A Combined Behavioural and Electrophysiological Examination of the Faulty Foraging Theory of Probability Matching

Zack William Ellerby

A thesis submitted for the degree of Doctor of Philosophy

University of Nottingham

Department of Psychology

November 2017

General Abstract

Given a repeated choice between two or more options with fixed, independent and identically distributed reward probabilities, overall pay-offs can be maximized by the exclusive selection of the option with the greatest likelihood of reward. The tendency to match response proportions to reward contingencies is suboptimal. Nevertheless, this behaviour is well documented. This thesis had two core objectives. First, it was aimed to ascertain the relative contributions of several existing accounts of probability matching in determining choice behaviour, particularly regarding the maximization versus diversification of choices. These accounts include failed pattern matching, driven by apophenia, and a heuristic-driven response that can be overruled with sufficient deliberation. Second, to then further address potential mechanisms underlying whichever factor was found to make the most substantive contribution to choice behaviour, through the combined application of behavioural and electrophysiological methods. Over a series of behavioural studies, the use of strategy over intuition proved to be the most consistent and substantial predictor of maximizing choices, in robust support of the heuristic account. Given this finding, the question emerges of why probability matching is the dominant intuitive response. One possibility is that matching represents the overextension of an evolutionarily stable foraging strategy, an account termed the 'Faulty Foraging Theory'. A simple simulated foraging task was designed to assess choice behaviour when reward schedules incorporate the factor of natural resource accumulation and depletion. The contrast in choice behaviour on this task with the standard probability matching preparation was consistent with the Faulty Foraging model. A series of electrophysiological studies were then conducted, in an attempt to uncover the putative illusory internal representation of reward accumulation on the standard task, which could drive suboptimal diversification of choices. This was ultimately unsuccessful. However, further corroborating electrophysiological evidence was obtained for the heuristic account, in the form of characteristic patterns of prefrontal activity relating to maximizing vs. diversification of choices.

Keywords: judgement under uncertainty, probability matching, apophenia, heuristics, Faulty Foraging Theory, electroencephalography, cognitive control

Acknowledgements

I would like to express my deepest gratitude to a number of people who have helped and influenced me in different ways throughout my time completing this thesis. My special thanks must go to my primary and secondary supervisors, Richard Tunney and Markus Bauer, without whose supervision and guidance this would not have been possible. I would also like to convey my great appreciation to my family: Olivia, Joshua, both my own and Olivia's parents, my sister, grandparents and many others for their unfailing practical and moral support. I must also thank my office mates, Richard James, Mike Long, Thomas Colvin, Laura Outhwaite, Sam Doehren and Rabi Al-Khaldi, with whom I have established numerous valued friendships over my time here.

In addition, the suggestions and insights provided by Peter Bibby regarding mediation analysis, and Claudia Danielmeier with respect to feedback related negativity, were greatly appreciated. Technical advice concerning the use of the biosemi EEG system, provided by both Walter van Heuven and Miguel Granja Espirito Santo, was also of great assistance to me, for which I am very grateful. I would also like to thank my internal assessors, Alistair Smith and Peter Chapman, and my external assessor Caren Frosch, for their time, support and helpful recommendations; along with many others, both within and beyond the School of Psychology, who have helped and supported me in various ways over the last three years.

Author's Declaration

The research submitted in this thesis is the author's own original work. Excluding excerpts of Chapter 4, it has not previously been submitted for a degree, in the same or different form, to this or any other institution. Data analysed in Chapter 4 was originally collected for the author's MSc thesis. However, an extended reanalysis was conducted of this electrophysiological data for the purposes of the present thesis, with the inclusion of additional data. The experiments described within Chapter 2 are published in: Ellerby, Z.W., & Tunney, R.J. (2017). The effects of heuristics and apophenia on probabilistic choice. *Advances in Cognitive Psychology*.

Zack Ellerby

Table of Contents

General Abstract **2**

Acknowledgements **3**

Author's Declaration **4**

Table of Contents **5**

Acronym Definitions and Glossary of Commonly Used Terms **9**

Chapter 1: General Introduction **12**

1.a. Probability matching **12**

1.a.i. Apophenia 14

1.a.ii. Heuristics 16

1.a.iii. Faulty Foraging Theory 20

1.a.iv. Alternative Accounts 24

1.b. Electrophysiology and the Neuroscience of Decision Making **29**

1.b.i. Action-Values 31

1.b.ii. Technique Specific Considerations of EEG 34

1.b.iii. Prefrontal Activity as a Proxy for Top-down Cognitive Control 37

1.b.iv. Feedback Related Negativity and Prediction Error 39

1.b.v. Neuroscience of Probability Matching 43

1.c. Thesis Summary **47**

Chapter 2: A Behavioural Assessment of Heuristic and Apophenia Accounts **50**

2.a. Introduction: An Overview of Experiments **50**

Experiment 1 ..... **51**

2.b. Method – Experiment 1 ..... **52**

2.b.i. Participants 52

2.b.ii. Design 52

2.b.iii. Stimuli 53

2.b.iv. Procedure 53

2.b.v. Questionnaire 54

2.c. General Results – Experiment 1 ..... **55**

2.c.i. Subjective estimates of reward probabilities 59

2.c.ii. Heuristics 60

2.c.iii. Apophenia 60

2.c.iv. Alternative factors 61

2.c.v. Win-Stay-Lose-Shift analysis 62

2.d. Discussion – Experiment 1 ..... **65**

Experiment 2 ..... **66**

2.e. Method – Experiment 2 ..... **66**

2.e.i. Participants 66

2.e.ii. Design 66

2.e.iii. Procedure 67

2.f. General Results – Experiment 2 ..... **67**

2.f.i. Subjective estimates of reward probabilities 68

2.f.ii. Heuristics & Apophenia 68

2.f.iii. Comparisons with Experiment 1 68

2.f.iv. Subjective estimates of reward probabilities 69

2.f.v. Heuristics 70

2.f.vi. Apophenia 70

2.g. Discussion – Experiment 2 ..... **71**

2.h. Mediation Analyses **71**
2.i. Assessment of Reliability of Within Group Effects **73**
2.j. Controlling for Variance in Perceived Reward Probabilities: Summary of a Parallel Analysis **74**
2.k. Assessment of Alternative Factors. **76**
2.l. General Discussion **77**

Chapter 3: A Behavioural Assessment of Faulty Foraging Theory on a Simple Simulated Foraging Task **83**
3.a. Introduction **83**
3.a.i. Further motivations and considerations 86
3.b. Method **88**
3.b.i. Participants 88
3.b.ii. Design 88
3.b.iii. Stimuli 89
3.b.iv. Procedure 89
3.b.v. Questionnaire 90
3.c. General Results **91**
3.c.i. Subjective estimates of reward probabilities 93
3.c.ii. Heuristics & Apophenia 93
3.d. Comparisons with Standard Condition of Experiment 1 ..... **94**
3.d.i. Subjective estimates of reward probabilities 96
3.d.ii. Heuristics 96
3.d.iii. Apophenia 96
3.e. Mediation Analyses **97**
3.f. Win-Stay-Lose-Shift **99**
3.g. Local maximising **101**
3.h. Discussion **103**

Chapter 4: Detecting Linear Modulation of Internally Represented Expected Values using Electroencephalography **108**
4.a. Introduction: **108**
4.b. Method **109**
4.b.i. Participants 109
4.b.ii. Equipment 110
4.b.iii. Design 111
4.b.iv. Stimuli 111
4.b.v. Procedure 112
4.c. Behavioural Results **113**
4.d. Electrophysiological Results **115**
4.d.i. Software 115
4.d.ii. Pre-processing 115
4.d.iii. Analysis 116
4.d.iv. Assessment of Linearity of Value-Coding 117
4.e. Discussion **120**

Chapter 5: An Electrophysiological Assessment of the Internal Representation of both Global and Local Expected Values on the SSFT **123**
5.a. Introduction **123**
5.b. Method **124**
5.b.i. Participants 124
5.b.ii. Equipment 124
5.b.iii. Design 127
5.b.iv. Stimuli 128
5.b.v. Procedure 128
5.b.vi. Questionnaire 129

5.c. Behavioural Results **130**
5.c.i. Subjective estimates of reward probabilities 131
5.c.ii. Heuristics & Apophenia 131
5.c.iii. Win-Stay Lose-Shift 132
5.c.iv. Local maximising 132
5.c.v. Behavioural Results Summary 133
5.d. Electroencephalographic Results **134**
5.d.i. Software 134
5.d.ii. Pre-processing 134
5.d.iii. Analysis 135
5.d.iv. Global Generation Probabilities 137
5.d.v. Local Expected Values 138
5.d.vi. Behavioural and Electrophysiological Model Fit Comparison 139
5.d.vii. Descriptive ERP Plots 140
5.e. Discussion **141**

Chapter 6: An Electrophysiological Assessment of Both Real and Putative Illusory Expected Value Related Activity on the Standard Probability Matching Task **148**
6.a. Introduction: **148**
6.a.i. Experimental Hypotheses 152
6.b. Method **153**
6.b.i. Participants 153
6.b.ii. Equipment 153
6.b.iii. Design 154
6.b.iv. Stimuli 155
6.b.v. Procedure 155
6.b.vi. Questionnaire 156
6.c. Behavioural Results **157**
6.c.i. Subjective estimates of reward probabilities 158
6.c.ii. Heuristics & Apophenia 159
6.c.iii. Win-Stay Lose-Shift 160
6.c.iv. Faulty Foraging Model Predictions 161
6.c.v. Mediation Analyses 163
6.c.vi. Behavioural Results Summary 166
6.d. Electroencephalographic Results **169**
6.d.i. Software 169
6.d.ii. Pre-processing 169
6.d.iii. Stimulus-locked Analysis 170
6.d.iv. A Priori Between Groups Prefrontal Contrasts - Stimulus-Locked 171
6.d.v. A Posteriori Between Group Contrasts – Stimulus-Locked 172
6.d.vi. Assessment of Model Interrelationship 173
6.d.vii. Simple Maximizing Model Related Activity – Stimulus-Locked 175
6.d.viii. Faulty Foraging Model Related Activity – Stimulus-Locked 176
6.d.ix Response and Feedback-locked Analysis 176
6.d.x. A Priori Prefrontal Contrasts – Feedback Locked 178
6.d.xi. Stated vs. Learned condition – Feedback Locked 178
6.d.xii. Max vs. Min and Win vs. Lose – Feedback Locked 178
6.d.xiii. Whole-scalp Analysis – Feedback Locked 183
6.d.xiv. Prediction Error Coding – Feedback Locked 189
6.d.xv. Assessment of Lateralization of Observed Effects – Feedback Locked 189
6.d.xvi. Descriptive ERP Plots 193
6.d.xvii. Within Group Descriptive ERP Plots 200
6.d.xviii. Electrophysiological Results Summary 205
6.e. Discussion **217**

Chapter 7: General Discussion 222
7.a. Summary of Key Findings 222
7.b. Potential Alternatives to Faulty Foraging Theory 227
7.c. Practical Implications and Potential Avenues of Future Research 231

References 234

Appendices 248
i. Appendix A: Wording of Questionnaire Items (with the yellow light as the high reward probability option) 248
ii. Appendix B: Task Instructions. Slide 1 249
iii. Appendix C: Task Instructions. Additional Reward Probability Information (only shown in Stated probabilities group) 250
iv. Appendix D: Task Instructions. Additional Pattern Information (only shown in 'Instruction' Apophenia manipulation) 250
v. Appendix E: Statistical Considerations Regarding Potential Violations of Assumptions of ANOVA 250

Acronym Definitions and Glossary of Commonly Used Terms

- **2AFC: Two-Alternative Forced-Choice** - A type of decision task in which repeated choices must be made between two options.
- **ACC: Anterior Cingulate Cortex** - The anterior part of the cingulate cortex, which is found in the medial portion of the cerebral hemispheres and considered part of the limbic lobe.
- **Apophenia:** The tendency to misperceive patterns or rules in random or unconnected sequences of events or stimuli.
- **BOLD: Blood Oxygen Level Dependent** - A contrast used in fMRI to generate brain images, which derives from the local ratio of oxygenated to deoxygenated blood.
- **CRT: Cognitive Reflection Test** - A 3-item task designed to assess participants' ability to deliberatively overrule their initial intuitions.
- **CPP: Centro-parietal Positive Potential** - A centralised ERP signal identified in humans, which relates to the integration of stimulus evidence.
- **Choice Diversification:** The act of not exclusively selecting a single option. This is incompatible with maximizing on a fixed probability and i.i.d reward schedule. Probability matching, overmatching and undermatching each represent different relative levels of choice diversification.
- **EEG: Electroencephalography** - A non-invasive electrophysiological measurement technique, by which electrical potentials are recorded from the scalp.
- **ERP: Event Related Potential** - A stereotyped electrophysiological response that is time-locked to a specific stimulus or event.
- **FFT: Faulty Foraging Theory** - A theory which proposes that probability matching represents the faulty overextension of an otherwise normative foraging strategy.
- **FRN: Feedback Related Negativity** - A relative negative inflection in post-feedback ERP, in relation to negative as compared to positive feedback.
- **fMRI: functional Magnetic Resonance Imaging** - A non-invasive technique that uses differences in relative blood flow within the brain to infer characteristics of neural activity.

- **Heuristic:** A cognitive shortcut, or rule of thumb. These can be used to avoid the investment of greater effortful processing, but sometimes lead to systematic biases in behaviour or choices.
- **IFD: Ideal Free Distribution** - The optimal group level distribution of foraging individuals over space, with the aim of maximizing resource exploitation at the group level and average resource gain for each individual.
- **i.i.d: Independent and identically distributed** - This reflects that the reward probabilities of each option do not depend upon either previous choices or previous outcomes in any manner, nor do they systematically fluctuate over time. All outcomes are drawn from the same probability distribution.
- **LIP: Lateral Intra-Parietal** - An area of cortex that has been found, in monkeys, to stochastically integrate evidence about perceptual stimuli up to a fixed threshold of neural activity, at which point the appropriate motor response is triggered.
- **Local Maximizing:** The exclusive selection of the option with the experienced higher probability of reward over a limited window of past trials. Choices can therefore diverge from those of an absolute maximizing strategy depending upon temporally local fluctuations in actual outcomes.
- **Matching Law:** A rule which predicts that choices are allocated in proportion not to the relative probabilities of reward, but to the actual ratio of previously obtained reward from each option. Though similar in name to probability matching, this law predicts probability maximizing on the standard fixed probability schedule.
- **Overmatching:** The selection of the high probability option at a higher rate than it's ratio of reward, but with some degree of choice diversification. For instance, splitting choices 80:20 when the actual ratio of reward is 70:30.
- **PFC: Pre-Frontal Cortex** - The portion of the cerebral cortex that covers the anterior part of the frontal lobe, often associated with top-down executive control.
- **Probability Matching:** The allocation of choices in proportion to the rates of reward of different options. For instance, splitting choices 70:30 when the actual ratio of reward is 70:30.

- **Probability Maximizing:** The exclusive selection of the high probability option on a fixed and i.i.d reward probability task.
- **PSP: Post-Synaptic Potential** - A change in electrical potential at the membrane of the post-synaptic terminal, arising due to chemical transmission of a signal between neurons, over the synaptic cleft.
- **Undermatching:** The selection of the high probability option at a lower rate than its ratio of reward. For instance, splitting choices 60:40 when the actual ratio of reward is 70:30.
- **VI: Variable Interval** - A reward schedule in which rewards are obtainable according to the time elapsed since the last reward was received, but where this amount of time is not fixed.
- **VR: Variable Ratio** - A reward schedule in which rewards are obtainable according to the number of attempts made to obtain them, but where this required number of attempts is not fixed.
- **SEF: Supplementary Eye-Fields** - An area of the primate cerebral cortex that subserves saccadic eye-movement based response preparation and control.
- **SMA: Supplementary Motor Area** - An area of the primate cerebral cortex that subserves motor response preparation and control.
- **SSFT: Simple Simulated Foraging Task** - A modified version of the standard probability matching task, which is designed both to superficially appear as a foraging situation and to incorporate the natural potential for unharvested rewards to accumulate over time into the reward schedule.
- **WSLS: Win-Stay Lose-Shift** - A one-trial-back choice strategy, in which if the previous trial was rewarded the same choice is made again, but if it was unrewarded then the next choice is switched to another alternative. The most extreme form of local maximizing.

Chapter 1: General Introduction

1.a. Probability matching

Multiple examples of apparently suboptimal human choice behaviour have been empirically documented over the last century (Kahneman, Slovic, & Tversky, 1982; Kahneman & Tversky, 2000; Shafir & LeBoeuf, 2002; Stanovich, 1999). One prominent example is probability matching, whereby people tend to select probabilistically rewarded responses in proportion to the relative likelihoods of reward, rather than the optimal strategy of always selecting the option with the highest reward probability (Shanks, Tunney, & McCarthy, 2002). For instance, if there was a dice with four red sides and two green sides, and it were to be rolled 10 times, the maximum expected number of correct predictions is 6.67, which can be achieved by exclusively predicting that it will land on red. However, most people will instead predict red on only around two-thirds of the rolls, and green on the other third. This will lead to an average of only 5.56 correct predictions.

Probability matching was first documented in rat behaviour under concurrent variable-ratio schedules of reinforcement (Brunswik, 1939). In humans it was first reported in a verbal conditioning paradigm (Grant, Hake, & Hornseth, 1951; Jarvik, 1951) followed by a series of studies in which participants made repeated choices between probabilistic and differentially rewarded options (Goodnow & Postman, 1955; Neimark, 1956). This behavioural tendency, towards the systematic and suboptimal diversification of choices, is now supported by a comprehensive empirical literature in humans (Vulkan, 2000). In other animals the matching of choice proportions to the amount of reward obtained from each option has been consistently shown on concurrent variable-interval schedules. This phenomenon originated the now well-established 'matching law', although observed departures from strict matching have since

required the inclusion of additional paremeters in what is termed the 'generalized matching law' (Davison & McCarthy, 1988; Herrnstein, 1961; Poling, Edwards, Weeden, & Foster, 2011; Vulkan, 2000). It is important to note that, on ratio schedules, the matching law predicts maximizing rather than probability matching, as discussed by Schulze, van Ravenzwaaij and Newell (2017). This prediction has been supported in pigeons, which have been shown, under experimental conditions, to learn to approach maximizing behaviour on concurrent variable-ratio schedules (Herrnstein & Loveland, 1975). In the field of ecology however there is an overwhelming empirical consensus that, in natural environments, animals tend to allocate foraging time in proportion to the amount of resources available in different areas (Weber, 1998). Whether this situation is generally better approximated by a VR schedule, in which rewards depend upon the number of attempts made to obtain them, or a VI schedule, where rewards depend instead upon time elapsed since they were last obtained, is open to question, though there are clearly other factors to take into account as well, such as reward persistence over time and the presence of competitors (see Schulze, van Ravenswaaij & Newell, 2015; 2017). This foraging pattern has been termed the Ideal Free Distribution (Fretwell & Lucas 1970, Fretwell 1972) and is well documented across a multitude of animal species.

Various explanatory accounts have been proposed for probability matching behaviour. Many, though not all, of these accounts fall into two main categories; one possibility is that participants believe that the sequence of outcomes is determined not by chance, but by some relationship to their previous choices or previous outcomes. Probability matching might therefore be based upon a misguided, though not necessarily irrational, search for a pattern in the reward sequence. I refer to this category of model as the Apophenia account. A second category of model suggests that the tendency to match response probabilities to reward probabilities is the result of a heuristic choice strategy

that can be overcome by the application of sufficient deliberative reasoning and deliberation. These accounts are outlined below.

1.a.i. Apophenia

One of the two best-supported accounts is that matching behaviour results from a misguided search for a pattern of sequential dependencies in the reward sequence. Such a pattern could allow for accurate prediction of future outcomes and thus better exploitation of rewards, but in reality does not exist and in practice leads to suboptimal choices. Apophenia is the tendency to perceive illusory patterns in random and unconnected events or stimuli (Conrad, 1958). In humans, this is an empirically well-documented phenomenon (Ayton & Fischer, 2004; Falk & Konold, 1997; Gilovich, Vallone, & Tversky, 1985). Moreover, a number of studies report that participants often perceive patterns in the sequences of outcomes in a standard probability learning task, even when none are present, and that this perception is associated with probability matching. Jarvik (1951) found that over half of participants believed a two-alternative forced-choice (2AFC) verbal-conditioning sequence to derive from some form of pattern. Goodnow (1955) found overmatching in a task framed as gambling, but that participants' behaviour reverted to probability matching when they were instructed to 'search for a principle'. Both Rubinstein (1959) and Peterson and Ulehla (1965) found that participants made significantly more maximizing choices when the randomness of the series had been clearly demonstrated, on card and dice based alternatives of the task respectively. Yellott (1969) introduced a 'noncontingent success' condition in the final block of a standard task, in which participants' choices were always rewarded regardless of the option selected. Under this condition participants generated deterministic and patterned 'superstitious' response sequences, with several participants spontaneously reporting having successfully identified a pattern.

Of course, if the likelihoods of reward were not independent and identically distributed (i.i.d) then the discovery of a predictable pattern in the reward sequence could allow for accurate prediction of future outcomes, potentially leading to greater success than the exclusive selection of the more likely option. In this case pattern matching could be the optimal response. This line of reasoning has led to the development of 'smart' accounts of probability matching. Gaissmaier and Schooler (2008) found that participants with a higher working memory capacity, who also probability matched in the first half of their task (in which there was no pattern) were more likely to discover a pattern that was inserted into the second half of the task. In the majority of studies however, rewards on the standard task *are* serially independent, so any attempt at pattern matching is detrimental to task performance. Nevertheless, participants are not usually told whether or not there is such a pattern in the sequence before the experiment. This means that although pattern-search behaviour is inevitably unsuccessful, it is based upon either a lack or misappraisal of stimulus information, rather than being an inherently irrational choice strategy. It may therefore be considered an overextension of an otherwise normative behaviour.

Of the relatively few recent studies that have tested the contribution of apophenia to probability matching, results have been somewhat inconsistent. Unturbe and Corominas (2007) found that participants who reported more complex rules in the sequence were more likely to probability match than those who did not, while Wolford, Newman, Miller, and Wig (2004) observed that increasing the probability of alternation between rewarded stimuli lead to the perception of greater randomness and, in turn, increased levels of maximizing.

By contrast, a follow-up study, designed specifically to test the 'selective pattern search hypothesis', found strikingly different results (Koehler & James, 2009). In this study participants completed both a learning and test phase. The test phase comprised

the standard 2AFC probability-matching task, framed as the repeated prediction of which of two possible colours of marbles would be drawn from a bag. The learning phase was split into two conditions: Participants either viewed a random sequence of 40 marbles being drawn from a bag one after another (30 green and 10 red), or were instead informed that a total of 30 green and 10 red marbles had been drawn from the bag, but did not observe the sequence in which they had been drawn. Probability matching, in the subsequent test phase, was found to be equally prevalent when participants had no exposure to the reward sequence. The authors concluded that probability matching was not *'readily attributed to a selectively applied pattern search strategy'* (Koehler & James, 2009, pg. 126). Moreover, another finding of this study was that 40% of participants, despite probability matching on the initial task, later acknowledged that a maximising strategy would have led to higher pay-offs when presented with both matching and maximising strategies in an explicit format. In light of this, the authors further expound their results as *'more consistent with probability matching as a mistake rooted in a fast, intuitive response that is not reliably overridden by a more effortful reconsideration'* (Koehler & James, 2009, pg. 126).

1.a.ii. Heuristics

The second category of model assumes that probability matching is a heuristic driven behaviour that can be overruled through deliberative reasoning. The putative role of heuristics in probability matching was illustrated by Kahneman and Frederick (in Gilovich, Griffin, & Kahneman, 2002). This paper introduced the concept of 'attribute substitution': that participants may naturally substitute in the answer to the simpler and more intuitively estimated question of how often each option is likely to win, in place of the more effortful processing required to determine the optimal response strategy in light of this information. This paper also linked probability matching to the 'representativeness' heuristic, proposing that even when they do not misinterpret the

question itself, participants naturally presume that responses should represent the pattern of expected results.

The heuristic account can also be viewed in the framework of dual system models of decision-making, System 1 and System 2; a terminology introduced by Stanovich and West (2000) and popularised by Kahneman (2011). System 1 refers to heuristic based decision processes that are fast, automatic, and relatively effortless, although prone to errors, whereas System 2 processes are slower and generally more accurate, but require much more effort and processing power. System 1 therefore generally determines the prepotent, or default response, so as to conserve resources where possible, while System 2 can be brought to bear when necessary.

A number of empirical results implicate System 1 in matching behaviour, and System 2 in deviation from it. Neimark and Shuford (1959) found that probability matching on a standard task shifted to overmatching when participants could refer back to the results of previous trials during the experiment, or when they were asked to explicitly estimate the reward frequencies of each option, prompting explicit analysis of the outcome probabilities. Similarly, Nies (1962) found significant overmatching when participants were made explicitly aware of stimulus reward probabilities, as did Fantino and Esfandiari (2002). Another manipulation in the latter study found significant overmatching in response to providing a strategy recommendation to another participant midway through the task, requiring the explicit formulation of such a strategy.

Individual differences in heuristic and analytic processing are related to proximal indicators of intelligence. For example, West and Stanovich (2003) found that participants who applied a maximizing strategy had significantly higher SAT scores than those who probability matched. This led them to suggest that the matching response is a non-normative cognitive shortcut that is fast, effortless and relatively intuitive; accounting for their results in that people higher in cognitive ability are more efficient at

deliberative reasoning and thus in identifying the maximizing response. Consistent with this, Koehler and James (2010) found that high scorers on the Cognitive Reflection Test (CRT), a test designed to assess participants' ability to deliberatively overrule their initial intuitions (Frederick, 2005), were much more likely to maximize (74%) than low scorers (11%). They also found both that many more participants subsequently endorsed maximizing as the best strategy than had actually applied it on the preceding task; and that when asked to consider the possible success of both matching and maximizing strategies before the task, participants were significantly more likely to maximize. In light of these results, the authors argue that matching is only the dominant response because it is more readily available than the maximizing strategy. They conclude that matching results from a failure to sufficiently engage in deliberation. Correspondingly, Kogler and Kuehberger (2007) overtly applied dual-process theories in their account of probability matching, or what they term the 'diversification bias'. They propose that System 1 automatically triggers the matching response on the basis of 'expectation matching', which is broadly equivalent to the representativeness heuristic. When this fails to be corrected by System 2 it results in matching behaviour. They found that introducing contextual cues for System 2, by describing the task as a statistical test rather than lottery, led to a decrease of diversification and a concurrent increase in maximising. These results mirror those of Goodnow (1955), who found a shift to significant overmatching when a standard task was re-framed as a gambling one. In addition, a number of studies have shown that increasing performance-based financial incentives, which may encourage participants to engage in effortful analytic processing, leads to a reduction in the number of participants who probability match (Brackbill, Starr, & Kappy, 1962; Shanks et al., 2002; Siegel & Goldstein, 1959).

However, as with the apophenia account, the evidence is not entirely consistent. Jones and Liverant (1960) found significantly greater maximizing in nursery (4-6) than

elementary age children (9-11), who tended to probability match. Derks and Paclisanu (1967) also found an overall positive relationship between age and matching across multiple age groups from nursery to college. These findings led to suggestions that it is maximizing that is the more basic associational response, driven by positive recency, while probability matching results from some form of cognitive control. In the latter paper, studies showing that behaviour eventually approaches maximizing over longer tasks, up to 1000 trials (Derks, 1962; Edwards, 1961), were also interpreted as being due to fatigue engendering more basic association-driven responses as the task progressed.

There are, in turn, issues that can be raised with these studies however. First, Jones and Liverant (1960) randomized outcomes without replacement over every ten trials. This entails that reward probabilities were not fixed or independent, and maximizing was not actually the optimal response. Derks and Paclisanu (1967) also randomized without replacement, although they determined reward sequences over a much longer window of 100 trials. Nevertheless, upon careful reading of their results it becomes apparent that the ostensible age-related decrease in maximizing was driven primarily by the choices of children in the nursery age group. In turn, this was greatly influenced by responses of the ten youngest participants within the nursery sample, of whom eight exclusively predicted hand A for the final 90-197 trials and one hand B for the final 96 trials, while another alternated between the two for the final 198 of 200 responses. Of these children however, only one could actually report which option had been more likely to provide a reward. This suggests that their choices may have been driven either by chance, some factor of experimental design, or inadvertent experimenter influence. In fact, although the number of individuals that probability matched increased with age, this was largely as a result of a corresponding decrease in undermatching. Response perseveration and, in general, progressively higher levels of group-level maximising were evident as age increased above kindergarten level.

1.a.iii. Faulty Foraging Theory

If matching truly is a heuristic response, which can be overruled through the exertion of top-down cognitive control, then the question remains as to why such a systematically suboptimal response pattern is dominant when acting intuitively. One compelling possible answer to this question stems from fundamental differences between the artificial decision tasks, upon which matching has been generally observed, and their more natural counterparts. The repeated choice between two options with fixed reward probabilities, that are each entirely independent of levels of exploitation, is a situation that is rare in nature. It seems therefore highly unlikely for sufficient selection pressure to have had the chance to occur in order for the optimal solution to this specific problem to become genetically instantiated. Moreover, to better approximate problems that were commonly faced over a sustained period of evolutionary ancestry, such as optimal time allocation when foraging, certain additional factors must be taken into account. Key examples of these include competition, generally with conspecifics, along with the potential for unharvested resources to accumulate and deplete over time. The latter may often occur in a manner contingent upon levels of exploitation, which can therefore provide a negative feedback mechanism. Furthermore, accounting for these additional factors can radically alter what constitutes an optimal response pattern.

First, in a social environment, once a certain threshold of rivals is reached that are in competition for a more lucrative food source, it becomes the better option to take advantage of a lesser but unchallenged resource (Gallistel 1990; Gigerenzer & Fiedler, 2007). Moreover, the optimal spread of competitors is in direct proportion to the size of each resource. While natural selection does not generally occur on the group level, the same principle is equally applicable at an individual level once it becomes prevalent in the population. In fact, diversification of choices confers a progressively greater competitive advantage in an environment in which competitors act as overmatchers or, as in the most

extreme case, maximisers. Therefore, in terms of competitive foraging behaviour, there is both a pressure for probability matching to become the dominant tendency and a means for it to remain so, as an evolutionarily stable strategy (see Gigerenzer, 1996; 2000; Skyrms, 1996; 1997).

There is already an extensive empirical literature on the topic of optimal foraging distributions within the field of ecology. Here, a concept termed the 'Ideal Free Distribution', or IFD, (see Fretwell & Lucas 1970, Fretwell 1972) has become well established. Foraging behaviour in accordance with the IFD has been documented across a wide range of species. These include mallards (Harper, 1982), sparrows (Gray, 1994), pigeons (Baum & Kraft, 1998; Bell & Baum, 2002), common cranes (Bautista, Alonso, & Alonso, 1995), rats (Tan et al., 2014), roe deer (Wahlstrom & Kjellander, 1995), white-tailed deer (Kohlmann & Risenhoover, 1997), guppy fish (Abrahams, 1989), cichlid fish (Godin & Keenleyside, 1984; Grand & Grant, 1994; Tregenza & Thompson, 1998), coho salmon (Grand, 1997), wood ants (Lamb & Ollason, 1993), dung flies (Blanckenhorn, Morf, & Reuter, 2000), and also bumblebees, both directly in terms of IFD (Dreisig, 1995) and in the form of analogous 'majoring and minoring' long-term foraging behaviour (Heinrich, 1979). For review, see Weber (1998).

The phenomenon of near optimal group-level matching behaviour under the influence of competition has also been experimentally demonstrated in humans. The majority of studies in this area did so upon a variety of modified simple zero-sum probabilistic choice tasks (Kraft & Baum, 2001; Kraft, Baum, & Burge, 2002; Madden, Peden, & Yamaguchi, 2002; Sokolowski, Tonneau, & Baque, 1999), while Goldstone and Ashpole (2004) utilised a more complex and realistic simulated foraging environment. A recent study has also highlighted individual-level matching behaviour as 'socially adaptive under competition', and sensitive to the strategy employed by simulated competitors (Schulze, van Ravenzwaaij, & Newell, 2015).

It is also important to note that although behaviour has been broadly consistent with the IFD across all of these studies, another common finding has been some degree of undermatching. This has been evident in the behaviour of both humans and other animals. Intriguingly, one study reported substantial undermatching when rewards were probabilistically distributed, while dividing points evenly between participants that had selected an option resulted in a shift to near IFD matching (Kraft et al., 2002).

As touched on before, there are pertinent environmental factors to take into account even when considering an individual in isolation. One such consideration is that many natural resources can become depleted, or even completely exhausted, through over-exploitation. In this case it is no longer optimal to exclusively take advantage of a single resource, even if it begins at a very high-value, but instead to spread foraging behaviours across various options. The optimal allocation depends upon the rates of resource replenishment, but is again broadly proportional to each option's rate of reward.

To demonstrate this point, consider the simplified example of two areas of forest, with one area being more fertile and just over twice as likely to produce an edible item (e.g. fruit, nuts or mushroom) over any given time period. In this instance an average of seven versus three items are generated in each area per day. A forager is able to harvest food from only one of these two areas each day, and aims to allocate their foraging time between the two areas in a manner that maximizes average long-term reward. The first day it is of course best to search within the more productive area. The next day is the same, with the first area still likely to contain more food, despite the fact that the expected number of food items within the less fertile area has now compounded over the two days to six. On the third day however, the sum of food expected to have grown in the unharvested region has reached nine. This surpasses the expected daily generation amount of the more productive area (seven); it is therefore optimal in this instance to forage within the lower fecundity area. The next day, it becomes much better

to search in the more productive area than ever before, as the expected items of food within this area are now fourteen, compared to only three within the less fertile region. Following this point however, the same pattern emerges of the less fertile area being the better option to forage within once every three days. Moreover, it is easy to slightly increase the complexity of this simple model so as to simulate satiety of the forager, through introducing a maximum limit of ten items of food that can be taken from each area per day. Once incorporating this factor, the outcome for optimal foraging between the two areas repeats at an exact 7:3 ratio, closely matching the rates of resource replenishment (see Table 1).

Table 1: Illustration of an optimal foraging sequence in a simplified two area model with resource generation rates of seven versus three and satiety limit of ten items per day.

Trial Number	Expected Items: Area 1	Expected Items: Area 2	Choice (Assumes optimal)
1	3	7	Area 2
2	6	7	Area 2
3	9	7	Area 1
4	3	14	Area 2
5	6	11	Area 2
6	9	8	Area 1
7	3	15	Area 2
8	6	12	Area 2
9	9	9	Area 1
10	3	16	Area 2
11	6	13	Area 2
12	9	10	Area 2
13	12	7	Area 1
14	5	14	Area 2
15	8	11	Area 2

16	11	8	Area 1
17	4	15	Area 2
18	7	12	Area 2
19	10	9	Area 1
20 (repeat 10)	3	16	Area 2
21 (repeat 11)	6	13	Area 2

Millions of generations of evolutionary ancestry have passed in which performance in natural foraging type situations would have been a dominant selection pressure. It seems clear that over such a timescale probability matching could have become established, through natural selection, as what could be considered the 'default' response pattern. By contrast, there has so far been a relatively negligible amount of time in which to adjust, at the genetic level, to the more rigid fixed probability schedules that have become more prevalent in modern day situations, such as those underlying common gambling tasks (e.g. dice, slot-machines, or online equivalents) and which form the basis of the standard 2AFC probability matching task. Together, this provides what I consider to be a natural and compelling account for how probability matching may have become established, and remain to this day, as the prepotent response. In this thesis, this account is termed the 'Faulty Foraging Theory'.

