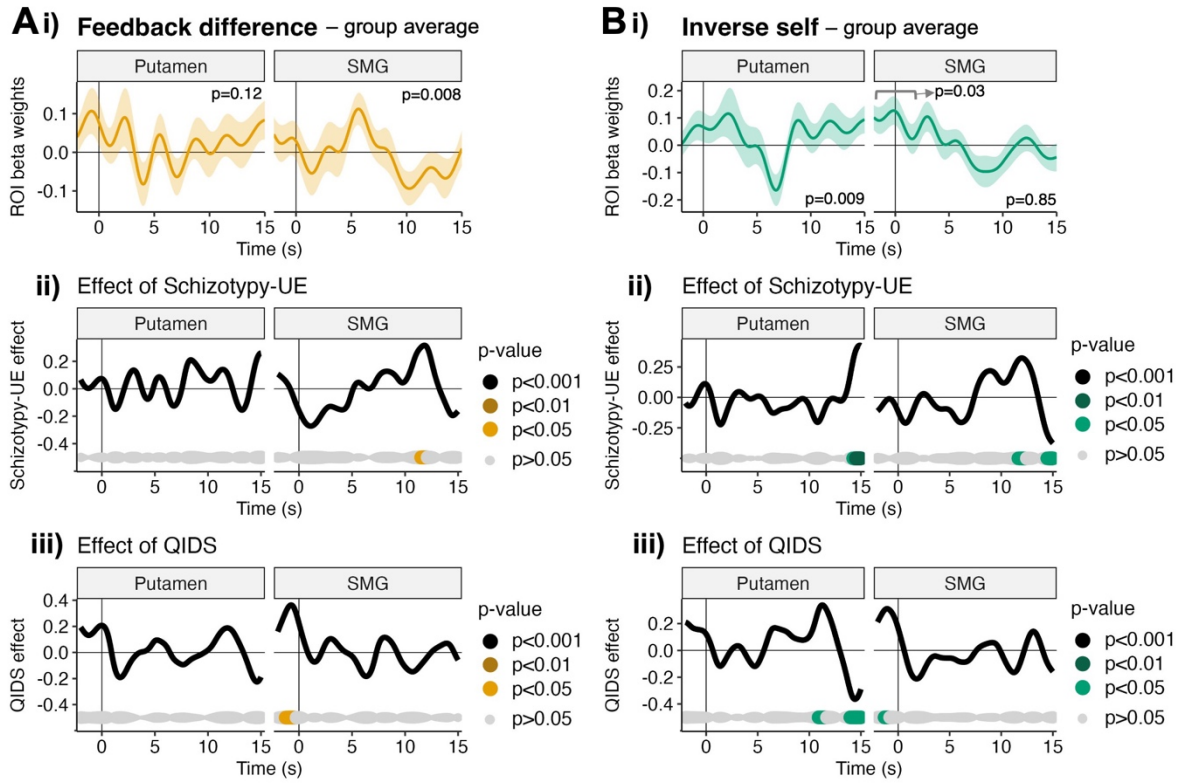


Figure 5.5. Time courses of feedback difference and inverse self in the SMG and putamen. Ai) We found that on a group level, SMG represents the feedback difference between normal and AD trials. **ii)** The SMG effect of feedback difference appeared to be modulated by schizotypy-UE but this effect was relatively late and short. **iii)** QIDS effects on SMG activity of the feedback difference were prior to outcome onset and we will therefore not interpret them. **Bi)** We found that the putamen was sensitive to the inverse of participants' self-performance. SMG showed, as in chapter 3, an early representation of the inverse self. Unlike in chapter 3, the later negative peak of the inverse self in SMG was not significant. **ii)** We found relatively late (more than 10 sec after outcome onset) effects of schizotypy-UE on putamen and SMG representations of the inverse self. Due to their timing, they require further investigation as to whether they are more related to the next trial's ratings, rather than the current outcome. **iii)** QIDS showed late effects on putamen activity and early effects (prior to outcome onset) on SMG activity. Time courses are time locked to the outcome phases and include switch trials during the Control-Other phase.

5.3.3 Effects of schizotypy-UE and QIDS on the neural correlates of self- and other-assigned PEs

In chapter 4, we presented behavioural results showing that participants with higher schizotypy-UE scores over-adjusted for their control when assigning PEs to themselves. They

assigned themselves too little credit when they had a low control. In this section, we examine the neural correlates of this effect. For this, we again used the time course analyses developed in the smaller MRI sample and included terms for individual schizotypy-UE and QIDS scores. As in chapter 3, this analysis focussed on normal trials during the Self-Other phase because this is when participants can split up tPEs to themselves and others according to their control.

In the smaller MRI sample in chapter 3, we found that SMG represented self- and other-assigned PEs (sPEs and oPEs) at the time of outcome (Figure 3.6 in chapter 3). In this large MRI sample, we did not find that either SMG or putamen were sensitive to sPE on a group level (sPE: SMG, $t(46)=-0.74$, $p=0.46$; putamen, $t(46)=-1.68$, $p=0.10$; Figure 5.6Ai). Schizotypy-UE appeared to modulate representations of sPEs in putamen but not SMG (peak of schizotypy-UE effect on sPE time course: putamen, $\beta=0.39$, $p=0.01$, 6.53 sec after outcome onset; Figure 5.6Aii). We did not find that sPE-related activity in the SMG is modulated by schizotypy-UE (Figure 5.6Aii), nor did QIDS modulate the putamen or SMG time courses of sPEs (Figure 5.6Aiii). On a group level, we found that oPEs are represented in the SMG, consistent with chapter 3, but not in the putamen (oPE: SMG, $t(46)=-3.09$, $p=0.003$; putamen, $t(46)=1.14$, $p=0.26$; Figure 5.6Bi). We did not find that these oPE time courses were modulated by schizotypy-UE (Figure 5.6Bii). QIDS modulated how the SMG represented the oPE (peak effect of QIDS on oPE time course in SMG: $\beta=-0.35$, $p=0.02$, 9.13 sec after outcome onset; Figure 5.6Biii).

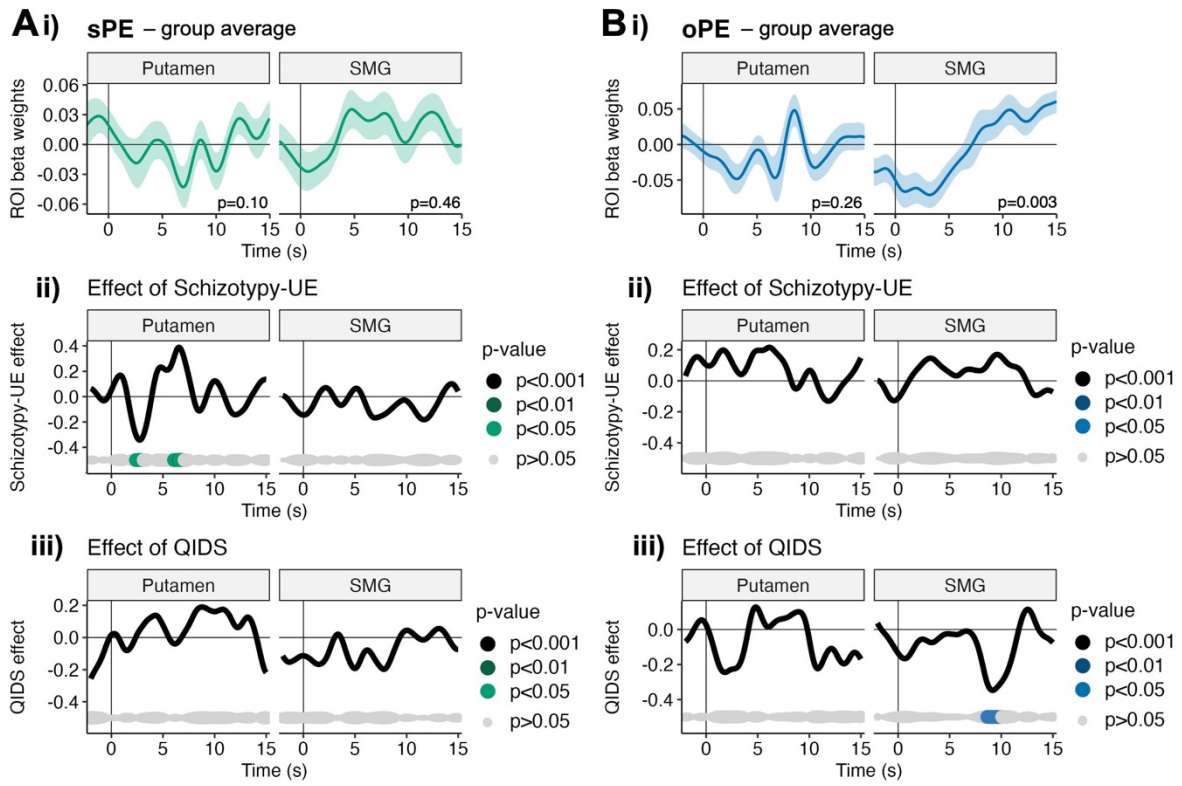


Figure 5.6. Time courses of self- and other-assigned PEs (sPE and oPE) in the SMG and putamen. **Ai)** Neither putamen nor SMG represented sPEs on a group level. **ii)** We found that schizotypy-UE appeared to modulate the sPE time courses in putamen but not SMG. **iii)** QIDS did not modulate sPE representations in putamen and SMG. **Bi)** oPEs were represented in the SMG, reproducing the result we found in chapter 3. **ii)** Schizotypy-UE did not modulate these representations. **iii)** QIDS appeared to modulate how sensitive SMG was to oPEs. Time courses are time locked to the outcome phase and are on normal trials in the Self-Other phase.

In chapter 3, we found that SMG tracks prediction errors in an agent-based manner (in terms of sPEs and oPEs), rather than by tracking the tPE and the known control in the Self-Other phase. Next, we tested whether, in this large MRI sample, our ROIs showed effects of this alternative coding scheme, and if these were modulated by schizotypy-UE or QIDS. As in chapter 3, we found that on a group level, SMG and putamen did not represent the tPE (tPE: putamen, $t(46)=-1.29$, $p=0.20$; SMG, $t(46)=-0.04$, $p=0.97$; Figure 5.7Ai). Schizotypy-UE appeared to modulate how the putamen represented tPEs, however this effect was very short and lasted for only 40 msec (peak schizotypy-UE effect on tPE time course in putamen: $\beta=0.31$,

p=0.04, 6.26 sec after outcome onset; Figure 5.7Aii). We found no effect of QIDS on the tPE time courses (Figure 5.7Aiii). On a group level and consistent with chapter 3, the known control was not represented in the SMG or putamen (control: SMG, t(46)=-1.78, p=0.08; putamen, t(46)=-1.78, p=0.08; Figure 5.7Bi). Schizotypy-UE and QIDS did not appear to strongly modulate the known control representations in putamen or SMG (Figure 5.7Bii-iii).

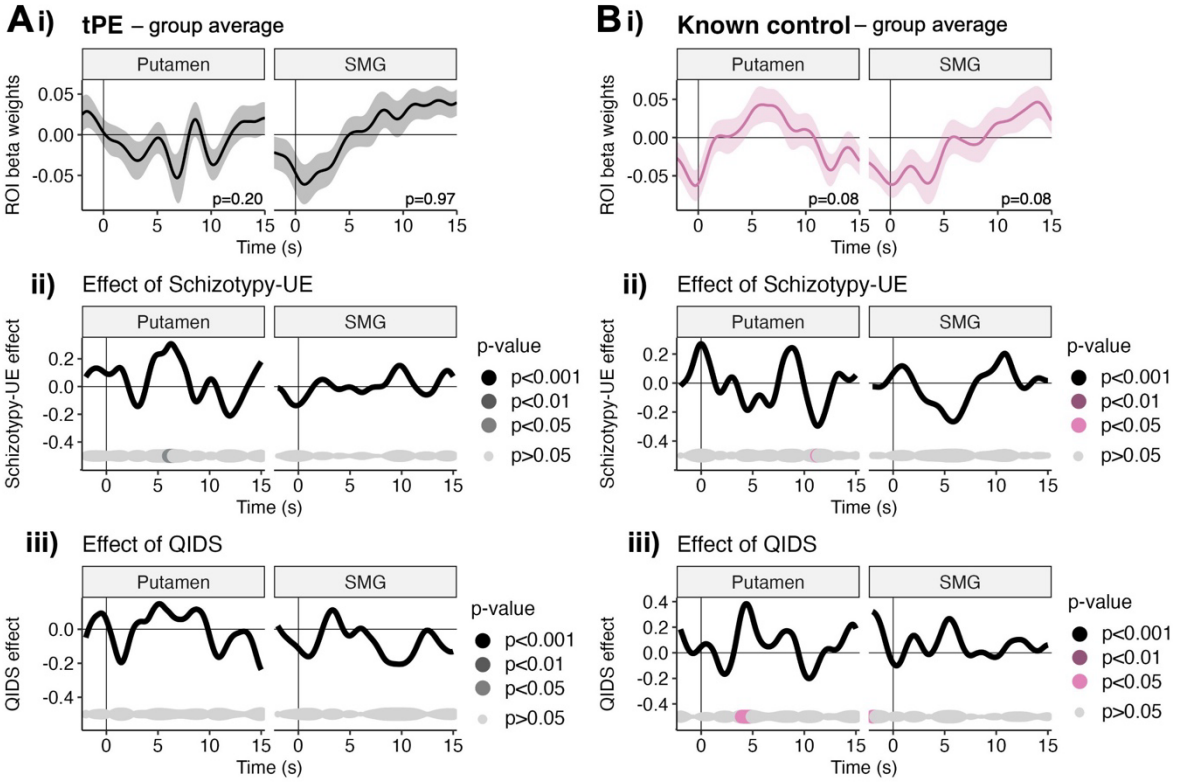


Figure 5.7. tPE and known control in the SMG and putamen during the outcome phase. Ai) Neither putamen nor SMG represented the tPE on a group level. **ii)** These time courses showed only very short changes by schizotypy-UE in the putamen. **iii)** No effects of QIDS on tPE time courses in putamen or SMG were observed. **Bi)** Consistent with chapter 3, we did not find that the SMG represented the known control during the Self-Other phase. Putamen was not sensitive to the known control. **ii)** These putamen and SMG control-related time courses were not strongly modulated by schizotypy-UE. **iii)** QIDS modulated how sensitive the putamen was to the known control but not the time course in the SMG. Time courses are time locked to the outcome phase and include normal trials in the Self-Other phase.

5.3.4 Effects of schizotypy-UE and QIDS on reward representations in the ventral striatum

Finally, we investigated whether schizotypy-UE and QIDS changed reward representations in the neural activity. In chapter 3, we found that the nucleus accumbens tracks feedback and feedback change, in the smaller MRI sample (Figure 3.9 in chapter 3). Therefore, here, we investigated whether we find this effect in the larger MRI sample and whether schizotypy-UE or QIDS modulate these representations. For this, we used the analyses developed in chapter 3 on the bilateral nucleus accumbens mask. During the outcome phase, we found that on a group level, as expected, the nucleus accumbens tracked the shown feedback and how much the feedback changed from the last to the current trial (feedback, $t(46)=2.14$, $p=0.04$, Figure 5.8Ai; feedback change, $t(46)=3.24$, $p=0.002$, Figure 5.8Bi). Next, we investigated whether these time courses were modulated by participants' schizotypy-UE or QIDS scores. Interestingly, we found that schizotypy-UE modulated both feedback and feedback change representations in the nucleus accumbens (peak effects of schizotypy-UE: feedback, $\beta=-0.35$, $p=0.02$, 7.97 sec after outcome onset, Figure 5.8Aii; feedback change, $\beta=0.40$, $p=0.008$, 10.22 sec after outcome onset, Figure 5.8Bii). We found that QIDS had an early effect on the feedback time course (peak effect of QIDS: feedback, $\beta=-0.36$, $p=0.02$, 3.46 sec after outcome onset, Figure 5.8Aiii).

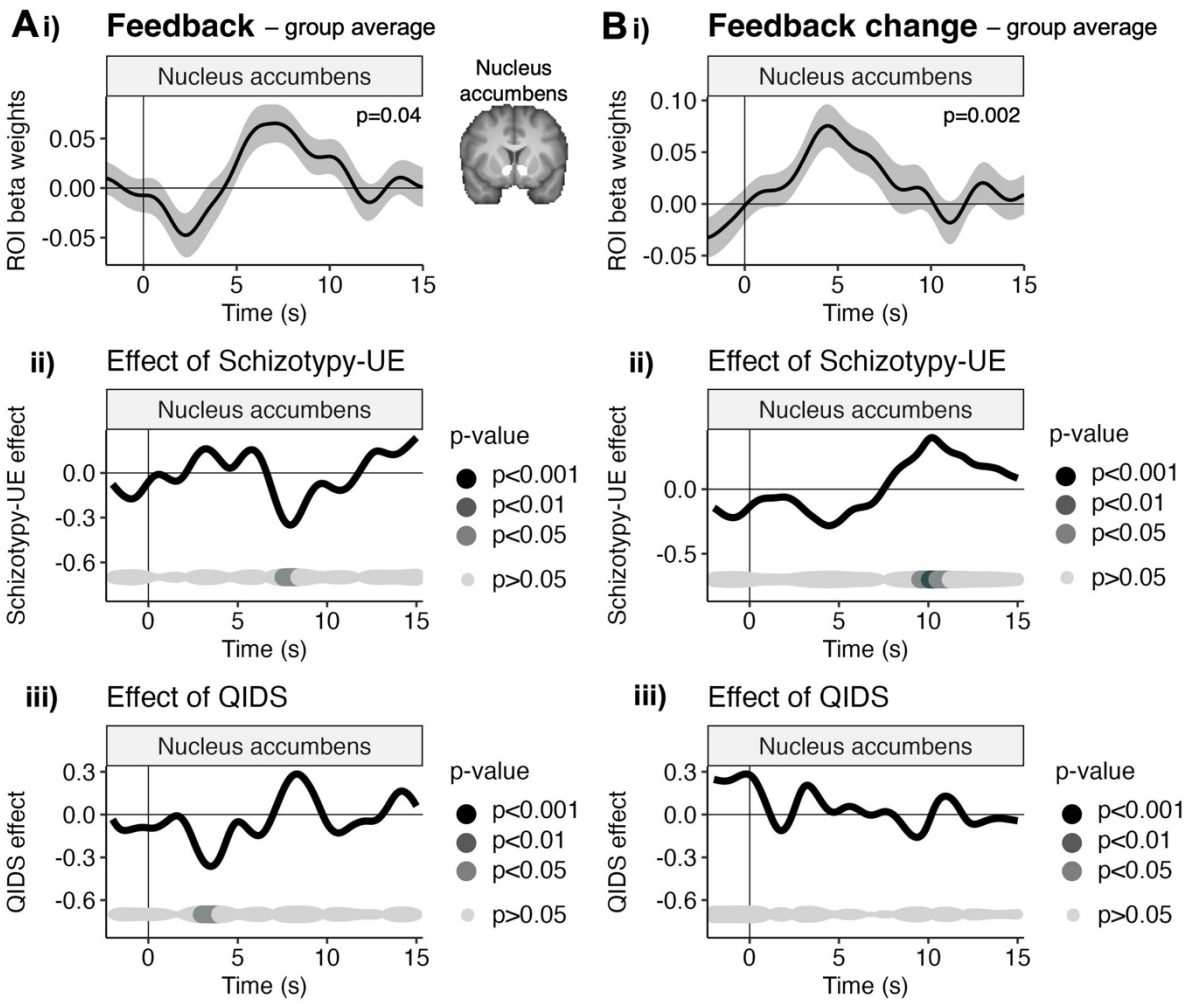


Figure 5.8. Time courses of feedback and feedback change in the nucleus accumbens during the outcome phase. Ai) Neural activity in the nucleus accumbens was sensitive to the feedback shown to participants during the outcome phase. We found that feedback representation was modulated by both **ii)** schizotypy-UE and **iii)** QIDS. **Bi)** Again consistent with our results in chapter 3, we found that the nucleus accumbens represented the trial-wise feedback change. This effect was modulated by **ii)** schizotypy-UE**, iii)** but not QIDs. Effects are time locked to the outcome phase and are on normal trials following normal trials across Self-Other and Control-Other phase.

5.4 Discussion

In this chapter, I presented findings on how the neural correlates of AD, control inference, and credit assignment vary as a function of schizotypal traits and depression scores. For this, I applied the analyses developed in chapter 3 to the full MRI sample and included terms for schizotypy-UE and QIDS. I found that on a group level, a set of regions including SMG signalled uncertainty and AD at the time of action generation, similar to the results presented in chapter 3. At the time of outcome and after whole-brain cluster-correction of the fMRI data analysis, I found that putamen activity for AD was decreased in individuals with higher schizotypy scores. Exploratory time course analyses on SMG and putamen revealed that putamen signalling of control seemed to increase with schizotypy.

One of the main neural findings of this chapter is that putamen activity to AD was decreased in participants with higher schizotypy scores. Previous research found people with schizophrenia show reduced activations to anticipated rewards in the ventral (Esslinger et al., 2012; Juckel et al., 2006; Nielsen et al., 2012) and dorsal striatum (Mucci et al., 2015). Similar results are found in first-degree relatives of patients with schizophrenia (Grimm et al., 2014). Here, I discovered reduced striatal responses following AD in individuals with schizotypal traits, suggesting that striatal activation to outcomes generated through active information seeking is modulated in schizotypy.

Exploratory time course analyses revealed that putamen activity signalled estimated control more strongly in people with higher schizotypal traits. Quite how this change emerged is not clear, however, research on learned helplessness in rats found that uncontrollable stress results in the activation of serotonergic neurons in the dorsal raphe nucleus (Grahn et al., 1999; Maswood et al., 1998) which also project to the striatum (Imai et al., 1986; Van der Kooy &

Hattori, 1980). It is possible that alteration in activity in this pathway could lead to the differences in striatal activity patterns that I found.