1.a.iv. Alternative Accounts

There also exist a number of additional putative explanations for probability matching behaviour. These will be more briefly summarized here, but their investigation is not the primary focus of this thesis. One of these is that people will optimistically deviate from an optimal rule for a strategy that could in theory get 100% correct, even though in practice this strategy will yield, on average, a lower reward (Arkes, Dawes, & Christensen, 1986). This account is of greatest relevance in situations in which reward

sequences are generated without replacement, as was done in a number of early studies in the area (Fiorina, 1971), but is no longer common practice. In this case it would only be possible to achieve a 100% success rate through selecting the options in proportion to their likelihood of reward, although a higher average score will still be obtained through the exclusive selection of the high probability option. A sole exception, in which diversification of choices could become genuinely optimal, must be made for cases in which reward probabilities are known and towards the end of the task there remain sufficiently more instances of the originally lower probability option than would have been expected by chance. This situation is rare, but can be of particular concern for tasks in which longer reward sequences are determined as a concatenation of multiple shorter sequences, that are in turn each generated without replacement over a relatively small number of trials (cf. Jones and Liverant, 1960).

When considering the standard probability matching preparation, in which rewards are fixed, randomised without replacement and i.i.d, a full run throughout the entire task of exclusive instances of the more likely option, although in itself still highly unlikely to occur, is not only possible, but in fact the single most likely run of trial-by-trial results. In this case, as rewards are independent, even if all preceding trials have been the high probability option, the same remains more likely to occur on each subsequent trial. Similarly, on the standard preparation, the exclusive selection of the more likely option is not only likely to provide the highest overall reward, but is also the most likely strategy to yield a 100% win rate. Nevertheless, even on the standard task participants routinely fail to intuitively understand this, and instead generally assume that a 70 vs. 30 percent split in choices is a more likely, or even necessary provision in order to obtain a 100% success rate. In fact, with respect to the standard task, this bias can be effectively equated to that of 'expectation-matching', rooted in the 'representativeness' heuristic (see Kahneman and Frederick, in Gilovich, Griffin, & Kahneman, 2002). This is true in the

sense that participants assume that a pattern of responses must be 'representative' of the expected outcome sequence in order to be fully successful. Some evidence of this factor providing a non-negligible contribution to choice behaviour, even in the standard preparation, can be drawn from a manipulation conducted by Fantino and Esfandieri (2002). In this study, participants' proportions of maximizing choices were increased when they were directly informed that achieving a 75% success rate was considered a 'perfect' score (the reward probability of the more likely option was .75). However, this information was found only to substantially increase maximizing when paired with explicit information about the actual reward probabilities themselves. It can also be argued that this account is highly interrelated with the apophenia account. Even if participants assume that they can only obtain 100% success through their choices matching the expected average distribution of outcomes, they can still only hope to do so in practice through some rule-based prediction of exactly how these outcomes will be distributed on a trial-by-trial basis. Similarly, although this might not be understood by many participants, the information that 75% is considered a perfect score also includes the implicit information that there will be no predictable pattern that could allow for predicting the trial-by-trial distribution at an above chance level.

One other distinct possible explanation is that it may simply be more exciting to choose the less likely option, and more satisfying to win having done so. Such a win might feel like 'beating the system', and this could give the less probable option an expected utility disproportionate to its simple likelihood of reward. Rubinstein (1959) attempted to test this account through splitting the less likely option into three sub-options. This had been done before, without reference to testing this specific effect and with conflicting results (Cotton & Rechtschaffen, 1958; Gardner, 1957; Neimark, 1956). In any case, it remains unclear how such a manipulation would give exclusive evidence for this specific account of matching. Taking a different approach, two further studies

found evidence that participants were more likely to pick the maximising option having correctly predicted the less likely option on the preceding trial than when they had incorrectly done so (Brackbill & Bravos, 1962; Brackbill et al., 1962). This finding was interpreted as evidence for greater utility having been obtained from correctly predicting the less probable option, though again whether or not this is the most appropriate interpretation is not entirely clear. In addition, a number of studies have independently found that providing increasing levels of monetary incentives for good performance on the task leads to increasing levels of overmatching (Brackbill et al., 1962; Shanks et al., 2002; Siegel & Goldstein, 1959). This can be interpreted as the increase in objective material utility of winning outweighing the putative subjective utility of 'skillfully' predicting the low probability option. However this effect could equally be interpreted as evidence for a heuristic-based account, by which it becomes more worthwhile to engage the effortful deliberation and cognitive control necessary to determine and assert the maximizing response as the stakes increase.

One final account of matching, which is perhaps the most distinct from the other explanations, is that participants do not truly probability match at all; rather they intuitively maximise, but do so according to statistical integration over a relatively short past window. Due to fluctuations in local outcome sequences leading to deviation from to the long-term probability distribution this leads, in practice, to diversification of choices and the emergence, in some cases, of apparent matching behaviour. There is evidence that participants are capable of both perception and reproduction of the sequential statistical structure of apparently random series of binary symbols (Hake & Hyman, 1953). The authors argued that this was best explained by participants' responding being driven by recent sequences of results, rather than any statistical rule, normative or otherwise. Moreover, Hinson and Staddon (1983) argued that 'hill-climbing' rules, such as momentary maximising, are better placed than the matching law

to explain the fact that animals can consistently choose the better of two options in simple tasks, while also accounting for apparent matching behaviour on extended tasks through the 'generalized matching law'.

More recent studies have provided more direct evidence for this account. Yu and Huang (2014) found that while average first-fixation behaviour on a visual-search equivalent of the standard task resembled the matching response in aggregate data, when analysed on local time-scales this behaviour much more closely resembled maximising. They argue that this is evidence for Bayesian inference as the basis of decisions, in their own words, that: '*matching-like stochasticity in human visual search is neither random nor heuristics-based, but attributable specifically to fluctuating beliefs about stimulus statistics.'* (Yu & Huang, 2014, pg. 275).

However, one key aspect of this account is that it does not truly predict probability matching. It is established upon the idea of a limited memory for past events, but only a one-trial-back memory limit (equivalent to a 'win-stay lose-shift' strategy) would directly match responses to reward probabilities. Any longer integration window will necessarily entail some degree of overmatching, with a sufficiently long window leading effectively to maximising. In fact, even relatively short windows would predict significant deviation from matching, for instance a 3-back window on a standard 2AFC task, with reward probabilities of .7 and .3, would anticipate an average proportion of .784 maximising choices. A 5-back window under the same circumstances would predict a proportion of .837 maximising responses, while a 7-back window, consistent with 'Miller's Law' of average working memory capacity (Miller, 1956), would entail a proportion of .874 maximising responses. Although slight overmatching is often found, such strong deviations from the matching response are inconsistent with a large proportion of the experimental literature. This suggests that momentary maximizing,

though an intriguing consideration, is not in itself sufficient to explain the observed phenomenon of probability matching.

1.b. Electrophysiology and the Neuroscience of Decision Making

Much of our modern understanding of the decision-making process at the neurophysiological-level is rooted in findings from perceptual decision-making research conducted in monkeys. In this area multiple studies have identified internal activation patterns which seem to code quantitatively for the stimulus evidence representing each option, in the form of a neural firing rate based salience signal (Gold & Shadlen, 2001; Huk & Shadlen, 2005; Roitman & Shadlen, 2002; Schall, 2001; Shadlen & Newsome, 2001). In monkeys, this activity has been measured predominantly from within the primate lateral intra-parietal cortex (LIP), but candidate analogue activity has also since been identified in humans, in the form of a centro-parietal positive potential (CPP) (Kelly & O'Connell, 2013; O'Connell, Dockree, & Kelly, 2012).

It is crucial to guard against over-interpretation of findings; as in any area of research, a common counsel is that correlation does not necessarily imply causation. In the area of neuroscience it is equally important to bear in mind that a simple correlation between neural activity and an external variable does not necessarily mean that this activity itself constitutes the internal representation of this variable. The chance of Type I error is a particular concern in the case of exploratory analyses in studies that lack specific and well-defined a priori hypotheses. However, in the case described here a cohesive base of empirical evidence has grown to be quite convincing. This is true in the sense that not only do populations of neurons within area LIP fire in a manner predictive of immediately subsequent saccades into their specific receptive fields, but measured activity in these populations appears to correlate with stimulus evidence that is indicative

of the appropriateness of their specific saccadic response (Gold & Shadlen, 2001). Further, this putative evidence integration appears to be parametrically modulated in accordance with the coherence of stimulus evidence, but to trigger a motor response once a common threshold of activity was reached, across all levels of stimulus coherence (Roitman & Shadlen, 2002). Behaviourally of course, this pattern leads to faster reaction times when stimulus evidence for one option is more coherent, and the choice is more clear-cut. However, in a fascinating addition to this body of research, it was also found that when trials were limited to a fixed duration, neural activity within area LIP would be maintained at the level of the decision threshold throughout this artificial delay period, until the motor response could be made (Shadlen & Newsome, 2001). Together, this evidence provides a quite compelling case that this measured activity does genuinely represent internal coding of some form of decision variable.

Perceptual decision-making (Heekeren, Marrett, Bandettini, & Ungerleider, 2004; Heekeren, Marrett, & Ungerleider, 2008) differs superficially in some respects quite greatly from value-based decision-making (Rangel, Camerer, & Montague, 2008). This is true in the sense that a perceptual decision generally comprises a relatively implicit judgement over stimulus properties, which may not even necessarily entail a specific response; though the experimental tasks upon which this form of decision making has been explored generally have required some response to indicate the judgement made. By contrast, value-based choice involves determining a particular preference, and is therefore linked intrinsically, although not necessarily immediately, with some form of overt response behaviour.

Nevertheless, despite this apparent gulf between core characteristics of these two forms of decision-making, empirical research suggests a high degree of functional equivalence, at least in the mechanism by which a decision is reached. At the cognitive level there is a wealth of evidence suggesting that equivalent stochastic accumulation-

integration based 'Drift Diffusion' or 'Random Walk' processes underpin each of these decision processes (Ashby, 1983; Basten, Biele, Heekeren, & Fiebach, 2010; Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Ratcliff, 1978; Ratcliff & McKoon, 2008; Ratcliff & Rouder, 1998; Ratcliff, Van Zandt, & McKoon, 1999; Smith, 2000; Wenzlaff, Bauer, Maess, & Heekeren, 2011). This is mirrored at the physiological level, with quantitative coding of reward values for each option, in the form of neural firing rates, having been shown to apply also to value-based choices. Moreover, this signal has been identified in the same region of the rhesus monkey cortex as perceptual stimulus driven values, area LIP (Sugrue, Corrado, & Newsome, 2004). This apparent functional equivalence of neural activity coding for perceptual and value-based decision variables has been described as a 'common currency', while a more recent review paper has aimed to more overtly reconcile these parallel threads of decision-making research and to further elucidate mechanisms that are held commonly between the two (Summerfield & Tsetsos, 2012).

1.b.i. Action-Values

The internal coding of decision related reward values has also been more directly demonstrated in a number of other brain areas, particularly within the striatum. A significant proportion of neurons within the macaque monkey ventral striatum have been shown to modulate their activity in a positive linear relationship with the expected magnitude of an anticipated appetitive stimulus (Cromwell & Schultz, 2003; Schultz, Apicella, Scarnati, & Ljungberg, 1992). Expanding upon this research, action-specific reward values were identified within the putamen and sub-regions of the caudate nucleus of the same species, with this activity being found to linearly represent the expected values of both selected and unselected options (Samejima, Ueda, Doya, & Kimura, 2005). This original finding was made with a decision task in which responses were made

through hand-movements, while a subsequent study demonstrated analogous action-value coding on a saccadic response task, within the oculomotor caudate nucleus (Lau & Glimcher, 2007). These seemingly cross-modal activity patterns, each reflecting the expected value of a specific motor response, have been termed 'action-values'. In addition, the latter study found that as each trial progressed, and particularly in the period immediately after a decision had been made, activity within the same region transitioned towards representing exclusively the reward value of the chosen option. This was paired with evidence for differential activity, within another subset of caudate nucleus neurons, which depended upon whether or not a reward was received. This common localisation of activity encoding the expected value of the chosen option and the actual outcome would lend itself to the efficient computation of a prediction error. Prediction error calculation is crucial for the purpose of evaluative processing that subserves the reinforcement learning process, and may be considered an embodiment of the variables r and Vt in the Rescorla-Wagner learning rule $Vt+1 = Vt + \alpha(r-Vt)$. Furthermore, this proposition is in close accordance with the findings of an fMRI based study conducted on human subjects, which showed activity representative of a prediction error signal in both ventral (ventral putamen and NAcc) and dorsal (anterior caudate) striatal regions (O'Doherty et al., 2004). Interestingly, this activity was framed as dissociable within an actor-critic framework, with the dorsal striatum (anterior caudate nucleus) seeming to code for prediction error on only the instrumental task (actor), whereas the ventral striatum (ventral putamen & NAcc) represented prediction error on both the instrumental learning and a yoked Pavlovian conditioning control task (critic).

Crucially for the purposes of the present thesis, evidence for action-values is not limited to the striatum. Other electrophysiological studies have highlighted action-value activity within modality specific motor effector regions that are more closely concerned with enacting the respective motor response. First, on a saccadic response task, 45% of

neurons within the supplementary eye field (SEF) were found to encode both the appropriate saccade direction and expected value of a particular response (So & Stuphorn, 2010). Intriguingly, this activity seemed to represent subjective expected values, in that it reflected a general behavioural bias in response preferences in favour of more risky options over sure alternatives. Second, Sohn and Lee (2007) demonstrated comparable activity within supplementary and presupplementary motor regions (SMA and pre-SMA) on a hand movement based task, with this activation representing a multiplicative combination of a given direction of movement with the temporally discounted expected value of enacting that response.

Both of these studies were conducted in rhesus monkeys. In order to assess whether these findings are applicable to the human brain it is necessary to apply non-invasive procedures. Wunderlich, Rangel, and O'Doherty (2009) conducted a quite comprehensive study with the aim of addressing this question, making use of BOLD-based fMRI. One advantage of this technique is that it allows for the simultaneous imaging of the entire brain, rather than only a specific region of interest that has been identified a priori. Of course, it is important in these cases to consider the potential for alpha-inflation when left uncontrolled for, which results from the greater number of comparisons being made and therefore higher chance of Type I error. This technique enabled the study to probe value related activation across a broad range of relatively disparate brain regions, for comparison with the findings of a number of separate electrophysiology studies. In this study, different response modalities were required to select between the two options, meaning that action-value coding for each response should be subserved by sufficiently distinct cortical regions as to be separable at the spatial resolution of fMRI. Results were widely congruent with the findings of previous primate electrophysiology studies (So & Stuphorn, 2010; Sohn & Lee, 2007). Whereas for saccadic responses action-values were identified in the pre-SEF, action-value consistent

activity for hand movements was evident within the SMA, within the hemisphere contralateral to the response hand. Value-related activity was also identified within lateral parietal cortex, in accordance with Sugrue et al. (2004), along with the ACC, in accordance with Kennerley, Dahmubed, Lara, and Wallis (2009) and Croxson, Walton, O'Reilly, Behrens, and Rushworth (2009) and the right dorsal putamen, in accordance with Samejima et al. (2005). This localisation of internal value representations to cortical regions, such as SMA, provides an encouraging basis for further investigation of this activity through the use of extra-cranial electrophysiological techniques such as EEG.

1.b.ii. Technique Specific Considerations of EEG

The reason that the findings of Wunderlich et al. (2009) form such an important prerequisite for the potential EEG based observation of action-values are predominantly technical, resulting from intrinsic physical characteristics of this technique. These are discussed here. Whereas the overwhelming majority of intra-cranial electrophysiology studies conducted in monkeys measure neural activity in the form of action potentials, activity recorded through EEG is driven primarily by post-synaptic potentials. There are a number of reasons for this. First, action potentials create two equal and opposite current dipoles, which effectively cancel each other out in terms of any signal that can be recorded remotely. The resulting quadrupolar field also diminishes with distance at an unfeasible rate for extracranial recording ($1/r^3$). Even if this weren't the case, a significant amount of temporal summation would still be necessary to provide a measurable signal, and as the duration of action potentials themselves is only around 1ms, such temporal summation is highly unlikely. Postsynaptic potentials (PSPs) however, do not suffer from these same issues. They generate a single current dipole, for which the field decreases according to distance as only $1/r^2$, rather than the $1/r^3$ for action potentials' quadrupolar field. Furthermore, PSPs last 10s of milliseconds, increasing the probability of temporal

summation by an order of magnitude. A massive amount of additive activity is nevertheless required. While a single PSP generates a current dipole along the dendrite of around 20fAm, EEG is sensitive to sources with strength of around 10nAm. This equates to the summation of about 500,000 local PSPs. Such aggregation is generally confounded in practice by spatiotemporal misalignment of PSPs, and thus current dipoles. Crucially however, the largely parallel arrangement of the apical dendrites of cortical pyramidal cells lends itself to additive net activity. PSPs in these cells therefore contribute a large proportion of the signal detectable through EEG. In light of this, the existence of empirical evidence for value-related activity within cortical areas forms a fundamental basis for the justification of investing time into an EEG based examination of this value-related activity.

As a non-invasive technique, EEG is similarly suitable for recording from human subjects as fMRI, and it also allows for simultaneous recording over the entire scalp. The use of EEG also offers the key theoretical benefit, relative to fMRI, that it retains the similarly high temporal resolution (in the order of milliseconds) as the targeted intra-cranial recordings that are generally only possible to conduct in non-human subjects.

There is a major trade-off to this, and that is of poor spatial resolution and a loss of sensitivity to deep sources. The electrical potentials to which EEG is sensitive are heavily diffused by the layers of tissue and bone between the cortical surface and the measurement electrodes on the scalp. The effects of the dural layer and skull itself block around 95% of this signal. In addition, neural activity patterns do not map directly onto the pattern of signals recorded at the scalp. Instead, the topography of activity observed extra-cranially depends upon multiple factors, including both the orientation and laminar profile of the neural source from which the signal originates. As an example, a positive or negative pole in the recorded electrical field does not necessarily represent a corresponding maxima or minima of neural activity directly beneath it. In fact, this

assumption would only be accurate given an entirely radial source. It is therefore impossible to interpret a source's position or orientation based upon single peak in voltage (Olejniczak, 2006). Equivalent factors underlie the polarity and position of a dipolar field. In this case, a measured field may be generated by a single tangentially oriented source. Alternatively, the same field may represent two independent radial sources, either with opposite activation profiles or opposite extracellular voltage field orientations, which could be driven by varying sites of dendritic PSPs (see Figure 1). The issue of multiple possible sources accounting for the same measured activity profile at the scalp is termed the inverse problem. Furthermore, the distortion of the signal across the scalp through all intervening layers means that an accurate source-localization model not only has to deal with the inverse problem, but also to accurately model and account for this distortion.

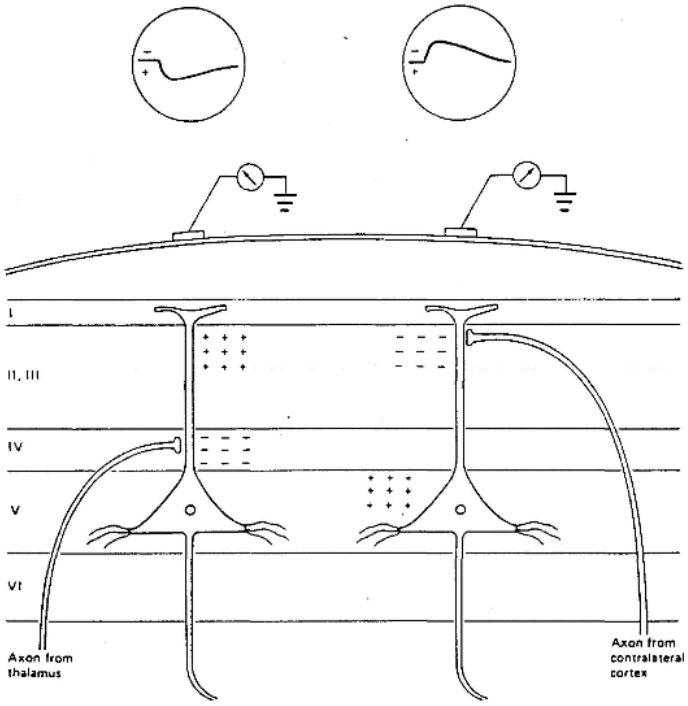


Figure 1: Illustration of variance in measured surface-activity at the scalp according to specific site of axo-dendritic synaptic input. Reproduced from Olejniczak (2006).

In practice these problems leave EEG as a technique with relatively poor spatial resolution. Nevertheless, with a paradigm in which putative value coding can be sufficiently spatially separated it should still be possible to discriminate internal representations of different response options. It is for this reason that a core component of the task designs used within this thesis will be the clear lateralization of responses for each option. It should also be possible to perform a less specific spatial characterisation of measured activity; for instance discriminating prefrontal from sensorimotor activation.

1.b.iii. Prefrontal Activity as a Proxy for Top-down Cognitive Control

There is an abundance of neuroimaging-based evidence for the involvement of multiple sub-regions of the PFC in the value-based decision making process (Rushworth, Noonan, Boorman, Walton, & Behrens, 2011). These include the dorsolateral prefrontal cortex (dlPFC) (Basten et al., 2010; J. N. Kim & Shadlen, 1999; S. Kim, Hwang, & Lee, 2008; Wallis & Miller, 2003), ventromedial prefrontal cortex (vmPFC) (Basten et al., 2010; Chib, Rangel, Shimojo, & O'Doherty, 2009; Glascher, Hampton, & O'Doherty, 2009; Lin, Adolphs, & Rangel, 2012) and orbitofrontal cortex (OFC) (Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; O'Doherty, 2004; Padoa-Schioppa & Assad, 2006; Peters & Buchel, 2010; Tremblay & Schultz, 1999; Wallis & Miller, 2003). Furthermore, an EEG study conducted by Polania, Krajbich, Grueschow, and Ruff (2014) discovered that gamma activity over frontal regions coded for value-based evidence accumulation, while fronto-parietal coupling of gamma oscillations related to the accuracy of value-based decisions.

Moreover, drawing upon the findings of a multitude of studies, a broad empirical consensus has become established that prefrontal activity is associated with top-down executive control over voluntary action (Brunia, 1999; Fuster, 2000; E. K. Miller, 2000;

E. K. Miller & Cohen, 2001; Passingham, 1993). In a similar vein, Wise, Murray, and Gerfen (1996) discuss links between the frontal cortex and basal ganglia as an integrated system involved in response learning. They propose that while these regions are linked, each subserves a distinct aspect of the process; with the basal ganglia working to potentiate already established rules according to reinforcement history, while the frontal cortex is required to replace old rules with new. To further substantiate this idea, there is behavioural evidence that prefrontal lesions lead to impairments in learning new rules, in both humans and other primates, on the Wisconsin card sorting task, and an equivalent analogue designed for non-humans (Dias, Robbins, & Roberts, 1996; Milner, 1963). Moreover, to even more directly illustrate the relevance of this putative function of prefrontal cortex, within a comprehensive review written by Earl Miller a paper is discussed that explicitly documents deficits in strategy use as an effect of frontal lobe damage (Shallice & Burgess, 1991). This study draws the conclusion that in patients with prefrontal damage, *'the basic elements of behaviour are intact but it seems that they are missing the flexibility to shift between different rules and so override prepotent responses to persist toward a goal.'* (Miller, 2000, p.61, Box 1). If the heuristic-account is found to best explain suboptimal diversification of choices in the early studies of the present thesis then this would further suggest that the prefrontal cortex likely subserves a function that is particularly important in establishing probability maximizing over an otherwise prepotent probability matching response. In light of this converging body of evidence, observed activity at prefrontal electrode sites could potentially offer a simple electrophysiological measure of the relative application of top-down cognitive control. Therefore, if behavioural evidence is found to support the heuristic account of probability matching, then measured prefrontal activity could be contrasted between trials in which maximizing vs. non-maximizing choices are made, to provide electrophysiological corroboration of this finding.

1.b.iv. Feedback Related Negativity and Prediction Error

Multiple studies have used EEG techniques to investigate event-related potentials, or ERPs (Sur & Sinha, 2009). An early example of this is the 'P300' which is a positive component 300ms following stimulus presentation, with an amplitude modulated by the level of salience or surprise of said stimulus (Ruchkin, Sutton, & Tueting, 1975; Sutton, Braren, Zubin, & John, 1965; Sutton, Tueting, Zubin, & John, 1967; Tueting, Sutton, & Zubin, 1970). Many more recent studies have identified a negative deflection in activity that corresponds to negative feedback and appears at around 250-300ms post-feedback (Gehring & Willoughby, 2002; Nieuwenhuis, Holroyd, Mol, & Coles, 2004). This potential has been termed feedback related negativity (FRN), and is generally most prominent over fronto-central electrodes, though further studies have suggested that it originates in the ACC (Hauser et al., 2014; Luu, Tucker, Derryberry, Reed, & Poulsen, 2003; Miltner, Braun, & Coles, 1997) and is conveyed from the mesencephalic dopamine system as the output of evaluative decision-processing within the basal ganglia (Holroyd & Coles, 2002).

These findings tie into a range of existing literature that documents post-feedback coding of the prediction error. This is simply the difference between expected and actual reward on any given trial, and is represented by the variables (r-Vt) in the Rescorla-Wagner learning rule $Vt+1 = Vt + \alpha(r-Vt)$ (Rescorla & Wagner, 1972). The prediction error is a crucial component of the reinforcement learning process as a positive feedback loop, allowing behaviour that *'runs off in the direction of the search'* (Richmond, 1980). This is true in the sense that the prediction error allows behaviour to 'home in' on the most rewarding option, even as reward schedules may change over time. Single-unit electrophysiology studies have found prediction error coding within the primate ACC (Amiez, Joseph, & Procyk, 2006; Matsumoto, Matsumoto, Abe, & Tanaka, 2007), while a range of human neuroimaging studies have found that activity levels in the

ventral and dorsal Striatum, OFC, anterior insula and ACC code for prediction error magnitude, valence or both (Bray & O'Doherty, 2007; Jensen et al., 2007; Jessup, Busemeyer, & Brown, 2010; Oliveira, Mcdonald, & Goodman, 2007; Valentin & O'Doherty, 2009).

Whether or not FRN relates to prediction error coding has not always been clear. Hajcak, Holroyd, Moser, and Simons (2005) found, on a probability task, that while measured P300 magnitude was higher for more unexpected feedback, the magnitude of the FRN was not significantly larger for unexpected than expected outcomes. Similarly, the same authors found that FRN did not significantly vary in relation to the relative magnitude of a withheld reward (Hajcak, Moser, Holroyd, & Simons, 2006). However, a number of other studies have instead indicated that the FRN does vary according to either magnitude or expectancy of reward, in a manner consistent with the prediction error. Both Holroyd, Nieuwenhuis, Yeung, and Cohen (2003) and Yasuda, Sato, Miyawaki, Kumano, and Kuboki (2004) found significantly greater FRN in relation to unexpected relative to expected outcomes. Cohen, Elger, and Ranganath (2007) found that varying reward contingencies modulated the FRN response in relation to wins, but not losses. Intriguingly, on a probabilistic reversal-learning task, Chase, Swainson, Durham, Benham, and Cools (2011) found that FRN coded for prediction error magnitude according to a reinforcement-learning model, but did not relate to unsurprising negative outcomes preceding rule-based choice reversals. By contrast, the same study found that P300 was higher for rule-based expected negative outcomes before rule-based choice-reversals than for unexpected ones that weren't followed by a choice-reversal. The authors conclude that FRN codes for prediction error and adjustment of values according to the reinforcement learning process, while P300 predicts explicit rule-based adjustment. On a probabilistic gambling task, Pfabigan, Alexopoulos, Bauer, and Sailer (2011) found that FRN coded for both negative and

positive reward prediction error, while P300 amplitude was largest after unexpected feedback irrespective of valence. The authors conclude that the FRN does represent a *'reinforcement-learning signal, which is detecting mismatch between internal and external representations indexed by the ACC to extract motivationally salient outcomes'* (Pfabigan et al., 2011, p. 1).

For a full and comprehensive review of the FRN literature see Walsh and Anderson (2012). Here the authors list 32 known studies that found FRN does code for prediction error, contrasted with seven studies that found it did not. Ten studies are listed localizing the neural origin of the FRN to the ACC, six studies to the PCC and three to the basal ganglia. It is also important to note that the observed global characteristics of the FRN can vary. For instance, Foti, Weinberg, Dien, and Hajcak (2011) report findings from a gambling task, in which participants made both monetary gains and losses. Here, FRN was observed at fronto-central sites as an overall positive potential, which was significantly retarded for non-rewards relative to rewards. In fact, studies have regularly reported post-feedback activity that does not show an overall negative inflection, but within which there is a clear depression in positive activity for negative outcomes relative to positive ones. It is this characteristic relative difference in activity that is most accurately interpreted as FRN. This position is further clarified in Walsh and Anderson (2012), who cite one of their own preceding studies as an example (Walsh & Anderson, 2011). They note that *'although waveforms are relatively more negative following losses, they do not literally drop below zero. This is because the FRN is superimposed upon the larger, positive-going P300, which is evoked by stimulus processing'* (Walsh & Anderson, 2012, p. 1874-5).

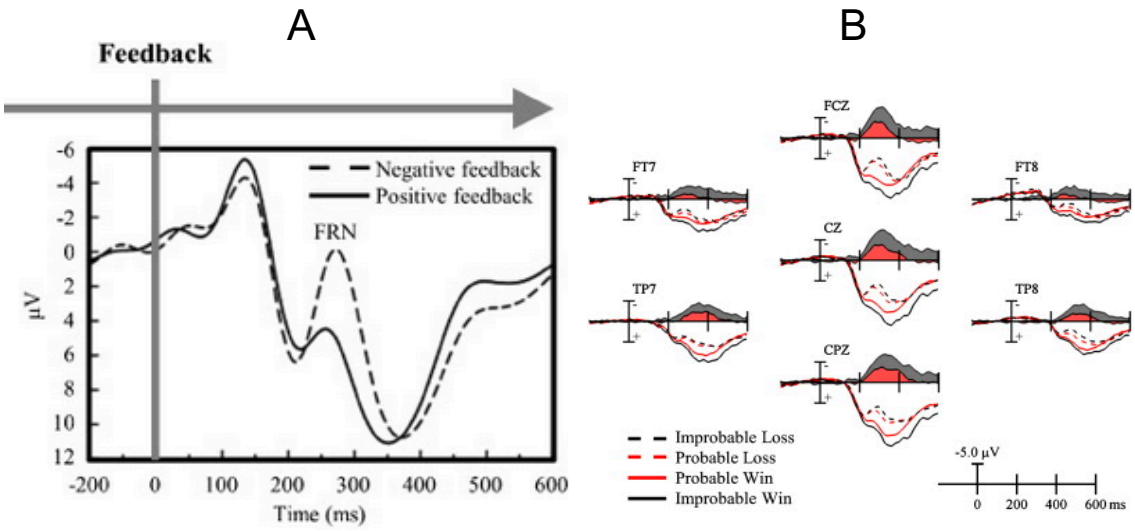


Figure 2: Illustrations of feedback-related negativity (FRN) as the difference in averaged ERP activity following positive vs. negative feedback. Both are presented in Walsh and Anderson (2012). Panel A is adapted from an original in Nieuwenhuis et al. (2002). Panel B is Adapted from an original in Walsh and Anderson (2011). Both use inverted y-axes, with negativity plotted high.

Further EEG-studies have also investigated post-feedback activity outside of the immediate context of the existing FRN literature. Mars et al. (2008) systematically varied stimulus probabilities over blocks of trials, concluding that magnitude of stimulus surprise was encoded in the amplitude of the fronto-central component of P300, which has been termed the P3a (Polich, 2007). Bennett, Murawski, and Bode (2015) found that stimulus-preceding negativity in the immediate pre-feedback period reflected uncertainty, while post-feedback P300 coded for belief update size. Another study revealed post-feedback prediction error coding had two components (Philiastides, Biele, Vavatzanidis, Kazzer, & Heekeren, 2010). These peaked at around 220 and 320ms respectively, though the first covered a broader time window, being discriminable from around 180 to 380ms post-feedback. The authors concluded that the earlier component codes mainly for

prediction error valence, while the later component predominantly represents prediction error magnitude.

The EEG research that is planned for the latter parts of the present thesis offers an opportunity to replicate findings from the studies discussed above. Further, the final study in particular offers the potential to assess any manner in which post-feedback decision processing may be biased or affected by the nature of the preceding choice, particularly in relation to whether it was a maximizing or non-maximizing response.

1.b.v. Neuroscience of Probability Matching

Relatively little neuroscience literature exists documenting studies conducted with the explicit aim of investigating probability matching, or in fact that even mention the topic. Of the few studies that have addressed this issue, whether more or less directly so, findings have been remarkably inconsistent. Congruent with studies discussed earlier in this chapter, prefrontal immaturity, resulting from heterochronous cortical development (Huttenlocher & Dabholkar, 1997) relates to a lack of cognitive control. Thompson-Schill, Ramscar, and Chrysikou (2009) discuss how this late development may be beneficial for learning, and then proceed to directly propose that probability matching correlates positively with prefrontal maturation. To support this assertion they cite Derks and Paclisanu (1967), in which higher levels of maximizing were found in children under the age of four. They go on to justify this relationship as 'brute strength competition' leading to domination of the 'most frequent form'; in essence, proposing that maximizing results naturally from the reinforcement learning process. However, this interpretation is clearly at odds with a multitude of subsequent studies, which have indicated that it is instead probability matching that is the prepotent response and maximizing that results from the application of greater cognitive control (Frederick, 2005; Koehler & James, 2009, 2010; Kogler & Kuehberger, 2007; West & Stanovich,

2003). Moreover, as discussed earlier, the apparent age-related decrease in maximizing observed by Derks and Paclisanu (1967) was driven primarily by the choices of the ten youngest participants within the nursery age group. These produced extremely patterned responses, but only one could actually report which option had been more likely to provide a reward. These choices therefore seem more likely to have been driven by some other extraneous or confounding factor. Furthermore, the increase in individual level probability matching in line with age largely reflected a decrease in undermatching, rather than a decrease in maximizing. In fact, at the group-level progressively higher levels maximising were found as age increased above kindergarten level. It is also worth noting, as highlighted by Fiorina (1971) that, as with many early studies in this area, reward sequences were randomised without replacement and determined manually by the experimenter, entailing that maximizing was not strictly the optimal response to the task. In light of these considerations, the interpretation and conclusions drawn from this study by Thompson-Schill et al. (2009) may be flawed. In fact when considering a broader body of evidence, the opposite hypothesis might be more appropriate. If maximizing results from the application of greater cognitive control, so as to overrule a heuristic-based matching response, it is therefore more likely to relate to greater prefrontal activity.