The time course of how schizotypy modulated control activity in the putamen corresponded with the time at which hemodynamic changes would be expected to peak. However, the individual difference analyses on the time courses presented in this chapter are preliminary and exploratory. Although they are extended over time, no control for autocorrelations in the BOLD response were made and the p-values reported were not corrected for multiple comparisons. In the future, I will use more robust and rigorous statistical tests to analyse the effects of schizotypy and depression scores on neural time courses. For this, I will employ permutation-based tests commonly used in EEG research (Maris & Oostenveld, 2007).

In this chapter, I reproduced many group level results from chapter 3. On a whole-brain level, I found that the SMG tracks uncertainty and AD during action generation in the larger MRI sample as it did in the smaller sample (chapter 3). Using ROI-based time course analyses, I could also reproduce the finding that the SMG tracked the difference between AD and normal feedback, as well as participants' self-estimates at the time of outcome. However, in the large MRI sample, the SMG did not seem to track the inferred control, control update or the self-assigned PE significantly on a group level. There are several possibilities why this might be the case. One possibility is that in the larger MRI sample, the inclusion of people with a wider range of schizotypy and depression scores lead to weaker group-level effects. This means that some of the group-level effects of the SMG are only present in the healthy sample included in the analyses chapter 3. Another possibility is that the SMG ROIs of chapters 3 and 5 were in slightly different parts of the anterior SMG. This was due to slightly different peak locations of the whole-brain effects in the small and large MRI sample. The SMG ROI in chapter 5 was located in a more superior part of the SMG than in chapter 3. It has previously been shown that different

portions of the SMG can be related to different functions such as reaching movements and behavioural changes based on observed outcomes (Mars et al., 2011).

I did not find significant whole-brain effects of symptoms of depression, which might not be surprising given that I did not find depression effects on behaviour in the previous chapter. As discussed before and given the literature on negative attributional biases in depression (attributing positive events to other causes and negative events to oneself), it might be beneficial to refine my neural analysis. For example, I might be able to identify different neural correlates of attributional biases to positive compared to negative feedback (or prediction errors) in depression.

Taken together, in this chapter, I investigated how depression and schizotypy alter the neural correlates of how people actively disambiguate outcomes, assign credit to themselves and others, and infer their control from observed outcomes. I found that schizotypy modulates neural representations of AD in the putamen. Exploratory analyses suggest that representations of the inferred control in the putamen are stronger in individuals with higher schizotypal traits.

CHAPTER 6 General Discussion

Key human abilities include inferring whether one's actions make any difference to the environment and estimating which events were caused by us or by other people. Understanding how people come to understand their level of control and their own performance compared to the performances of others is the focus of this thesis. In chapter 2, I presented a novel task developed to probe how people do so. I found that people learned about themselves and others by assigning prediction errors to themselves and others in proportion to the level of control they estimated for each actor. In addition, they established their level of control by disambiguating whether they themselves or the other actor caused feedback by varying their own behaviour and performance on purpose. In other words, they sometimes made mistakes on purpose to see whether this had any impact on the outcome. This allowed them to reduce uncertainty and assign credit better. They used this strategy in particular to estimate their level of control. I demonstrated that participants' learning behaviour resembled one exhibited by an optimal Bayesian learning model. In chapter 3, I showed that the supramarginal gyrus carried the neural correlates of these behavioural patterns in the healthy brain. In chapters 4 and 5, I examined how depression and schizotypy are associated with interindividual differences in credit assignment, both on a behavioural and neural level. I found that people scoring higher in schizotypy overadjusted for their level of control when assigning themselves (but not others) prediction errors. In addition, higher schizotypy was associated with a more optimal control estimation. Neural activity related to active disambiguation (AD) in the putamen varied as a function of schizotypy, and appeared to represent the estimated control for participants with higher schizotypy scores.

6.1 Inferring control and learning about self and others

In chapter 2, I demonstrated that people inferred their control over ambiguous outcomes by making deliberate mistakes on some occasions and trying their best on others. By comparing the difference this makes to the feedback and taking into account their best possible performance, participants were able to estimate their level of control over the feedback. In chapter 4, I reproduced this group-level effect in the full MRI sample. This result suggests that correctly estimating controllability not only involves comparing action outcomes in the environment, but also monitoring one's own actions and overall performance. On an abstract level, this is reminiscent of how control is formulated in the contexts of learned helplessness, depressive realism and schizophrenia. According to Maier and Seligman (1976) and the Delta P rule (Allan, 1980), control is estimated from the difference in the likelihood of an outcome given an action versus no action. Ligneul (2021) argued that in addition, other states in the environment need to be taken into account. Frith et al. (2000a) proposed that a sense of agency is derived from the difference between the expected and experienced action outcome. In my thesis, I argue that in social contexts, people can estimate their level of control from outcomes following different actions and strategically performing poorly. In addition, it is necessary to take into account how "good" one's actions are, that is one's own skill level.

More generally, AD shares features with information-seeking and causal intervention in related domains of behaviour. AD helped participants assign credit better, and participants were rewarded for how accurately they assigned credit in my task. Since AD aided task completion, it can be considered a form of instrumental information-seeking. AD shares features with causal interventions because causal interventions are purposeful changes to a variable that can help determine the causal relationships in the environment (Pearl, 2009; Sloman, 2005; Steyvers et al., 2003). Such causal interventions also help disambiguate unobserved causes (Kushnir et al., 2003), and AD is an instance of an intervention aimed at learning about hidden causes in

my task. However, in my task, participants did not simply choose between options with different information content (information-seeking) or manipulated the value of some external variable (causal intervention), but instead they actively explored by performing poorly in order to test their level of control.

In chapter 2, I identified a second learning mechanism by which participants learned about themselves and others. This learning mechanism was independent of AD and was employed when participants had an accurate estimate of their own level of control. On non-AD trials, participants split up prediction errors to a self- and other-assigned PE, proportional to their level of control. This suggests that participants' estimates of their control influenced how they adjusted their beliefs about themselves and others. Again, I found similar group-level results in the larger MRI sample analysed in chapter 4. Previous work has found that credit assignment is modulated by attention (Akaishi et al., 2016; Leong et al., 2017), and that generally, the informational value of outcomes influences learning (Behrens et al., 2007; Pearce & Hall, 1980). The valence of outcomes can lead to biased learning and perceptions about self and others (Dorfman et al., 2019; Miller & Ross, 1975).

6.2 The neural correlates of control and causal learning

In chapter 3, I found that the SMG is a critical brain region for tracking control and credit assignment. The SMG tracked AD and uncertainty during action generation, and was involved in all components of inferring control from AD and non-AD, as well as self- and other-assigned prediction errors. The unique role of the SMG in these processes of credit assignment and control learning might be surprising given past work on the role of frontal regions in decision-making and learning. Past research has highlighted the role of area 47/12o, or the lateral orbitofrontal cortex, in credit assignment (e.g. Folloni et al., 2021; Jocham et al., 2016; Noonan et al., 2017; Walton et al., 2010). Area 47/12o is highly interconnected with the anterior

cingulate cortex and both mediate information seeking and credit assignment (e.g. Jezzini et al., 2021; Monosov & Rushworth, 2022; Sallet et al., 2020; Trudel et al., 2021). There are several key differences between my studies and this previous work that might explain why different brain regions track credit assignment in different contexts. Firstly, I studied credit assignment in a social environment while previous work often focused on credit assignment in non-social contexts. Cooperating with others requires people to monitor their own actions, and the SMG is a region that has been linked to action monitoring and the abstract representation of actions in terms of their meaning and purpose (e.g. Desmurget et al., 2009; Peeters et al., 2009; Peeters et al., 2013). In addition, previous work suggests that the SMG is a key region for establishing the sense of agency during social interactions (Farrer et al., 2003; Farrer & Frith, 2002; Ohata et al., 2020; Sirigu et al., 1999). Secondly, previous work studied credit assignment of binary rewards to choices (e.g. Jocham et al., 2016; Walton et al., 2010), while I investigated credit assignment of continuous feedback to one's own or others' performance and level of control.

6.3 Interindividual differences in the behavioural and neural signatures of control and causal learning

In chapter 4, I presented preliminary results on how depression and schizotypy affect performance and control estimates. Participants with higher schizotypy scores had lower initial beliefs about their own control and were worse at judging their own performance. They overadjusted for their level of control and gave themselves too little credit when they estimated they had low control. In addition, they relied more on feedback differences and performance estimates to infer their control. These results suggest that individuals with higher schizotypy scores relied more on external information (feedback and prediction errors) to learn about their own performance and control. This supports findings that individuals with high schizotypy and patients with schizophrenia rely more on external cues to derive their sense of agency (Moore

et al., 2021; Synofzik et al., 2010; Voss et al., 2017). In addition, I found that schizotypy modulated learning about self and control, but not about other players. This aligns with previous work on biased self- and control-related information processing in schizotypy and schizophrenia, such as self-serving biases in schizophrenia that may protect self-esteem (Bentall et al., 2001; Bentall et al., 1994). Such self-serving biases describe attributional biases to positive versus negative outcomes. Therefore, in my task, analysing credit assignment after positive compared to negative outcomes and how this might be biased in schizotypy could be beneficial in the future.

On a neural level, in chapter 5, I found that schizotypy was associated with a deactivation of AD-related putamen activity, and the putamen appeared to signal control in individuals with higher schizotypy scores. In addition, in my study, schizotypy appeared to modulate reward-related activity in the ventral striatum. Previous research showed that schizophrenia is related to differential activation of the ventral striatum to reward prediction errors (Schlagenhauf et al., 2014) and anticipated rewards (Esslinger et al., 2012; Juckel et al., 2006; Mucci et al., 2015; Nielsen et al., 2012). One drawback of the time course analysis on the neural data presented in chapter 5 is that the clinical effects were not controlled for multiple comparisons or the autocorrelations of the BOLD response. Therefore, in the future, the preliminary and exploratory time course analyses presented in chapter 5 need to be more rigorously tested using more robust analysis techniques such as permutation tests used in EEG research.

More generally, my findings link to the theoretical models introduced in the general introduction of this thesis (chapter 1) in two ways. Firstly, I found that on a group level in both the small and large MRI samples (chapters 2 and 4), people assigned themselves and others portions of social prediction errors in relation to their perceived level of control. This means that people learned about themselves and others by attributing outcomes, which I quantified as prediction

errors. In addition, I found that schizotypy was related to differences in attributing prediction errors to the self. This finding confirms my hypothesis that concepts from reinforcement learning can be used to mathematically describe causal attributions and attributional biases. In other words, prediction error-guided learning can be used as a mathematical formulation of causal outcome attributions.

Secondly, my findings on how people infer their level of control based on AD have important parallels to reinforcement learning, the comparator model (Frith et al., 2000a, 2000b), learned helplessness (Seligman, 1974), the Delta P rule (Allan, 1980), and the work by Ligneul (2021). All of these frameworks suggest that to infer control, an agent needs to compare action outcomes to some degree. For example, the comparator model posits that a sense of agency arises through an internal action monitoring process that compares experienced and predicted action outcomes. In reinforcement learning, beliefs are updated via prediction errors, which involve comparing predictions with outcomes. The frameworks of learned helplessness, the Delta P rule, and the model by Ligneul suggest that control is inferred by comparing the outcome likelihoods when an action occurs versus when no action happens. In my thesis, I presented studies suggesting that in social environments, a similar comparison process might be at play. In my studies, I found that people infer their control by changing their behaviour (performing well versus AD) and observing the changes this makes to outcomes while accounting for their own skill level (chapters 2 and 4). While this finding complements previous controllability accounts, it extends them in important ways. Previous frameworks did not consider how one's perception of one's own skill interacts with control beliefs. In addition, previous work focussed on the presence and absence of actions, while my findings highlight that varying one's own behaviour in a deliberate way and introspection into this are crucial to estimating control in social environments.