Another study designed to investigate the phenomenon of probability matching from a neuroscience perspective was conducted by Wolford, Miller, and Gazzaniga (2000). Interestingly, this was done in the context of apophenia. The authors assume that pattern search is the key determinant of probability matching, citing Yellott (1969) as evidence for this, although ultimately they proceed to state that it is their own view that matching is driven by a mistaken attempt to determine a non-existent pattern. This is then used as a justification to substitute matching behaviour as a proxy for pattern search. It is this topic that the authors seem primarily interested in attributing to a specific brain area, in this case the left hemisphere. Such an unequivocal claim may not

be entirely justified, a recent study designed to assess the putative link between probability matching and apophenia failed to replicate this (Koehler & James, 2009). Nonetheless, irrespective of whether or not a valid link exists, the findings of this study can be interpreted in the context of the matching behaviour itself.

The study comprised twin experiments. In the first, two split-brain patients each completed concurrent probabilistic choice tasks. These were presented in a manner that their choices would be determined, in theory, by either the left or right hemisphere independently. Results showed lower deviation from probability matching in choices that were apparently made by the left rather than the right hemisphere. However, there was a clear confound here in that different probability ratios were used for the task presented to each side, with contingencies of .8 vs. .2 for the right visual field (left hemisphere), and .7 vs. .3 for the left visual field (right hemisphere). It is unclear why this design was chosen, as this seems an obvious potential confound when comparing choices across tasks. The results were also based upon a sample size of only two participants, which again does not inspire confidence in their reliability. In the second experiment, patients with focal frontal or prefrontal lesions, lateralized to either the right or left hemisphere only, completed the same task as in the first experiment. Here though, results were not split according to the hemifield to which stimuli were presented, but were instead compared between subjects. These indicated that the four patients with right-hemisphere damage showed lower deviation from matching than the single patient with left-hemisphere damage. The authors conclude this is evidence for impairment in pattern search resulting from left-hemisphere damage, though it could perhaps be more directly argued as impairment in Strategy-based maximizing in those with right-hemisphere damage. In any case, I would argue that the serious proposal of any interpretation of a between groups difference where the sample size of a group is one is inappropriate.

A neuroimaging-based follow-up study was conducted five years later which explicitly attempted to determine brain activity relating to probability matching (Miller, Valsangkar-Smyth, Newman, Dumont, & Wolford, 2005). The primary aim of this study was reconciliation of previous findings, summarised as *'the left hemisphere of split-brain patients frequency matches, but the right maximizes'* (Miller et al., 2005, p.1), with a number of subsequent neuroimaging studies that seemed consistently to suggest that it was right, rather than left, hemisphere activity that was more closely associated with predicting and guessing (Elliott & Dolan, 1998; Elliott, Rees, & Dolan, 1999; Huettel, Mack, & McCarthy, 2002; Schubotz & von Cramon, 2002; Volz, Schubotz, & von Cramon, 2003). Percentage signal change in BOLD activity was assessed between a probabilistic prediction task and a detection-only control task, across a number of brain regions. Results fell generally in line with the preceding neuroimaging studies, rather than the split-brain paper discussed previously (Wolford et al., 2000), though with such small sample sizes in the original study it is unsurprising that these results did not replicate. This point is acknowledged in the discussion, which notes that the small number of split-brain patients tested may not have been representative of the general population.

In the latter study (Miller et al., 2005), due to an unfortunate lack of maximizing participants (only one vs. 18 categorised as matchers), statistical power to investigate differences in activity relating to probability maximizing vs. non-maximizing choice behaviour was limited. Nevertheless, results did indicate a negative correlation in two brain regions between % signal change on the prediction task and a ratio-based metric designed to represent deviation from matching. These areas were the right medial frontal gyrus and right angular gyrus. However, this relationship was found only to be statistically significant after averaging measured signal change across both regions, while the same relationship was found to be non-significant in any of three other areas considered alone. Further, the observed *p*-value (.047) was dubiously close to the alpha

criterion of .05, when considering that the paper neglects to mention whether any corrections were made for multiple comparisons in this particular analysis. The reported *F*-ratio indicates that they were not. Fittingly, the discussion section of this paper opens with the comment *'a common error in decision-making is to put forward a causal relationship when the evidence is inadequate or indicates a random relationship'* (Miller et al., 2005, p.8). This is certainly important to bear in mind.

In summary, there exist few studies that have attempted to draw links between maximizing vs. matching behaviour and characteristic neural activity. Moreover, of those that do so, the evidence is far from reliable or consistent enough to draw a convincing conclusion. It is hoped therefore that the investigation planned for the latter stages of the current thesis, to assess electrophysiological activity relating to probability matching and predictions of Faulty Foraging Theory, will provide some informative foundational insights in this regard.

1.c. Thesis Summary

There were two core components to the plan of this thesis. First to determine the relative contributions that several of the models described earlier might make in determining choice behaviour, particularly in terms of engendering maximizing versus diversification of choices. Second, to explicitly assess mechanisms underlying whichever of these is found to make the most substantive contribution to choice behaviour, through the combined application of behavioural and electrophysiological methods.

In practice, Chapter 2 documents two behavioural experiments that were designed to address the first of these issues. These experiments predominantly focus on the heuristic and apophenia-based accounts, given the weight of existing evidence in support of these explanations. Here, the heuristic account is found to provide the principal basis for determining matching behaviour. In light of this, Chapter 3

documents the development and application of a simple simulated foraging task (SSFT). This was designed to test Faulty Foraging Theory as a potential underlying explanation for why probability matching is the prepotent response when acting intuitively. Here the pattern of behavioural effects is found to be broadly consistent with the predictions of Faulty Foraging Theory.

It was then planned to move forward with the use of non-invasive electrophysiological methods (EEG) to further probe putative underlying determinants of matching vs. maximizing choice behaviour. Chapter 4 documents a targeted reanalysis of data collected in a previous EEG study, which was necessary to determine the feasibility of the planned experimental paradigm for the remaining two experiments. Following this, Chapter 5 documents the application of EEG recording in conjunction with the SSFT. Results indicate robust evidence for the internal representation of global resource generation rates. This is paired with statistically significant, although somewhat less substantive, evidence for the internal representation of local expected values. These take into account trial-by-trial variance in reward availability between the two options, which depends upon previous levels of exploitation.

Chapter 6 is the final experimental chapter. This documents the subsequent extension of EEG methods to the standard probability matching preparation, in a larger study. A number of specific experimental hypotheses are made. Here, a stricter cluster-permutation based approach is applied in the EEG analysis, to better correct for potential inflation of the Type 1 error rate resulting from multiple comparisons. Here no evidence was found for the internal representation of illusory reward accumulation within the unchosen option, a hypothesis derived from Faulty Foraging Theory. However, an inverse pattern of relationships was evident between maximizing and prefrontal activity within the pre-choice and post-feedback periods. In addition, a dual-component post-feedback ERP was identified, with characteristics that were consistent

with generally linear prediction error coding, coupled with a seemingly disproportionate subjective utility for successful non-maximizing choices. Chapter 7 comprises a general discussion and theoretical interpretation of the key findings from throughout the thesis.

Chapter 2: A Behavioural Assessment of Heuristic and Apophenia Accounts

2.a. Introduction: An Overview of Experiments

The accounts of probability matching discussed in Chapter 1 have often been considered in isolation, with the majority of studies designed specifically to test one or another potential explanation alone. However, the models need not be mutually exclusive. It is quite possible that multiple factors work simultaneously, playing additive, counteractive or interactive roles in determining the behavioural response; most likely with different examples having more or less influence on choice behaviour depending on the specific situation or task. The core aim of the studies documented in this chapter is to assess the relative contribution of multiple factors of probability matching. This is done by considering the two principal and most credible accounts together; according to bodies of existing empirical literature these are the apophenia and heuristic accounts. Further, a paradigm is applied that combines both experimental manipulation and self-report measures in order to further elucidate pathways of effect.

Two factors were designed to manipulate both participants' use of intuition versus strategy in determining their choices on the task, and the extent to which participants believed there to be a predictable pattern of sequential dependencies in the reward sequence. Apophenia is tested by contrasting the choice behaviour, along with self-reported levels of apophenia, of participants who are explicitly informed that there are no predictable patterns to be found in the reward sequence with participants who are left to discover this for themselves. The heuristic account of probability matching is tested by contrasting the choice behaviour, along with self-reported use of intuition versus strategy, of participants who are explicitly told the reward contingencies with participants for whom they are learned through experience.

One previous study (Fantino & Esfandiari, 2002) did also include experimental conditions designed to manipulate both strategy usage and pattern belief, by informing participants of the reward probabilities and that 75% was considered a perfect score respectively. However, the latter manipulation is quite indirect, with a degree of inference required by participants to interpret it as related to the potential presence of a pattern. Moreover, the assumption that either of these manipulations operated through the mechanisms that were predicted remains untested. The studies reported here include a post-experimental questionnaire to assess the extent to which participants adopted an intuitive versus strategic approach to the task, and their level of belief regarding whether or not there had been a predictable pattern in the sequence. This is designed to provide an additional, more direct, measure of the efficacy of each of our experimental manipulations. Crucially, when combined with the behavioural data in a mediation-analysis, this can provide additional evidence of the putative underlying mechanisms of any observed effects. The questionnaire was also designed to ascertain individuals' subjective estimates of the reward contingencies. Our primary intention here was to determine whether participants were well calibrated to the actual probabilities of reward, but this also enables a comparative analysis of probability matching behaviour in which inter-individual differences in probability estimates are controlled for (cf. Koehler & James, 2009).

Experiment 1

Experiment 1 was designed to assess the relative contribution of apophenia and heuristic processing to probability matching, through manipulating the information provided to the participants at the start of the task. If probability matching results from apophenia then explaining to participants that the sequences of outcomes were random, other than any overall difference in reward probabilities, should lead to an increase in

maximizing choices compared to participants who received no such explanation. It was also predicted that explicitly providing participants with the outcome probabilities prior to the task would facilitate a strategic approach, reducing heuristic based decision-making and increasing maximizing behaviour.

2.b. Method – Experiment 1

2.b.i. Participants

One hundred and twenty members of the University of Nottingham community volunteered to take part in this experiment. Of these, 37 were male and 83 were female. Their ages ranged from 18 to 52 (M=23.27, SD=4.91).

Participants were paid an inconvenience allowance with a value contingent upon the choices that they made during the experiment. A maximizing strategy would accumulate an average total of £5.88, while matching would lead to an average pay-off of £4.87. Performance at chance would lead to an expected payment of £4.20.

2.b.ii. Design

This was a 2x2x42 mixed model design with Probability (Stated vs. Learned) and Pattern (Standard vs. No-pattern) as between-subjects factors and 10-trial Block (1 to 42) as a within subjects factor. The basic experimental manipulations relate to the instructions that participants received (see Appendix B). In the Stated Probability conditions the participants were instructed that the outcome probabilities were 70% vs. 30%, and in the Learned Probability conditions the participants were not informed of the outcome probabilities. Participants in the No-pattern condition were told that there were no patterns in the sequences of outcomes and that these were entirely independent of the choices that they made, and participants in the Standard Pattern condition did not receive these instructions.

Which of the two colours was the more likely outcome was counterbalanced between participants, to account for any bias in favour of a particular colour. All participants were randomly assigned to an experimental condition.

2.b.iii. Stimuli

A plain white background was maintained throughout the task. A black central fixation cross was used between trials. Choice stimuli consisted of images of a blue and a yellow light bulb (see Figure 3). These were positioned equidistant from the centre and the left or right edge of the screen, with the side to which each colour stimulus was presented randomised between trials. The experimental procedure was designed and coded within Psychopy (Peirce, 2007).

2.b.iv. Procedure

The task was an iterated two-alternative forced choice, in which repeated predictions were made between whether a blue or a yellow light bulb would flash on, for each of 420 trials. There was no initial training period. Task instructions that were provided to participants are shown in the Appendices. Each trial began with a black central fixation cross on a white background. The duration of the fixation varied randomly between 1000ms and 2000ms. When considered together with the 500ms blank screen transitions immediately preceding and following the fixation cross, this gives a stochastically variable inter-trial-interval of between two and three seconds. This was followed by the appearance of the two light bulbs on the left and right of the computer screen. The right-left position of the two coloured bulbs varied randomly between trials and was not predictive of reward. This ensures that the specific stimulus-response mapping changes repeatedly and unpredictably throughout the task, which is an important consideration for future versions of the task if it is to be run in conjunction

with EEG recording. Participants were explicitly told in the instructions to make their choices on the basis of the colour rather than the position of the bulb, which was random (see Appendix B). Immediately upon presentation of the stimuli participants were free to make their predictions, by pressing the 'q' or 'p' keys on a standard QWERTY keyboard to indicate the stimulus to the left or right side of the screen respectively. Shortly after each prediction (500ms) one of the two light bulbs would be illuminated. This graphic was accompanied by either the word 'WIN' or 'LOSE' shown to the centre of the screen. Feedback remained on-screen for 1500ms. Participants received a £0.02 reward for each correct prediction, while a running total of winnings was shown at the top of the screen throughout the task. The probability that each light would be illuminated on any one trial, and thus the reward contingencies of the two options, was fixed throughout the task. One colour bulb had a reward probability of 0.7 and the other of 0.3. Average trial duration, including response time, was approximately 5000ms. Average total task duration was therefore around 35 minutes. In order to reduce the effects of boredom, the experiment was split into three equal sized trial-sets. These were punctuated by messages informing participants of their progression through the task. Each of these break screens remained indefinitely, until participants indicated that they were ready to continue.

2.b.v. Questionnaire

At the end of the experiment the participants completed a questionnaire (see Appendix A) in which they were asked to estimate the probability of reward for each light bulb (0% to 100%); the extent to which they believed there to have been a predictable pattern in the reward sequence (1 to 5); and the extent to which they relied upon intuition or strategy in making their predictions (1 to 5).

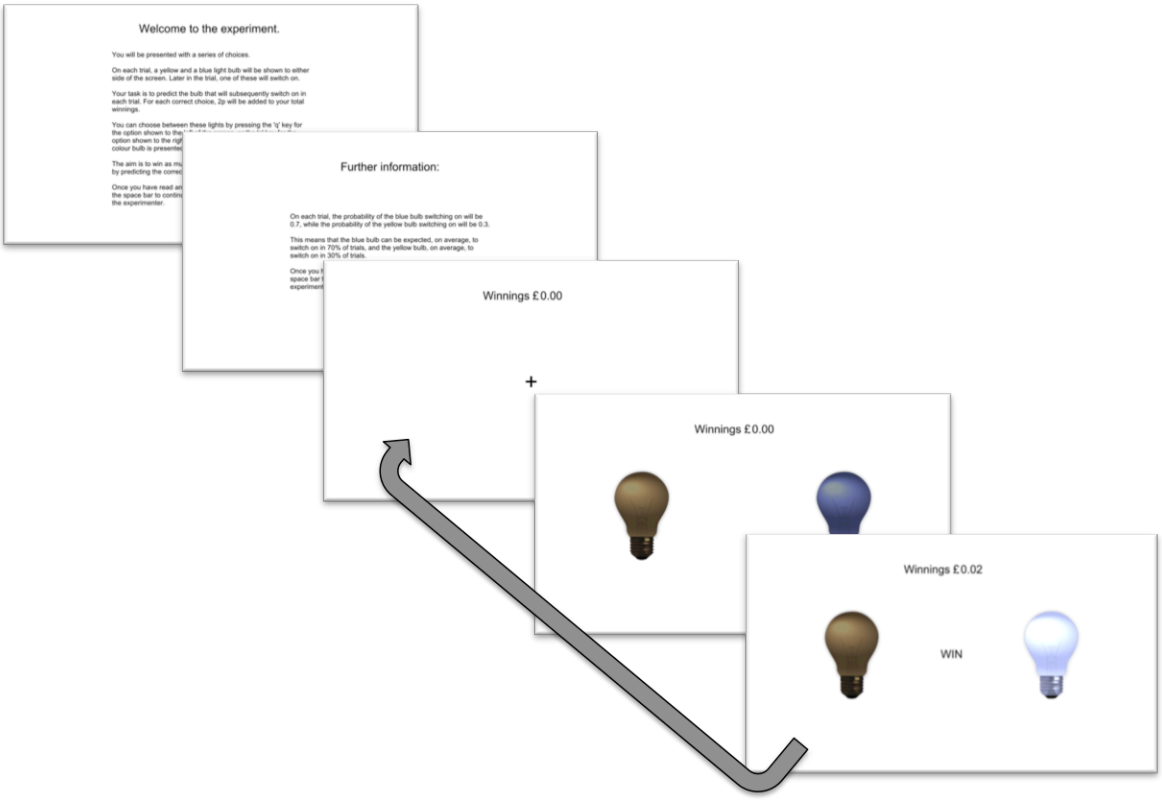


Figure 3: Example trial structure for Experiment 1 (Standard, Stated reward probabilities condition).

2.c. General Results – Experiment 1

The proportion of maximizing responses for each 10-trial block are shown for each condition in Figure 4A. Distributions of individual proportions of maximizing choices across the entire task are shown in Figure 5, Panels D-E. These data were entered into a 2x2x42 mixed model ANOVA with Probability (Stated vs. Learned) and Pattern instruction (Standard vs. No-Pattern) as between-subjects factors and Block (1 through 42) as a within-subjects factor. A main effect of Block ($F_{19.127,\ 2218.687}$ = 13.797, MS_e = .039, p < .001, η_p^2 = .106) and a significant linear effect of Block ($F_{1,\ 116}$ = 87.088, MS_e = .099, p < .001, η_p^2 = .429) indicated that participants learned to allocate more responses to the maximizing alternative as the experiment progressed. A main effect of

Probability ($F_{1, 116}$= 37.815, MS_e = .583, p < .001, η_p^2 = .246) indicated that when participants were told the probability of reward they tended to allocate more responses to the more likely outcome than when they had to learn it from experience alone. A reliable interaction between Block and Probability ($F_{19.127, 2218.687}$ = 3.089, MS_e = .039, p < .001, η_p^2 = .026) and a reliable linear interaction between the Block and Probability ($F_{1, 116}$ = 15.099, MS_e = .099, p < .001, η_p^2 = .115) suggested that stating the probabilities affected the rate at which participants learned to allocate responses to the more likely outcome. However, there was no effect of Pattern instruction (F < 1.0) and no significant interaction between Probability and Pattern instruction ($F_{1, 116}$= 3.298, MS_e = .583, p = .072, η_p^2 = .028). There was also neither an interaction nor a linear interaction between Block and Pattern instruction (F < 1.0). These indicate that informing the participants that there was no pattern to be found in the sequence of outcomes did not influence their choices. There were no 3-way interactions (F < 1.0).

The interaction between Block and Probability, as shown in Figure 4A, indicates a lower group-level learning effect in the Stated probabilities condition. This might represent a lower learning rate across all participants in this condition. Alternatively, it may be explained by a sub-set of participants beginning the task either at or near maximizing behaviour, having reasoned that this is the optimal strategy from the additional probability information given to them before the task began. Figure 5 shows histograms of individual-level proportions of maximizing choices in the initial 10-trials of the task for each the Standard condition (Panel A) and the No-pattern condition (Panel B). These clearly show that more participants in the Stated probabilities condition maximized from the outset of the experiment.

In order to assess experimental effects exclusively upon steady-state responding, a separate 2x2 between subjects ANOVA was conducted, with factors of Probability

condition and Pattern instruction, and the dependent variable of proportion of maximizing choices made over the final third of the task (from trial 280 onwards). This revealed a main effect of Probability ($F_{1, 116}$= 15.964, MS_e = .018, p < .001, η_p^2 = .121), but neither a main effect of Pattern instruction (F < 1.0), nor a reliable interaction effect ($F_{1, 116}$= 3.005, MS_e = .018, p = .086, η_p^2 = .025). Distributions of individual proportions of maximizing choices over the final third of the task are shown in Figure 5, Panels G-H.

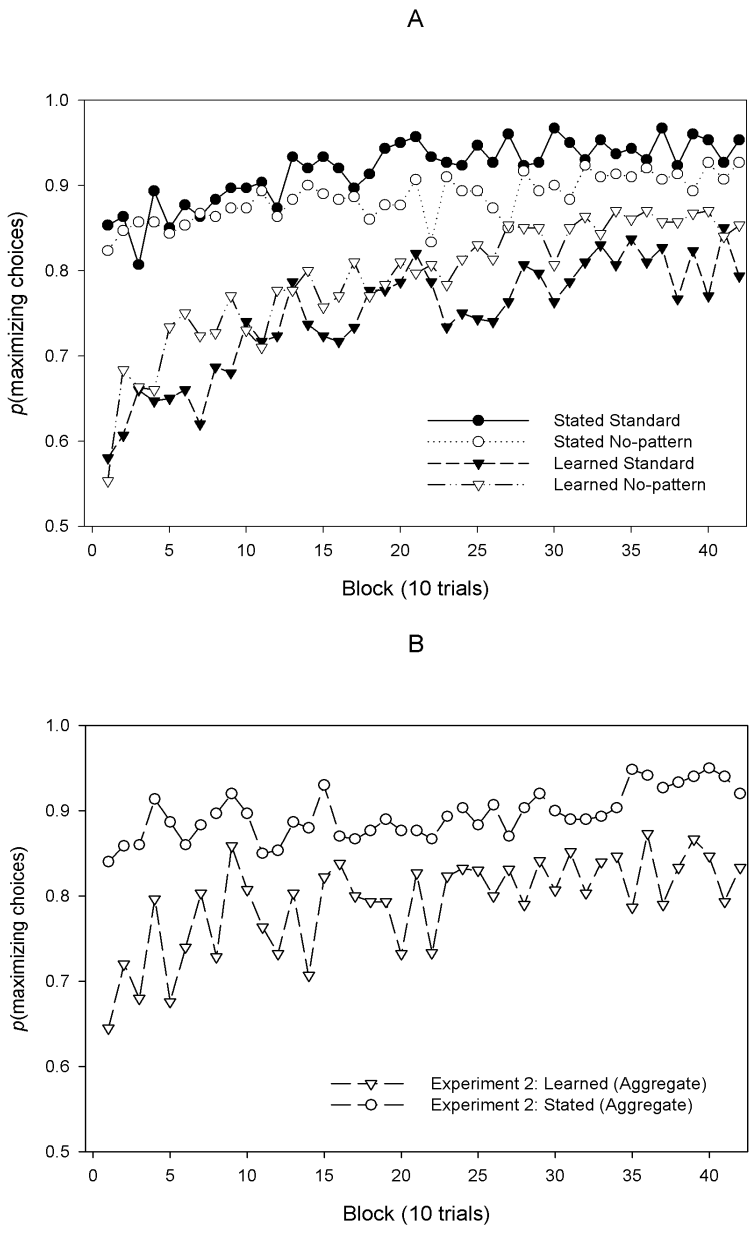
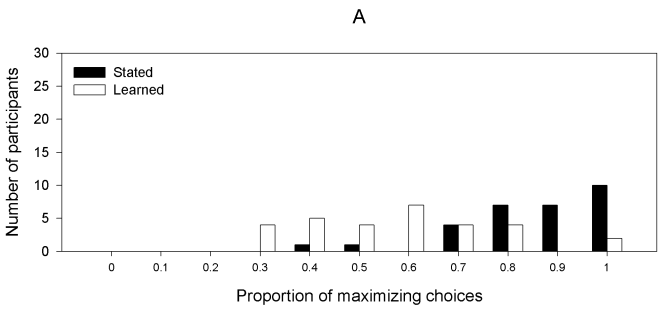
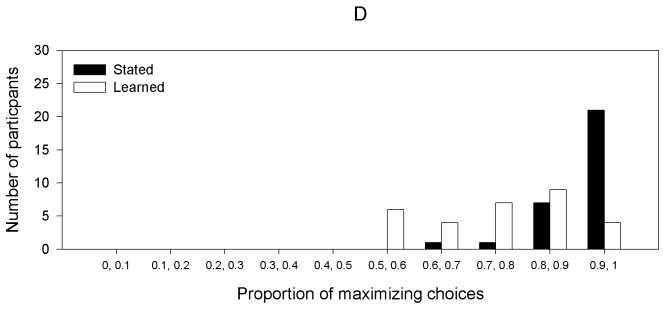
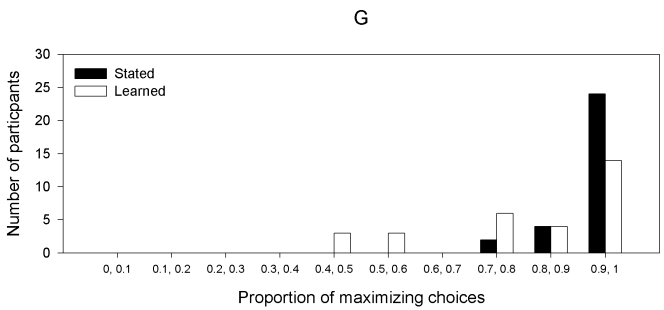
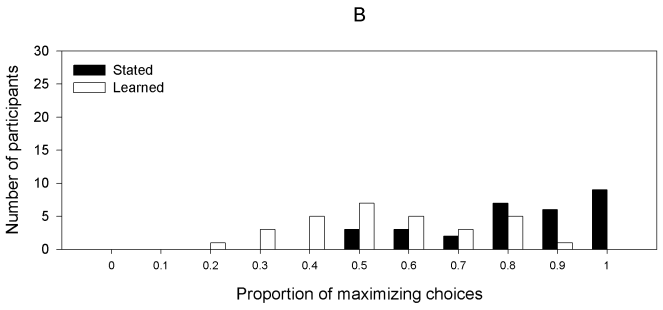


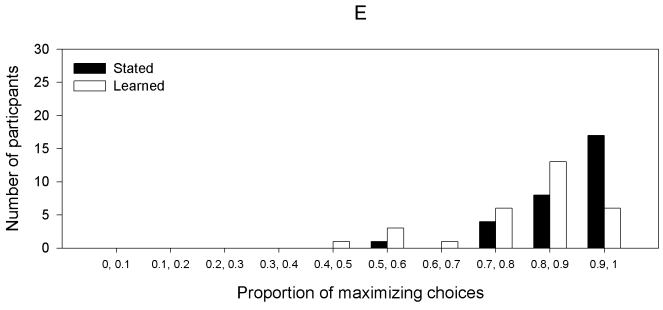
Figure 4: Showing the proportions of maximizing choices over the course of Experiment 1 (Panel A) and Experiment 2 (Panel B). Each block consisted of 10 trials.

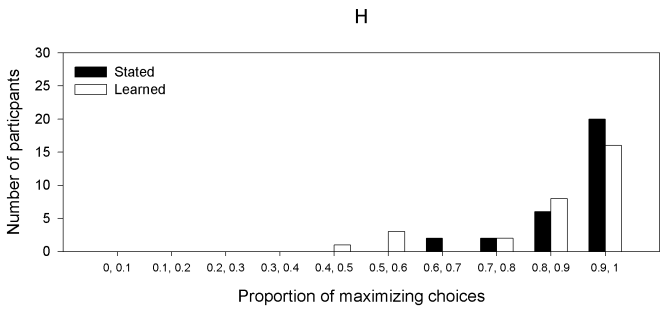

A
Stated
Learned
Number of participants
0 5 10 15 20 25 30
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Proportion of maximizing choices

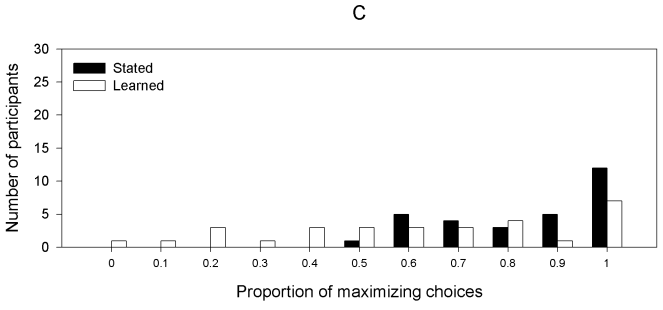

D
Stated
Learned
Number of particpants
0 5 10 15 20 25 30
0, 0.1 0.1, 0.2 0.2, 0.3 0.3, 0.4 0.4, 0.5 0.5, 0.6 0.6, 0.7 0.7, 0.8 0.8, 0.9 0.9, 1
Proportion of maximizing choices

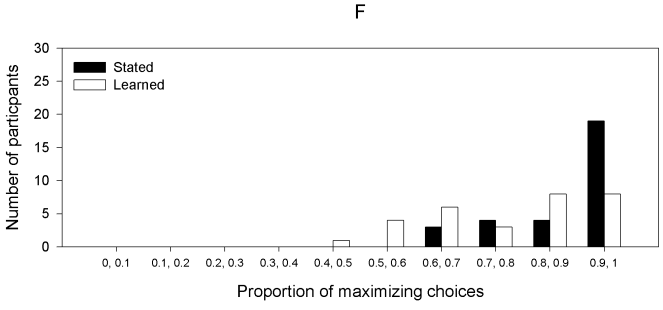

G
Stated
Learned
Number of particpants
0 5 10 15 20 25 30
0, 0.1 0.1, 0.2 0.2, 0.3 0.3, 0.4 0.4, 0.5 0.5, 0.6 0.6, 0.7 0.7, 0.8 0.8, 0.9 0.9, 1
Proportion of maximizing choices


B
Stated
Learned
Number of participants
0 5 10 15 20 25 30
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Proportion of maximizing choices


E
Stated
Learned
Number of particpants
0 5 10 15 20 25 30
0, 0.1 0.1, 0.2 0.2, 0.3 0.3, 0.4 0.4, 0.5 0.5, 0.6 0.6, 0.7 0.7, 0.8 0.8, 0.9 0.9, 1
Proportion of maximizing choices


H
Stated
Learned
Number of particpants
0 5 10 15 20 25 30
0, 0.1 0.1, 0.2 0.2, 0.3 0.3, 0.4 0.4, 0.5 0.5, 0.6 0.6, 0.7 0.7, 0.8 0.8, 0.9 0.9, 1
Proportion of maximizing choices


C
Stated
Learned
Number of participants
0 5 10 15 20 25 30
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Proportion of maximizing choices


F
Stated
Learned
Number of particpants
0 5 10 15 20 25 30
0, 0.1 0.1, 0.2 0.2, 0.3 0.3, 0.4 0.4, 0.5 0.5, 0.6 0.6, 0.7 0.7, 0.8 0.8, 0.9 0.9, 1
Proportion of maximizing choices

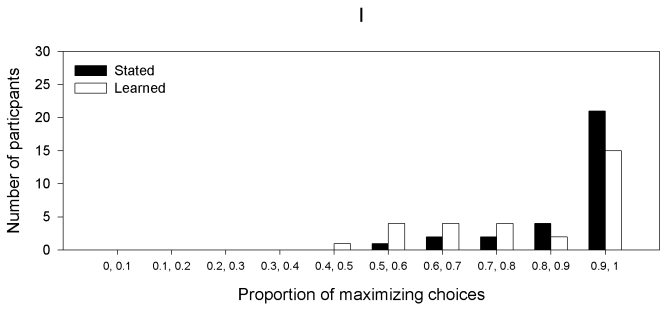

I
Stated
Learned
Number of particpants
0 5 10 15 20 25 30
0, 0.1 0.1, 0.2 0.2, 0.3 0.3, 0.4 0.4, 0.5 0.5, 0.6 0.6, 0.7 0.7, 0.8 0.8, 0.9 0.9, 1
Proportion of maximizing choices

Figure 5: Histograms showing the number of participants who made different proportions of maximizing choices. Panels A to C show responding in the first 10-trial block, Panels D to F show the average proportions of maximizing responses across the whole task, and Panels G to I show average proportions of steady state maximizing choices over the final third of the task. Panels A, D, and G show the Standard condition; Panels B, E, and H show the No-pattern condition; and Panels C, F, and I - Experiment 2.

2.c.i. Subjective estimates of reward probabilities

At the end of each experiment the participants were asked to say what they thought the outcome probabilities were. These data are shown in Table 2. To determine what effect the instructions had on participants' representation of the outcome probabilities these subjective estimates of the outcomes were entered into a 2x2 ANOVA with Probability (Stated vs. Learned) and Pattern information (Standard vs. No-pattern) as between-subjects factors. There was neither an effect of Probability ($F_{1, 116}$ = 1.891, MS_e = .006, p = .172, η_p^2 = .016), nor of Pattern information ($F_{1, 116}$ = 1.761, MS_e = .006, p = .187, η_p^2 = .015), nor a significant interaction ($F_{1, 116}$ = 2.771, MS_e = .006, p = .099, η_p^2 = .023).

Table 2: Subjective estimates of the outcome probability of the high probability option, by condition and experiment.

	Pattern instruction					
	Standard		No-pattern		Experiment 2	
	m	se	m	se	m	se
Stated	.685	.010	.690	.010	.692	.009
Learned	.727	.020	.686	.013	.651	.017

2.c.ii. Heuristics

Participants were asked to rate the extent to which they relied on strategy or intuition to make their decisions (see Table 3). These data were entered into a 2x2 ANOVA with Probability and Pattern information as between-subject factors. A main effect of Probability ($F_{1, 116}$ = 4.803, MS_e = 0.999, p = .030, η_p^2 = .040) showed that participants in the Stated probability conditions reported greater strategy use than those in the Learned conditions. There was no effect of Pattern instruction (F<1.0), and no interaction between the two ($F_{1, 116}$ = 1.201, MS_e = 0.999, p = .855, η_p^2 = .010).

Table 3: Subjective reports of strategy use by condition and experiment (scores range from 1 = pure intuition to 5 = pure strategy).

	Pattern instruction					
	Standard		No-pattern		Experiment 2	
	m	se	m	se	m	se
Stated	3.800	0.176	3.633	0.212	4.000	0.198
Learned	3.200	0.188	3.433	0.149	3.600	0.189

2.c.iii. Apophenia

Participants were also asked to rate their belief that the outcome sequences contained a pattern (see Table 4). A third 2x2 ANOVA revealed neither a significant main effect of Probability condition ($F_{1, 116}$ = 1.904, MS_e = 1.751, p = .170, η_p^2 = .016) nor, surprisingly, of Pattern instruction (F < 1.0). There was no significant interaction between the two factors ($F_{1, 116}$ = 1.542, MS_e = 1.751, p = .217, η_p^2 = .013).

Table 4: Subjective reports of belief that outcome sequences contained a pattern by condition and experiment (scores range from 1=strongly disagree, to 5 = strongly agree).

	Pattern instruction					
	Standard		No-pattern		Experiment 2	
	m	se	m	se	m	se
Stated	2.000	0.209	2.267	0.235	2.267	0.203
Learned	2.633	0.256	2.300	0.263	2.700	0.204

2.c.iv. Alternative factors

As well as questions regarding strategy use and apophenia, participants were asked questions relating to two alternative accounts discussed in the general introduction. First, whether or not they believed it was necessary to select the lower probability option on some trials to have any possibility of obtaining 100% correct predictions. Second, whether or not they agreed that it was sometimes worth deliberately selecting the lower probability option because it is more satisfying to win having done so. These self-report variables are here termed 'Diversify for 100%' and 'Diversify for satisfaction'. Descriptive statistics are detailed in Tables 5 and 6.

Table 5: Agreement that diversification of choices is necessary to obtain 100% success, by condition and experiment (scores range from 1=strongly disagree, to 5 = strongly agree).

	Pattern instruction					
	Standard		No-pattern		Experiment 2	
	m	se	m	se	m	se
Stated	4.367	0.195	4.533	0.157	3.500	0.291
Learned	3.700	0.272	4.433	0.201	4.333	0.211

Table 6: Agreement that greater satisfaction is an adequate motivation for choice diversification, by condition and experiment (scores range from 1=strongly disagree, to 5 = strongly agree).