6.4 Limitations, ongoing and future work

The interindividual difference analyses presented in chapters 4 and 5 were exploratory and can be refined in the future. For example, I did not find any effects of depression on credit assignment. One potential reason might be that my task or analyses are not sensitive to cognitive biases in depression. Public and Patient Involvement (PPI) in clinical research aims to improve research quality and relevance by engaging patients and the public (National Institute for Health Research, 2014). In the future, more sensitive analyses or tasks might be developed from discussing patients' experiences in the task. Depressed individuals often attribute negative events to themselves and positive events to external factors (Abramson et al., 1989; Beck, 2008), while non-depressed individuals show the opposite pattern and attribute positive outcomes to themselves (Dorfman et al., 2019; Miller & Ross, 1975; Zamfir & Dayan, 2022). In my analyses, I did not compare credit assignment after positive or negative outcomes. In the future, I could explore how depression affects credit assignment and control perceptions after positive compared to negative feedback or prediction errors.

There are other differences between my task and previous work linking control and symptoms of depression. Previous work on depressive realism and the Delta P rule highlighted that the depressive realism effect is modulated by the degree to which participants experience situations when neither actions nor outcomes occur (Msetfi et al., 2005). In contrast, in my task, participants always needed to perform an action (press a button) in the games to observe the feedback afterwards. Even to remove their own contribution from the feedback, participants had to perform a (poor) action (AD trial). It could be the case that this requirement for participants to always perform some type of action is why we did not find different control estimations in people with higher depressive symptoms. In the future, the task could be adapted so that AD is not achieved by performing poorly, but by not responding at all. For example, the AD feedback could be shown when participants do not press a button within a

fixed amount of time (e.g. 5 sec) during the games. It would be interesting to explore if this might lead to different control estimations in people with and without depressive symptoms.

In my task, participants could remove their own contribution from the feedback by performing AD trials, i.e. make a deliberate mistake during the games. Previous work has found that the depressive realism effect is mediated by the process of action generation itself; people with depressive symptoms initiate fewer actions and therefore believe they have little control over events (Blanco et al., 2012). Indeed, when participants were instructed to respond with a high frequency, the depressive realism effect was reduced (Blanco et al., 2012). In my MRI sample, every participant performed AD during the task, and they received a hint during the instructions to do so (in order to make the neural data analysis possible). This might be why I did not find associations between depressive symptoms and the use of AD in chapter 4. In my DPhil, I collected a large online sample (described in further detail in the next paragraph), in which participants did not receive this hint in the task instruction. This data will allow me to assess how AD and AD-guided control learning might be altered in depression.

One way to follow-up on my behavioural findings and refine my analyses is to test them in a larger online sample. In the field of computational psychiatry, online data collection has gained popularity because large data sets have enough statistical power to disentangle the contributions of heterogeneous and comorbid symptoms (rather than disorders) to specific behaviours (Gillan & Daw, 2016). For this, factor analyses are typically used to develop transdiagnostic psychiatric dimensions, and they have been shown to better relate to behavioural biases than more traditional questionnaire sum scores (e.g. Gillan et al., 2016). Another advantage of large online samples is that it is possible to use one sample (discovery sample) to explore many different analyses and then preregister the results, and then to test the developed analyses and hypothesis in a second sample (confirmation sample), thereby

increasing the transparency and rigour of research. During my DPhil, I developed an online version of my task and collected two large online samples (discovery sample, n=400; confirmation sample, n=800). All participants completed my credit assignment task and afterwards filled in a similar battery of psychiatric questionnaires. The discovery data set will allow me to develop new behavioural analyses that are sensitive to, for example, symptoms of depression. I will then be able to apply the refined analysis to the behavioural data of my MRI study, and as a second step, identify the neural correlates of these behaviours in the fMRI data. A disadvantage of online studies is that participants are unobserved and data quality can be lower than in lab-based studies. I addressed this in my online sample by including rigorous data quality checks. The results from the online study, together with my MRI sample which is fully characterised with clinical interviews and psychiatric questionnaires, will allow me to draw firm conclusions on how the mechanisms of control and causal learning are linked to psychiatric symptoms.

In the MRI study presented in this thesis, I prioritised the sample characteristics in terms of depression symptoms rather than schizotypy. For example, I recruited participants based on their depression (and not schizotypy) scores, and volunteers who experienced psychotic or associated symptoms were excluded to reduce comorbidities. While this optimised the sample characteristics in terms of depression, it may have restricted interindividual variation in schizotypy scores. In the future, it would be interesting to replicate the behavioural schizotypy-UE results in my large online sample.

It has previously been found that specifically schizotypy-UE, in particular, can be related to behaviours such as goal-directed deficits (Gillan et al., 2016). It would be interesting to understand whether my results of schizotypy-UE and credit assignment are specific to the unusual experiences dimensions, or whether they relate to other dimensions of schizotypy as

well. I will be able to test this in the large online study, in which participants filled in a more extensive questionnaire including questions about bizarre experiences, perceptual anomalies and persecutory ideation (Capra et al., 2017).

In chapter 3, I presented the neural correlates of how participants actively disambiguated feedback, assigned credit and learned about their control. However, evidence from fMRI data is only correlational and therefore, it is not possible to know whether the SMG was causally important for these cognitive processes. For this, it is necessary to conduct studies using causal manipulations such as transcranial magnetic stimulation (TMS). In the future, it would be interesting to disrupt SMG activity with TMS and have participants do my credit assignment task afterwards. This will enable me to investigate whether SMG activation causes certain behavioural or mental processes in the task; for example whether disrupting SMG activation changes how people think about what they themselves are causally responsible for. This will assess the neural mechanisms identified in this thesis in a complementary manner.

In my thesis, I investigated the mechanisms of control and causal inference in social environments where people cooperate with others. Using AD, people inferred their level of control over feedback (chapter 2). In addition, people assigned social prediction errors to themselves and others according to their perceived level of control. SMG activity mediated these processes (chapter 3). In the future, it would be interesting to investigate whether these processes are socially specific or unfold similarly in non-social settings. A non-social version of my task could be developed by changing the cover story. Instead of instructing participants to play together with others and the feedback representing their own and someone else's performance, participants could be instructed to play the games alone, with the feedback partially based on chance. Here, reframing the other player's performance as chance would substitute the social elements of the task by a non-social element. Creating such a non-social

task condition, which should be tightly matched to the social condition (for a similar approach, see Trudel et al., 2023), would allow me to assess which of my findings are socially specific and which are domain-general. I hypothesize that in a non-social environment, people would still perform AD to disambiguate feedback. However, they might use this knowledge differently if they cooperate with someone else versus when there is chance involved.

6.5 Summary

In sum, I used a novel task together with neuroimaging and computational methods to investigate how humans assign credit and learn about their control in social interactions. I identified specific behavioural strategies and computational mechanisms that supported these processes and that helped people attribute ambiguous outcomes to themselves and others. I found that a specific region in the parietal cortex, the supramarginal gyrus, carried different types of inference-related signals as well as exploratory action representations that guided learning. Schizotypy was associated with behavioural differences in outcome attributions to oneself and one's own control, and neural signals in the striatum tracked differences in the processing of exploratory outcomes related to schizotypy.

References

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological review*, *96*(2), 358.

Adams, R. A., Huys, Q. J., & Roiser, J. P. (2016). Computational psychiatry: towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery & Psychiatry*, *87*(1), 53-63.

Akaishi, R., Kolling, N., Brown, J. W., & Rushworth, M. (2016). Neural mechanisms of credit assignment in a multicue environment. *Journal of Neuroscience*, *36*(4), 1096-1112.

Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, *15*(3), 147-149.

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, *108*(4), 441.

Alloy, L. B., Abramson, L. Y., Whitehouse, W. G., Hogan, M. E., Tashman, N. A., Steinberg, D. L., Rose, D. T., & Donovan, P. (1999). Depressogenic cognitive styles: Predictive validity, information processing and personality characteristics, and developmental origins. *Behaviour research and therapy*, *37*(6), 503-531.

Ang, Y.-S., Lockwood, P., Apps, M. A., Muhammed, K., & Husain, M. (2017). Distinct subtypes of apathy revealed by the apathy motivation index. *PLoS One*, *12*(1), e0169938.

Asai, T. (2016). Self is “other”, other is “self”: poor self-other discriminability explains schizotypal twisted agency judgment. *Psychiatry research*, *246*, 593-600.

Bang, D., Moran, R., Daw, N. D., & Fleming, S. M. (2022). Neurocomputational mechanisms of confidence in self and others. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-31674-w

Barnett, L., Barrett, A. B., & Seth, A. K. (2009). Granger causality and transfer entropy are equivalent for Gaussian variables. *Physical review letters*, *103*(23), 238701.

Barrett, L., Dunbar, R., & Lycett, J. (2002). *Human evolutionary psychology*. Princeton University Press.

Beck, A. T. (2008). The evolution of the cognitive model of depression and its neurobiological correlates. *American Journal of Psychiatry*, *165*.

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical psychology review*, *8*(1), 77-100.

Behrens, T. E., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. (2008). Associative learning of social value. *Nature*, *456*(7219), 245-249.

Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature neuroscience*, *10*(9), 1214-1221.

Benassi, V. A., & Mahler, H. I. (1985). Contingency judgments by depressed college students: sadder but not always wiser. *Journal of personality and social psychology*, *49*(5), 1323.

Bennett-Levy, J. (2003). Mechanisms of change in cognitive therapy: The case of automatic thought records and behavioural experiments. *Behavioural and Cognitive Psychotherapy*, *31*(3), 261-277.

Bennett-Levy, J., Westbrook, D., Fennell, M., Cooper, M., Rouf, K., & Hackmann, A. (2004). Behavioural experiments: Historical and conceptual underpinnings. *Oxford guide to behavioural experiments in cognitive therapy*, 1-20.

Bentall, R. P., Corcoran, R., Howard, R., Blackwood, N., & Kinderman, P. (2001). Persecutory delusions: a review and theoretical integration. *Clinical psychology review*, *21*(8), 1143-1192.

Bentall, R. P., Kinderman, P., & Kaney, S. (1994). The self, attributional processes and abnormal beliefs: towards a model of persecutory delusions. *Behaviour research and therapy*, *32*(3), 331-341.

Berry, D. A., & Fristedt, B. (1985). Bandit problems: sequential allocation of experiments (Monographs on statistics and applied probability). *London: Chapman and Hall*, *5*(71-87), 7-7.

Blakemore, S.-J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of cognitive neuroscience*, *11*(5), 551-559.

Blakemore, S.-J., Wolpert, D., & Frith, C. (2000). Why can't you tickle yourself? *Neuroreport*, *11*(11), R11-R16.