	Pattern instruction					
	Standard		No-pattern		Experiment 2	
	m	se	m	se	m	se
Stated	2.567	0.266	3.600	0.218	2.233	0.252
Learned	3.100	0.222	2.667	0.251	2.533	0.248

2.c.v. Win-Stay-Lose-Shift analysis

As an additional variable of interest, proportions of choices over the entire task that were consistent with a Win-Stay-Lose-Shift (WSLS) strategy were calculated for each participant. These data are shown in Table 7. Figure 6 shows the distributions of individual-level behaviour consistent with a WSLS strategy for both the Standard and No-pattern conditions (Panels A and B respectively). A 2x2 ANOVA was conducted to assess differences in WSLS proportions between conditions. This revealed no significant main effect of either Probability ($F_{1, 116} = 3.121$, $MS_e = .005$, $p = .080$, $\eta_p^2 = .026$), or Pattern Instruction ($F_{1, 116} = 3.802$, $MS_e = .005$, $p = .054$, $\eta_p^2 = .032$), nor a significant interaction ($F < 1.0$).

Table 7: Proportions of Win-Stay-Lose-Shift consistent choices by condition.

	Pattern instruction			
	Standard		No-pattern	
	m	se	m	se
Stated	.694	.009	.660	.011
Learned	.662	.016	.647	.013

Following this, a linear regression revealed a significant positive effect of reported use of a strategy over intuition on raw proportions of WSLS consistent choices (b = .021, t_{118} = 3.483, p < .001), with strategy use explaining 9% sample variance (R^2 = .093).

However, it is important to consider here that a strict maximizing strategy will lead to a baseline proportion of .7 WSLS consistent choices (comprising 100% 'stay' choices of which 70% will be rewarded 'wins'). An entirely WSLS independent probability matching strategy is calculated to lead to an expected baseline of .58 WSLS consistent choices. As maximizing increases the proportion of WSLS congruent choices will also increase, independently of whether or not WSLS is actively pursued as an approach. This is a concern due to the large discrepancies in maximizing behaviour between participants, in particular the systematic divergence between Stated and Learned probability conditions and the potentially confounding relationship between maximizing and Strategy use.

In order to adjust for this issue, expected baseline proportions of WSLS consistent choices were calculated for each participant based upon their proportion of maximizing choices, over the entire task. These were then subtracted from actual proportions of WSLS consistent choices as a measure of deviation in WSLS like choice patterns over and above what would be expected by chance. These data are shown below in Table 8. Distributions of individual-level proportions of adjusted WSLS consistent choices are shown for both Standard and No-pattern conditions in Figure 6, panels C-D. Another 2x2 ANOVA was then conducted to assess differences in WSLS deviations from baseline between conditions. This revealed significant main effects of both Probability ($F_{1, 116}$ = 11.574, MS_e = .002, p < .001, η_p^2 = .091) and Instruction ($F_{1, 116}$ = 8.287, MS_e = .002, p = .005, η_p^2 = .067), but no significant interaction (F< 1.0). This

indicates that when participants are told that there is no pattern or told the reward probabilities they tend to make fewer choices that are consistent with a WSLS strategy. The histograms suggest that only a few participants across all conditions made substantively more WSLS consistent choices than their expected baseline. Each of these effects seems therefore to result predominantly from a modest shift across a wider range of participants towards making slightly fewer WSLS consistent choices when aware of the reward probabilities or informed that there is no predictable pattern in the sequence.

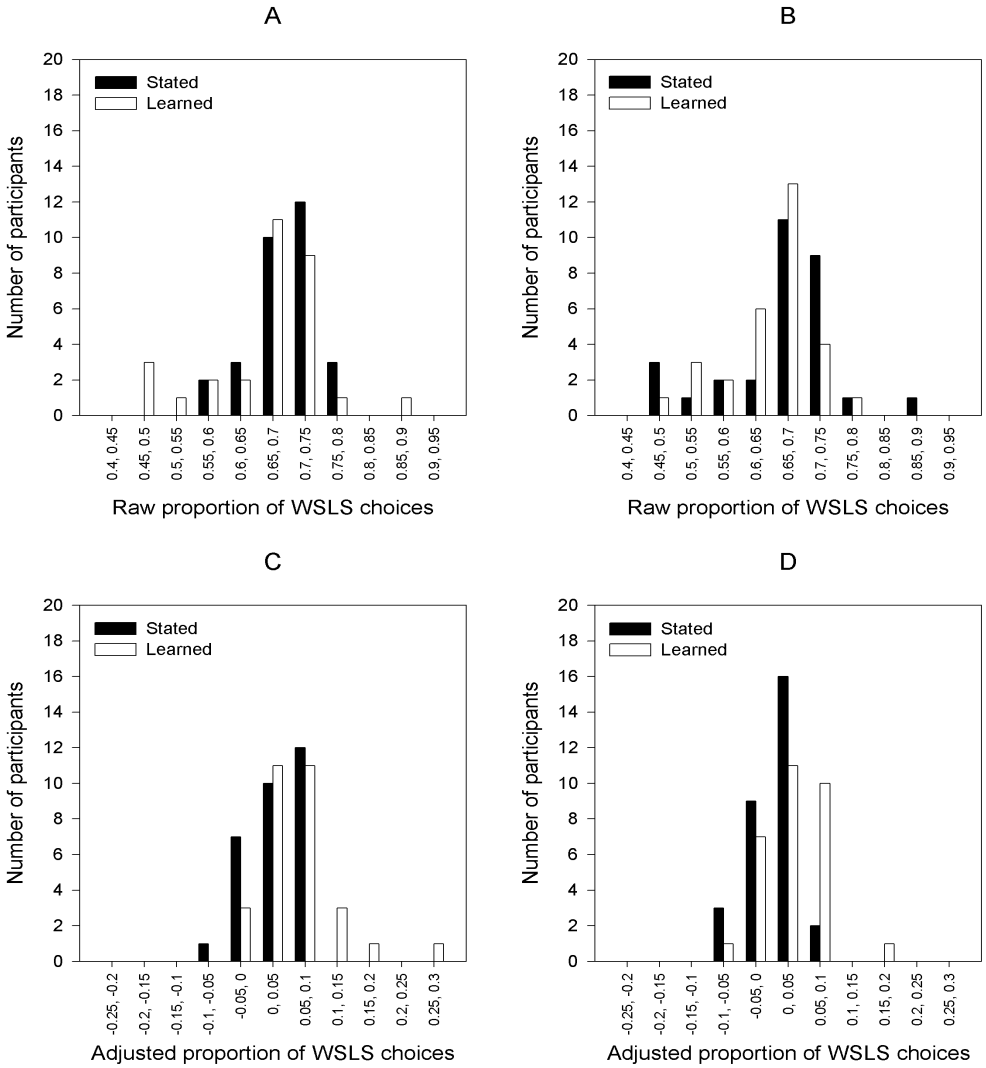


Figure 6: Histograms showing the number of participants and their choices that were consistent with a raw (Panels A & B) and adjusted (Panels C & D) Win-Stay Lose-Shift (WSLS) strategy in Experiment 1: Panels A & C Standard condition & Panel B & D No-pattern condition.

Table 8: Deviation in proportions of Win-Stay-Lose-Shift consistent choices from baseline, by condition.

	Pattern instruction			
	Standard		No-pattern	
	m	se	m	se
Stated	.027	.008	.006	.007
Learned	.062	.011	.031	.009

Following this, a linear regression analysis was conducted to assess the effect of strategy use on this adjusted WSLS variable. This revealed no significant effect of reported strategy use (b = -.007, t_{118} = -1.396, p = .165, R^2 = .016). Interestingly, although non-significant, the trend is for WSLS choices being associated with a more intuitive rather than strategic approach to the task.

2.d. Discussion – Experiment 1

The results from Experiment 1 show that stating the probabilities of reward results in both a greater proportion of maximizing responses and an increase in the use of strategy over intuition, compared to when participants are left to acquire knowledge of the reward probabilities through experience alone. However, informing participants that there were no patterns in the reward sequences, and that they were independent events, neither significantly affected proportions of maximizing choices nor reduced the tendency toward apophenia. These effects were equivalent whether considering choices made over the entire task or steady-state responding over only the final third of trials. The persistence of the positive effect of stating outcome probabilities on maximizing choices post-learning suggests that this reflects a genuine difference in approach to making choices, rather than resulting from the more prosaic difference in information available at the outset of the task. Experiment 2 introduces a potentially stronger

manipulation to reduce the effect of apophenia, by concealing the outcome of each prediction over each 10 trial block.

Experiment 2

In Experiment 1 it was observed that telling participants about the probability of the outcomes affected the predictions that they made. However, telling the participants that there were no patterns to be discovered in the sequence of outcomes, and that the outcomes were independent of the choices that they made, had no significant effect on their predictions. Moreover, although this manipulation was designed to reduce apophenia, no effect was actually observed upon participants' perception of randomness in the sequence of outcomes. In Experiment 2 a potentially stronger manipulation to reduce apophenia is attempted, in the form of concealing the outcome of each prediction and presenting participants instead only with a proportion correct at the end of each 10-trial block. In this manner it is hoped to preclude the possibility of observing a pattern in the trial-by-trial reward sequence, and therefore reduce or eliminate apophenia.

2.e. Method – Experiment 2

2.e.i. Participants

Sixty members of the University of Nottingham community volunteered to take part in this experiment. Of these, 23 were male and 37 were female. Their ages ranged from 17 to 35 (M=22.60, SD=2.97). The methods of recruitment and of calculating individual participants' inconvenience allowances were identical to Experiment 1.

2.e.ii. Design

This was a 2x42 mixed model design with Probability (Stated vs. Learned) as a between-subjects factor and Block (1 to 42) as a within-subjects factor. The basic experimental preparation was identical to Experiment 1 with the exception that the

participants were only given the outcome of their predictions at the end of each 10-trial block.

2.e.iii. Procedure

The procedure was identical to Experiment 1 with the exception that participants did not see the illumination of a light bulb after each trial. Instead feedback was given in written format after 10 trials.

2.f. General Results – Experiment 2

The proportions of maximizing responses for each 10-trial block are shown for each condition in Figure 4B. The distribution of choices for the first 10-trials are shown in Figure 5 Panel C. This suggests, just as in Experiment 1 (Panels A & B), that participants tended to allocate more of their initial responses the maximizing choice when they are told the reward probabilities. These data were entered into a 2x42 mixed model ANOVA with Probability (Stated vs. Learned) as a between-subjects factor and Block (1 through 42) as a within-subjects factor. A main effect of Block ($F_{14.625, 848.246}$ = 2.574, MS_e = .103, p < .001, η_p^2 = .042)[1] and a significant linear effect of Block ($F_{1, 58}$ = 19.877, MS_e = .098, p < .001, η_p^2 = .255) indicated that participants learned to allocate more responses to the maximizing alternative as the experiment progressed. A main effect of Probability ($F_{1, 58}$ = 10.761, MS_e = .775, p < .002, η_p^2 = .157) indicated that when participants were told the probabilities of reward they tended to allocate more responses to the more likely outcome than when they had to learn them from experience alone. There was no interaction between Block and Probability (F<1.0).

To isolate the effect of Probability condition on steady-state responses, an independent samples t-test was conducted. This found that the positive effect of Stated over Learned reward probabilities on proportion of maximizing choices remained over

the final third of the task (t_{58} = 2.458, p = .017). This effect was therefore again found not to result from a discrepancy between groups during the learning phase, but to indicate a more substantive difference in decision making between the two conditions.

2.f.i. Subjective estimates of reward probabilities

Subjective estimates of outcome probabilities are shown in Table 2. An independent samples t-test revealed that participants in the Learned probability condition estimated the outcome probabilities to be lower than the Stated probability condition (t_{58} = 2.191, p = .032).

2.f.ii. Heuristics & Apophenia

Self-reports of strategic choices and belief that the outcome sequences contained patterns are shown in Tables 3 and 4. These did not differ between groups (t_{58} = 1.146, p = .149; t_{58} = 1.505, p = .138).

2.f.iii. Comparisons with Experiment 1

To assess the efficacy of the aggregate feedback manipulation introduced in Experiment 2, these results were compared to the Standard condition of Experiment 1.

The proportion of maximizing responses for each 10-trial block were entered into a 2x2x42 mixed model ANOVA with Probability (Stated vs. Learned) and Pattern information (Standard vs. Experiment 2) as between-subjects factors and Block (1 through 42) as a within-subjects factor. A main effect of Block ($F_{18.233,\ 2115.021}$ = 7.314, MS_e = .062, p < .001, η_p^2 = .059)[1] and a significant linear contrast of Block ($F_{1,\ 116}$ = 58.200, MS_e = .118, p < .001, η_p^2 = .334) indicated that participants learned to allocate more responses to the maximizing option as the experiment progressed. A main effect of Probability ($F_{1,\ 116}$= 39.338, MS_e = .657, p < .001, η_p^2 = .253) indicated that when

participants were told the probability of reward they tended to allocate more responses to the more likely outcome than when they had to learn it from experience alone. Although there was no reliable interaction between Block and Probability ($F_{18.233, 2115.021}$ = 1.461, MS_e = .062, p = .094, η_p^2 = .012), modelling the interaction between Block and Probability as a linear contrast did reveal a significant effect ($F_{1, 116}$ = 6.906, MS_e = .118, p = .010, η_p^2 = .056). This suggests that stating the probabilities again affected the rate at which participants learned to allocate responses to the more likely outcome. However, there was no effect of Pattern information (F < 1.0). There were also no significant interactions between Pattern information and Probability ($F_{1, 116}$= 1.515, MS_e = .657, p = .221, η_p^2 = .013) or Block ($F_{18.233, 2115.021}$ = 1.292, MS_e = .062, p = .181, η_p^2 = .011), nor a reliable linear interaction between Pattern information and Block ($F_{1, 116}$ = 1.192, MS_e = .118, p = .277, η_p^2 = .010). These indicate that concealing the trial-by-trial outcome sequence did not influence participants' choices. There were no 3-way interactions (F < 1.0).

In order to again assess experimental effects exclusively upon steady-state responding, a separate 2x2 between subjects ANOVA was conducted, with factors of Probability condition and Pattern information, and the dependent variable of proportion of maximizing choices made over the final third of the task. This revealed a main effect of Probability condition ($F_{1, 116}$= 20.538, MS_e = .020, p < .001, η_p^2 = .150), but neither a main effect of Pattern information (F < 1.0), nor an interaction effect (F < 1.0).

2.f.iv. Subjective estimates of reward probabilities

Subjective estimates of outcome probabilities are shown in Table 2. A 2x2 ANOVA was conducted with Probability (Stated vs. Learned) and Pattern information (Standard vs. Experiment 2) as between-subjects factors. There was no effect of

Probability ($F < 1.0$). However, a main effect of Pattern information ($F_{1, 116} = 5.685$, $MS_e = .006$, $p = .019$, $\eta_p^2 = .047$) suggested that participants in Experiment 2 estimated the outcome probabilities to be lower than those in the Standard condition of Experiment 1. A reliable interaction between Probability and Pattern information ($F_{1, 116} = 8.226$, $MS_e = .006$, $p = .005$, $\eta_p^2 = .066$) indicated that stating the reward probabilities moderated the difference in estimates between the Standard condition of Experiment 1 and the aggregate feedback condition of Experiment 2.

2.f.v. Heuristics

Self-reported use of strategy vs. intuition (see Table 3) was entered into a 2x2 ANOVA with Probability and Pattern information as between-subject factors. A main effect of Probability ($F_{1, 116} = 7.085$, $MS_e = 1.059$, $p = .009$, $\eta_p^2 = .058$) showed that participants in the Stated probability conditions reported greater strategy use than those in the Learned conditions. There was no significant effect of Pattern information ($F_{1, 116} = 2.550$, $MS_e = 1.059$, $p = .113$, $\eta_p^2 = .022$), and no interaction between the two ($F < 1.0$).

2.f.vi. Apophenia

Self-reported beliefs that the outcome sequences contained a pattern (see Table 4) were entered into a third 2x2 ANOVA. A significant main effect of Probability condition ($F_{1, 116} = 5.923$, $MS_e = 1.441$, $p = .016$, $\eta_p^2 = .049$) indicated that participants in the Learned probability conditions reported significantly higher levels of apophenia than those in the Stated conditions. Surprisingly, as in Experiment 1, there was neither an effect of Pattern information ($F < 1.0$) nor an interaction between the two factors ($F < 1.0$).

2.g. Discussion – Experiment 2

Experiment 2 attempted to use a potentially stronger manipulation to reduce the effect of apophenia in probability learning. However, the effect of concealing the outcomes over each 10-trial block was similarly ineffective as simply telling participants that there was no pattern in the sequence. It can be inferred from this that participants were just as capable of internally generating a perception of a sequential pattern as they were of mistakenly deriving one from an observed outcome sequence. No WSLS analysis was conducted on these data because the participants could not see the outcomes of individual trials and could not therefore base a subsequent choice on the result of any single preceding trial.

The next section documents the use of mediation analysis to determine the extent to which the observed effect of Probability condition on participants' choices was mediated by its effects on strategy usage and apophenia.

2.h. Mediation Analyses

Experiments 1 and 2 each found a significant effect of Probability condition on participants' proportion of maximizing choices. Experiment 1 also found a significant effect of Probability condition on self-reported use of Strategy. Although, when considered alone, the same effect did not reach significance within the sixty participants in Experiment 2, the observed effect was both in the same direction and of an equivalent size to the combined effect observed across the Experiment 1 conditions. When considering the same Aggregate feedback group together with the Standard condition of Experiment 1, the increase in power found the effect of Probability condition on levels of strategy use to be once again statistically significant. By contrast, no effects of Pattern instruction or Pattern information, in Experiment 1 or 2 respectively, were found upon either strategy use or apophenia.

To investigate the extent to which the effect of Probability condition on choice behaviour was mediated by its influence on participants taking a strategic versus intuitive approach to the task and their levels of apophenia, a Mediation Analysis was conducted. This used the PROCESS macro for SPSS (Hayes, 2012, 2013). Data was concatenated across all three apophenia conditions, due to the lack of effect of either Pattern instruction (Experiment 1) or Aggregate feedback (Experiment 2) on either choice behaviour, levels of strategy use or apophenia. For this analysis, proportions of maximizing choices were calculated for the final third of the task, with the aim of accounting for learning effects. Mediation model No. 6 was used, with a single independent variable and two mediator variables. Effects were calculated for each of 10,000 bootstrapped samples. Model results are presented in Figure 7 (Variable coding is as follows: Stated = 0 Learned = 1, Intuition 1 – 5 Strategy, Belief no Pattern 1 – 5 Belief was Pattern, Proportion of Maximizing Choices 0 – 1).

The direct effect of Probability condition on strategy usage was found to be significant (t_{178}= -2.628, p = .009), with this predictor accounting for 4% of the sample variance (R^2= .037). The direct effect of Probability condition on apophenia was found to be non-significant (t_{177}= 1.187, p = .237), while the direct effect of strategy use on apophenia was found to be significant (t_{177}= -4.306, p < .001), with these two predictors accounting for 11% of the sample variance (R^2= .114). The direct effects of Probability condition (t_{176}= -3.776, p < .001), and strategy use (t_{176}= 7.351, p < .001), on proportion of maximizing choices were each found to be significant, whereas the direct effect of apophenia was not (t_{176}= -.309, p = .758). These three predictors accounted for 34% of the sample variance (R^2= .341).

The unstandardized indirect effects of Probability condition on proportion of maximizing choices were as follows: Through intuition versus strategy use (-.400)(.068)= -.027, with bias corrected 95% confidence intervals ranging from -.051 to -.008,

indicating statistical significance of this effect at α =.05. Through apophenia (.216)(-.002)= -.0005, with bias corrected 95% confidence intervals ranging from -.009 to .003, indicating non-significance of this effect at α =.05. Through first strategy use then apophenia (-.400)(-.378)(-.002)= -.0004, with bias corrected 95% confidence intervals ranging from -.004 to .002, indicating that this effect was also non-significant at α =.05.

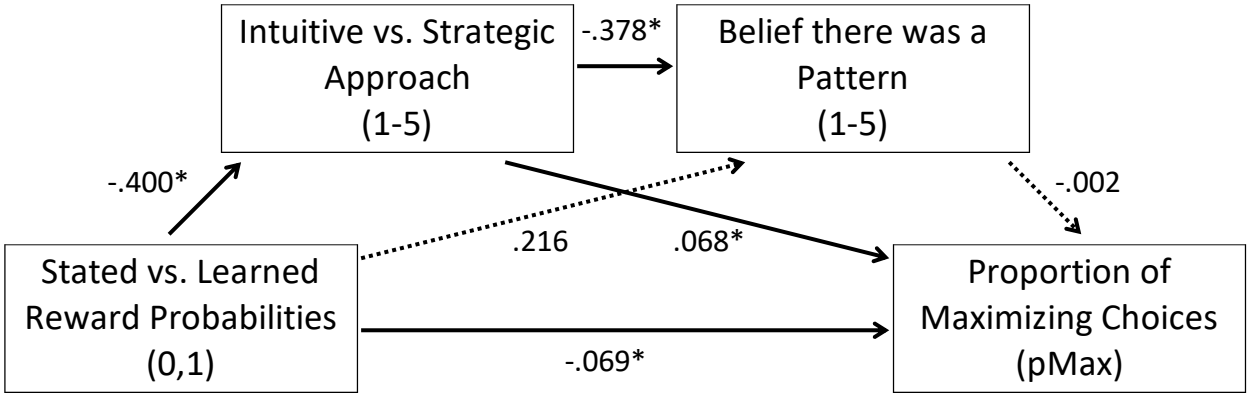


Figure 7: Mediation model for all participants. Unstandardized regression coefficients shown with asterisk if significant. Solid and dotted lines indicate significant and non-significant effects respectively. α = .05.

2.i. Assessment of Reliability of Within Group Effects

In the mediation analysis data from all three apophenia conditions were concatenated. In light of this, it remains unclear whether these effects are consistent between individual experimental conditions. To address this issue, a series of multiple regression analyses were conducted, for each experimental group. Predictor variables were self-reported strategy use and apophenia, and the dependent variable was maximizing choices over the final third of the task. These revealed a significant positive effect of strategy use on proportions of maximizing choices within each experimental condition, but no significant effect of apophenia in any experimental condition. Results are shown in Table 9.

Table 9: Effects of Strategy use and Apophenia, split by experimental condition.

Experimental Condition	Strategy Use				Apophenia			
	b		*p*		*b*		*p*	
	Learned	Stated	Learned	Stated	Learned	Stated	Learned	Stated
Standard	.077	.032	.019*	.012*	-.009	-.015	.688	.158
Instruction	.086	.064	.017*	<.001*	.034	.013	.088	.265
Aggregate	.072	.078	.011*	<.001*	-.049	.001	.058	.939

(*b*-values unstandardized, *p*-values uncorrected, asterisks represent significance at α=.05)

2.j. Controlling for Variance in Perceived Reward Probabilities: Summary of a Parallel Analysis

In principle, probability matching may be accurately defined as the selection of each alternative in proportion to its subjectively perceived likelihood of reward, even where this differs from an option's objective reward probability. Since the participants in the Learned probabilities conditions are never explicitly told the reward contingencies, it follows that their perceived likelihoods may differ from the actual values. Although group average estimates of these probabilities are generally well calibrated to actual probabilities, greater variance is seen in both the choice behaviour and probability estimates of participants in the Learned probability conditions. Separate linear regressions reveal a reliable effect of perceived reward probabilities on proportions of maximizing choices within the Learned (b = .775, t_{88} = 4.638, p < .001, R^2 = .196) but not the Stated (b = .202, t_{88} = 0.973, p = .333, R^2 = .011) reward probability conditions. This supports the hypothesis that the disproportionate variance in proportions of maximizing choices when probabilities are learned is largely driven by individual differences in perceptions of the reward probabilities.

At least one previous study has controlled for individual differences in perceived reward contingencies (Koehler & James, 2009). A parallel analysis was conducted to

investigate the effects of controlling for this variance in a similar manner. A new measure of choice behaviour was created for the purposes of this analysis: pMax – pEst. This measure accounts for between subject variance in perception of reward probabilities through simply subtracting each participant's subjective estimate of the reward probability of the more likely option from their proportion of maximizing choices over the final third of the task. Group averages are shown in Table 10.

Table 10: Deviation of choices from probability estimates within each experimental group.

Experimental Condition	pMax - pEst	
	Learned	Stated
Standard	.078 (.027)	.259 (.015)
Instruction	.168 (.025)	.219(.023)
Aggregate	.170 (.027)	.224 (.019)

(Standard Error shown in brackets)

The pattern of results with respect to the effect of Apophenia manipulations is different from the analysis in which individual variance in perceptions of reward probabilities was not accounted for. In the present case, and treating the three conditions together, while there was no significant main effect of Apophenia condition on deviation from matching $F(2,174)<1$, a significant interaction between Probability and Apophenia conditions was evident $F(2,174)=4.948$, $p=.008$, $\eta_p^2=.054$. This effect is driven by the relatively low apparent deviation from matching in the Learned Standard group relative to the Learned probabilities conditions of each Apophenia manipulation. By contrast, no sign of such an interaction effect on choice behaviour involving either Pattern instruction or Pattern information was apparent in the primary analysis of Experiment 1 or 2, in which variance in estimates was not adjusted for.

Similarly, parallel mediation analyses indicated a significant negative effect of apophenia on apparent deviation from matching (pMax-pEst) within the Standard condition only (t_{56}= -3.219, p = .002). However, it is clear from our primary analysis that this result does not reflect any effect of apophenia upon the proportions of maximizing choices. The alternative explanation is that this is an illusory effect driven instead by an unforeseen relationship between apophenia and probability estimates. The presence of this relationship was confirmed, through correlation analysis, within the Standard (r=.342, N=60, p=.007) but neither Instruction (r=.148, N=60, p=.260) nor Aggregate feedback (r=-.166, N=60, p=.206) conditions.

Reasons for this relationship are unclear. I am unaware of any other studies that have found an equivalent effect and it may represent nothing more than a Type 1 error. However, these results suggest that raw proportions of maximizing choices are, on balance, the more reliable measure. A majority of previous studies have used this measure. However discrepancies highlighted from running this parallel analysis do raise concerns, and if replicated may prove a substantial consideration for any studies that wish to adjust for variance in perceived reward probabilities when determining matching behaviour.

2.k. Assessment of Alternative Factors.

Although these experiments focused on the two predominant accounts of matching behaviour, they were also designed to investigate the potential contribution of alternative factors that were discussed in the general introduction. To do so, participants' self-reported agreement with two statements was assessed: First, that it was necessary to select the lower probability option on some trials to have any possibility of obtaining 100% success, and second, that it was sometimes worth deliberately selecting the lower probability option because it is more satisfying to win having done so. These self-report

variables are here termed 'Diversify for 100%' and 'Diversify for satisfaction' (see Tables 5 and 6 respectively).

Following the primary analyses, the self-report data pertaining two these two factors were entered into another simultaneous multiple regression analysis, together with self-reported Intuition vs. Strategy use and levels of Apophenia. Together the four factors accounted for a significant proportion of the sample variance in proportions of steady state maximizing choices $F(4,175) = 17.895$, $p < .001$, $R^2 = .290$, $R^2_{Adjusted} = .274$). With respect individual factors, this analysis revealed a significant effect of Strategy use ($b = .073$, $t(179) = 7.601$, $p < .001$), but no significant effect of either Apophenia ($b = -.005$, $t(179) = -0.577$, $p = .565$), Diversify for 100% ($b = -.002$, $t(179) = -0.327$, $p = .744$) or Diversify for satisfaction ($b = .006$, $t(179) = 0.829$, $p = .408$).

2.1. General Discussion

These studies examined the effects of apophenia and heuristics, along with alternative factors, on choice behaviour in a probability learning task. Experiment 1 manipulated the information that participants received in order to reduce the effects of apophenia and heuristic decision making. Experiment 2 used a potentially stronger manipulation, namely concealing trial-by-trial outcomes, to reduce apophenia. Self-report measures of belief that the sequence had contained a pattern, and of intuitive or strategic approaches to the task were also collected.

The results of Experiment 1 indicated higher learning rates, unsurprisingly, in the Learned probabilities condition. Relatively lower learning levels when probabilities were Stated are limited by the ceiling of maximizing behaviour coupled with a higher initial level of maximizing responses, given the possibility of determining the optimal strategy with reward contingency information before the task began. This pattern of results is congruent with those of Newell, Koehler, James, Rakow, and van Ravenzwaaij (2013), in

which an interaction is described between higher initial maximizing driven by 'top-down' information and more slowly increasing levels resulting from 'bottom-up' feedback-based learning.

Overall, the explicit statement of outcome probabilities led participants to make a significantly greater proportion of maximizing choices than when probabilities required learning through experience. Similar effects have been reported by both Nies (1962) and Fantino and Esfandiari (2002). Steady state analyses revealed this to remain the case even in the post-learning period of the task. In fact, strikingly, the average proportion of maximizing choices over the final 10 of the 420 trials, for participants who had to learn reward contingencies through experience alone (m = .793 se = .040), remained lower than the equivalent proportion on the first 10 trials for participants who were explicitly informed of reward probabilities before the task (m = .853 se = .028). This suggests that there may be some qualitative impediment to reaching the maximizing response through experience alone. It also extends a finding of Newell and Rakow (2007), that there is a negative effect of experience relative to description after 60 trials.

Participants in the Stated probabilities conditions were also found to be significantly more likely to take a strategic, rather than intuitive approach to the task. By contrast, explicitly telling participants that the reward sequence would be random, with no predictable pattern, neither significantly affected choice behaviour or levels of apophenia. This lack of behavioural effect of stressing the randomness of the sequence seems inconsistent with the findings of both Rubinstein (1959) and Peterson and Ulehla (1965). However, the absence of any effect upon even self-reported levels of apophenia suggests that written instruction might simply not be sufficient to persuade participants of the randomness of the sequence.

In addition, proportions of choices within each condition of Experiment 1 that were predictable by a simple Win-Stay-Lose-Shift strategy were assessed, while

controlling for the confounding effect of raw proportions of maximizing choices. This revealed that participants made significantly fewer WSLS consistent choices when probabilities were Stated before the task, and also when they were instructed that there would be no predictable pattern in the reward sequence. No significant relationship was found between Strategy-use and proportions of WSLS consistent choices, while the trend was for WSLS choice behaviour to be related to an Intuitive rather than Strategic approach to the task. This result seems to run contrary to a finding of Otto, Taylor, and Markman (2011), that WSLS behaviour was less prevalent in participants in a dual-task situation that was designed to compromise executive resources.

In Experiment 2 outcome sequences were concealed from participants, reward feedback was provided only in aggregate format, over 10-trial blocks. It was hypothesised that this would be a stronger method of reducing apophenia than simply informing participants that there would be no pattern. However, comparing this dataset with the standard condition of Experiment 1 revealed that aggregating feedback was also ineffective, both at influencing choice behaviour and reducing levels of apophenia. The absence of any effect here is surprising, particularly considering that, in an almost identical manipulation Gao and Corter (2015) did find an effect of presenting grouped rather than single trials. It is instead more consistent with Koehler and James (2009), who also observed no effect of providing outcome information in aggregate format, although the present study masked trial-by-trial outcomes that were concurrent with responses, rather than the outcomes of an independent pre-task learning phase.

In a novel development with respect to the findings of previous studies that have investigated the effects of heuristics or apophenia on probability matching, a mediation analysis was applied to more clearly delineate the pathways of action of any observed effects. This revealed a significant indirect effect of stating reward probabilities, via facilitation of strategy use over intuition, on choice behaviour. This finding corroborates

the mechanism by which this experimental manipulation was hypothesised to take effect. These findings are consistent with a heuristic based account of probability matching, by which matching is an intuitive response and maximizing depends upon more rational deliberation, as espoused by Kogler and Kuehberger (2007) and Koehler and James (2010). This finding also further discredits an assumption of Thompson-Schill et al. (2009), that maximizing relates from 'brute strength competition' and probability matching relates to prefrontal maturation linked with greater cognitive control. In fact, the findings of the present study would suggest the opposite hypothesis: that, as maximizing results from the application of greater cognitive control, in the form of applying a Strategy to overrule the more intuitive matching response, it is therefore most likely to relate to greater prefrontal activity.

This mediation analysis also indicated that not only were attempted manipulations of apophenia ineffective, but that apophenia itself had no substantive effect on choice behaviour. This fails to replicate results of Unturbe and Corominas (2007). It also fails to support a core assumption made by Wolford et al. (2000), that matching behaviour can be used as a proxy for pattern search. It is important to note that the questionnaire method used in the present study can only account for post-experimental levels of apophenia. It is possible that some participants did at times during the task believe there to be a pattern in the sequence, and subsequently learned that this was not the case. This is an interesting consideration, and may merit further study in itself. However, as the effects of apophenia were compared with steady state responding it is ultimately steady state levels of apophenia that are the most appropriate corresponding measure. Post-experimental questioning should be suited to capture this with a sufficient degree of accuracy.

One further intriguing result of the mediation analysis was that it highlighted a significant negative relationship between strategy use and apophenia, despite pattern

search being a seemingly strategic approach. This is likely due to a substantial sub-set of participants who applied a highly deliberative approach in determining that the maximizing strategy was optimal, and who also realized that this was incompatible with the presence of a predictable pattern in the reward sequence.

To emphasize the reliability of these findings, a series of separate multiple regression analyses revealed that the significant positive effect of Strategy use on steady state proportions of maximizing choices was evident within each of the six experimental conditions, across both experiments. By contrast, there was no significant effect of reported apophenia evident in any of these conditions. A further analysis was also conducted to assess the potential contribution of two further alternative factors, in conjunction with Apophenia and Strategy use. These were that participants may believe diversification of choices is necessary to have any chance of achieving 100% success, and that deliberate selection of the lower probability bulb may be attractive due to a higher subjective utility being obtained from successfully predicting it. Consistent with the previous results, this analysis revealed a significant and substantial effect of Strategy use on proportion of steady state maximizing choices, but no such effect for Apophenia or either of the two additional factors.

A parallel analysis was also conducted in which individual differences in participants' estimates of the reward contingencies was adjusted for, in a similar manner to that of Koehler and James (2009). This analysis revealed remarkably different results than the primary analysis with respect to the apparent effects of apophenia on choices. However, further investigation revealed that these effects were driven by a perplexing relationship between levels of apophenia and estimates of reward probabilities, which was present only within the Standard task condition.

In summary, results failed to replicate previous findings that have implicated apophenia as a key determinant of probability matching. They also do not provide any

evidence in support of two less prominent alternative factors. By contrast, they do indicate a robust effect of intuition versus strategy use, and support the conclusion that this is the primary factor behind a predisposition toward matching or maximizing behaviour respectively. Furthermore, a mediation analysis provided statistical verification of the hypothesis that the increase in maximizing behaviour when reward probabilities are Stated is driven largely by the promotion of a more Strategic approach to the task. These findings are consistent with a heuristic based account, by which the matching response is prepotent when the decision-maker acts intuitively, but may be overruled if the maximizing strategy is recognized. Additional findings included an unforeseen relationship between self-reported levels of apophenia and perceived reward probabilities, higher levels of WSLS consistent choices when participants began the task naive of either reward contingencies or the statistical independence of outcomes and the absence of any reliable relationship between WSLS and overall strategy use.