Blanco, F., Matute, H., & M, A. V. (2012). Mediating role of activity level in the depressive realism effect. *PLoS One*, *7*(9), e46203. https://doi.org/10.1371/journal.pone.0046203

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

Boorman, E. D., O'Doherty, J. P., Adolphs, R., & Rangel, A. (2013). The behavioral and neural mechanisms underlying the tracking of expertise. *Neuron*, *80*(6), 1558-1571.

Boorman, E. D., Rajendran, V. G., O'Reilly, J. X., & Behrens, T. E. (2016). Two anatomically and computationally distinct learning signals predict changes to stimulus-outcome associations in hippocampus. *Neuron*, *89*(6), 1343-1354.

Bradley, B. P., Mogg, K., & Williams, R. (1995). Implicit and explicit memory for emotion-congruent information in clinical depression and anxiety. *Behaviour research and therapy*, *33*(7), 755-770.

Bromberg-Martin, E. S., & Monosov, I. E. (2020). Neural circuitry of information seeking. *Current Opinion in Behavioral Sciences*, *35*, 62-70.

Brooks, J. C., Beckmann, C. F., Miller, K. L., Wise, R. G., Porro, C. A., Tracey, I., & Jenkinson, M. (2008). Physiological noise modelling for spinal functional magnetic resonance imaging studies. *Neuroimage*, *39*(2), 680-692.

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Cannon, T. D., Kaprio, J., Lönnqvist, J., Huttunen, M., & Koskenvuo, M. (1998). The genetic epidemiology of schizophrenia in a Finnish twin cohort: a population-based modeling study. *Archives of General Psychiatry*, *55*(1), 67-74.

Capra, C., Kavanagh, D. J., Hides, L., & Scott, J. G. (2017). Current CAPE-15: A measure of recent psychotic-like experiences and associated distress. *Early intervention in psychiatry*, *11*(5), 411-417.

Carlsson, A. (1978). Does dopamine have a role in schizophrenia? *Biol Psychiatry*.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1).

Cavada, C., & Goldman-Rakic, P. S. (1989a). Posterior parietal cortex in rhesus monkey: I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections. *Journal of Comparative Neurology*, *287*(4), 393-421.

Cavada, C., & Goldman-Rakic, P. S. (1989b). Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *Journal of Comparative Neurology*, *287*(4), 422-445.

Chambon, V., Pacherie, E., Barbalat, G., Jacquet, P., Franck, N., & Farrer, C. (2011). Mentalizing under influence: abnormal dependence on prior expectations in patients with schizophrenia. *Brain*, *134*(12), 3728-3741.

Chambon, V., Théro, H., Vidal, M., Vandendriessche, H., Haggard, P., & Palminteri, S. (2020). Information about action outcomes differentially affects learning from self-determined versus imposed choices. *Nature Human Behaviour*, *4*(10), 1067-1079.

Chan, S. W., Goodwin, G. M., & Harmer, C. J. (2007). Highly neurotic never-depressed students have negative biases in information processing. *Psychological medicine*, *37*(9), 1281-1291.

Charpentier, C. J., Bromberg-Martin, E. S., & Sharot, T. (2018). Valuation of knowledge and ignorance in mesolimbic reward circuitry. *Proceedings of the National Academy of Sciences*, *115*(31), E7255-E7264.

Chen, J., Paciocco, J. U., Deng, Z., & Culham, J. C. (2023). Human neuroimaging reveals differences in activation and connectivity between real and pantomimed tool use. *Journal of Neuroscience*, *43*(46), 7853-7867.

Claridge, G. E. (1997). *Schizotypy: Implications for illness and health*. Oxford University Press.

Clark, L., Chamberlain, S. R., & Sahakian, B. J. (2009). Neurocognitive mechanisms in depression: implications for treatment. *Annu Rev Neurosci*, *32*, 57-74. https://doi.org/10.1146/annurev.neuro.31.060407.125618

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *Journal of Neuroscience*, *18*(18), 7426-7435.

Cromwell, H. C., & Schultz, W. (2003). Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *Journal of neurophysiology*, *89*(5), 2823-2838.

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC medicine*, *11*, 1-8.

Daw, N. D., O'doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876-879.

Deichmann, R., Gottfried, J. A., Hutton, C., & Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage*, *19*(2), 430-441.

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

Dombrovski, A. Y., Szanto, K., Clark, L., Reynolds III, C. F., & Siegle, G. J. (2013). Reward signals, attempted suicide, and impulsivity in late-life depression. *JAMA psychiatry*, *70*(10).

Dorfman, H. M., Bhui, R., Hughes, B. L., & Gershman, S. J. (2019). Causal inference about good and bad outcomes. *Psychological science*, *30*(4), 516-525.

Dorfman, H. M., Tomov, M. S., Cheung, B., Clarke, D., Gershman, S. J., & Hughes, B. L. (2021). Causal inference gates corticostriatal learning. *Journal of Neuroscience*, *41*(32), 6892-6904.

Dunbar, R. I., & Shultz, S. (2007). Evolution in the social brain. *Science*, *317*(5843), 1344-1347.

Ebert, J. P., & Wegner, D. M. (2010). Time warp: Authorship shapes the perceived timing of actions and events. *Consciousness and cognition*, *19*(1), 481-489.

Esslinger, C., Englisch, S., Inta, D., Rausch, F., Schirmbeck, F., Mier, D., Kirsch, P., Meyer-Lindenberg, A., & Zink, M. (2012). Ventral striatal activation during attribution of stimulus saliency and reward anticipation is correlated in unmedicated first episode schizophrenia patients. *Schizophrenia research*, *140*(1-3), 114-121.

Ettinger, U., Mohr, C., Gooding, D. C., Cohen, A. S., Rapp, A., Haenschel, C., & Park, S. (2015). Cognition and brain function in schizotypy: a selective review. *Schizophrenia bulletin*, *41*(suppl_2), S417-S426.

Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., & Jeannerod, M. (2003). Modulating the experience of agency: a positron emission tomography study. *Neuroimage*, *18*(2), 324-333.

Farrer, C., & Frith, C. D. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage*, *15*(3), 596-603.

First, M. B., Williams, J. B., Karg, R. S., & Spitzer, R. L. (2016). SCID-5-CV: Structured clinical interview for DSM-5 disorders: Clinician version. *(No Title)*.

Fischer, A. G., Bourgeois-Gironde, S., & Ullsperger, M. (2017). Short-term reward experience biases inference despite dissociable neural correlates. *Nature Communications*, *8*(1), 1690.

Fisher, R. (1934). Statistical methods for research workers.

Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The Obsessive-Compulsive Inventory: development and validation of a short version. *Psychological assessment*, *14*(4), 485.

Folloni, D., Fouragnan, E., Wittmann, M. K., Roumazeilles, L., Tankelevitch, L., Verhagen, L., Attali, D., Aubry, J.-F., Sallet, J., & Rushworth, M. F. (2021). Ultrasound modulation of macaque prefrontal cortex selectively alters credit assignment–related activity and behavior. *Science advances*, *7*(51), eabg7700.

Fourneret, P., Vignemont, F. d., Franck, N., Slachevsky, A., Dubois, B., & Jeannerod, M. (2002). Perception of self-generated action in schizophrenia. *Cognitive neuropsychiatry*, *7*(2), 139-156.

Fried, E. I., & Nesse, R. M. (2015). Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC medicine*, *13*, 1-11.

Frith, C. D., Blakemore, S.-J., & Wolpert, D. M. (2000a). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1404), 1771-1788.

Frith, C. D., Blakemore, S.-J., & Wolpert, D. M. (2000b). Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain research reviews*, *31*(2-3), 357-363.

Gabry, J. (2021). cmdstanr: R Interface to'CmdStan'. *(No Title)*.

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in cognitive sciences*, *4*(1), 14-21.

Gariepy, G., Honkaniemi, H., & Quesnel-Vallee, A. (2016). Social support and protection from depression: systematic review of current findings in Western countries. *The British Journal of Psychiatry*, *209*(4), 284-293.

Garrett, N., González-Garzón, A. M., Foulkes, L., Levita, L., & Sharot, T. (2018). Updating beliefs under perceived threat. *Journal of Neuroscience*, *38*(36), 7901-7911.

Garrett, N., Sharot, T., Faulkner, P., Korn, C. W., Roiser, J. P., & Dolan, R. J. (2014). Losing the rose tinted glasses: neural substrates of unbiased belief updating in depression. *Frontiers in human neuroscience*, *8*, 639.

Gillan, C. M., & Daw, N. D. (2016). Taking psychiatry research online. *Neuron*, *91*(1), 19-23.

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *elife*, *5*, e11305.

Gooding, D. C., Tallent, K. A., & Matts, C. W. (2005). Clinical status of at-risk individuals 5 years later: further validation of the psychometric high-risk strategy. *Journal of abnormal psychology*, *114*(1), 170.

Gottlieb, J., Cohanpour, M., Li, Y., Singletary, N., & Zabeh, E. (2020). Curiosity, information demand and attentional priority. *Current Opinion in Behavioral Sciences*, *35*, 83-91.

Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J., & Steele, J. D. (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain*, *134*(6), 1751-1764.

Grahn, R. E., Will, M. J., Hammack, S. E., Maswood, S., McQueen, M. B., Watkins, L. R., & Maier, S. F. (1999). Activation of serotonin-immunoreactive cells in the dorsal raphe nucleus in rats exposed to an uncontrollable stressor. *Brain Research*, *826*(1), 35-43.

Grimm, O., Heinz, A., Walter, H., Kirsch, P., Erk, S., Haddad, L., Plichta, M. M., Romanczuk-Seiferth, N., Pöhland, L., & Mohnke, S. (2014). Striatal response to reward anticipation: evidence for a systems-level intermediate phenotype for schizophrenia. *JAMA psychiatry*, *71*(5), 531-539.

Gur, R. C., Erwin, R. J., Gur, R. E., Zwil, A. S., Heimberg, C., & Kraemer, H. C. (1992). Facial emotion discrimination: II. Behavioral findings in depression. *Psychiatry research*, *42*(3), 241-251.

Gureckis, T. M., & Markant, D. B. (2012). Self-directed learning: A cognitive and computational perspective. *Perspectives on Psychological Science*, *7*(5), 464-481.

Haber, S. N. (2016). Corticostriatal circuitry. *Dialogues in clinical neuroscience*, *18*(1), 7-21.

Haber, S. N., & Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*, *35*(1), 4-26.

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*(4), 196-207.

Halahakoon, D. C., Kieslich, K., O'Driscoll, C., Nair, A., Lewis, G., & Roiser, J. P. (2020). Reward-processing behavior in depressed participants relative to healthy volunteers: A systematic review and meta-analysis. *JAMA psychiatry*, *77*(12), 1286-1295.

Halverson, T. F., Orleans-Pobee, M., Merritt, C., Sheeran, P., Fett, A.-K., & Penn, D. L. (2019). Pathways to functional outcomes in schizophrenia spectrum disorders: Meta-analysis of social cognitive and neurocognitive predictors. *Neuroscience & Biobehavioral Reviews*, *105*, 212-219.

Harmer, C. J., O'Sullivan, U., Favaron, E., Massey-Chase, R., Ayres, R., Reinecke, A., Goodwin D Phil, F. P., Guy M, & Cowen, P. J. (2009). Effect of acute antidepressant administration on negative affective bias in depressed patients. *American Journal of Psychiatry*, *166*(10), 1178-1184.

Harrow, M., Hansford, B. G., & Astrachan-Fletcher, E. B. (2009). Locus of control: Relation to schizophrenia, to recovery, and to depression and psychosis—A 15-year longitudinal study. *Psychiatry research*, *168*(3), 186-192.

Heilman, K. M., Rothi, L. J., & Valenstein, E. (1982). Two forms of ideomotor apraxia. *Neurology*, *32*(4), 342-342.

Hiroto, D. S., & Seligman, M. E. (1975). Generality of learned helplessness in man. *Journal of personality and social psychology*, *31*(2), 311.

Hirvonen, J., van Erp, T. G., Huttunen, J., Aalto, S., Någren, K., Huttunen, M., Lönnqvist, J., Kaprio, J., Hietala, J., & Cannon, T. D. (2005). Increased caudate dopamine D2 receptor availability as a genetic marker for schizophrenia. *Archives of General Psychiatry*, *62*(4), 371-378.

Hunt, L., Daw, N., Kaanders, P., MacIver, M., Mugan, U., Procyk, E., Redish, A., Russo, E., Scholl, J., & Stachenfeld, K. (2021). Formalizing planning and information search in naturalistic decision-making. *Nature neuroscience*, *24*(8), 1051-1064.

Huys, Q. J., & Dayan, P. (2009). A Bayesian formulation of behavioral control. *Cognition*, *113*(3), 314-328.

Huys, Q. J., Renz, D., Petzschner, F., Berwian, I., Stoppel, C., & Haker, H. (2016). German translation and validation of the cognitive style questionnaire short form (CSQ-SF-D). *PLoS One*, *11*(3), e0149530.

Imai, H., Steindler, D., & Kitai, S. (1986). The organization of divergent axonal projections from the midbrain raphe nuclei in the rat. *Journal of Comparative Neurology*, *243*(3), 363-380.

Jeong, H., Taylor, A., Floeder, J. R., Lohmann, M., Mihalas, S., Wu, B., Zhou, M., Burke, D. A., & Namboodiri, V. M. K. (2022). Mesolimbic dopamine release conveys causal associations. *Science*, *378*(6626), eabq6740.

Jezzini, A., Bromberg-Martin, E. S., Trambaiolli, L. R., Haber, S. N., & Monosov, I. E. (2021). A prefrontal network integrates preferences for advance information about uncertain rewards and punishments. *Neuron*, *109*(14), 2339-2352. e2335.

Jocham, G., Brodersen, K. H., Constantinescu, A. O., Kahn, M. C., Ianni, A. M., Walton, M. E., Rushworth, M. F., & Behrens, T. E. (2016). Reward-guided learning with and without causal attribution. *Neuron*, *90*(1), 177-190.

Juckel, G., Schlagenhauf, F., Koslowski, M., Wüstenberg, T., Villringer, A., Knutson, B., Wrase, J., & Heinz, A. (2006). Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage*, *29*(2), 409-416.

Kaanders, P., Nili, H., O'Reilly, J. X., & Hunt, L. (2021). Medial frontal cortex activity predicts information sampling in economic choice. *Journal of Neuroscience*, *41*(40), 8403-8413.

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of artificial intelligence research*, *4*, 237-285.

Kashdan, T. B., Disabato, D. J., Goodman, F. R., & McKnight, P. E. (2020). The Five-Dimensional Curiosity Scale Revised (5DCR): Briefer subscales while separating overt and covert social curiosity. *Personality and individual differences*, *157*, 109836.

Keren, H., O'Callaghan, G., Vidal-Ribas, P., Buzzell, G. A., Brotman, M. A., Leibenluft, E., Pan, P. M., Meffert, L., Kaiser, A., & Wolke, S. (2018). Reward processing in depression: a conceptual and meta-analytic review across fMRI and EEG studies. *American Journal of Psychiatry*, *175*(11), 1111-1120.

Kety, S. S., Rosenthal, D., Wender, P. H., Schulsinger, F., & Jacobsen, B. (1976). Mental illness in the biological and adoptive families of adopted individuals who have become schizophrenic. *Behavior Genetics*, *6*, 219-225.

Khalighinejad, N., Garrett, N., Priestley, L., Lockwood, P., & Rushworth, M. F. S. (2021). A habenula-insular circuit encodes the willingness to act. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-26569-1

Kinderman, P., & Bentall, R. P. (1996). A new measure of causal locus: the internal, personal and situational attributions questionnaire. *Personality and individual differences*, *20*(2), 261-264.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science*, *273*(5280), 1399-1402.

Króliczak, G., & Frey, S. H. (2009). A common network in the left cerebral hemisphere represents planning of tool use pantomimes and familiar intransitive gestures at the hand-independent level. *Cerebral Cortex*, *19*(10), 2396-2410.

Kumar, P., Goer, F., Murray, L., Dillon, D. G., Beltzer, M. L., Cohen, A. L., Brooks, N. H., & Pizzagalli, D. A. (2018). Impaired reward prediction error encoding and striatal-midbrain connectivity in depression. *Neuropsychopharmacology*, *43*(7), 1581-1588.

Kumar, P., Waiter, G., Ahearn, T., Milders, M., Reid, I., & Steele, J. D. (2008). Abnormal temporal difference reward-learning signals in major depression. *Brain*, *131*(Pt 8), 2084-2093. https://doi.org/10.1093/brain/awn136

Kushnir, T., Gopnik, A., Schulz, L., & Danks, D. (2003). Inferring hidden causes. Proceedings of the annual meeting of the cognitive science society,

Lagnado, D. A., & Sloman, S. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(4), 856.

Lange, K., Kühn, S., & Filevich, E. (2015). " Just another tool for online studies"(JATOS): An easy solution for setup and management of web servers supporting online studies. *PLoS One*, *10*(6), e0130834.

Lavallé, L., Dondé, C., Gawęda, Ł., Brunelin, J., & Mondino, M. (2021). Impaired self-recognition in individuals with no full-blown psychotic symptoms represented across the continuum of psychosis: A meta-analysis. *Psychological Medicine*, *51*(16), 2864-2874.

Lenzenweger, M. F. (2011). *Schizotypy and schizophrenia: The view from experimental psychopathology*. Guilford Press.

Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., & Niv, Y. (2017). Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron*, *93*(2), 451-463.

Leotti, L. A., & Delgado, M. R. (2011). The inherent reward of choice. *Psychological science*, *22*(10), 1310-1318.

Liebowitz, M. R. (1987). Social phobia. *Modern problems of pharmacopsychiatry*, *22*(141), e173.

Ligneul, R. (2021). Prediction or causation? Towards a redefinition of task controllability. *Trends in cognitive sciences*, *25*(6), 431-433.

Ligneul, R., Mainen, Z. F., Ly, V., & Cools, R. (2022). Stress-sensitive inference of task controllability. *Nature Human Behaviour*, *6*(6), 812-822.

Lindner, A., Thier, P., Kircher, T. T., Haarmeier, T., & Leube, D. T. (2005). Disorders of agency in schizophrenia correlate with an inability to compensate for the sensory consequences of actions. *Current Biology*, *15*(12), 1119-1124.

Lockwood, P. L., Apps, M. A., & Chang, S. W. (2020). Is there a 'social'brain? Implementations and algorithms. *Trends in cognitive sciences*, *24*(10), 802-813.

Lockwood, P. L., Wittmann, M. K., Apps, M. A. J., Klein-Flügge, M. C., Crockett, M. J., Humphreys, G. W., & Rushworth, M. F. S. (2018). Neural mechanisms for learning self and other ownership. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-07231-9

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*, 1122-1135.

Maier, S. F., & Seligman, M. E. (1976). Learned helplessness: theory and evidence. *Journal of Experimental Psychology: General*, *105*(1), 3.

Maier, S. F., & Seligman, M. E. (2016). Learned helplessness at fifty: Insights from neuroscience. *Psychological review*, *123*(4), 349.

Majchrowicz, B., & Wierzchoń, M. (2018). Unexpected action outcomes produce enhanced temporal binding but diminished judgement of agency. *Consciousness and cognition*, *65*, 310-324.

Malenka, R. C., Angel, R. W., Hampton, B., & Berger, P. A. (1982). Impaired central error-correcting behavior in schizophrenia. *Archives of General Psychiatry*, *39*(1), 101-107.

Mancinelli, F., Roiser, J., & Dayan, P. (2021). Internality and the internalisation of failure: Evidence from a novel task. *PLoS computational biology*, *17*(7), e1009134. https://doi.org/10.1371/journal.pcbi.1009134

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods*, *164*(1), 177-190.

Mars, R. B., Jbabdi, S., Sallet, J., O'Reilly, J. X., Croxson, P. L., Olivier, E., Noonan, M. P., Bergmann, C., Mitchell, A. S., & Baxter, M. G. (2011). Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. *Journal of Neuroscience*, *31*(11), 4087-4100.

Martin, D. J., Abramson, L. Y., & Alloy, L. B. (1984). Illusion of control for self and others in depressed and nondepressed college students. *Journal of personality and social psychology*, *46*(1), 125.

Mason, O., Linney, Y., & Claridge, G. (2005). Short scales for measuring schizotypy. *Schizophrenia research*, *78*(2-3), 293-296.

Maswood, S., Barter, J. E., Watkins, L. R., & Maier, S. F. (1998). Exposure to inescapable but not escapable shock increases extracellular levels of 5-HT in the dorsal raphe nucleus of the rat. *Brain Research*, *783*(1), 115-120.

Mata, I., Gilvarry, C. M., Jones, P. B., Lewis, S. W., Murray, R. M., & Sham, P. C. (2003). Schizotypal personality traits in nonpsychotic relatives are associated with positive symptoms in psychotic probands. *Schizophrenia bulletin*, *29*(2), 273-283.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, *38*(2), 339-346.

McFarland, N. R., & Haber, S. N. (2000). Convergent inputs from thalamic motor nuclei and frontal cortical areas to the dorsal striatum in the primate. *Journal of Neuroscience*, *20*(10), 3798-3813.

McGuire, P., David, A., Murray, R., Frackowiak, R., Frith, C., Wright, I., & Silbersweig, D. (1995). Abnormal monitoring of inner speech: a physiological basis for auditory hallucinations. *The Lancet*, *346*(8975), 596-600.

Meehl, P. E. (1962). Schizotaxia, schizotypy, schizophrenia. In *Schizophrenia* (pp. 21-46). Routledge.

Melis, A. P., & Raihani, N. J. (2023). The cognitive challenges of cooperation in human and nonhuman animals. *Nature Reviews Psychology*, *2*(9), 523-536.

Mellor, C. S. (1970). First rank symptoms of schizophrenia: I. the frequency in schizophrenics on admission to hospital II. Differences between individual first rank symptoms. *The British Journal of Psychiatry*, *117*(536), 15-23.

Mill, J. S. (1843). Philosophy of Scientific Method.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, *82*(2), 213.

Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the IRE*, *49*(1), 8-30.

Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. (2014). Unwritten rules: virtual bargaining underpins social interaction, culture, and society. *Trends in cognitive sciences*, *18*(10), 512-519. https://doi.org/10.1016/j.tics.2014.05.010

Mogg, K., Bradley, B. P., & Williams, R. (1995). Attentional bias in anxiety and depression: The role of awareness. *British journal of clinical psychology*, *34*(1), 17-36.

Monosov, I. E. (2024). Curiosity: primate neural circuits for novelty and information seeking. *Nature Reviews Neuroscience*, *25*(3), 195-208.

Monosov, I. E., & Rushworth, M. F. (2022). Interactions between ventrolateral prefrontal and anterior cingulate cortex during learning and behavioural change. *Neuropsychopharmacology*, *47*(1), 196-210.

Moore, J. W., Tang, T. C., & Cioffi, M. C. (2021). Schizotypy and the vicarious experience of agency. *Psychology of Consciousness: Theory, Research, and Practice*, *8*(1), 65.

Moritz, S., Thompson, S. C., & Andreou, C. (2014). Illusory control in schizophrenia. *Journal of Experimental Psychopathology*, *5*(2), 113-122.

Moscarello, J. M., & Hartley, C. A. (2017). Agency and the calibration of motivated behavior. *Trends in cognitive sciences*, *21*(10), 725-735.

Msetfi, R. M., Kumar, P., Harmer, C. J., & Murphy, R. A. (2016). SSRI enhances sensitivity to background outcomes and modulates response rates: A randomized double blind study of instrumental action and depression. *Neurobiology of learning and memory*, *131*, 76-82.

Msetfi, R. M., Murphy, R. A., & Simpson, J. (2007). Depressive realism and the effect of intertrial interval on judgements of zero, positive, and negative contingencies. *Quarterly journal of experimental psychology*, *60*(3), 461-481.

Msetfi, R. M., Murphy, R. A., Simpson, J., & Kornbrot, D. E. (2005). Depressive realism and outcome density bias in contingency judgments: the effect of the context and intertrial interval. *Journal of Experimental Psychology: General*, *134*(1), 10.

Mucci, A., Dima, D., Soricelli, A., Volpe, U., Bucci, P., Frangou, S., Prinster, A., Salvatore, M., Galderisi, S., & Maj, M. (2015). Is avolition in schizophrenia associated with a deficit of dorsal

caudate activity? A functional magnetic resonance imaging study during reward anticipation and feedback. *Psychological medicine*, *45*(8), 1765-1778.

Mueller, H., Betz, L. T., & Bechdolf, A. (2021). A comprehensive meta-analysis of the self-serving bias in schizophrenia spectrum disorders compared to non-clinical subjects. *Neuroscience & Biobehavioral Reviews*, *120*, 542-549.

Murphy, P., Bentall, R. P., Freeman, D., O'Rourke, S., & Hutton, P. (2018). The paranoia as defence model of persecutory delusions: a systematic review and meta-analysis. *The Lancet Psychiatry*, *5*(11), 913-929.

Murphy, R. A., Vallée-Tourangeau, F., Msetfi, R., & Baker, A. G. (2005). Signal-outcome contingency, contiguity, and the depressive realism effect.

Na, S., Blackmore, S., Chung, D., O'Brien, M., Banker, S. M., Heflin, M., Fiore, V. G., & Gu, X. (2022). Computational mechanisms underlying illusion of control in delusional individuals. *Schizophrenia research*, *245*, 50-58.

National Institute for Health Research. (2014). Patient and public involvement in health and social care research: A handbook for researchers. In: National Institute for Health Research London, UK.

Ng, T. H., Alloy, L. B., & Smith, D. V. (2019). Meta-analysis of reward processing in major depressive disorder reveals distinct abnormalities within the reward circuit. *Translational psychiatry*, *9*(1), 293.

Niebuhr, R. (1943). The serenity prayer. *Bulletin of the Federal Council of Churches*, *1*(1), 1.

Nielsen, M. Ø., Rostrup, E., Wulff, S., Bak, N., Lublin, H., Kapur, S., & Glenthøj, B. (2012). Alterations of the brain reward system in antipsychotic naïve schizophrenia patients. *Biol Psychiatry*, *71*(10), 898-905.

Niu, G.-f., Shi, X.-h., Yao, L.-s., Yang, W.-c., Jin, S.-y., & Xu, L. (2023). Social exclusion and depression among undergraduate students: the mediating roles of rejection sensitivity and social self-efficacy. *Current Psychology*, *42*(28), 24198-24207.

Noonan, M. P., Chau, B. K., Rushworth, M. F., & Fellows, L. K. (2017). Contrasting effects of medial and lateral orbitofrontal cortex lesions on credit assignment and decision-making in humans. *Journal of Neuroscience*, *37*(29), 7023-7035.

Norbury, A., Hauser, T. U., Fleming, S. M., Dolan, R. J., & Huys, Q. J. (2024). Different components of cognitive-behavioral therapy affect specific cognitive mechanisms. *Science advances*, *10*(13), eadk3222.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*(5669), 452-454.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329-337.

Ohata, R., Asai, T., Kadota, H., Shigemasu, H., Ogawa, K., & Imamizu, H. (2020). Sense of agency beyond sensorimotor process: decoding self-other action attribution in the human brain. *Cerebral Cortex*, *30*(7), 4076-4091.

Orban, G. A., & Caruana, F. (2014). The neural basis of human tool use. *Frontiers in Psychology*, *5*, 310.

Overmier, J. B., & Seligman, M. E. (1967). Effects of inescapable shock upon subsequent escape and avoidance responding. *Journal of comparative and physiological psychology*, *63*(1), 28.

Papakostas, G. I. (2006). Dopaminergic-based pharmacotherapies for depression. *European Neuropsychopharmacology*, *16*(6), 391-402.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review*, *87*(6), 532.

Pearl, J. (1999). Graphs, Causality, and Structural Equation Models. In *Research Methodology in the Social, Behavioural and Life Sciences: Designs, Models and Methods*. SAGE Publications, Ltd.

Pearl, J. (2009). *Causality*. Cambridge university press.

Pedrelli, P., Blais, M. A., Alpert, J. E., Shelton, R. C., Walker, R. S., & Fava, M. (2014). Reliability and validity of the Symptoms of Depression Questionnaire (SDQ). *CNS spectrums*, *19*(6), 535-546.

Peeters, R., Simone, L., Nelissen, K., Fabbri-Destro, M., Vanduffel, W., Rizzolatti, G., & Orban, G. A. (2009). The representation of tool use in humans and monkeys: common and uniquely human features. *Journal of Neuroscience*, *29*(37), 11523-11539.

Peeters, R. R., Rizzolatti, G., & Orban, G. A. (2013). Functional properties of the left parietal tool use region. *Neuroimage*, *78*, 83-93.

Pike, A. C., & Robinson, O. J. (2022). Reinforcement learning in patients with mood and anxiety disorders vs control individuals: A systematic review and meta-analysis. *JAMA psychiatry*, *79*(4), 313-322.

Pinkham, A. E., Penn, D. L., Green, M. F., Buck, B., Healey, K., & Harvey, P. D. (2014). The social cognition psychometric evaluation study: results of the expert survey and RAND panel. *Schizophrenia bulletin*, *40*(4), 813-823.

Pisauro, M. A., Fouragnan, E. F., Arabadzhiyska, D. H., Apps, M. A. J., & Philiastides, M. G. (2022). Neural implementation of computational mechanisms underlying the continuous trade-off between cooperation and competition. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-34509-w

Quinless, F. W., & Nelson, M. A. M. (1988). Development of a measure of learned helplessness. *Nursing Research*, *37*(1), 11-15.

Raab, H. A., Foord, C., Ligneul, R., & Hartley, C. A. (2022). Developmental shifts in computations used to detect environmental controllability. *PLoS computational biology*, *18*(6), e1010120.

Raine, A. (2006). Schizotypal personality: neurodevelopmental and psychosocial trajectories. *Annu. Rev. Clin. Psychol.*, *2*(1), 291-326.

Reader, S. M., Hager, Y., & Laland, K. N. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1017-1027.

Ree, M. J., French, D., MacLeod, C., & Locke, V. (2008). Distinguishing cognitive and somatic dimensions of state and trait anxiety: Development and validation of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA). *Behavioural and Cognitive Psychotherapy*, *36*(3), 313-332.

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and social psychology bulletin*, *27*(2), 151-161.

Rouault, M., Dayan, P., & Fleming, S. M. (2019). Forming global estimates of self-performance from local confidence. *Nature Communications*, *10*(1), 1141.

Rouault, M., Weiss, A., Lee, J. K., Drugowitsch, J., Chambon, V., & Wyart, V. (2022). Controllability boosts neural and cognitive signatures of changes-of-mind in uncertain environments. *Elife*, *11*, e75038.

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, *15*(8), 549-562.

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., Markowitz, J. C., Ninan, P. T., Kornstein, S., & Manber, R. (2003). The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry*, *54*(5), 573-583.

Rushworth, M. F., Ellison, A., & Walsh, V. (2001a). Complementary localization and lateralization of orienting and motor attention. *Nature neuroscience*, *4*(6), 656-661.

Rushworth, M. F., Nixon, P. D., Renowden, S., Wade, D. T., & Passingham, R. E. (1997). The left parietal cortex and motor attention. *Neuropsychologia*, *35*(9), 1261-1273.

Rushworth, M. F., Paus, T., & Sipila, P. K. (2001b). Attention systems and the organization of the human parietal cortex. *Journal of Neuroscience*, *21*(14), 5262-5271.

Rutledge, R. B., Dean, M., Caplin, A., & Glimcher, P. W. (2010). Testing the reward prediction error hypothesis with an axiomatic model. *Journal of Neuroscience*, *30*(40), 13525-13536.

Sallet, J., Mars, R. B., Noonan, M., Andersson, J. L., O'reilly, J., Jbabdi, S., Croxson, P. L., Jenkinson, M., Miller, K. L., & Rushworth, M. F. (2011). Social network size affects neural circuits in macaques. *Science*, *334*(6056), 697-700.

Sallet, J., Noonan, M. P., Thomas, A., O'Reilly, J. X., Anderson, J., Papageorgiou, G. K., Neubert, F. X., Ahmed, B., Smith, J., & Bell, A. H. (2020). Behavioral flexibility is associated with changes in structure and function distributed across a frontal cortical network in macaques. *PLoS biology*, *18*(5), e3000605.

Santini, Z. I., Koyanagi, A., Tyrovolas, S., Mason, C., & Haro, J. M. (2015). The association between social relationships and depression: A systematic review. *Journal of affective disorders*, *175*, 53-65.

Sapp, S. G., & Harrod, W. J. (1993). Brief locus of control scale. *Psychological Reports*.

Sato, A., & Yasuda, A. (2005). Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition*, *94*(3), 241-255.

Schlagenhauf, F., Huys, Q. J., Deserno, L., Rapp, M. A., Beck, A., Heinze, H.-J., Dolan, R., & Heinz, A. (2014). Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *Neuroimage*, *89*, 171-180.

Scholl, J., Kolling, N., Nelissen, N., Browning, M., Rushworth, M. F., & Harmer, C. J. (2017). Beyond negative valence: 2-week administration of a serotonergic antidepressant enhances both reward and effort learning signals. *PLoS biology*, *15*(2), e2000756.