Chapter 3: A Behavioural Assessment of Faulty Foraging Theory on a Simple Simulated Foraging Task

3.a. Introduction

In the previous chapter, the contributions of multiple putative factors in determining maximizing versus diversification of choices were assessed on the standard task, in which reward probabilities are fixed and i.i.d. Results indicated that taking an intuitive versus strategic approach to the task is the predominant factor behind matching versus maximising behaviour. This is consistent with a heuristic-based account of probability matching; in which matching is the more readily available response. In this case probability matching is dominant when acting intuitively, but may be overruled upon recognition and application of the maximising strategy, through sufficient deliberation and subsequent cognitive control. By contrast, no evidence was found in support of apophenia-based, or two other accounts of probability matching.

However, as discussed in Chapter 1, while such a heuristic account may give a superficial explanation of probability matching, it does not really get to the root of the issue. After hundreds of millions of years of natural selection it could be expected that a generally optimal rule would have emerged as the default response. This begs the question of how such an apparently suboptimal tendency as probability matching could have managed to both become established, and remain to this day, as the dominant intuitive strategy. As also discussed in Chapter 1, one proposal is that probability matching might reflect the expression of a generally optimal decision criterion for the purposes of time allocation in foraging situations, which has become established over an extended period of evolutionary ancestry. In humans this rule is therefore expressed predominantly when the actor behaves in accordance with their intuitions. Hereafter, this account will be termed the 'Faulty Foraging Theory' (FFT). Under this interpretation, humans do in fact have what has been generally the most adaptive response pattern

readily and naturally available. This leads to the dominance of this choice strategy when alternatives lack availability, as discussed by Koehler and James (2010). Therefore, the expression of probability matching behaviour when confronted with the standard 2AFC task, in which reward contingencies are fixed and i.i.d, might represent an overextension of what has been an otherwise normative behaviour.

As humans however, perhaps uniquely, we also possess a level of analytic ability that allows us to logically calculate the instrumentally maximizing response for a particular task. This is coupled with a level of top-down cognitive control that allows us to override our instinctive response in situations where it is recognised that this would not result in maximum pay-offs. However, to do so requires a degree of effortful consideration to determine the optimal strategy, coupled with a degree of effortful control to assert this strategy. This effort is not always invested, particularly when the stakes are low. This results in a spread of participants between predominantly matching and maximizing response behaviours, with the primary predictor for where a participant's behaviour will fall on this spectrum being the degree of deliberation that they put into deciding their choices. In this manner, FFT is eminently compatible with heuristic-based accounts of probability matching, which are supported both by the findings of the previous experimental chapter and a comprehensive body of existing empirical literature (Koehler & James, 2010; Kogler & Kuehberger, 2007; West & Stanovich, 2003). Moreover, FFT further extends these accounts through providing a theoretical basis for the somewhat counterintuitive (but empirically well-supported) finding that probability matching seems to have become established as the naturally most readily available and intuitively prepotent behaviour, despite, on the basic level, being a seemingly more computationally complex response pattern than maximizing.

A number of previous studies have tested whether human matching behaviour falls in line with the Ideal Free Distribution (IFD, see Chapter 1) when under the

influence of competition (Kraft & Baum, 2001; Kraft et al., 2002; Madden et al., 2002; Schulze et al., 2015; Sokolowski et al., 1999). By contrast, at the time that this experiment was conducted, no published study had assessed the equivalent behavioural predictions for a purely individual situation, in which the optimality of probability matching derives instead purely from the prospect of resource accumulation. Here, we can consider an individual's choice allocations to be distributed over time in accordance with the IFD, rather than the more conventional application of the IFD to the spread of a group of competitors' choices over space. One study did include a similar assessment in Rhesus monkeys (Sugrue et al., 2004). This comprised a 'dynamic foraging task', whereby choices were made between two options with varying ratios of reinforcement, and which allowed for the accumulation of assigned rewards over time. This study found a close approximation of choice proportions to reward probabilities on this task.

The present study was designed to test whether this finding can be extended to human behaviour. It will also be informative to assess whether the finding of undermatching, as observed in a number of competition-based studies, is replicated. To do so, choice behaviour will be examined on a simple simulated foraging task (SSFT). This task modifies the standard probability matching preparation to incorporate the accumulation and depletion of resources over time into the underlying reward schedule. Task instructions and choice stimuli are also changed to superficially denote a foraging situation, however the task still represents a simple 2AFC task. Analysis will not solely comprise the examination of behaviour on the SSFT. Participants' choice behaviour on this task will then be compared and contrasted with that on the standard probability matching preparation. An investigation into the effects of manipulation of reward probability information and self-reported intuition vs. strategy will also be conducted, as was previously done for the standard task. This pattern of effects will also be contrasted across the two tasks.

It is hypothesised that participants will generally probability match on the SSFT, possibly with some degree of undermatching. It is also predicted that choices on the foraging task will no longer show any effect of either stated vs. learned reward probabilities, or of intuition vs. strategy usage. When considering the two tasks together these should manifest as significant interaction effects, between probability condition and task and also between strategy use and task.

In an important recent development, since the time of conducting the present experiment, another study has been published with an analogous theoretical motivation (Schulze et al., 2017). Here, the authors similarly acknowledge that reward-persistence in an unchosen option will make diversification an effective choice strategy. The study then compared the effects of 'lingering' rewards on participants' actual patterns of choice behaviour. This study compared a 'reward-hold' condition, in which an unharvested reward would persist until collected, to a 'no-hold' condition, in which it wouldn't. This manipulation is functionally very similar to the incorporation of the accumulation of unharvested resources in the present study's SSFT, with the key difference being that only a single reward could be held in each option at any given time, irrespective of how long it had remained unexploited. By contrast, on the SSFT resource accrual was allowed to continue indefinitely, in line with each option's respective resource generation probability. This means that there could be more than a single reward held in the 'bank' at an option that had not been chosen for multiple trials. The findings of these two studies will be compared in the discussion section of this Chapter.

3.a.i. Further motivations and considerations

There were two additional reasons to conduct this study. The first is purely pragmatic, it was planned to later apply this paradigm in conjunction with EEG recording, in an attempt to measure internal expected value coding, and where the

representation of value accumulation over time in the unchosen option would be of particular interest. It is necessary to test whether this SSFT is suitably clear and understandable for participants in a more simple behavioural study prior to implementing it together with EEG.

Second, it will be important to assess whether an alternative heuristic-based 'win-stay lose-shift' (WSLS) strategy, by which participants could simply repeat their previous choice if it was successful and switch to the other option if unsuccessful, may turn out to dominate behaviour on this task. As a simple decision rule, WSLS may be an attractive strategy due to the reduced effort that it requires compared to the repeated recalculation of expected value of each option on each trial. Despite its simplicity, on this task a WSLS strategy could approximate optimal outcomes. Any loss will indicate having fully exploited the stored resources of a given option, therefore reducing it to the base probability of reward accumulation and prompting the need to switch to selection of the other option. This approach may seem to lead to slightly suboptimal choices, as it involves waiting for an unsuccessful forage rather than allowing for more precise anticipatory switching between alternatives. Nevertheless, in practice this would not negatively impact overall winnings apart from at the very end of the task, as the delay in switching would be offset by the prospect of further accumulation in the unchosen option, extending the possible sequence of successful forages on the next run of choices.

Assessment of the prevalence of WSLS use on this task is important for three reasons. First it will be theoretically informative. Second, WSLS would appear identical to standard probability matching in long run, as the proportion of choices of each option will remain dependent on their relative rates of reward. It should therefore be considered a possible confound, as participants appearing to probability match may in fact be doing so solely as a result of a qualitatively different choice-rule. A third concern raised by the possible use of WSLS would be that, as choices would be being made based on a simple

rule rather than a calculated value judgement, this may entail that there will be no clear internal representation of expected values. This could render the planned future EEG version of the task futile.

3.b. Method

3.b.i. Participants

Sixty participants completed the SSFT. These were recruited in an opportunity sample, drawn primarily from the student population at the University of Nottingham. Of these, 17 were male and 43 female. Their ages ranged from 17 to 44 (M=24.03, SD=4.74). In the standard task group, there were a further 60 participants, recruited in the same manner but in a previous experimental session. In this group 16 participants were male and 44 female, with an age range of 18 to 33 (M=22.62, SD=3.09).

Participants were paid an inconvenience allowance with a value contingent upon the choices that they made during the task. Optimal exploitation of resources would accumulate average total winnings of £5.88. The experiment was approved by the University of Nottingham ethics committee, according to BPS guidelines.

3.b.ii. Design

This experiment took a 2x42 design, with Probability condition (Stated vs. Learned) as a between subjects factor and Block (1 to 42) as a within subjects factor. These data would also be contrasted with the standard condition of Experiment 1 in a 2x2x42 design, with Task as an additional between subjects factor. In the Stated Probability condition participants were informed that the two patches of land would accumulate resources probabilistically, at the different rates of 0.49 vs. 0.21. In the Learned Probability condition participants were not provided with any information about reward accumulation rates.

Which of the two colours had the higher rate of resource accumulation was counterbalanced between participants, to account for any bias in favour of a particular colour. All participants were randomly assigned to an experimental condition.

3.b.iii. Stimuli

A plain white background was maintained throughout the task. A black central fixation cross was used between trials. Stimuli consisted of images of two identical patches of woodland, with a blue and a yellow signpost to distinguish between them (see Figure 8). These were positioned equidistant from the centre and the left or right edge of the screen, with the side to which each colour stimulus was presented randomised between trials. The experimental procedure was designed and coded within Psychopy (Peirce, 2007).



Figure 8: Example choice stimuli for Experiment 3 (SSFT).

3.b.iv. Procedure

The procedure was in many ways identical to Experiments 1 and 2 (see previous Chapter). Two principal changes were made to the task. The first of these was superficial, the SSFT was explicitly described as a foraging task, and choices were made between two patches of land instead of two light bulbs. Similarly feedback was not presented in the form of a light switching on, but was instead given in the form of a green tick or a red cross that was overlaid on the image of the patch of land that had been chosen (see Figure 9). The second, more substantive, change was to the underlying reward schedule.

The probabilities of a new resource appearing in each patch were fixed throughout the task, at 0.49 and 0.21 for the high and low accumulation rate options respectively. However, uncollected rewards were allowed to persist and accumulate over time. Actual reward contingencies therefore varied over the task, in a manner that depended upon the participants' preceding choices. The accumulation rates of both patches were set so that they were proportional to the reward probabilities on the standard task, but when considered together sum to 0.7. This means that an optimal strategy could expect to obtain a reward on an average of 70% of trials, maintaining maximum pay-offs on the SSFT as equivalent to those of the standard probability matching task. Otherwise, the task remained an iterated two-alternative forced choice, in which repeated decisions were made between the two patches of land over 420 trials, with the aim of maximizing total reward over the entire task. All stimulus durations, inter-trial and inter-stimulus intervals remained the same. Total average trial length remained 5000ms and average overall task duration remained approximately 35 minutes.

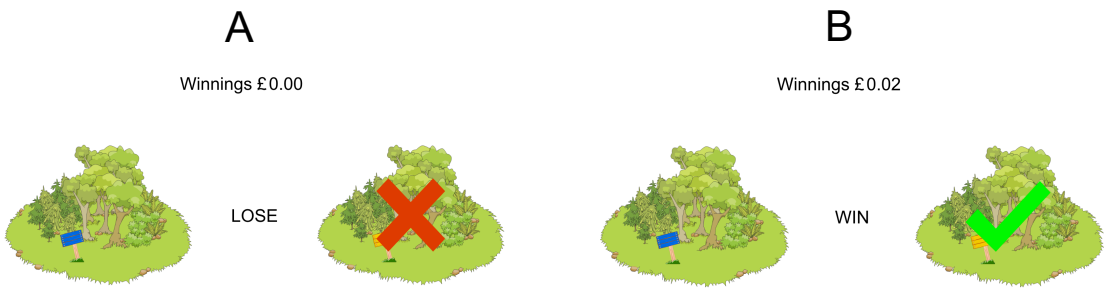


Figure 9: Example feedback stimuli for Experiment 3 (SSFT). Panel A shows an unsuccessful forage attempt. Panel B shows a successful forage attempt.

3.b.v. Questionnaire

Participants again completed a post-experimental questionnaire. Precise wording was changed to keep it consistent with the foraging based task. Participants were asked

to estimate the proportion of trials in which a new resource had appeared in each colour patch (0% to 100%); the extent to which they relied upon intuition or strategy in making their predictions (1 to 5); and the extent to which they believed there to have been a predictable pattern in the reward sequence (1 to 5).

3.c. General Results

The proportions of selection of the high resource generation rate option for each 10-trial block are shown for each probability condition in Figure 10. These data were entered into a 2x42 mixed model ANOVA with Probability condition (Stated vs. Learned) as a between subjects factor and Block (1 to 42) as a within subjects factor. This found no significant main effect of Block ($F_{16.644, 965.375} = 1.091$, $MS_e = .076$, $p = .358$, $\eta_p^2 = .018$)[1] and no significant linear effect of Block ($F_{1, 58} < 1.0$), indicating that participants did not significantly increase their proportions of selections of the maximizing alternative as the experiment progressed. There was also no main effect of Probability condition ($F_{1, 58} = 3.093$, $MS_e = .170$, $p = .084$, $\eta_p^2 = .051$) suggesting that participants who had been explicitly informed of the probabilistic accumulation rates did not allocate significantly more responses to the patch with the higher accumulation rate than those who had to learn these from experience alone. There was neither a significant interaction ($F_{16.644, 965.375} < 1.0$) nor a linear interaction ($F_{1, 58} < 1.0$) between Block and Probability.

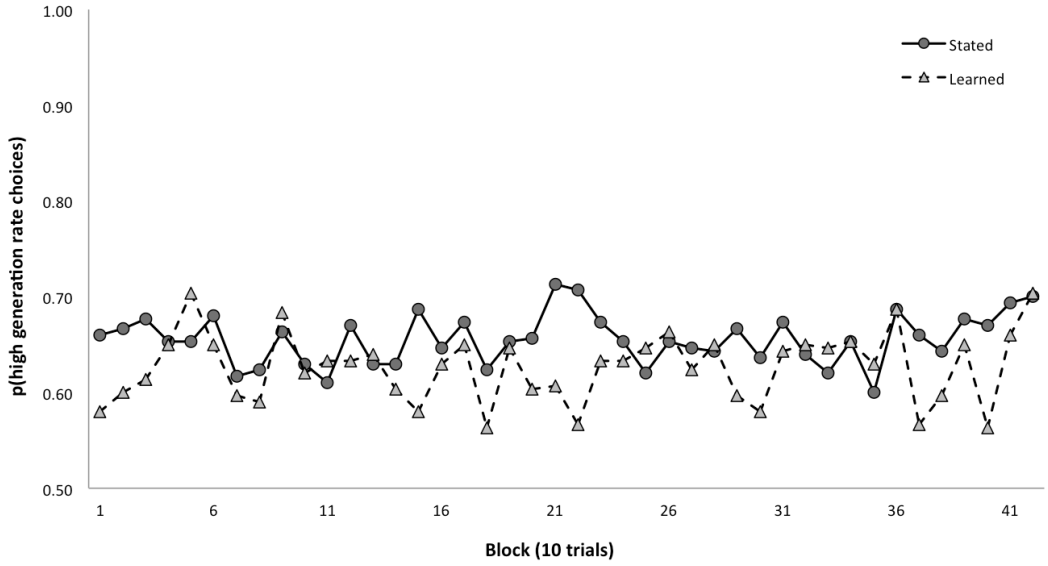


Figure 10: Proportions of selection of the high resource generation rate option for each block and probability condition.

Averaging steady state proportions of high accumulation rate choices over the final third of the task revealed values of .659 for the Stated condition and .630 for Learned condition with 95% CIs of .022 and .036 respectively (see Table 11). These indicate that there was no significant difference between the two groups, but average behaviour within each group significantly undermatched the actual ratio of reward probabilities.

Table 11: Steady state proportions of maximizing and high resource generation rate choices for the Standard task and SSFT respectively.

	Task			
	Standard		SSFT	
	m	se	m	se
Stated	.944	.021	.659	.011
Learned	.805	.032	.631	.019

3.c.i. Subjective estimates of reward probabilities

Subjective estimates of the relative resource generation rates of the two options were calculated through determining the proportion of total estimated generation rate accounted for by the high accumulation rate option (arH^{est} / arH^{est} + arL^{est}). Group average are shown in Table 12. Average reward probability estimates from the standard condition of Experiment 1 are shown for comparison purposes. Of participants who completed the SSFT, an independent samples t-test revealed no significant difference between relative estimated accumulation rates of those within the Learned and Stated probability conditions (t_{58} = .316, p = .753). These estimates also did not significantly differ from actual relative accumulation rates within any condition.

Table 12: Subjective estimates of outcome probabilities and relative accumulation rates by probability condition and task.

	Task			
	Standard		SSFT	
	m	se	m	se
Stated	.685	.010	.680	.012
Learned	.727	.020	.674	.017

3.c.ii. Heuristics & Apophenia

Self-reports of strategic choices and belief that the outcome sequences contained patterns are shown in Tables 13 and 14. On the SSFT, these also did not differ between Probability groups (t_{58} = .897, p = .373; t_{58} = -.923, p = .360).

Table 13: Subjective reports of strategy use by condition and task (scores range from 1 = pure intuition to 5 = pure strategy).

	Task			
	Standard		SSFT	
	m	se	m	se
Stated	3.800	0.176	3.467	0.205
Learned	3.200	0.188	3.233	0.157

Table 14: Subjective reports of belief that outcome sequences contained a pattern, by condition and task (scores range from 1=strongly disagree, to 5 = strongly agree).

	Task			
	Standard		SSFT	
	m	se	m	se
Stated	2.000	0.209	3.133	0.227
Learned	2.633	0.256	3.400	0.177

3.d. Comparisons with Standard Condition of Experiment 1

Following this analysis, data from the SSFT was directly compared with equivalent data from the standard condition of Experiment 1, in which participants had completed an analogous 2AFC probability matching task, but where reward probabilities were fixed and i.i.d. Selection proportions of the higher accumulation rate option on the SSFT are shown in Figure 10. Selection proportions of the higher fixed probability option in the Standard condition of the fixed probability task are shown in Figure 4A (see Chapter 2). These data were entered into 2x2x42 mixed model ANOVA with both Probability condition (Stated vs. Learned) and Task (Standard vs. SSFT) as between subjects factors and with Block (1 to 42) as a within subjects factor. This revealed both a significant main effect of Block ($F_{20.539, 2382.530} = 3.806$, $MS_e = .049$, $p < .001$, $\eta_p^2 = .032$)[1] and a significant linear effect of Block ($F_{1, 116} = 30.162$, $MS_e = .088$, $p < .001$, $\eta_p^2 = .206$),

indicating that participants overall learned to allocate a significantly higher proportion of choices to the higher probability options (of either reward or resource accumulation) as the experiment progressed. There were also significant main effects of both Probability condition ($F_{1, 116}$ = 35.657, MS_e = .354, p < .001, η_p^2 = .235) and Task ($F_{1, 116}$ = 131.215, MS_e = .354, p < .001, η_p^2 = .531). These indicate that, overall, participants who had received explicit information regarding the probability rates chose the high probability option more often than those who had to learn this information for themselves and that participants in the standard task group allocated significantly more choices to the maximizing option than participants in the SSFT group allocated to the high accumulation rate option.

This analysis also reveals theoretically informative interaction effects. First between Block and Task, there was both a significant interaction ($F_{20.539, 2382.530}$ = 3.364, MS_e = .049, p < .001, η_p^2 = .028) and a significant linear interaction ($F_{1, 116}$ = 20.956, MS_e = .088, p < .001, η_p^2 = .153), representing significantly greater learning on the standard task than the SSFT. Second, there was also a significant interaction between Probability condition and Task ($F_{1, 116}$ = 18.042, MS_e = .354, p < .001, η_p^2 = .135), with the effect of stating reward probabilities on participants' choices being significantly greater on the standard than the foraging task. When viewing these interaction effects in conjunction with the findings of the preceding analysis, which was conducted upon behaviour solely within the SSFT group, it is evident that the main effects of both Block and Probability condition were driven by differences within the Standard condition. Between Block and Probability condition there was neither a significant interaction ($F_{20.539, 2382.530}$ = 1.212, MS_e = .049, p = .231, η_p^2 = .010) nor a linear interaction ($F_{1, 116}$ = 2.801, MS_e = .088, p = .097,

η_p^2 = .024). There was neither a three-way interaction ($F_{20.539, 2382.530}$ < 1.0), nor a three-way linear interaction ($F_{1, 116}$ = 2.767, MS_e = .088, p = .099, η_p^2 = .023).

3.d.i. Subjective estimates of reward probabilities

Subjective estimates of outcome probabilities and relative resource accumulation rates are shown in Table 12. A 2x2 ANOVA was conducted with Probability (Stated vs. Learned) and Task (Standard vs. SSFT) as between-subjects factors. There was no significant effect of either Probability ($F_{1, 116}$ = 1.338, MS_e = .007, p = .250, η_p^2 = .011) or Task ($F_{1, 116}$ = 3.690, MS_e = .007, p = .057, η_p^2 = .031), nor a reliable interaction between the two ($F_{1, 116}$ = 2.539, MS_e = .007, p = .114, η_p^2 = .021).

3.d.ii. Heuristics

Self-reported use of strategy vs. intuition (see Table 13) was entered into a 2x2 ANOVA with Probability and Task as between-subject factors. A main effect of Probability ($F_{1, 116}$ = 5.189, MS_e = 1.004, p = .025, η_p^2 = .043) showed that participants in the Stated probability conditions reported greater strategy use than those in the Learned conditions. There was no significant effect of Task ($F_{1, 116}$ < 1.0), and no interaction between the two ($F_{1, 116}$ = 1.005, MS_e = 1.004, p = .318, η_p^2 = .009).

3.d.iii. Apophenia

Self-reported beliefs that the outcome sequences contained a pattern (see Table 14) were entered into another 2x2 ANOVA. This revealed both a significant main effect of Probability condition ($F_{1, 116}$ = 4.204, MS_e = 1.445, p = .043, η_p^2 = .035), indicating that participants in the Learned probability conditions reported significantly higher levels of

apophenia than those in the Stated conditions, and of Task ($F_{1, 116}$ = 18.736, MS_e = 1.445, p < .001, η_p^2 = .139) indicating that participants that had completed the SSFT reported significantly higher levels of apophenia than those who had completed the standard task. There was no interaction between the two factors ($F_{1, 116}$ < 1.0).

3.e. Mediation Analyses

Independent mediation analyses were conducted within both the Standard and SSFT conditions, using the PROCESS macro for SPSS (Hayes, 2012, 2013). The primary purpose of these analyses was to compare and contrast between the two tasks the relative effects of the Probability manipulations and, in particular, self-reported intuition vs. strategy use and apophenia on choice behaviour. For this analysis, proportions of maximizing choices were calculated for the final third of the tasks. This was done with the aim of accounting for learning effects within the Standard condition and also within the SSFT group for consistency, despite no evidence of learning on this task. Mediation model No. 6 was used, with a single independent variable and two mediator variables. Effects were calculated for each of 10,000 bootstrapped samples. Model results are presented in Figures 11 and 12 (Variable coding is as follows: Stated = 0 Learned = 1, Intuition 1 – 5 Strategy, Belief no Pattern 1 – 5 Belief was Pattern, Proportion of Maximizing Choices 0 – 1).

For participants who completed the standard task the direct effect of Probability condition on strategy usage was found to be significant (t_{58}= -2.332, p = .023), with this predictor accounting for 9% of the sample variance (R^2= .086). The direct effect of Probability condition on apophenia was found to be non-significant (t_{57}= 1.168, p = .248), while the direct effect of strategy use on apophenia was found to be significant (t_{57}= -2.566, p = .013), with these two predictors accounting for 16% of the sample variance (R^2= .157). The direct effects of Probability condition (t_{56}= -3.045, p = .004),

and strategy use (t_{56}= 3.373, p = .001), on proportion of maximizing choices were each found to be significant, whereas the direct the effect of apophenia was not (t_{56}= -.897, p = .374). These three predictors accounted for 39% of the sample variance (R^2= .394).

The unstandardized indirect effects of Probability condition on proportion of maximizing choices were as follows: Through intuition versus strategy use (-.600)(.056)= -.034, with bias corrected 95% confidence intervals ranging from -.074 to -.008, indicating statistical significance of this effect at α =.05. Through apophenia (.385)(-.012)= -.005, with bias corrected 95% confidence intervals ranging from -.034 to .005, indicating non-significance of this effect at α =.05. Through first strategy use then apophenia (-.600)(-.413)(-.012)= -.003, with bias corrected 95% confidence intervals ranging from -.019 to .003, indicating that this effect was also non-significant at α =.05.

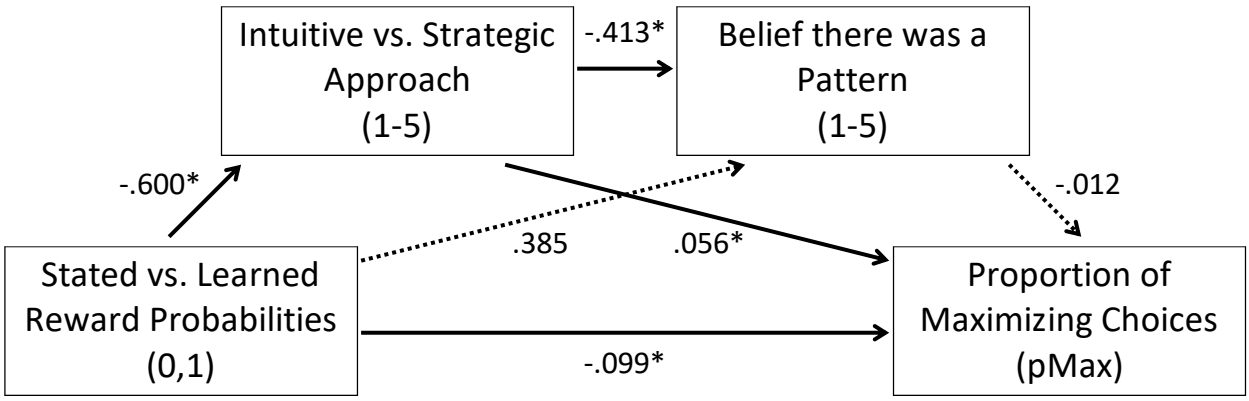


Figure 11: Mediation model for standard task condition. Unstandardized regression coefficients shown with asterisk if significant. Solid and dotted lines indicate significant and non-significant effects respectively. α=.05.

For participants who completed the SSFT, all effects were found to be non-significant. The direct effect of Probability condition on strategy usage (t_{58}= -.233, p = .373), accounted for 1% of the sample variance (R^2= .014). The direct effects of Probability condition (t_{57}= 1.130, p = .263), and strategy use (t_{57}= 1.696, p = .095), on

apophenia together accounted for 6% of the sample variance (R^2= .062). The direct effects of Probability condition (t_{56}= -1.302, p = .198), strategy use (t_{56}= .980, p = .331), and apophenia (t_{56}= 1.147, p = .256) on proportions of high accumulation rate choices together accounted for 8% of the sample variance (R^2= .076).

The unstandardized indirect effects of Probability condition on proportion of high accumulation rate choices were as follows: Through intuition versus strategy use (-.233)(.011)= -.003, with bias corrected 95% confidence intervals ranging from -.025 to .025. Through apophenia (.324)(.012)= .004, with bias corrected 95% confidence intervals ranging from -.002 to .023. Through first strategy use then apophenia (-.233)(.244)(.012)= -.001, with bias corrected 95% confidence intervals ranging from -.009 to .0003. All indirect effects were also non-significant at α =.05.

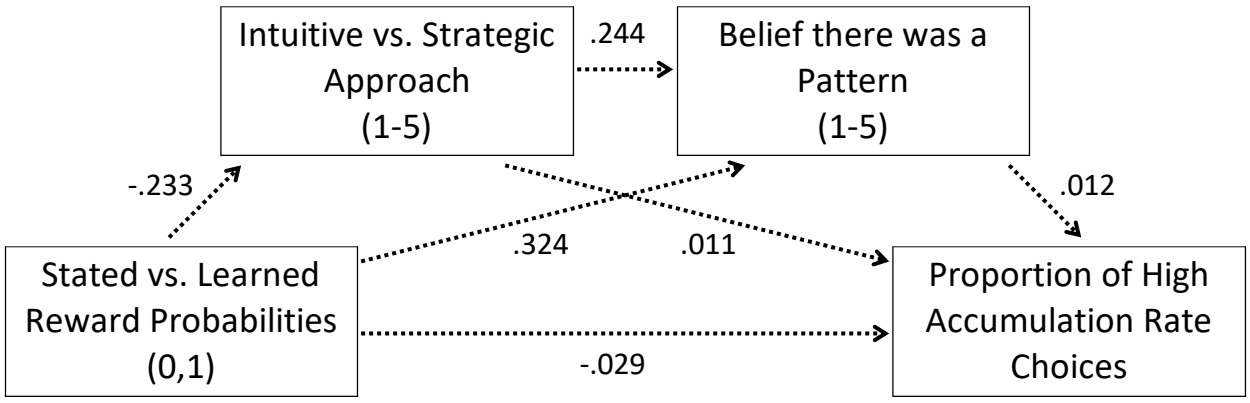


Figure 12: Mediation model for SSFT condition. Unstandardized regression coefficients shown with asterisk if significant. Solid and dotted lines indicate significant and non-significant effects respectively. α=.05.

3.f. Win-Stay-Lose-Shift

The proportion of trials that could be predicted by a win-stay lose-shift strategy was also assessed, for participants within both the SSFT and standard tasks (see Table 15). This proportion was also calculated relative to that which would be expected to arise from optimal choice behaviour across each task. This base value differed between the

two tasks. On the standard task a maximizing strategy consists of exclusive selection and reselection of the .7 probability option, with an average 70% win rate. The method of calculating base rate for the SSFT is demonstrated in Table 16. A trial and response sequence is determined for the optimal strategy and whether this represents a stay or shift is compared with the probability of having won or lost on the preceding trial. This can be used to calculate a base rate for the proportion of choices that would be expected to match WSLS predictions under an optimal strategy.

Table 15: Proportion of choices predicted by WSLS strategy, and deviation of this proportion from that expected by an optimal (non-WSLS) strategy on the task, by probability condition and task.

	Task					
	Standard (base .7)			SSFT (base .62)		
	m	m-dev	se	m	m-dev	se
Stated	.694	-.006	.011	.627	.007	.022
Learned	.661	-.039	.017	.685	.065	.023

Table 16: Demonstration of trial sequence and expected values on the SSFT with optimal choices (asterisk indicates selection). Note that from trial 3 onwards the sequence repeats every three trials.

EV-Left	EV-Right	Previous Prob	Stay/Switch	Sum
.21	.49*	N/A	N/A	
.42	.49*	.49 Win	Stay	
.63*	.49	.51 Lose	Switch	Average
.21	.98*	.37 Lose	Switch	WSLS:
.42	.49*	.98 Win	Stay	.62
.63*	.49	.51 Lose	Switch	Average
.21	.98*	.37 Lose	Switch	WSLS:
.42	.49*	.98 Win	Stay	.62

Average proportions of WSLS consistent choices, relative to the baseline for each task, were entered into a 2x2 ANOVA, with Probability condition (Stated vs. Learned) and Task (Standard vs. SSFT) as between-subjects factors. There was no effect of Probability condition ($F_{1,116} < 1.0$). However, there was a significant effect of Task ($F_{1,116} = 9.831$, $MS_e = .010$, $p = .002$, $\eta_p^2 = .078$), indicating that participants in the SSFT made a higher proportion of WSLS like choices, relative to the amount that would result from an optimal strategy, than those on the standard task. There was also a reliable interaction between the two ($F_{1,116} = 5.928$, $MS_e = .010$, $p = .016$, $\eta_p^2 = .049$), This reflects that, on the SSFT, those in the learned accumulation rate condition made more WSLS like choices than those in the stated condition, whereas on the standard task the opposite was true.

3.g. Local maximising

Local expected values were modelled on a trial-by-trial basis throughout the task. The expected value of each option on the SSFT, according to all information that was available to the participant, was calculated to be an option's resource generation rate multiplied by the number of trials since it was last seen to be exhausted (through a failed forage attempt), minus the number of times that it had been successfully foraged from since it was last observed as exhausted. This subtraction following a successful forage was not allowed to cause an expected value to drop below a baseline of zero, entailing that each option's expected value at the time of a choice could never fall below it's probability of resource generation.

The difference between the expected values of the option presented to the right and left-hand sides of the screen was calculated for each trial, as the variable evR(t)-evL(t). These differences were then split into six bins, ranging progressively from the

greatest left-hand relative expected values, to greatest right-hand relative expected values: <-1, -1 to -0.5, -0.5 to 0, 0 to 0.5, 0.5 to 1 and >1. The mean proportion of actual choices of the option presented to the right of the screen was then calculated for trials that fell into each of these six bins. These data, averaged across all participants, are shown in Figure 13.

In order to determine the statistical reliability of this relationship, linear regression slopes were calculated for the relationship between the local right-left expected value difference and actual proportions of right-hand choices, for each individual participant. This was done separately within each Probability condition. In the Stated probabilities condition, the mean unstandardized coefficient was .058, with 95% confidence intervals of .052, indicating a statistically significant positive relationship between local expected value and choices at $\alpha = .05$. By contrast, within the Learned probabilities condition the mean unstandardized coefficient was -.013, with 95% confidence intervals of .023, indicating no such reliable relationship. These confidence intervals also indicate a statistically significant difference in turn between the regression slopes of these two conditions, with choices in the Stated probabilities group relating significantly more highly with local expected values.

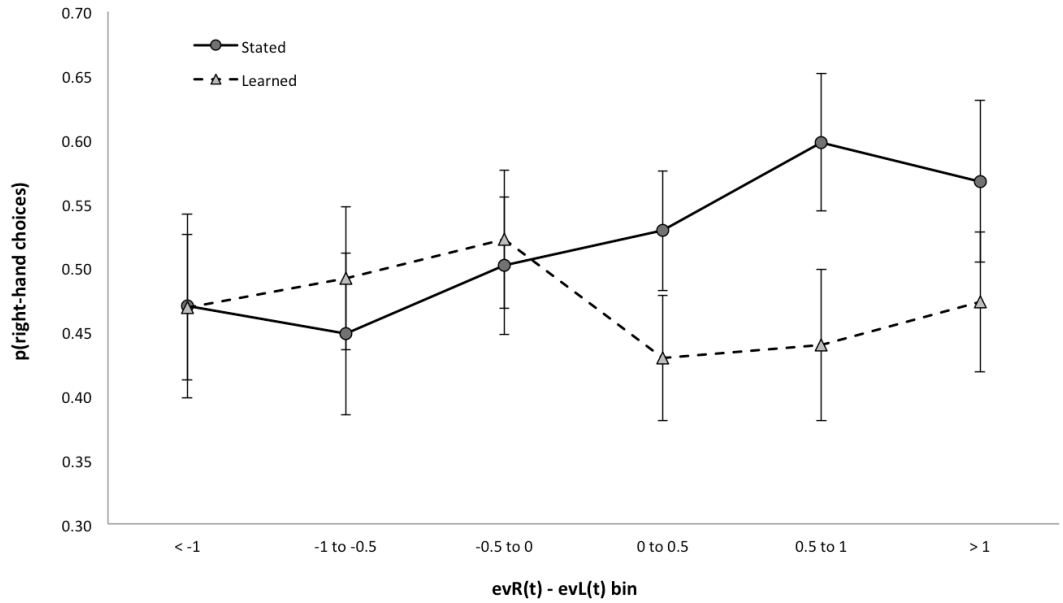


Figure 13: Illustrating the relationship between local right-left expected value differences and actual right hand choice proportions within each SSFT condition. Error bars show 95% confidence intervals.

3.h. Discussion

This study investigated human behaviour on a simple simulated foraging task, which was designed to incorporate the factor of resource accumulation over time into the standard 2AFC probability matching preparation. Behaviour on this task was then compared with that on the standard version, along with assessing differences in how a number of other factors relate to choice behaviour, and each other, across the two tasks. Viewing behaviour on the SSFT alone revealed no significant learning over the task, along with no effect on choice behaviour of explicitly stating the reward accumulation rates. For participants who completed the SSFT, stating reward probabilities also had no effect on either subjective estimates of accumulation rates, use of strategy over intuition, or levels of apophenia. A tendency towards group-level undermatching was also evident, this is congruent with the findings of a number of previous studies.

Direct comparison of these results with those of the standard task revealed a stark contrast between effects over the two tasks. Simple main effects were found indicating substantially lower proportions of high resource generation rate choices on the SSFT relative to proportions of maximizing choices on the standard fixed probabilities task. Moreover, significant interactions were found involving task type and both learning and the effects of stating reward probabilities. These findings are indicative of robust effects of each of these factors within the standard condition, paired with their absence on the SSFT. In fact, these effects were so strong in the standard task condition as to result in significant main effects when considering both groups together, despite not being present to any substantive degree within the SSFT condition. Furthermore, while there was no significant difference in overall levels of strategy usage between the two tasks, mediation analysis revealed that the effect of intuition vs. strategy use on choice behaviour that was present on the standard task also disappeared on the SSFT. In line with this, the indirect effect on choices of stating reward probabilities, via facilitation of strategy use, was also not present on the SSFT. This evidence indicates that while on the standard task participants behaved differently depending upon whether they adopted a predominantly intuitive or strategic approach, on the SSFT the degree to which a more deliberative strategy is applied no longer has any substantive influence on choice behaviour. This finding is consistent with Faulty Foraging Theory, which predicts that probability matching is both optimal on tasks in which reward accumulation is accounted for, such as the SSFT, and also prepotent when participants act intuitively. Therefore, according to this model, participants on the SSFT who act either intuitively or strategically will converge upon the same optimal matching response. By contrast, on the standard task, in which probabilities are fixed, taking an intuitive versus strategic approach will lead to a systematic divergence in choice behaviour.

Incidentally, levels of apophenia were found to be significantly higher on the SSFT than the standard task, as well as being higher in the learned than stated conditions. This makes sense, as on the SSFT rewards were contingent in some manner upon previous choices, and therefore not entirely i.i.d. Nevertheless, as in the standard task, levels of apophenia were found to have no notable effect on choice behaviour.

The extent to which participants' choices conformed to a win-stay lose-shift strategy was also investigated, including consideration of the expected proportion of WSLS like choices given an optimal non-WSLS based approach to each task. This revealed higher proportions, relative to expectations, of WSLS like choices on the SSFT than the standard task. This was coupled with an interaction effect, whereby there were fewer WSLS-like choices in the learned than the stated condition on the standard task, but the contrary was true on the SSFT. Furthermore, proportions of WSLS-like choices were significantly lower than expected on the learned standard condition, and significantly higher than expected on the learned SSFT condition, when compared with group-level baselines based upon optimal responding for each task. However, WSLS-like choices fell well within the expected range for the Stated probabilities conditions of both the standard and foraging tasks, despite a WSLS approach approximating optimal outcomes on the SSFT. Moreover, another analysis revealed that the relationship between local expected values and actual choices was only statistically significant within the Stated probabilities condition. For planned future research involving the SSFT, it is desired to discourage the excessive and potentially confounding use of a WSLS strategy, while engendering participants understanding of the true nature of the task. The latter may be indicated by a positive relationship between choices and optimal local expected values. Together these results strongly suggest that, for the purposes of future study, reward probabilities should be provided to participants prior to any SSFT-based task.

As described in the introduction, a study has been published since completion of the present experiment that has very similar theoretical motivations (Schulze et al., 2017). The findings of this study can provide a useful comparative measure. Schulze et al. found neither a significant group-level effect of reward retention on average choice proportions, nor any reliable group-level interaction effects. Nevertheless, further probing of distributions of individual participant level choices suggested that substantial underlying differences were in fact present between the two reward schedules. In the 'reward-hold' condition, the modal peak response was near optimal diversification of choices. By contrast, individual choice proportions demonstrated a relatively platykurtic distribution within the 'no-hold' condition. Here, while the modal response was maximizing, a tail of participants extended substantially beyond optimal levels of diversification, for a small subset of participants even reaching so far as chance levels of responding. In raw numbers, effective diversification was 6.94 times more prevalent in the 'reward-hold' condition, whereas maximizing was 9.63 times more prevalent in the 'no-hold' condition. The authors therefore argue that their results do provide clear evidence for qualitatively different response patterns between the two conditions, despite the lack of apparent group-level differences. Results of the present study are broadly congruent with these findings, although they did also uncover a clear behavioural difference between choices on the SSFT and the standard task at the group-level. There was one key difference between these two studies, which merits clear reiteration. This was that while in the 'reward-hold' condition of Schulze et al. unharvested rewards could persist over time, they were not allowed to accumulate as they were on the SSFT. This entails that while the two tasks are analogous in one core aspect, with some degree of diversification becoming optimal on each task, the specific level of optimal diversification differs between them. This difference may explain the apparent

differences in group-level effects that were observed between the two tasks, relative to choice behaviour on the standard probability matching preparation.

To summarise, the results of this experiment were consistently as hypothesised, in a manner that is compatible with both an existing study and Faulty Foraging Theory. A new simple simulated foraging task was developed. This was done through making superficial modification of the standard 2AFC probability matching task, along with modelling the persistence and accumulation of unharvested resources within the task's reward schedule. Contrasting findings between the two tasks with respect to learning, the effect of stating reward probabilities, and self-reported strategy use combine to highlight that the effect of deliberation over intuition in engendering maximizing behaviour is particular to the standard fixed probability task. When the potential for uncollected resources to accumulate over time is incorporated into the reward schedule there is no longer any effect of overall levels of strategy use on choice behaviour. Rather, there is a convergence of choice patterns across participants, with those acting intuitively now behaving statistically equivalently to those that take a more deliberative approach. More pragmatically, results also indicated that when applying the SSFT in future research it will be best to provide stated outcome probabilities to participants before the task, so as both to minimise potentially confounding WSLS use, and maximize understanding of the underlying reward schedules and levels of optimal local choice distributions.

Chapter 4: Detecting Linear Modulation of Internally Represented Expected Values using Electroencephalography

4.a. Introduction:

The first section of the thesis comprised three studies. In the first two studies, robust evidence was found indicating that probability matching is a predominantly heuristic driven approach on the standard fixed-probability task, which relates to the insufficient consideration or application of strategy. The third study assessed behaviour on a simple simulated foraging task. The observed pattern of discrepancies in both behaviour and pathways of effects between the two-tasks were consistent with the Faulty Foraging Theory (FFT) of probability matching, by which matching behaviour is hypothesised to be an overextension of an evolutionarily stable foraging strategy. Following these studies, the ultimate aim of this thesis is now to combine behavioural and electrophysiological methods to identify correlates of matching vs. maximizing choice behaviour, and also to further probe the possible underlying effect of FFT in engendering the suboptimal diversification of choices. The former will include the potential for electrophysiological verification of the heuristic account, primarily though the use of prefrontal activity as proxy for top-down cognitive control. The aim to achieve the latter is through uncovering evidence for the internal illusory representation of reward accumulation in the unselected option, in line with the reward schedule on the SSFT, but even within the fixed probability task, where this resource accrual is not reflected in the actual reward schedule of the task.

In order to achieve this, it is required to first to pragmatically determine whether it is feasible to measure the internal representation of reward values with sufficient accuracy to discern any such illusory accumulation in subjective expected value. It is further necessary to do so using the planned experimental paradigm, which combines extra-cranial electrophysiological recording with a simple 2AFC choice task. In the

current chapter a reanalysis of an existing data-set is described, which was designed to address this objective.

As stated in the author's declaration, the data analysed in this chapter was originally collected during completion of my MSc, in 2013. It remains so far unpublished. This experiment comprised a value-based decision making task with concurrent EEG recording. This was originally planned in a theoretically unrelated manner to the topics of either probability matching or foraging like behaviour patterns. The original aim of the study was the identification of candidate action-values signals in the human cortex, using EEG. This was coupled with the delineation of more specific temporal characteristics of these signals than was evident from preceding fMRI based research (Wunderlich et al., 2009). The experimental design was similar in a number of respects to the standard probability matching task, in that repeated choices were made between two probabilistically rewarded options on each trial, each with a different but fixed and i.i.d. likelihood of reward. However, although the task comprised a two-alternative forced choice on each trial, in this study there was a pool of three options, from which two would be drawn and presented on each trial. These had low (0.3), intermediate (0.5) and high (0.7) reward contingencies respectively. Data originally collected for this study were reanalysed for the purposes of the present thesis. Corrected data was also reintroduced from one additional participant, whose data had been excluded from the original analysis due to a systematic technical fault in trigger value recording.

4.b. Method

4.b.i. Participants

A total of 18 participants took part in the study. Two were excluded from the electrophysiological analysis, having performing at or below chance. Participants were recruited in an opportunity sample, and predominantly comprised a mixture of

undergraduate and postgraduate students studying at the University of Nottingham. Precise demographic data for all participants was unfortunately no longer available by the time of this reanalysis. However, of the participants included in the final analysis approximately equivalent numbers were male and female, while ages ranged from 19 to 31. Participants were paid a fixed inconvenience allowance to compensate for their time taking part in the study.

4.b.ii. Equipment

Stimuli were presented through a 24 inch iiyama Pro Lite X2472HD-B1 monitor, with a 8ms latency. This was hardwired to a counterpart Stone PC-1103 stimulus presentation computer, under control of the experimenter and from which the task was run.

EEG data were acquired using a DBPA-1 digital amplifier (Sensorium Inc., VT, USA) in an electrically shielded room. An Easycap manufactured cap and electrically conductive gel system was used. Each cap recorded from 117 Ag/AgCl scalp electrodes and was linked to the amplifier through a Hypertronics CAGE50541 NPJY19/14KMH2/TH connector; electrode layout is illustrated in Figure 14. All channels were referenced to the left mastoid and chin ground throughout recording. Input impedance was adjusted at <25-75kOhm at each electrode.

Responses were made through the left and right-most keys of a Cedrus RB-730 response box, hardwired through a parallel port for minimal transmission latency, to another Stone PC-1103 computer dedicated to data acquisition.

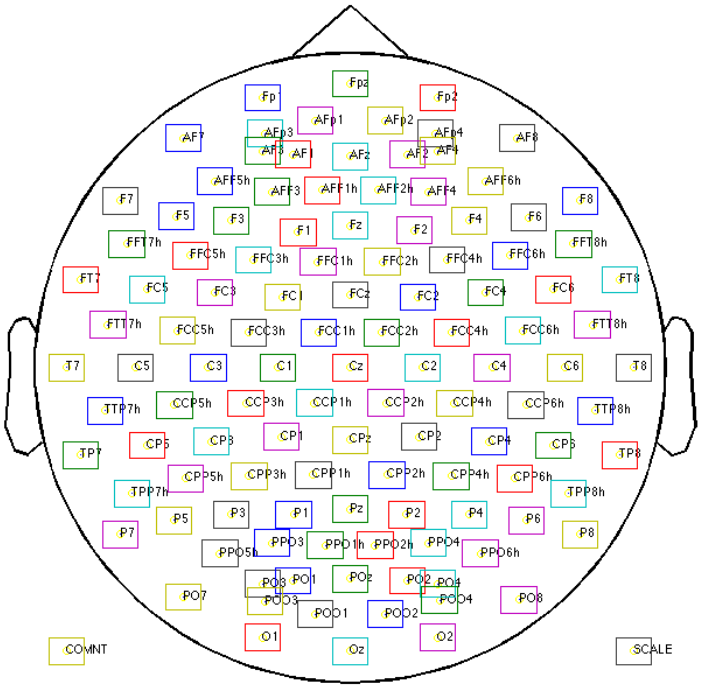


Figure 14: Showing layout of Easycap 117-electrode system.

4.b.iii. Design

In this experiment there was no between subjects experimental manipulation, all participants completed the same task, with the same instructions and information provided. Participants were not informed of the outcome probabilities of any of the three options. These were 0.3, 0.5 and 0.7. For both behavioural and EEG-based analyses, the primary independent variable was the right minus left expected value difference on each trial: probR(t)-probL(t).

4.b.iv. Stimuli

This task was framed as a repeated gamblers choice between football teams playing numerous independent league matches across an extended football season. Each trial represented a single round of matches, with each team playing a different unnamed

opponent. Three teams were selected all from the same league and with approximately similar recent historical performance at the time of the experiment; these were Aston Villa, Swansea City and Southampton. Each team was represented in the task by the club crest. An example stimulus pairing is shown in Figure 15. Stimuli were positioned equidistant from the centre and the left or right edge of the screen respectively. Which of the possible six stimulus pairings were presented on each trial was randomly determined for each trial, without replacement. The side to which each stimulus was presented also randomly varied between trials. Together, these factors preclude participants making an anticipatory decision regarding which side to choose prior to stimulus presentation. The experimental procedure was designed and coded within Psychopy (Peirce, 2007).

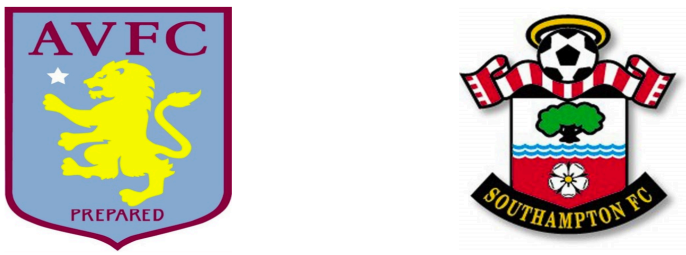


Figure 15: Showing an example pairing of choice stimuli.

4.b.v. Procedure

The task was a modified two-armed bandit problem, in which participants made repeated choices between two probabilistically rewarded options, over a total of 360 trials. This was similar in many respects to the standard probability matching task design. Each trial was preceded by a randomly variable inter-trial-interval of between two and three seconds, during which a fixation cross was presented for the majority. The two randomly selected choice stimuli were then presented to either side of the screen, until participants indicated their choice, using their left or right index finger. In this case

responses were made by pressing the left or rightmost keys on a Cedrus RB-730 response box. Following a brief interval, feedback was presented, indicating whether or not the prediction had been successful. Average trial length, including response time was approximately 5000ms, giving a total task duration of around 30 minutes.

For each participant, this study was preceded by preparation for EEG recording, in a number of stages. First, head size was measured and a suitable cap selected. A central reference point was located and marked on the scalp to ensure accurate positioning of the cap. The headcap was then fitted, ensuring a firm but comfortable fit. Following this, electrically conductive gel was applied, with electrodes attached to the face and scalp. Participants were then moved into the electrically shielded room in which they would complete the task and electrodes were connected to the amplifier. Finally, input impedance were assessed for each electrode, to identify any poor connections, and any channels that showed an impedance of greater than 75mV were repaired.

Participants were then presented with the task instructions. At this point participants were also given the additional verbal instruction to try and remain as still as possible throughout the task, so as to minimise movement related noise in the recording. They were informed that there would be two breaks, in which they could stretch or take a drink if necessary. Once participants indicated that they were ready, the door to the electrically shielded room was closed, and they proceeded to complete the task.

4.c. Behavioural Results

A simple behavioural analysis revealed that choices on this task were broadly in line with those found in Chapter 2 of this thesis. The most appropriate comparison can be made between choices made on trials in the present study between the high and low probability options (0.7 vs. 0.3) against choices of those within the learned probabilities condition of the standard probability matching preparation. This revealed a mean average

proportion of .82 high probability choices (SE = .04). Descriptive statistics are shown in Figure 16. This represents group-level overmatching on choices between the options with the greatest expected value difference, coupled with a significant degree of suboptimal diversification of choices.

In addition, a one-way repeated measures ANOVA, with the independent variable of the difference in lateralized reward likelihood: probR(t)-probL(t), and the dependent variable of proportions of right-hand choices, revealed a statistically significant effect of probR(t)-probL(t) on the proportion of right-hand choices $F_{1.170,\ 18.722} = 51.747$, $p < .001$, $\eta 2 = .764$. Post-hoc pairwise comparisons revealed statistically significant differences between each of the four value difference conditions (p<0.001, Bonferroni corrected).

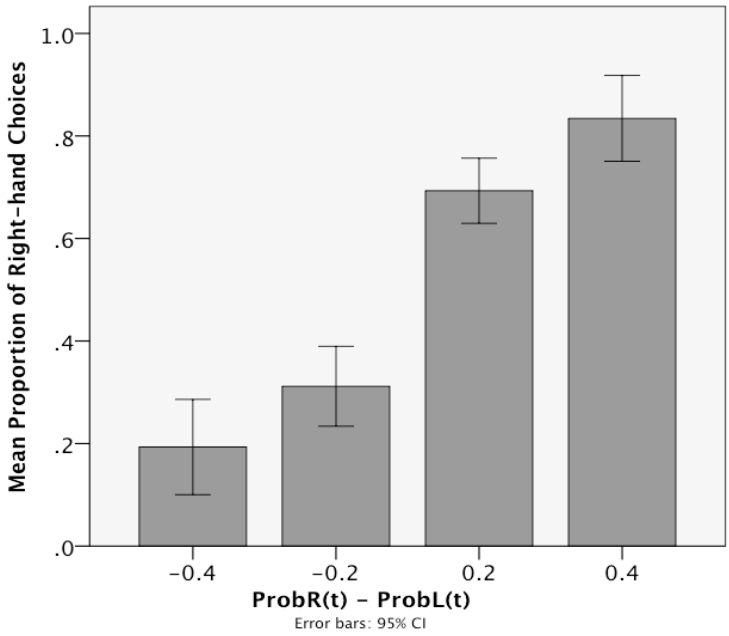


Figure 16: Showing average proportions of right-hand choices for each of the four conditions of right minus left expected value difference.

4.d. Electrophysiological Results

4.d.i. Software

EEG data were preprocessed and analysed in Matlab (Mathworks, MA, USA) using a combination of custom code and built-in functions of the 'Fieldtrip' software toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011).

4.d.ii. Pre-processing

First, single trials were identified and extracted through segmenting the continuous data set according to trigger values that were initialized by stimulus presentation. Time windows of interest were defined as running from 500ms prior to stimulus presentation until 1500ms following stimulus presentation. Next, data were put through a three-stage noise removal process. This called the field trip function 'ft_rejectvisual' to first visually identify and completely remove any trials with excessively high noise levels. The small number of trials excluded at this stage were generally outliers to an order of magnitude higher than the mean, indicative of an extreme source of task unrelated activity, such as a head movement or yawn. Lateral eye-movement and eye-blink related noise within the remaining trials was then targeted and removed. This was done through plotting the primary four components, identified through a principal components analysis (PCA), onto separate two-dimensional topographic renderings of the scalp. Visual identification was then made of the components with topographic distributions characteristic of both eye-movements and eye-blinks, which were then removed using the FieldTrip function 'ft_rejectcomponent'. Finally, the cleaned data set was run once again through the 'ft_rejectvisual' function, in this stage any trials that remained over the more stringent objective variance threshold of $1000\mu V^2$, following the removal of eye movement related noise, were also excluded from the final data-set. On the rare occasion that an individual channel was found to have provided consistently

noisy data throughout the task, this channel was also removed and its data were instead interpolated from neighbouring channels. Following this noise removal process, data were downsampled to 100Hz and referenced to the average of all scalp electrodes. Data were then also passed through a 10Hz low-pass filter, to remove higher frequency oscillatory dynamics that were not of interest for this ERP-based analysis.

4.d.iii. Analysis

A mass univariate regression analysis (Groppe, Urbach, & Kutas, 2011), was then conducted on the predictor variable of the trial specific expected value difference between the two options: probR(t) – probL(t) and the dependent variable of measured activation at each sensor and time point within the time-window of interest (500ms pre-stimulus to 1500ms post-stimulus presentation), averaged first across all trials and then all participants. This revealed two apparent temporal components of value-related activity, the first evident from around 250 to 550ms following stimulus presentation, and the second from 850-950ms. Results of this analysis are plotted below in Figures 17 and 18. Time-channel charts are shown to facilitate visual identification of time-points and electrodes at which the relationship between measured activity and either the lateral difference in global resource generation rates, or local expected values, was found to be statistically significant. Topographic plots better illustrate the position of value-related activity at specific timepoints of interest.

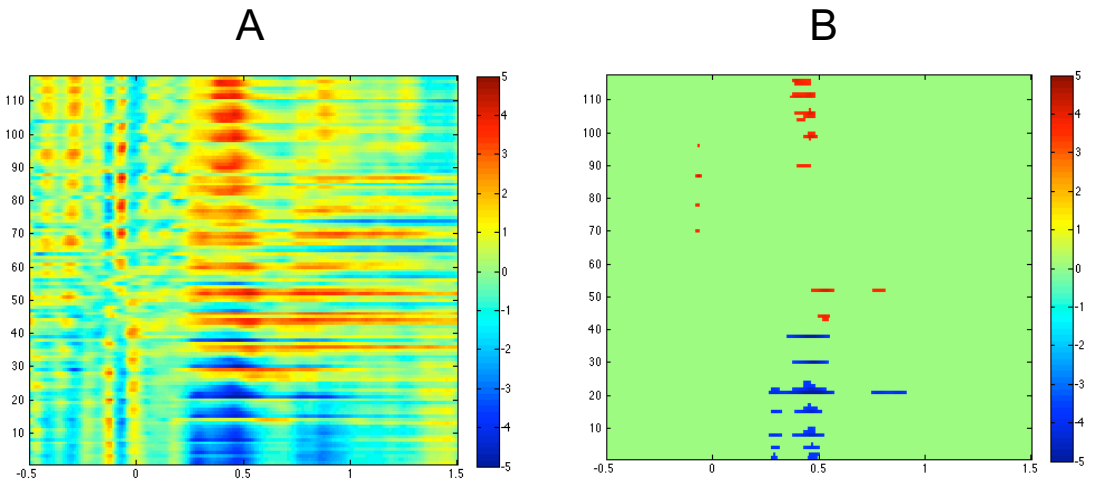


Figure 17: Time-channel charts showing activity correlating with probR(t)-probL(t) at each electrode (y- axis) and across all intra-trial time-intervals (x-axis). Panel A is unmasked. B masks all values at $t < 3.5$ (df = 15), $p > .0032$ (two-tailed, uncorrected). Colorbars represent t-values.

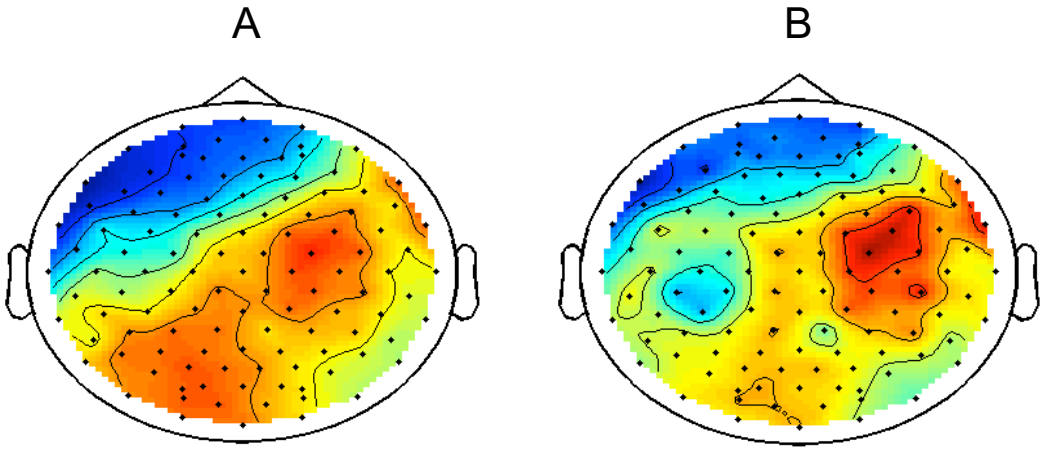


Figure 18: Topographic plots showing the correlation of surface activity with probR(t) – probL(t) across the scalp. Red indicates positive and blue negative. A is average activity from 250-550ms following stimulus presentation and B from 750-950ms.

4.d.iv. Assessment of Linearity of Value-Coding

If measured activity can be demonstrated to increase as a linear monotonic function of expected value then this could be used as a metric to determine a specific level of internal value representation. This is a crucial component of the design of the

studies planned for the remainder of this thesis. In order to address this question, the following analysis was conducted.

Three electrodes were selected, as exemplars for each of the three regions of interest that were identified during the early component, due to having demonstrated the most robust relationships with the expected value difference. These were FCC4h for the right sensorimotor region, PPO3 for the left parieto-occipital region and F7 for the left prefrontal region. Measured activity at each of these electrodes was then plotted against probR(t) – probL(t) for each of six possible stimulus combinations (see Figure 19 for demonstration). Values were then concatenated across each of the possible levels of expected value difference and plotted separately for each of the two temporal components (see Figure 20).

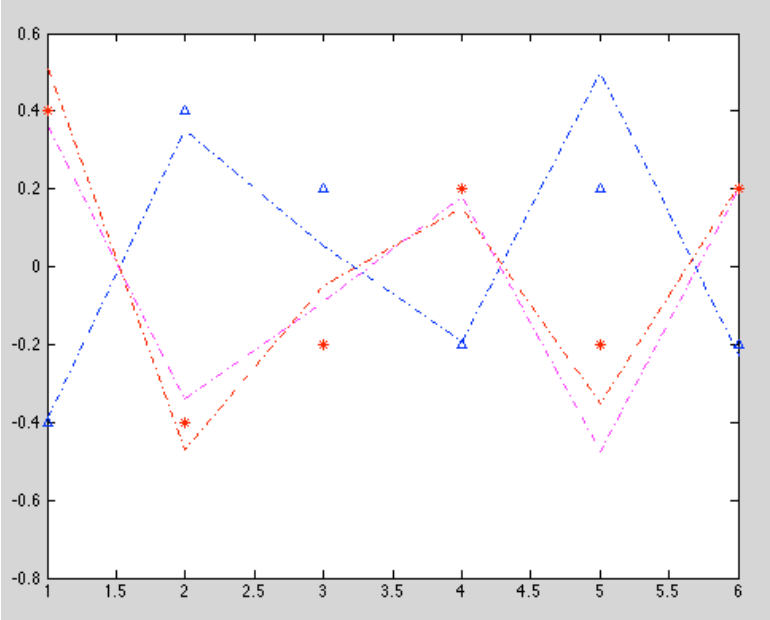


Figure 19: Peak early component activity (500-550ms) plotted against the lateral difference in expected reward probabilities for each of the six possible stimulus combinations. Red stars represent probR(t) – probL(t) and blue triangles represent probL(t) – probR(t). The red line represents activity at electrode FCC4h, the magenta at electrode PP03, the blue at electrode F7. Activity for each electrode was independently

scaled to differences in reward probabilities as follows: FCC4h activity *35000, +0.1, PP03 activity *35000, -1.6, F7 activity *10000, +0.6.

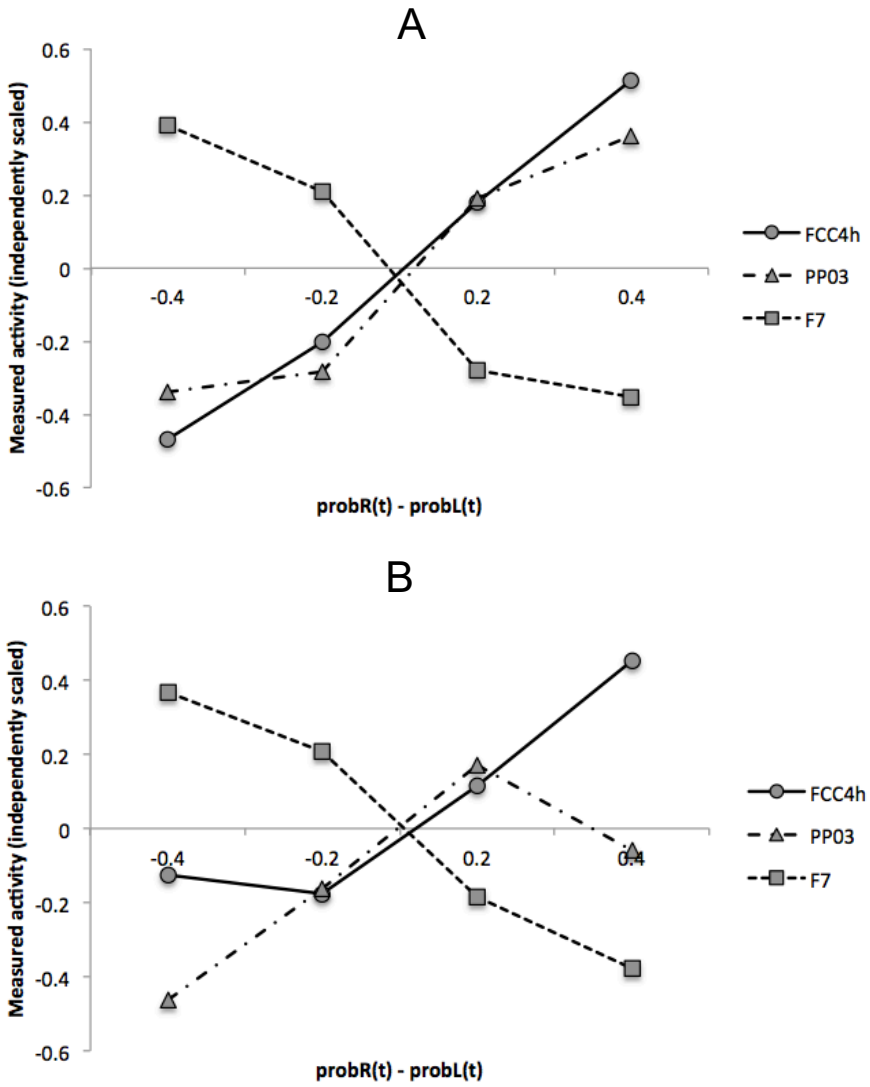


Figure 20: Relationships between scalp activity at each of the three exemplar electrodes and the four levels of probR(t) – probL(t). Panel A shows peak early component activity (500-550ms). Panel B shows peak late component activity (750-800ms).

4.e. Discussion

Though spatial resolution is poor with EEG, for reasons discussed in Chapter 1, the observed topographies of the measured value-related scalp activity were at least consistent with previous research. Activity that significantly positively correlated with probR(t) – probL(t) was present in clusters of both right sensorimotor and left parieto-occipital electrodes and a negative pole was evident over left prefrontal electrodes. This topography may potentially reflect two distinct underlying dipoles, one being primarily left lateralized and oriented on the anterior-posterior axis, while the other is oriented perpendicularly on the mediolateral axis and central on the anterior posterior-axis. Though this is unfeasible to ascertain for certain. The second of these putative spatial components is more clearly evident in the late temporal component (750-950ms), though this is primarily due to the lesser prevalence of the putative left lateralized A-P oriented dipole relative to throughout the earlier 250-550ms component. Overall, value-related activity was most clear within the early component, with a greater number of electrodes showing significant value-related activation at this time.

These results can be considered a baseline map of what may be expected for EEG activity recorded on a standard value-based decision task. Previous perceptual decision making studies have provided evidence for the coding of evidence for each option in the form of monotonically increasing firing rates (Huk & Shadlen, 2005; O'Connell et al., 2012; Shadlen & Newsome, 2001). The findings of the mass regression analysis, as shown in Figures 17 and 18 enable the determination of electrodes and time points that show an overall relationship with the difference in reward probabilities between the two options presented on each trial. They do not however allow for a certain conclusion to be drawn as to whether or not the activity identified encodes value in a strictly linear manner. To address this question, a further analysis was conducted, the results of which are plotted in Figures 19 and 20.

From these charts it is apparent that there are stimulus-locked time points at which a quite clear linear monotonic relationship exists between measured scalp activity at certain sensors and the right minus left expected value difference. In this case, during the early component, this relationship was particularly striking at electrode FCC4h (positioned over right anterior sensorimotor area), in which the relationship appears nearly perfectly linear. This finding is consistent with previous studies that have identified action-value representation within supplementary motor regions, both in monkeys (Sohn & Lee, 2007) and humans (Wunderlich et al., 2009). A good linear fit was also evident, over the same time period, at sensors located over each of the other regions of interest: PP03 for left posterior parietal and F7 for left prefrontal. However, at each of these sensors the relationship trended slightly towards resembling a cube root function (levelling off at each extreme). The measured relationship at F7 was also inverted relative to both FCC4h and PP03. During the late component, activity at electrode F7 was found to be roughly equivalent to that from 500-550ms. Intriguingly, electrodes FCC4h and PP03 show similar, but in a sense inverted patterns of activity to each other. At each sensor, measured activity increased as a linear function of expected value difference for three of the four possible values of probR(t) – probL(t), while at one extreme the fit was poor. For FCC4h this was the greatest negative value (-0.4), for which average activity was higher than at the less extreme negative value. By contrast, for PP03 this was the greatest positive value (0.4), at which average activity was lower than at the less extreme positive value. This mirror pattern is intriguing, but it may reflect a mere coincidence.

To conclude, this study detailed a targeted reanalysis of a pre-existing data set. The primary aim was to re-purpose data from a previous study as a form of pilot for the planned experimental paradigm of the remainder of the thesis. More specifically, the plan was to determine the feasibility of measuring the internal linear representation of expected value through extra-cranial electrophysiological recording, on a simple 2AFC

value-based decision making task. Results revealed stimulus-locked scalp activity changes that closely reflected the right minus left expected value difference. This relationship peaked at approximately 500-550ms post stimulus and was evident at sensors positioned over several regions, though most prominently over right sensorimotor, left parietal and left prefrontal areas. This finding provides a broad proof of principle for planned future studies. The identification of linear modulation of scalp activity establishes a marker of internal value coding, the characteristics of which can be further evaluated in future research, in relation to both matching vs. maximizing choice responses and the predictions of Faulty Foraging Theory.

Chapter 5: An Electrophysiological Assessment of the Internal Representation of both Global and Local Expected Values on the SSFT

5.a. Introduction

In Chapter 3 a simple simulated foraging task (SSFT) was established as a modification of the standard 2AFC probability matching paradigm. This task incorporates into the reward schedule a simple model of the natural process of resource accumulation and depletion over time, depending upon the level of exploitation of each option. The inclusion of this variable results in probability matching becoming the optimal strategy. Comparisons between the SSFT and the standard fixed probability task revealed that the introduction of resource accumulation to the reward schedule caused a shift in the choice behaviour of participants acting deliberatively away from maximizing, towards diversification of choices and probability matching, in line with those that acted more intuitively. This effect was evident in terms of both significantly higher group-level diversification, despite no change in average levels of strategy use, paired with the elimination of the effect of strategy use on choice behaviour. Overall, the pattern of behavioural results was found to be consistent with the Faulty Foraging model.