Scholl, J., Kolling, N., Nelissen, N., Wittmann, M. K., Harmer, C. J., & Rushworth, M. F. (2015). The Good, the Bad, and the Irrelevant: Neural Mechanisms of Learning Real and Hypothetical Rewards and Effort. *J Neurosci*, *35*(32), 11233-11251.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593-1599.

Selemon, L., & Goldman-Rakic, P. (1985). Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *Journal of Neuroscience*, *5*(3), 776-794.

Selemon, L. D., & Goldman-Rakic, P. S. (1988). Common cortical and subcortical targets of the dorsolateral prefrontal and posterior parietal cortices in the rhesus monkey: evidence for a distributed neural network subserving spatially guided behavior. *Journal of Neuroscience*, *8*(11), 4049-4068.

Seligman, M. E. (1972). Learned helplessness. *Annual review of medicine*, *23*(1), 407-412.

Seligman, M. E. (1974). *Depression and learned helplessness*. John Wiley & Sons.

Seligman, M. E., & Beagley, G. (1975). Learned helplessness in the rat. *Journal of comparative and physiological psychology*, *88*(2), 534.

Seligman, M. E., & Maier, S. F. (1967). Failure to escape traumatic shock. *Journal of experimental psychology*, *74*(1), 1.

Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2005). Evidence for sensory prediction deficits in schizophrenia. *American Journal of Psychiatry*, *162*(12), 2384-2386.

Shergill, S. S., White, T. P., Joyce, D. W., Bays, P. M., Wolpert, D. M., & Frith, C. D. (2014). Functional magnetic resonance imaging of impaired sensory prediction in schizophrenia. *JAMA psychiatry*, *71*(1), 28-35.

Shultz, S., Noë, R., McGraw, W. S., & Dunbar, R. (2004). A community–level evaluation of the impact of prey behavioural and ecological characteristics on predator diet composition. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1540), 725-732.

Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the society for research in child development*, 1-51.

Sirigu, A., Daprati, E., Pradat-Diehl, P., Franck, N., & Jeannerod, M. (1999). Perception of self-generated movement following left parietal lesion. *Brain*, *122*(10), 1867-1874.

Skilbeck, L., Spanton, C., & Roylance, I. (2020). Helping clients 'restart their engine'–use of in-session cognitive behavioural therapy behavioural experiments for engagement and treatment in persistent depression: a case study. *The Cognitive Behaviour Therapist*, *13*, e5.

Sloman, S. A. (2005). *Causal models : how people think about the world and its alternatives*. Oxford University Press.

Smith, R., Taylor, S., Wilson, R. C., Chuning, A. E., Persich, M. R., Wang, S., & Killgore, W. D. (2022). Lower levels of directed exploration and reflective thinking are associated with greater anxiety and depression. *Frontiers in Psychiatry*, *12*, 782136.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., & Flitney, D. E. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23*, S208-S219.

Snaith, R. P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., & Trigwell, P. (1995). A scale for the assessment of hedonic tone the Snaith–Hamilton Pleasure Scale. *The British Journal of Psychiatry*, *167*(1), 99-103.

Song, M., Baah, P. A., Cai, M. B., & Niv, Y. (2022). Humans combine value learning and hypothesis testing strategically in multi-dimensional probabilistic reward learning. *PLoS computational biology*, *18*(11), e1010699.

Spence, S. A., Brooks, D. J., Hirsch, S. R., Liddle, P. F., Meehan, J., & Grasby, P. M. (1997). A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control). *Brain: a journal of neurology*, *120*(11), 1997-2011.

Spiering, L., Trier, H., O'Reilly, J., Kolling, N., Wittmann, M., Rushworth, M., & Scholl, J. (2024). Controllability and cause in human collaboration. *BioRxiv*, 2024.2010. 2001.615833.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive science*, *27*(3), 453-489.

Stoll, F. M., Fontanier, V., & Procyk, E. (2016). Specific frontal neural dynamics contribute to decisions to check. *Nature Communications*, *7*(1), 11990.

Sun, X., Zhang, Y., Niu, G., Tian, Y., Xu, L., & Duan, C. (2021). Ostracism and problematic smartphone use: the mediating effect of social self-efficacy and moderating effect of rejection sensitivity. *International Journal of Mental Health and Addiction*, 1-14.

Surguladze, S. A., Young, A. W., Senior, C., Brébion, G., Travis, M. J., & Phillips, M. L. (2004). Recognition accuracy and response bias to happy and sad facial expressions in patients with major depression. *Neuropsychology*, *18*(2), 212.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Suzuki, S., Harasawa, N., Ueno, K., Gardner, L., Justin, Ichinohe, N., Haruno, M., Cheng, K., & Nakahara, H. (2012). Learning to Simulate Others' Decisions. *Neuron*, *74*(6), 1125-1137. https://doi.org/10.1016/j.neuron.2012.04.030

Synofzik, M., Thier, P., Leube, D. T., Schlotterbeck, P., & Lindner, A. (2010). Misattributions of agency in schizophrenia are based on imprecise predictions about the sensory consequences of one's actions. *Brain*, *133*(1), 262-271.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin*, *103*(2), 193.

Thorndike, E. L. (1933). A proof of the law of effect. *Science*, *77*(1989), 173-175.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675-691. https://doi.org/10.1017/s0140525x05000129

Tomassini, V., Jbabdi, S., Klein, J. C., Behrens, T. E., Pozzilli, C., Matthews, P. M., Rushworth, M. F., & Johansen-Berg, H. (2007). Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. *Journal of Neuroscience*, *27*(38), 10259-10269.

Toni, I., & Passingham, R. E. (1999). Prefrontal-basal ganglia pathways are involved in the learning of arbitrary visuomotor associations: a PET study. *Experimental brain research*, *127*, 19-32.

Touchette, H., & Lloyd, S. (2004). Information-theoretic approach to the study of control systems. *Physica A: Statistical Mechanics and its Applications*, *331*(1-2), 140-172.

Trudel, N., Lockwood, P. L., Rushworth, M. F., & Wittmann, M. K. (2023). Neural activity tracking identity and confidence in social information. *elife*, *12*, e71315.

Trudel, N., Scholl, J., Klein-Flügge, M. C., Fouragnan, E., Tankelevitch, L., Wittmann, M. K., & Rushworth, M. F. (2021). Polarity of uncertainty representation during exploration and exploitation in ventromedial prefrontal cortex. *Nature Human Behaviour*, *5*(1), 83-98.

Van der Kooy, D., & Hattori, T. (1980). Dorsal raphe cells with collateral projections to the caudate-putamen and substantia nigra: a fluorescent retrograde double labeling study in the rat. *Brain Research*, *186*(1), 1-7.

Vázquez, C. (1987). Judgment of contingency: cognitive biases in depressed and nondepressed subjects. *Journal of personality and social psychology*, *52*(2), 419.

Venkatasubramanian, G., Jayakumar, P. N., Keshavan, M. S., & Gangadhar, B. N. (2011). Schneiderian first rank symptoms and inferior parietal lobule cortical thickness in antipsychotic-naive schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(1), 40-46.

Voss, M., Chambon, V., Wenke, D., Kühn, S., & Haggard, P. (2017). In and out of control: brain mechanisms linking fluency of action selection to self-agency in patients with schizophrenia. *Brain*, *140*(8), 2226-2239.

Voss, M., Moore, J., Hauser, M., Gallinat, J., Heinz, A., & Haggard, P. (2010). Altered awareness of action in schizophrenia: a specific deficit in predicting action consequences. *Brain*, *133*(10), 3104-3112.

Walton, M. E., Behrens, T. E., Buckley, M. J., Rudebeck, P. H., & Rushworth, M. F. (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron*, *65*(6), 927-939. https://doi.org/10.1016/j.neuron.2010.02.027

Waltz, J. A., Wilson, R. C., Albrecht, M. A., Frank, M. J., & Gold, J. M. (2020). Differential effects of psychotic illness on directed and random exploration. *Computational Psychiatry (Cambridge, Mass.)*, *4*, 18.

Wegner, D. M., Sparrow, B., & Winerman, L. (2004). Vicarious agency: experiencing control over the movements of others. *Journal of personality and social psychology*, *86*(6), 838.

Weiss, A., Chambon, V., Lee, J. K., Drugowitsch, J., & Wyart, V. (2021). Interacting with volatile environments stabilizes hidden-state inference and its brain signatures. *Nature Communications*, *12*(1), 2228.

Wen, W., & Imamizu, H. (2022). The sense of agency in perception, behaviour and human–machine interactions. *Nature Reviews Psychology*, *1*(4), 211-222.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, *143*(6), 2074.

Wittmann, M. K., Fouragnan, E., Folloni, D., Klein-Flügge, M. C., Chau, B. K., Khamassi, M., & Rushworth, M. F. (2020). Global reward state affects learning and activity in raphe nucleus and anterior insula in monkeys. *Nature Communications*, *11*(1), 3771.

Wittmann, M. K., Kolling, N., Akaishi, R., Chau, B. K., Brown, J. W., Nelissen, N., & Rushworth, M. F. (2016a). Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nature Communications*, *7*(1), 12327.

Wittmann, M. K., Kolling, N., Faber, N. S., Scholl, J., Nelissen, N., & Rushworth, M. F. (2016b). Self-Other Mergence in the Frontal Cortex during Cooperation and Competition. *Neuron*, *91*(2), 482-493. https://doi.org/10.1016/j.neuron.2016.06.022

Wittmann, M. K., Lockwood, P. L., & Rushworth, M. F. S. (2018). Neural Mechanisms of Social Cognition in Primates. *Annual Review of Neuroscience*, *41*(1), 99-118. https://doi.org/10.1146/annurev-neuro-080317-061450

Wittmann, M. K., Trudel, N., Trier, H. A., Klein-Flügge, M. C., Sel, A., Verhagen, L., & Rushworth, M. F. S. (2021). Causal manipulation of self-other mergence in the dorsomedial prefrontal cortex. *Neuron*, *109*(14), 2353-2361.e2311. https://doi.org/10.1016/j.neuron.2021.05.027

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880-1882.

World Health Organization (2017). Depression and other common mental disorders: global health estimates.

Worsley, K. J. (2001). Statistical analysis of activation images. *Functional MRI: An introduction to methods*, *14*(1), 251-270.

Zamfir, E., & Dayan, P. (2022). Interactions between attributions and beliefs at trial-by-trial level: Evidence from a novel computer game task. *PLoS computational biology*, *18*(9), e1009920. https://doi.org/10.1371/journal.pcbi.1009920

Zhang, W.-N., Chang, S.-H., Guo, L.-Y., Zhang, K.-L., & Wang, J. (2013). The neural correlates of reward-related processing in major depressive disorder: a meta-analysis of functional magnetic resonance imaging studies. *Journal of affective disorders*, *151*(2), 531-539.

Zhou, S.-Y., Suzuki, M., Takahashi, T., Hagino, H., Kawasaki, Y., Matsui, M., Seto, H., & Kurachi, M. (2007). Parietal lobe volume deficits in schizophrenia spectrum disorders. *Schizophrenia research*, *89*(1-3), 35-48.

Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron*, *42*(3), 509-517.