In addition to this, in Chapter 4 a reanalysis was presented of already existing EEG data. Here it was determined that incremental linear modulation of the internal representation of expected values can be discriminated using EEG recording on a simple 2AFC decision task. The present study was designed as an extension of this experimental paradigm. Here, participants will complete the SSFT together with concurrent electroencephalographic recording (EEG), with the aim being the identification and characterisation of internal expected value coding on this task. These representations can be examined on a global scale, in terms of the fixed resource generation rates of each option. Crucially however, the core characteristic of the SSFT is the possibility for unharvested resources to both persist and accrue over time. If internal expected values

are coded accurately, then on this task it should also be possible to observe the veridical internal representation of this resource accumulation over time within an unchosen option. This will allow for a parallel investigation of local expected values, which vary on a trial-by-trial basis according to a combination of the resource generation rates and the allocations of participants' preceding choices.

5.b. Method

5.b.i. Participants

A total of twenty participants took part in the study. Two participants were excluded due to technical issues encountered during EEG recording, leaving eighteen in the final analysis. Participants were recruited in an opportunity sample, and predominantly comprised a mixture of undergraduate and postgraduate students studying at the University of Nottingham. Of the participants included in the final analysis, 7 were male and 11 female. Ages ranged from 19 to 39 (M=23.33, SD=5.85).

Participants were paid an inconvenience allowance comprising both a fixed and variable component. All participants received a basic amount of £4, to cover the inconvenience of the EEG set-up. In addition, participants could increase this payment by an amount contingent upon the choices they made during the task. Optimal exploitation of resources would accumulate average additional winnings of £5.88, for a total payment of £9.88. The experiment was approved by the University of Nottingham ethics committee, according to BPS guidelines.

5.b.ii. Equipment

The experiment was conducted within a sound attenuated Faraday cage (dim. 2.5 x 2.0 x 2.3 m) to isolate recording equipment from extraneous sources of electrical noise. All electrical equipment within this cage ran on direct current, to eliminate 50Hz

electrical noise, this was either mains current filtered to 12 or 48vdc, or battery power. Stimuli were presented through a 23.6inch VIEWPixx monitor (3D full version VPX-VPX-2005C, 1920x1080, 120Hz, 48vdc, firmware 37) with an advertised latency of zero and pixel response time of 1ms. This monitor was connected to a Mac Pro stimulus presentation computer, using extended desktop, through an optical Thunderbolt 2 cable, via a Thunderbolt 2 dock inside the cage. The actual mean display lag of this system was empirically determined to be 8.76ms.

EEG data was acquired using ActiView software, through a battery powered BioSemi ActiveTwo system. The BioSemi headcaps used each had 64 active scalp electrodes; electrode layout is illustrated in Figure 21. These incorporate an initial stage of signal amplification within the electrodes themselves, which allows the system to better tolerate high impedance conditions without picking up significant interference at the electrode wires. Electrode voltage offset from the common mode, rather than input impedance, was used to identify electrodes with poor contact, with each channel adjusted to <20mv offset from the common mode voltage before the task began. The Biosemi system is reference-free during recording.

Responses were made through a Cedrus RB-730 response box. This was hardwired to the stimulus presentation computer, via USB through the Thunderbolt 2 dock, to enable PsychoPy to record the response. It was also hardwired to a Cedrus StimTracker outside the Faraday cage, through a TTL/Ethernet shielded cable, which was grounded to the Faraday cage. This allowed responses to also be sent directly to the BioSemi ActiveTwo. A diagram of the EEG recording set-up is shown in Figure 22.

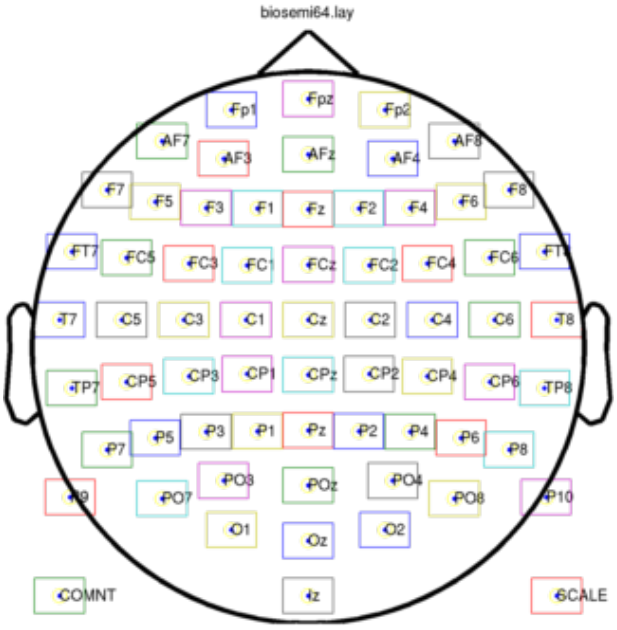


Figure 21: Showing layout of biosemi 64-electrode system.

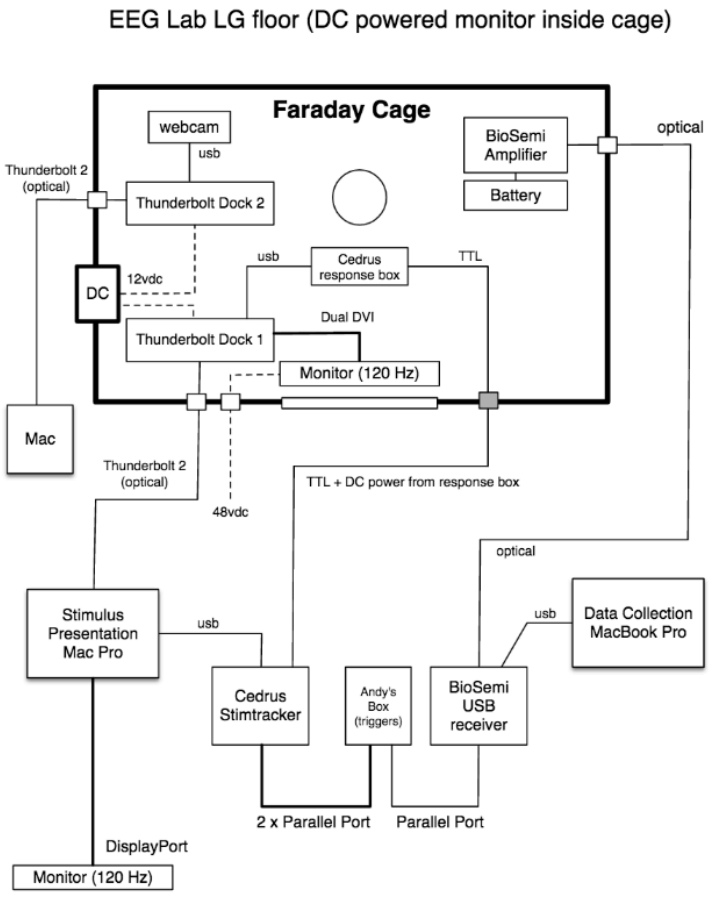


Figure 22: Circuit diagram illustrating layout of the EEG recording system. Created by Walter van Heuven.

5.b.iii. Design

In this experiment there was no between subjects experimental manipulation. For the behavioural analysis, the primary independent variable was block (42 level). For EEG analysis, the primary independent variable was the difference in expected value between the two options on any given trial of the task. This difference could be judged in two manners. First, in terms of the resource generation probabilities of each colour patch of land, which were fixed throughout the task. Second, in terms of local expected values, modelled at the trial-by-trial level. These were determined by a multiplicative combination of the generation rates and the number of trials that had elapsed since each option had last been explored, and therefore varied over the task depending upon previous choices. Dependent variables were proportions of choices of each option and measured voltage at each EEG sensor.

For this experiment all participants received explicitly stated information regarding the accumulation rates of each option before the task began. This was decided based upon findings of the previous experiment, in which only participants within the Stated condition were informed of the resource generation rates for each option. It was found that providing this information resulted in both a lower proportion of choices consistent with a win-stay lose-shift strategy, to the point that this did not significantly differ from an expected baseline level, as well as a stronger positive relationship between choices and local expected values. All participants were therefore informed that the two patches of land would generate resources stochastically, at the differing rates of 0.49 vs. 0.21, and to which of the two options each of these rates applied. They were also instructed that resources would remain in place, and accumulate over time, until they were harvested, while only one resource could be taken on each trial.

Which of the two colours generated resources at the higher rate was counterbalanced between participants, so as to account for any bias in favour of a particular colour. Participants were randomly assigned between these conditions.

5.b.iv. Stimuli

The task itself was the same 'simple simulated foraging task' (SSFT) as used in the previous chapter, all stimuli were both the same and presented in an identical manner. A plain white background was maintained throughout, with a black central fixation cross being presented between trials. Stimuli consisted of images of two identical patches of woodland, with a blue and a yellow signpost to distinguish between them (see Figure 8). These were positioned equidistant from the centre and the left or right edge of the screen respectively. The side to which each colour stimulus was presented randomly varied between trials, so as to preclude participants making an anticipatory decision regarding which side to choose prior to stimulus presentation. The experimental procedure was designed and coded within Psychopy (Peirce, 2007).

5.b.v. Procedure

The task procedure was identical to that of the behavioural only SSFT experiment described in Chapter 3. This was an iterated two-alternative forced choice, wherein repeated decisions were made regarding in which of the two patches of land to forage for each of 420 trials. The aim was to maximize total pay-offs over the entire task, through the distribution of choices in a manner that reduces the selection of a foraging area with completely exhausted resources to as few trials as possible.

In this study however the task was preceded by preparation for EEG recording, in a number of stages. First, head size was measured and a suitable cap selected. A central reference point was located and marked on the scalp to ensure accurate positioning of

the cap. The headcap was then fitted, ensuring a firm but comfortable fit. Following this, electrically conductive gel was applied, with electrodes attached to the face and scalp. Participants were then moved into the faraday cage and electrodes were connected to the amplifier. Finally, electrode offsets were assessed, to identify any poor connections, and any channels that showed a common mode offset of greater than 20mV were repaired.

Participants were then presented with the task instructions. At this point participants were also given the additional verbal instruction to try and remain as still as possible throughout the task, so as to minimise movement related noise in the recording. They were informed that there would be two breaks, in which they could stretch or take a drink if necessary. Once the participant was ready, the door to the faraday cage was closed, and they proceeded to complete the task. The only difference relative to the behavioural only version of the task was that instead of using the 'q' and 'p' keys on a standard keyboard, participants made their choice by pressing the left or rightmost keys on a Cedrus RB-730 response box. Responses were still made using the left or right index finger.

Once the task was finished each participant completed the same questionnaire as was used in the version of the task described in Chapter 3. Participants were removed from the faraday cage before all electrodes and the electrode cap were removed, prior to completing the post-experimental questionnaire. Participants were at this point also given the opportunity to wash and dry their hair if they wished to do so.

5.b.vi. Questionnaire

The post-experimental questionnaire was identical to that used for the preceding behavioural only version of the task, described Chapter 3. Key questions probed subjective estimates of the proportion of trials in which a new resource had appeared in each colour patch (0% to 100%); the extent to which participants had relied upon

intuition or strategy in making their predictions (1 to 5); and the extent to which they believed there to have been a predictable pattern in the reward sequence (1 to 5).

5.c. Behavioural Results

The proportions of selection of the high resource generation rate option, for each of 42 10-trial blocks, are shown in Figure 23. Linear regression slopes were calculated for each participant. This revealed an unstandardized coefficient of 0.0005 with 95% confidence intervals of 0.0011, indicating no statistically significant learning effect over the task at an α criterion of .05. The average steady state proportion of high generation rate choices, over the final third of the task, was .669, with 95% confidence intervals of .023. This indicates that, at the group level, there was slight but statistically significant undermatching (at $\alpha = .05$) from the actual ratio of resource generation rates between the two options (.7). This proportion of steady state high generation rate choices did not significantly differ from that recorded within the Stated condition of the behavioural only version of the task (M = .659).

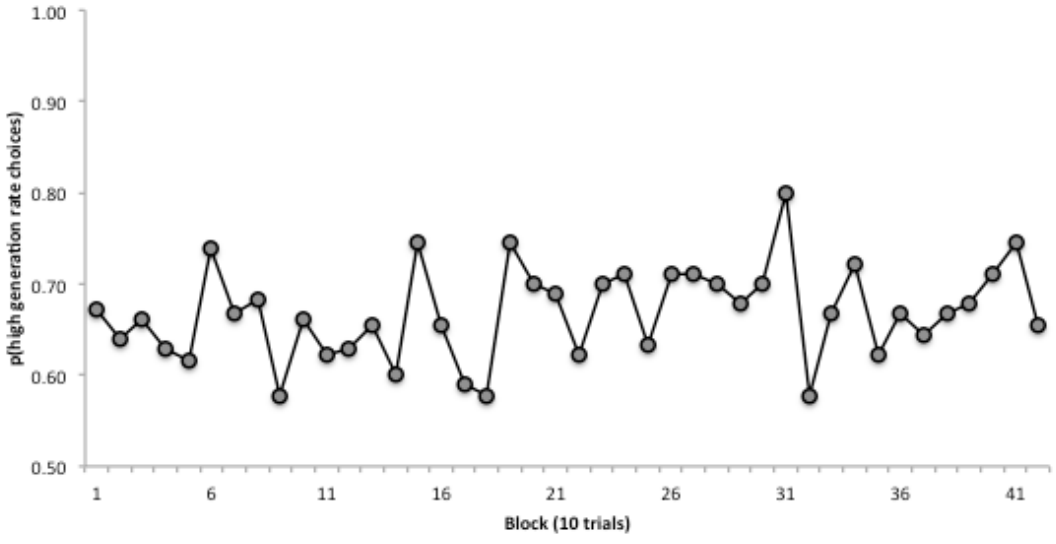


Figure 23: Proportions of selection of the high resource generation rate option.

5.c.i. Subjective estimates of reward probabilities

Subjective estimates of the relative resource generation rates of the two options were again calculated through determining the proportion of total estimated generation rate accounted for by the high accumulation rate option (arH^{est} / arH^{est} + arL^{est}). The average subjective estimate of relative accumulation rates was .690, with 95% confidence intervals of .019. This indicates no significant difference between the estimated and actual ratios of resource generation rates between the two options. The average estimate of those in the Stated condition of the previous study (.680) also falls well within these 95% confidence intervals.

5.c.ii. Heuristics & Apophenia

Average self-reported use of intuition vs. strategy use over the task and belief that the outcome sequences contained a pattern are shown below in Table 17. This table includes equivalent values from the Stated condition of the previous behavioural only version of the experiment. There were no significant differences between these two instances.

Table 17: Subjective reports of strategy use and belief that outcome sequences contained a pattern (scores range from 1 = pure intuition to 5 = pure strategy, 1 = strongly disagree, to 5 = strongly agree).

	Task and condition			
	EEG task (Stated only)		Non-EEG task (Stated condition)	
	m	se	m	se
Strategy use (1-5)	3.417	0.218	3.467	0.205
Pattern belief (1-5)	2.750	0.316	3.133	0.227

5.c.iii. Win-Stay Lose-Shift

The proportion of trials consistent with a purely win-stay lose-shift strategy was determined for each participant in the same manner as in the behavioural only version of this experiment. This proportion was also compared against the expected proportion of WSLS consistent choices that would arise from an optimal response pattern over the task. This base rate is .62, as calculated in the Chapter 3. The group mean proportion of WSLS consistent choices was .645, with a standard error of .023 and 95% confidence intervals of .044. This indicates that the proportion of WSLS consistent choices did not significantly differ from either the expected base rate, or the proportion observed in the Stated condition of the behavioural only version of the task (M = .627).

5.c.iv. Local maximising

Local expected values were again modelled on a trial-by-trial basis throughout the task. As described in Chapter 3, the expected value of each option, according to all information that was available to the participant, was an option's resource generation rate multiplied by the number of trials since it was last seen to be exhausted (through a failed forage attempt), minus the number of times that it had been successfully foraged from since it was last observed as exhausted. This subtraction following a successful forage was not allowed to cause an expected value to drop below a baseline of zero, meaning that each option's expected value at the time of a choice could never fall below it's probability of resource generation.

The difference between the expected values of the option presented to the left and right of the screen was calculated for each trial, as the variable evR(t)-evL(t). These differences were then split into 6 bins: <-1, -1 to -0.5, -0.5 to 0, 0 to 0.5, 0.5 to 1 and >1. The mean proportion of actual choices of the option presented to the right of the screen was then calculated for trials that fell into each of these six bins. These data, averaged

across all participants, are shown in Figure 24. To determine the reliability of this relationship, linear regression slopes were calculated between the local right minus left expected value difference and actual proportions of right-hand choices, for each individual participant. This revealed a mean unstandardized coefficient of .065, with 95% confidence intervals of .058. The linear relationship between the right minus left model predicted local expected value difference and actual proportions of right-hand choices was therefore statistically significant at $\alpha = .05$. This value also did not significantly differ from that observed in the Stated probabilities condition of the behavioural only version of the task (M = .058).

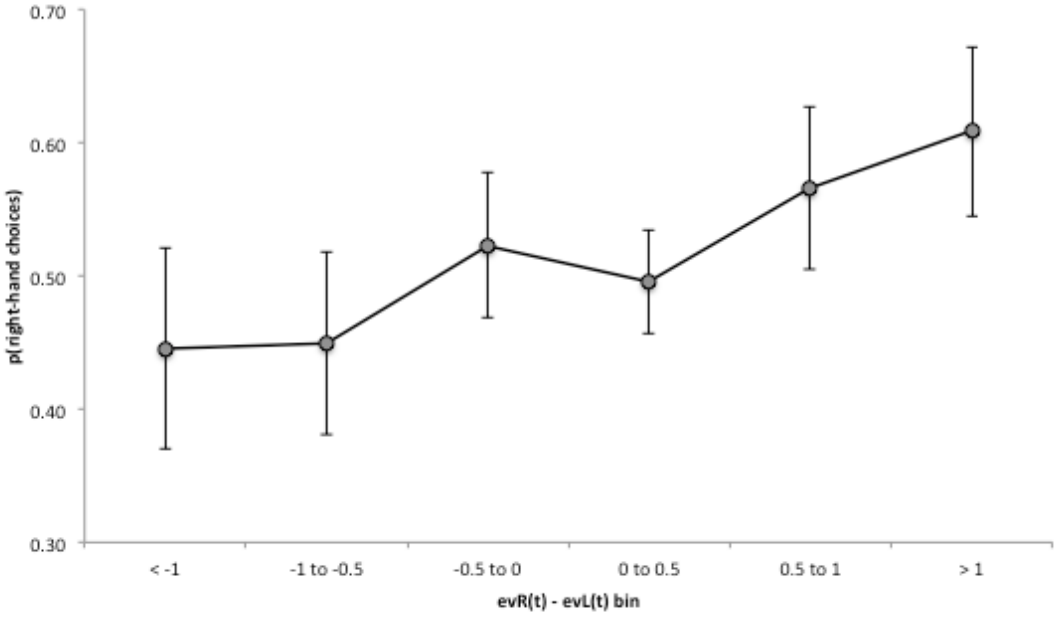


Figure 24: Illustrating the relationship between local right-left expected value differences and actual right hand choice proportions. Error bars show 95% confidence intervals.

5.c.v. Behavioural Results Summary

Behavioural results closely resembled those recorded in the preceding behavioural only version of the task, described in Chapter 3. No significant differences were found here regarding proportions of steady state high generation rate choices, estimates of reward probabilities, self-reported strategy use or apophenia, proportions of

WSLS consistent choices, or the relationship between trial-by-trial model predictions and actual choices. Proportions of WSLS consistent choices were again found not to differ significantly from the expected baseline, suggesting that this was not a commonly employed approach to the task. Modelling of trial-by-trial expected values again showed a significant positive relationship with actual behaviour, suggesting that local representations of value may well be internally represented as a basis for choices. Moreover, even where behaviour diverges from these model predictions it is possible that participants may realise and internally model the accumulation of unharvested rewards without expressing this behaviourally, through their immediate selection. In other words participants may sometimes 'save up' resources in one area for later exploitation, though this would be a risky strategy towards the end of the task.

5.d. Electroencephalographic Results

5.d.i. Software

EEG data were preprocessed and analysed in Matlab (Mathworks, MA, USA) using a combination of custom code and built-in functions of the 'Fieldtrip' software toolbox (Oostenveld et al., 2011).

5.d.ii. Pre-processing

Data were preprocessed in the same manner as in the reanalysis described in Chapter 4. First, trials were segmented from the continuous data-set using trigger values that were initialized by stimulus presentation. The time-window of interest for each trial was defined as running from 500ms prior to stimulus onset until 1500ms post-stimulus onset. Next, a three-stage noise-removal process was applied. Here, the reject-visual function was called to visually identify and reject trials with excessively high levels of noise, resulting from task unrelated motor activity. A principal components analysis

(PCA) was then used for the purposes of identification and rejection of components relating to both vertical and lateral eye-movements. The reject-visual function was then called again and any remaining trials with variance greater than the more stringent objective threshold of $1000\mu V^2$, following the removal of eye movement related noise, were then also excluded from the final data-set. Data were then downsampled to 100Hz and re-referenced to the average of all electrodes. As the current analysis focused on identifying event related potentials (ERPs) rather than oscillatory activity, a low pass filter was then applied at 10Hz.

5.d.iii. Analysis

This analysis includes an examination of activity relating to two sets of model predictions. First, a simple model that always predicts the option with the higher resource generation rate; these are fixed throughout the task and will therefore be referred to as the 'global' expected values of each option. Second, the trial-by-trial expected values of each option, which derive from all task-related information available to participants on any given trial. These fluctuate as the task progresses, depending upon the recent levels of exploitation of each area, and are therefore referred to as 'local' expected values. The global values model makes the simple prediction of either the right or the left option, depending only upon which side the high generation rate option is presented to on a given trial. The size of the difference between the right and left resource generation rates, grR(t) – grL(t), is fixed, with only the direction of prediction varying between trials. By contrast, model predictions of the local expected value model can vary in both direction and magnitude, in accordance with the local expected value difference, evR(t) – evL(t).

As in the behavioural analysis, the predictions of the local expected values model were split into six bins, in which evR(t) – evL(t) was <-1, -1 to -0.5, -0.5 to 0, 0 to 0.5, 0.5

to 1 and >1. Each trial could also fall into one of two predictions of the global values model. In practice, this resulted in each trial being grouped into one of twelve bins, according to the predictions of both the global and local expected value models. Average activity for each of the 64 scalp electrodes and 200 timepoints was then calculated across all trials that fell within each of these twelve bins. An independent mass regression analyses (Groppe et al., 2011) was then conducted for each model, using the built-in matlab function 'regstats'. The independent variable was the predicted right minus left expected value difference of the model and the dependent variable was average activity at each electrode and timepoint. In this manner, regression coefficients (beta values) were obtained for each of 12800 points of interest (64x200 electrodes x timepoints). Following this, one sample t-tests against zero were conducted, comparing the beta values of all 18 participants (df = 17) at each of these points.

Results of this analysis, for both global reward generation probabilities and local expected values, were plotted onto a series of time-channel charts and topographic plots. These allows for visual identification of electrodes and time-points at which the relationship between measured activity and either the lateral difference in global resource generation rates, or local expected values, was found to be statistically significant. These are shown below in Figures 25-28.

5.d.iv. Global Generation Probabilities

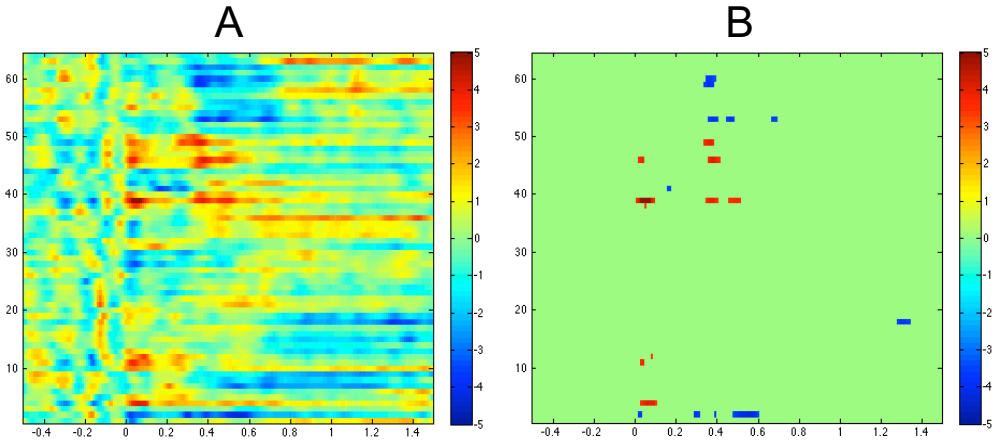


Figure 25: Time-channel charts showing activity at each electrode and timepoint correlating with grR(t)-grL(t) at each electrode (y- axis) and across all intra-trial time-intervals (x-axis). Panel A is unmasked. Panel B masks all values at $t < 3.4$, $p > .0034$ (two-tailed uncorrected). Colorbars represent *t*-values.

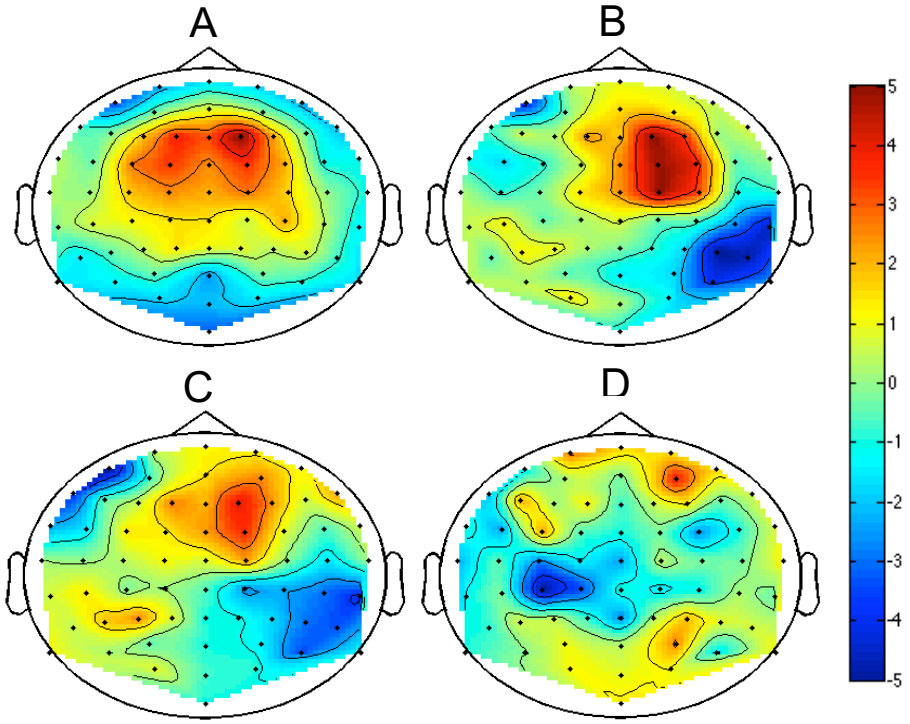


Figure 26: Topographic plots showing location of sensors at which activity was highly related with grR(t)-grL(t), at different time-windows of interest. Panel A shows 20-80ms post-stimulus, B 350-400ms, C 450-600ms and D 1280-1340ms. Colorbar represents *t*-values.

5.d.v. Local Expected Values

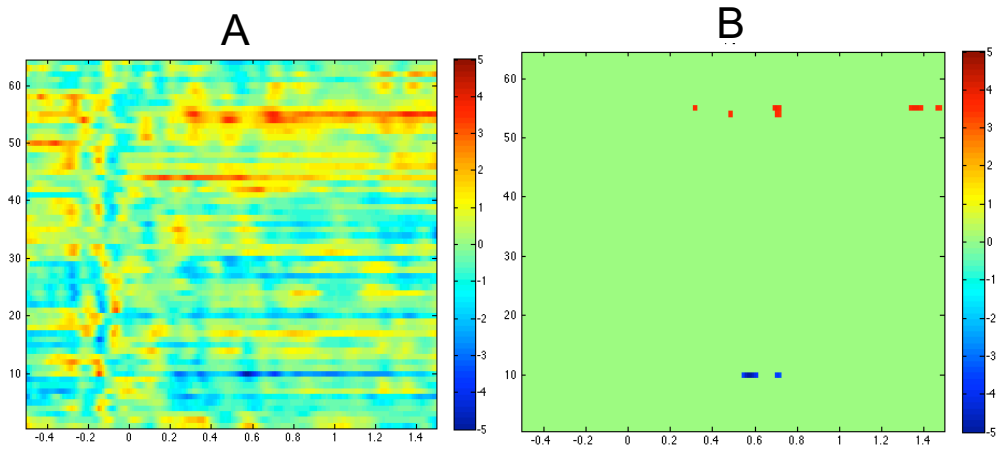


Figure 27: Time-channel charts showing regression coefficients of activity at each electrode and timepoint against the difference in calculated local expected values of the two options, evR(t)-evL(t). Panel A is unmasked, B masks all values at $t < 3.4$, $p > .0034$ (uncorrected). Colorbars represent *t*-values.

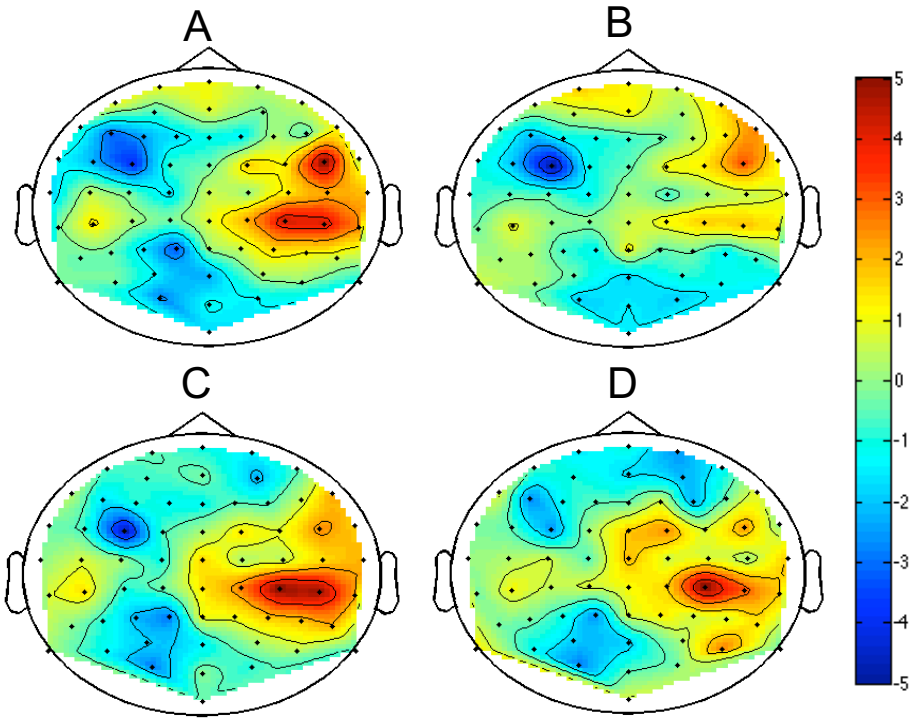


Figure 28: Topographic plots showing location of sensors at which activity was highly related with evR(t)-evL(t), at different time-windows of interest. Panel A shows 300-500ms post-stimulus, B 540-610ms, C 690-730ms and D 1330-1490ms. Colorbar represents *t*-values.

5.d.vi. Behavioural and Electrophysiological Model Fit Comparison

Following this, another analysis was conducted to compare the relationship between the fit of the local expected value model to both behaviour and electrophysiology at key sensors and timepoints of interest. The aim of this was to provide a further assessment of the validity of this apparent internal coding of local expected values. The rationale was that, if differences in observed scalp activity do genuinely reflect internal coding of local expected values, then these should also relate to differences in behavioural choices in accordance with the local expected value model.

For this analysis, specific electrodes and timepoints of interest were identified at which activity seemed most consistent with bilateral action-value coding of local expected values. These were the two electrodes FC3 and FC6, positioned over left and right anterior sensorimotor regions respectively, both at 290 and 350ms post-stimulus presentation. Two simple linear regression analyses were conducted, one for each of these timepoints. Each used the electrophysiological relationship with the predictions of the local expected value model, for each participant, as the predictor variable, and the corresponding behavioural relationship with the same model predictions as the outcome variable. Both of these relationships were represented in the form of *b*-values. As bilateral coding for the right minus left expected value difference was predicted, a priori, to be represented in an inverted pattern between the two hemispheres, the difference between relationships at the left and right hemisphere electrodes was calculated, by subtracting the *b*-values observed at FC6 from those at FC3. This creates a variable that is more representative of the overall fit of the bilateral predictions of the model than the relationship at either electrode alone. A significant effect was found at both the 290ms ($F_{1,16} = 7.125$, $MS_e = .012$, $p = .017$, $R^2 = .308$) and 350ms ($F_{1,16} = 6.869$, $MS_e = .012$, $p = .019$, $R^2 = .300$) timepoints.

5.d.vii. Descriptive ERP Plots

A series of ERP charts are shown below. These illustrate lateralized differences in scalp activity relating to the local expected value model, at the same two sensors assessed above (FC3 and FC6).

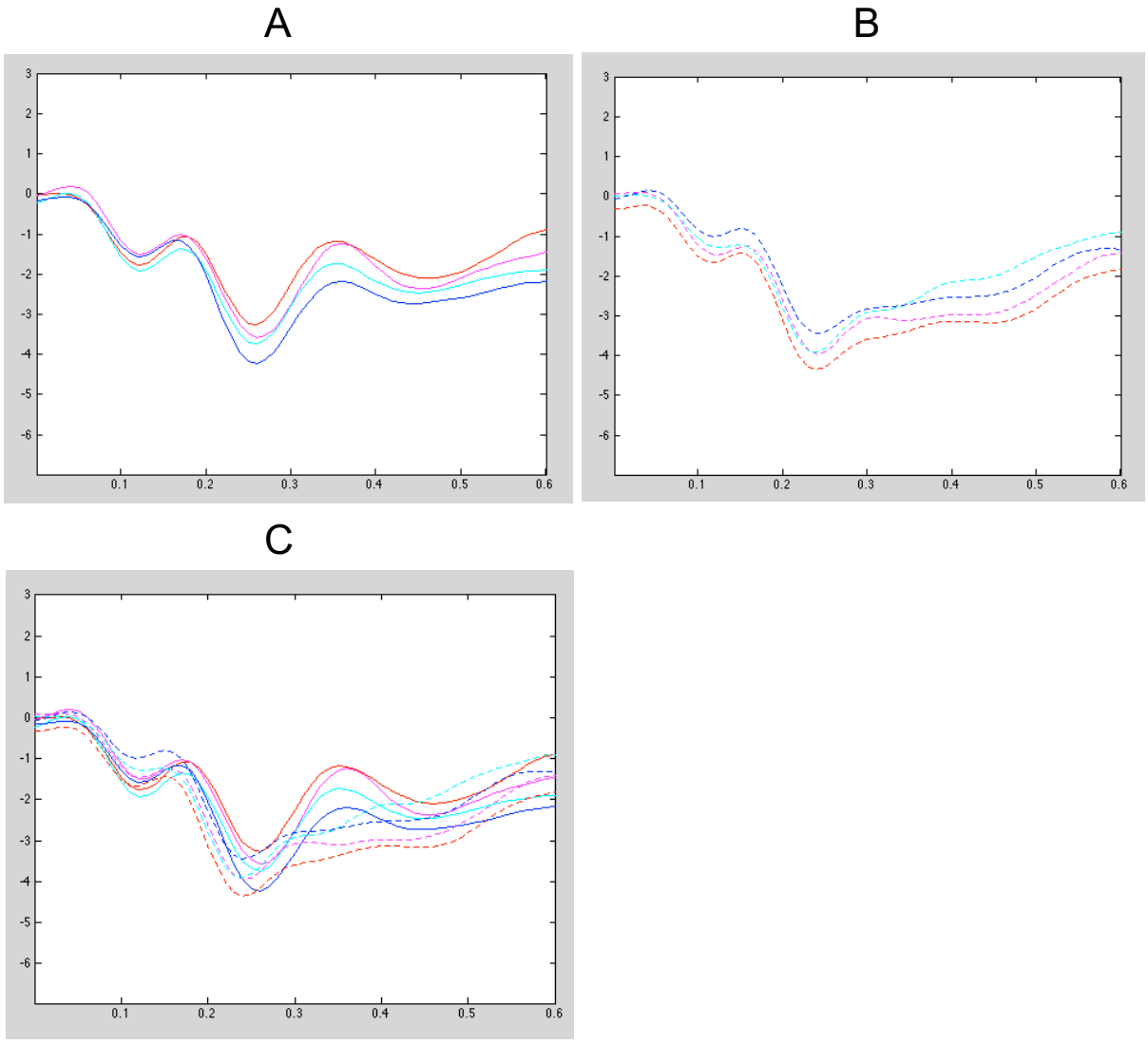


Figure 29: Charts plotting measured average activity (y-axis) over time (x-axis), within the four more extreme levels of evR(t)-evL(t). Time-window runs from 0-600ms following stimulus presentation. Panel A (solid lines) represents activity at sensor FC3. B (dashed lines) represents activity at sensor FC6. C shows overlaid activity traces for both sensors. The four bins of evR(t)-evL(t) are represented by line colour. Red = <-1, Magenta = -1 to -0.5, Cyan = 0.5 to 1, Blue = >1.

5.e. Discussion

In this study participants completed the same simple simulated foraging task as was developed in the previous chapter, but in this instance while under concurrent EEG recording. Analyses were conducted upon both behavioural and electrophysiological data. All participants were explicitly informed of the resource generation rates of each option before beginning the task.

As expected, behavioural results were in close accordance with those of participants that had completed the equivalent Stated condition of the behavioural only SSFT experiment, as documented in the Chapter 3. Again, there was no evidence of significant learning over the task, while a slight but statistically significant level of undermatching was evident at the group-level. Estimates of the relative resource generation rates of the two options did not significantly differ from either the actual ratio, or from those given by participants in the previous experiment. Self-reported levels of both strategy use and apophenia also did not significantly differ from those reported in the previous experiment. The observed proportion of choices congruent with a WSLS-based strategy was also consistent with the proportions seen in the stated probabilities group of the previous study, and did not significantly differ from that of the expected base-rate for the task. As in the stated condition of behavioural only version, a statistically significant positive relationship was found between the calculated right minus left difference in local expected value and the actual likelihood of choosing the right hand option on each trial. The coefficient of this relationship also did not significantly differ from the previous study.

Electrophysiological analysis allowed for the identification of internal activity that was correlated with the difference between both global and local reward value; in the form of the overall resource generation rates, which were fixed across the entire task, and the locally modelled expected values, which were updated on a trial-by-trial basis. This

ERP analysis highlighted both specific electrodes and stimulus-locked timepoints of particular interest.

Activity relating to the fixed resource generation rates was relatively robust. Two seemingly distinct temporal components were evident, each expressed at multiple adjacent electrodes. The first of these emerges surprisingly soon after stimulus presentation, running from around 20-80ms post-stimulus. The topography is characterised by a peak in activity fronto-central sensors, although this activity seems bilateral it is greatest at the right hemisphere electrode F2. The later component emerges a few hundred milliseconds later, from 350-400ms post-stimulus. This involves two more clearly dissociable clusters of electrodes, in a seemingly dipolar pattern. This component is much more clearly right lateralized, with the positive pole centered over right sensorimotor regions, extending from electrode F2 down across the more posterior sensors FC2 and C2. The negative pole is expressed at a group of electrodes (P6, P8, TP8) positioned over a broadly right parieto-occipital region. More isolated activity over left prefrontal cortex, almost exclusively at electrode AF7, was also found to highly negatively correlate with the difference in resource generation rates during this 350-400ms time-window. Similar activity was evident at this sensor again from 450-600ms, by which time comparable activity had also spread to the nearby electrodes F7, FT7 and FC5. Negatively grR(t)-grL(t) related activity was also briefly evident much later (1280-1340ms) at sensor CP3, and adjacent sensors C3 and CP1. These are positioned over left sensorimotor regions. This time period was more than double the average response time, and would be occurring generally soon following presentation of reward feedback, so this activity is almost certainly unrelated to response preparation.

This analysis also revealed variance in electrophysiological activity that was highly related to the calculated local expected values of each option, although there were substantially fewer instances of this than for the resource generation rates. Nevertheless,

between 300 and 500ms post-stimulus, activity with a strong positive correlation with local expected values was evident at the adjacent electrodes CP4 and CP6, along with the nearby FC6. All three of these are positioned over right sensorimotor cortex. Activity with a significant inverse relationship with evR(t)-evL(t) was also present at the contralateral electrode FC3. This activity peaked from 540-610ms post-stimulus, though was apparent to a lesser extent from as early as 300ms. Local expected value related activity was also evident at sensor FC3 from 690-730ms, paired with a simultaneous re-emergence of the strong positive relationship with evR(t)-evL(t) at sensors CP4 and CP6. Positively correlated activity was then evident again much later still at electrode CP4 alone, from 1330-1490ms, though this is well beyond the average response time.

Overall, electrophysiological results of the present study were broadly congruent with the existing evidence, both in terms of times and electrode locations of interest. The sensors that demonstrated the most robust evidence of value-related activity, in terms of both global generation rates and local expected values, were those positioned over right sensorimotor cortex. It is not feasible to assuredly infer underlying sources of activity measured at extra-cranial sensors. Nevertheless, this finding is at least consistent with previous literature, which has suggested action values for motor responses are represented in supplementary motor regions (Sohn & Lee, 2007; Wunderlich et al., 2009). Value-related activity, inversely correlated with resource generation rates, was also identified at electrodes positioned over left-prefrontal cortex. This is consistent with findings of a number of other studies (Basten et al., 2010; H. R. Heekeren et al., 2004; S. Kim et al., 2008; Wenzlaff et al., 2011). Nevertheless, when considering grR(t)-grL(t) related activity, observations were not entirely congruent with an interpretation of activity as coding directly for action-values. As resource generation rates were fixed throughout the task, there were only two possible values for grR and grL. It follows that grR(t)-grL(t) should have a perfect positive relationship with grR and an equivalent

negative relationship with grL. In this case, if action values are internally represented then they should be evident for both options. In this case it would be expected to observe a corresponding body of activity, with an inverse relationship to that seen in the right hemisphere, mirrored in the left hemisphere. This would reflect the generation rate of the alternative. Although it is inappropriate to conclusively interpret a null result, particularly in a study with relatively low statistical power, the present study simply did not find evidence for this. Activity at sensorimotor electrodes that was significantly related to grR(t)-grL(t) was limited to the right hemisphere only.

On another note, this apparent right lateralization of grR(t)-grL(t) related activity seems to be inconsistent with a series of previous findings, which have indicated instead a left lateralized dominance for activity related to motor planning of goal-directed movements (Frey, 2008; Haaland, Harrington, & Knight, 2000; Janssen, Meulenbroek, & Steenbergen, 2011; S. G. Kim et al., 1993; Schluter, Rushworth, Passingham, & Mills, 1998; Urbano et al., 1998). On the other hand, as noted in a review conducted by Serrien, Ivry, and Swinnen (2006), there is evidence that this lateralization is versatile, with the possibility of each hemisphere providing more particular input depending upon the specifics of the situation. For instance right hemisphere structures are involved in attentional orienting within and between both right and left hemispaces (Mesulam, 1999) and may be more heavily involved in tasks where the allocation of spatial attention between peripheral locations is of particular importance (Corbetta, Miezin, Shulman, & Petersen, 1993). This is somewhat true of the SSFT.

Focusing on the temporal characteristics, there was some consistency between local and global value-related activity within the present study. The most prominent activity relating to each fell between 300 and 500ms post-stimulus. This was also consistent with the most robust temporal component of value-related activity identified in the re-analysis of an earlier study described in Chapter 4. However, when contrasting

against these earlier findings, the overall amount of value-related activation identified in the current study was markedly lower. Further the later component of value-related activity, running from 850-950ms, was also simply not evident in the present experiment.

On the other hand, one previously unobserved component of grR(t)-grL(t) related activity was evident in the results of the present study. This appeared at a markedly earlier timepoint than anything identified in the preceding study, running from as early as 20-80ms following stimulus presentation; it was evident bilaterally at anterior sensorimotor sensors, though with a slight right hemispheric bias. To observe such an early component of seemingly value-related activity is surprising, as an ERP with such a short latency would generally be automatically considered an exogenous response. These inherently reflect purely sensory processing, predominantly of a stimulus' physical characteristics (Sur & Sinha, 2009). Although it seems unlikely, the identification of this activity may therefore be considered potentially novel evidence for some level of value processing within the earliest stages of perceptual processing. Considering this, the fact that this early activity only relates to the fixed generation rates of each option is less surprising, due to the lesser cognitive involvement needed to keep track of these fixed values over the task, relative to doing the same for the continually updated local expected value difference. Moreover, as there are only two values for the resource generation rates, this early activity may represent some early basis of attentional orienting to the more salient option. Alternatively, it is possible that it may simply reflect a Type 1 error. As with all unexpected findings, replication should be a prerequisite for serious speculation.

Returning to the evaluation of the evidence as potential action-value coding, when considering local expected value rather than resource generation rate related activity an intriguing candidate is provided by the inverse relationships of measured activation at electrodes FC3 and FC6 with evR(t)-evL(t). These sensors are positioned almost directly contralateral to each other, over anterior sensorimotor regions, and

demonstrate temporally overlapping evR(t)-evL(t) related activity between 200 and 600ms. This pattern of inversely related activity at these two electrodes is most reliable at around 300-350ms, and is therefore broadly consistent in terms of both time and location with what would be expected for action-value representation on a hand-movement based task. In addition, inter-individual variance in the size of the difference in regression coefficients between activity at these two sensors and evR(t)-evL(t), at both 290 and 350ms post-stimulus, was found in turn to significantly relate to the corresponding between subject variance in the strength of behavioural relationship with the same model predictions. Nonetheless, there are several flaws in this interpretation. First, the peaks of precisely simultaneous evR(t)-evL(t) related activity between these two electrodes are somewhat sporadic. In fact, the overall spread of the relationship appears slightly temporally offset, being around 100ms later in both beginning and end at the left-hemisphere electrode FC3. The two electrodes are also not exactly contralateral to each other, with FC3 being positioned slightly more centrally than FC6. Most importantly, it is simply not appropriate to confidently assert a finding that relies upon somewhat isolated patterns of activation, which are reflected in activity at only very few sensors within each hemisphere. Once again, replication is required before any certain conclusions are drawn.

To summarise, behavioural results were in close accordance with the results of the equivalent Stated condition of the previous SSFT-based experiment. Further, EEG analysis allowed for the identification of value-related electrophysiological activity. Variance in activity was found that related significantly to the difference between fixed resource generation rates of the two options, and also the local expected value difference, which was modelled on an individual trial level. The relationship was more robust for the former than the latter of these models. This is likely related to the relatively weak corresponding behavioural accordance between locally modelled expected values and choice behaviour. Both temporal and broadly inferred spatial characteristics of measured

value-related potentials were broadly congruent with the findings of previous studies. Primary epochs and areas of interest for both global and local value coding ran from 300-500ms post-stimulus presentation, at sensors positioned over lateralised sensorimotor regions. More surprisingly, activity relating to global resource generation rates in this time-period was almost exclusively right lateralized. This was inconsistent with the predicted characteristics of action-value representation, which should comprise paired inverse bilateral value representations. By contrast, a theoretically plausible candidate for action-value coding was identified in relation to locally modelled expected values, though this was evident at only a few electrodes.

Chapter 6: An Electrophysiological Assessment of Both Real and Putative Illusory Expected Value Related Activity on the Standard Probability Matching Task

6.a. Introduction:

The previous chapter documents a study that investigated the internal representation of reward variables during a simple simulated foraging task (SSFT). This focused on value-related activity measured over sensorimotor regions, which was broadly consistent with action-value coding. In this task unharvested rewards were allowed to accrue over time, in a stochastic manner, as determined by the different but fixed resource generation probabilities of each option. This study found clear evidence of internal value-representations that distinguished between the two options according to these global resource generation rates. This was coupled with further evidence, although to a less substantive extent, for the internal modelling of rewards in line with locally calculated expected values, as derived from an FFT-based model, which accounts for the persistence and accumulation of unharvested rewards over time.

The present study was designed to uncover a putative illusory internal representation of resource accumulation over time on the standard fixed probability task, in which unchosen resources do not actually persist. According to the Faulty Foraging Theory, it is hypothesised that such a misperception may provide an underlying explanation for the apparently suboptimal diversification of choices that has been documented in both the existing empirical literature and the previous behavioural studies of this thesis (see Chapter 2). This is undertaken by applying EEG recording concurrent with the behavioural task, as in the previous study, but now using the standard probability matching preparation, in which probabilities of reward are fixed and i.i.d. Again, the primary focus of investigation will be putative lateralized action-value coding. Although EEG lacks in spatial specificity this is most likely to be evident at electrodes positioned over motor effector regions, in this case anterior sensorimotor areas. For this

analysis, subjective local expected values will be calculated from an equivalent predictive model as that used to calculate local expected values in the preceding SSFT-based study, in accordance with predictions of the Faulty Foraging Theory.

As discussed in Chapter 1, the existing body of neuroscience literature relating to probability matching is both relatively sparse and inconsistent. The EEG-based experimental methodology applied here will allow for the exploration of a number of additional putative aspects of probability matching in terms of their underlying electrophysiology. For instance, the evidence of this thesis so far strongly suggests that suboptimal diversification is an intuitive, heuristic-driven response, while maximising relies upon top-down control, to overrule this otherwise prepotent behaviour. In light of the broad empirical consensus associating prefrontal activity with top-down executive control, as discussed in Chapter 1, one can expect to observe at least some choice-related activity changes over prefrontal electrodes in the present study. The prediction of particular interest however is of a putative electrophysiological parallel of the Intuition vs. Strategy effect documented in previous behavioural studies. When considering the relevant literature discussed in Chapter 1, this seems likely to be expressed as higher pre-choice prefrontal activity in relation to maximizing choices. In addition, this may be mirrored by an inverse effect of higher post-choice compensatory prefrontal activity following a non-maximizing choice, though this hypothesis is more tentative.

One other intriguing prospect afforded by the design of the present study is the potential to assess post-feedback activity, and how this relates to both reward expectation and actual reward. Previous studies have extensively documented feedback-related coding, often in terms of the prediction error. This is discussed in Chapter 1. The present study offers an opportunity to further investigate post-feedback reward related activity in the context of probability matching. First, it will be interesting in itself to see if previously observed decision variables, such as feedback-related negativity and P300-

based prediction error coding, are replicated in the present study. Further, if these effects do prove replicable under the current paradigm then it may be highly informative to investigate precisely how this activity is modulated by whether or not it was preceded by a maximizing choice.

One aspect in which such an analysis has the potential to be particularly interesting relates to an account of probability matching behaviour that was discussed briefly early in the thesis, but has since received little attention. This account suggests that matching behaviour is substantially contributed to by the deliberate selection of an option that is known to have a lower probability of reward, due to a potentially higher subjective utility being obtained from the perceived 'skilful' prediction of this lower probability option (Brackbill & Bravos, 1962; Brackbill et al., 1962; Rubinstein, 1959). In fact, some early studies, which found increasing levels of maximizing in relation to corresponding increases in performance-based monetary incentives, hypothesised that this effect was a result of the higher payment for correct choices directly counteracting, and in some cases outweighing the higher potential subjective utility obtainable through picking the low probability option (Brackbill et al., 1962; Siegel & Goldstein, 1959). The effect of higher payment however can, and more recently generally has, been explained as instead increasing the incentive to engage more effortful cognitive processing in determining and applying the optimal solution to the task. In the first study of the present thesis (see Chapter 2) it was found that a substantial proportion of participants did endorse sometimes deliberately selecting to worse option because it was more satisfying to win having done so, this does seem to be a notable factor. However, subsequent multiple regression analysis indicated that self-reported endorsement of this view had a relatively negligible impact on actual proportions of maximizing choices, at least when considered together with other more influential factors such as the use of an intuitive vs. strategic approach to the task. Nevertheless, it has been difficult to

disentangle these two explanations, in terms of which is the most appropriate interpretation of the behavioural effects of relative payment on a probability matching task. The present paradigm offers the potential to measure electrophysiological reward-value signals that emanate from within the brain. This is arguably a much more direct measure of subjective utility than can be inferred solely from the resulting patterns of behaviour, and therefore offers a unique opportunity to potentially observe any disproportionate subjective utility that may be attained from selecting the low probability option.

To summarise, there are a number of key motivations for the present study. The first of these is to bring the Faulty Foraging Theory arc of this thesis to a logical conclusion, through assessing whether or not it is possible to observe a subjective illusory representation of reward accumulation in the unchosen option on the standard probability matching task. If this is the case then it should provide clear evidence in support of the Faulty Foraging model of probability matching, suggesting that matching is driven by an intuitive response, which is rooted in optimal foraging behaviour and naturally overextended on the fixed probability task. Other motivations include the more general assessment of internal value and reward-related activity. This will include whether activations consistent with action-value representation will again be evident over motor effector regions, along with the potential for replication of feedback related effects such as FRN and P300-based prediction error coding. More specifically, it will be interesting to assess whether the account of maximizing as resulting from the engagement of top-down cognitive control will be reflected in higher prefrontal activity preceding maximizing relative to non-maximizing choices. It will also be highly informative to assess the specific characteristics of potentially FRN or prediction error related post-feedback activity, in particular how this may be modulated in relation to whether it is preceding by a maximizing or non-maximizing choice.

6.a.i. Experimental Hypotheses

Drawing upon evidence from both existing literature and the findings of previous studies within the current thesis, a number of specific experimental hypotheses are made.

First, behavioural results should generally replicate previous findings. More maximizing is expected in the Stated than the Learned reward probabilities condition. Higher self-reported strategy use is also expected in the Stated condition, as well as Strategy use in turn positively relating to proportions of maximizing choices. An inverse pattern of equivalent effects is also hypothesised in relation to proportions of choices that match the predictions of the FFT-based model.

Second, it is expected to see some level of electrophysiological evidence of the same. Prefrontal activity can be considered a proxy for top-down executive control. Therefore the more specific predictions can be made of generally higher pre-choice prefrontal activity in the Stated condition, higher pre-choice prefrontal activity preceding maximizing choices, as well as potentially higher compensatory prefrontal activity in the post-feedback period, following both non-maximizing choices and negative feedback.

Third, activity that is consistent with action-value coding is expected at sensors over motor and possibly premotor areas, as was evident in the preceding study. A core proposal of the current study is the potential detection of the internal representation of illusory accumulation of reward value in the unchosen option, in accordance with the FFT-based model. This will most likely be expressed within the putative action-value activity over motor effector regions and should be particularly evident within the Learned probabilities condition and in those who behaviourally probability match.

Fourth, results are expected to replicate previous studies that have shown feedback related negativity and feedback-locked prediction error coding at fronto-central electrodes. Moreover, if this activity is observed with a high enough degree of clarity then it may potentially be used to ascertain whether or not there is any electrophysiological

indication of disproportionately greater subjective utility being obtained for correct predictions of the low probability option than would be expected according to the magnitude of the prediction error alone.

6.b. Method

6.b.i. Participants

A total of thirty-two participants took part in this study. Participants were recruited in an opportunity sample; this mainly comprised a mixture of undergraduate and postgraduate students studying a variety of courses at the University of Nottingham. Of these participants, 12 were male and 20 female. Ages ranged from 19 to 31 (M=21.56, SD=2.38).

Participants were paid in the same manner as the previous study. This was in the form of an inconvenience allowance that comprised both a fixed and a variable component. All participants received a basic amount of £4, to cover the inconvenience of the EEG set-up. In addition, participants could increase this payment by an amount contingent upon the choices they made during the task. Maximizing behaviour would accumulate average additional winnings of £5.88, for a total payment of £9.88. while Probability matching would lead to an average additional pay-off of £4.87, for a total of £8.87. Performance at chance would lead to an average additional payment of £4.20, for a total of £8.20. The experiment was approved by the University of Nottingham ethics committee, according to BPS guidelines.

6.b.ii. Equipment

The experiment was conducted within the same environment, using the same experimental equipment as in the preceding study, documented in Chapter 5. The only difference was that for this study, stimuli were presented through a 23inch Samsung

monitor (S23A700D, 1920x1080, 120Hz, 14vdc) with an advertised response time of 2ms (GTG). This monitor was connected to the same Mac Pro stimulus presentation computer, again using extended desktop, through an optical Thunderbolt 2 cable, via a Thunderbolt 2 dock inside the Faraday cage. The average mean display lag of this system was empirically determined to be 27.34ms. A diagram of the EEG recording set-up is shown in Figure 22 (previous chapter); note that the monitor for the present study ran on 14vdc rather than 48vdc.

6.b.iii. Design

For the purposes of the behavioural analysis, this experiment was a 2x42 mixed design, with the between subjects independent variable of Stated vs. Learned reward probabilities, and the within subjects variable of block. As in the behavioural only version of this experiment (see Chapter 2) participants in the Stated Probability group were instructed that the outcome probabilities were 70% vs. 30%, whereas those in the Learned Probability group were not informed of the outcome probabilities, and therefore had to learn them through experience.

Similarly, differences in measured scalp activity according to Probability condition will be assessed in the electrophysiological analysis. This analysis will comprise both stimulus and feedback-locked analyses, each with a number of additional predictor variables. Stimulus-locked analysis will focus on two further independent variables. First, the actual fixed right minus left expected value difference of the two options. Second, a putative illusory expected value difference derived from an FFT-based model, in which rewards are misperceived to accumulate over time in the unchosen option, in the same manner as the model used in the previous SSFT based study (see Chapter 5). According to this model, local expected values are calculated through a multiplicative combination of adjusted reward contingencies and the number of trials that had elapsed since each

option had last been explored, these values therefore varied over the task depending upon previous choices.

In the feedback-locked analysis variance in scalp activity will be mainly assessed in relation to four alternative independent variables: whether the high probability option had been presented to the right or left hand side of the screen, whether or not the right or left hand option had been selected, whether or not a maximizing choice had been made on the immediately preceding choice and whether the choice was or was not rewarded.

6.b.iv. Stimuli

The task itself was the same standard probability matching preparation as used in Experiment 1 (see Chapter 2), with all stimuli presented in an identical manner. A plain white background was maintained throughout, with a black central fixation cross being presented between trials. Stimuli were images of two unlit light bulbs, one blue and the other yellow (see Figure 3). These were positioned equidistant from the centre and the left or right edge of the screen respectively. The side to which each colour stimulus was presented randomly varied between trials, so as to prevent participants from making an anticipatory decision regarding which side to choose prior to stimulus presentation. The experimental procedure was designed and coded within Psychopy (Peirce, 2007).

6.b.v. Procedure

The task procedure was identical to that of the behavioural only version of this task described in Chapter 2. Precisely, this was an iterated two-alternative forced choice, predicting whether either a blue or a yellow light bulb would turn on in each of 420 trials in order to maximize total reward over the entire task. Participants were randomly assigned between experimental conditions. Preparation for EEG recording was

conducted in the same manner as described in Chapter 5. Electrode offsets were assessed and repaired to a common mode offset of <20mV.

Participants were shown the same task instructions as described in Chapter 2. Only those in the Stated probabilities condition were informed of the reward contingencies of each option; this was done in both probabilistic and percentage-based format. All participants were verbally instructed to remain as still as possible throughout the task, to minimise movement related noise. Participants were informed that there would be two breaks, in which they could stretch or take a drink if necessary. Upon indicating that they were ready, the door to faraday cage was closed, and participants were instructed to begin the task. As in the preceding EEG study (see Chapter 5), participants made their choice by pressing the left or rightmost keys on a Cedrus RB-730 response box, using their left or right index finger.

Once the task was finished, each participant completed the same questionnaire as followed the behavioural only version of the task. Participants were removed from the faraday cage before all electrodes and the electrode cap were removed, prior to completing the post-experimental questionnaire. Participants were at this point also given the opportunity to wash and dry their hair if they wished to do so.

6.b.vi. Questionnaire

The post-experimental questionnaire was identical to the equivalent behavioural only version of the task, described in Chapter 2. This asked participants to estimate the proportion of trials in which each colour bulb had switched on (0% to 100%); the extent to which they had relied upon intuition or strategy in making their predictions (1 to 5); and the extent to which they believed there to have been a pattern in the reward sequence (1 to 5).

6.c. Behavioural Results

Average proportions of maximizing choices, for each experimental condition, are shown below over 42 10-trial blocks (Figure 30). Linear regression slopes were calculated for each participant. Within the Stated probabilities condition this revealed an unstandardized coefficient of 0.0007 with 95% confidence intervals of 0.0029, indicating no statistically significant learning effect at an α criterion of .05. By contrast, within the Learned probabilities condition the unstandardized coefficient was 0.0050 with 95% confidence intervals of 0.0040, indicating that there was a statistically significant learning effect at an α criterion of .05. A between subjects t-test also revealed that the learning slope was significantly greater in the Learned than Stated probabilities condition $t(30)=3.36$, $p=.002$.

The average proportion of maximizing choices over the entire task was .902 in the Stated probabilities condition and .764 in the Learned probabilities condition, with 95% confidence intervals of .030 and .052 respectively. These indicate statistically significant overmatching at the group level (at $\alpha = .05$) within both groups. A between subjects t-test confirmed that proportions of maximizing choices were significantly higher, over the entire task, within the Stated probabilities condition $t(30)=4.33$, $p<.001$. Proportions of maximizing choices over the final third of the task only were also calculated to assess steady state responding. The proportion of steady state maximizing choices within the Stated and Learned probability conditions were .902 and .823, with 95% confidence intervals of .037 and .066. These again indicate significant group-level overmatching at $\alpha = .05$. Another between subjects two-tailed t-test reveals no significant difference between steady state proportions of maximizing choices between the two groups $t(30)=1.97$, $p=.059$. However, if the clear a priori directional prediction is taken into account that it would be proportions of maximizing choices within the Stated group that

will be higher, then it can be justified to apply a one-tailed test. In this case the between group difference is considered statistically significant (p=.029).

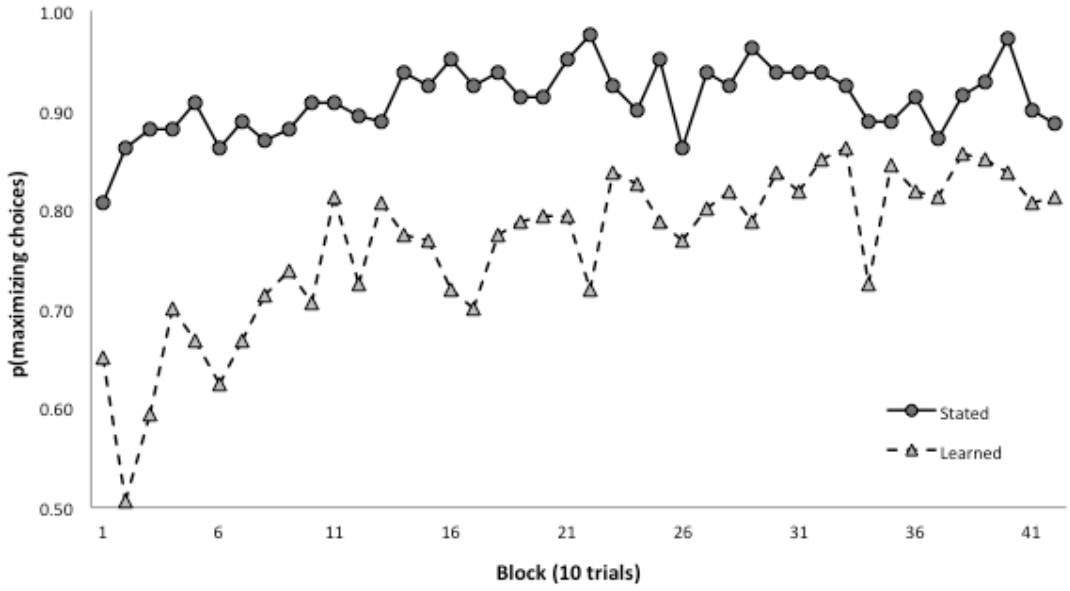


Figure 30: Proportions of maximizing choices for each block and probability condition.

6.c.i. Subjective estimates of reward probabilities

The average subjective estimate of the incidence of occurrence of the higher probability option was .691 in the Stated probabilities condition and .698 in the Learned probability condition, with 95% confidence intervals of .043 and .048. These indicate no significant group-level divergence of estimated from actual reward probabilities within either condition. The estimates in each condition of all previous studies in which the same fixed probability task was used also fall well within these 95% confidence intervals. These are shown below, in Table 18.

Table 18: Subjective estimates of reward probabilities across each non-aggregate feedback fixed probability task condition.

	Pattern instruction					
	Standard		No-pattern		Exp5 Std EEG	
	m	se	m	se	m	se
Stated	.685	.010	.690	.010	.691	.022
Learned	.727	.020	.686	.013	.698	.025

6.c.ii. Heuristics & Apophenia

Self-reported use of intuition relative to strategy use over the task and belief that the outcome sequences contained patterns are shown below in Tables 19 and 20. These tables include equivalent values from previous behavioural only versions of the experiment. Between subjects t-tests revealed no significant differences between Stated and Learned probabilities conditions for either Intuition vs. Strategy use $t(30)=0.00$, $p=1.00$ or Belief there was a pattern in the reward sequence $t(30)=1.56$, $p=.126$.

Table 19: Subjective reports of strategy use across each non-aggregate feedback fixed probability task condition (scores range from 1 = pure intuition to 5 = pure strategy).

	Intuition vs. Strategy					
	Standard		No-pattern		Exp5 Std EEG	
	m	se	m	se	m	se
Stated	3.800	0.176	3.633	0.212	3.313	0.272
Learned	3.200	0.188	3.433	0.149	3.250	0.187

Table 20: Subjective reports of belief that outcome sequences contained a pattern across each non-aggregate feedback fixed probability task condition (scores range from 1=strongly disagree, to 5 = strongly agree).

	Belief there was a pattern					
	Standard		No-pattern		Exp5 Std EEG	
	m	se	m	se	m	se
Stated	2.000	0.209	2.267	0.235	2.703	0.250
Learned	2.633	0.256	2.300	0.263	3.125	0.341

6.c.iii. Win-Stay Lose-Shift

The proportion of trials consistent with a win-stay lose-shift approach was determined for each participant in the same way as in the previous two studies (see Chapters 3 & 5). This proportion was also compared against the expected proportion of WSLS consistent choices that would arise from a maximizing strategy. This base rate is .7, with all choices being stay and, on average, .7 of these being win trials. The group mean proportion of WSLS consistent choices within the Stated probabilities condition was .697, and within the Learned probabilities condition .692, with respective 95% CIs of .012 and .033. These indicate that the average group-level proportions of WSLS consistent choices did not significantly differ from the expected base rate in either condition.

However, as discussed previously, while a maximizing strategy will lead to a proportion of .7 WSLS consistent choices, an entirely WSLS independent probability matching strategy will lead to an expected baseline of .58 WSLS consistent choices. As maximizing increases so too will the proportion of WSLS congruent choices, independently of whether or not WSLS is actively pursued as an approach. This is a concern due to the large discrepancies in maximizing behaviour between participants, and particularly so when considering the systematic divergence between Stated and Learned probability conditions and the potentially confounding relationship between maximizing and Strategy use that was seen in previous studies.

In order to adjust for this issue, expected baseline proportions of WSLS consistent choices were calculated for each participant based upon their proportion of maximizing choices over the entire task. These were then subtracted from actual proportions of WSLS consistent choices as a measure of deviation in WSLS like choice patterns over and above what would be expected by chance. Comparative results, across all fixed probability task conditions, are shown in Table 21. These results indicate

significantly higher than baseline proportions of WSLS consistent choices within both probability conditions of the present study. The proportion of WSLS like choices is also significantly higher in the Learned than Stated condition. These results are most similar to the findings of the Standard condition of the non-EEG task, in which proportions of WSLS like choices were also significantly higher than chance in both experimental conditions. Nonetheless, although higher than chance, average proportions of WSLS choices were still below .7 in each group, indicating that this was likely not a dominant strategy used by participants.

Table 21: Deviation in proportions of Win-Stay-Lose-Shift consistent choices from baseline, across each non-aggregate feedback fixed probability task condition.

	WSLS deviation from baseline					
	Standard		No-pattern		Exp5 Std EEG	
	m	se	m	se	m	se
Stated	.027	.008	.006	.007	.037	.008
Learned	.062	.011	.031	.009	.086	.012

6.c.iv. Faulty Foraging Model Predictions

Putative illusory FFT-based expected values were modelled on a trial-by-trial basis throughout the task. This was done in the same manner as the equivalent genuine local expected values had been in the SSFT version of the experiment, assuming illusory resource generation rates with an equivalent ratio to the actual reward contingencies, as in the preceding SSFT study (.49 vs. .21). Precisely, these now illusory FFT-based expected values were calculated through the multiplication of option's imagined resource generation rate by the number of trials since it was last subject to a failed selection (at which point any expected value would reset to zero), followed by the subtraction of the number of times in which it had been successfully selected since the same time. This

subtraction following an unsuccessful prediction was never allowed to suppress an expected value below a baseline of zero, entailing that each option's illusory expected value at the time of a choice could never be below its illusory resource generation rate.

The difference between the FFT-model derived expected values of the option presented to the left and right of the screen, FFTevR(t)-FFTevL(t), was calculated for each trial. These differences were then split into 6 bins: <-1, -1 to -0.5, -0.5 to 0, 0 to 0.5, 0.5 to 1 and >1. The mean proportion of choices of the option presented to the right of the screen, averaged across all participants within each condition, is shown for trials that fell into each of these six bins, in Figure 31. To determine the reliability of these relationships, linear regression slopes were calculated between the right minus left FFT-model predicted illusory expected value difference and actual proportions of right-hand choices, for each individual participant. The mean unstandardized coefficient was -.284 in the Stated probabilities condition and -.152 in the Learned probabilities condition, with 95% CIs of .047 and .050 respectively. These indicate a significant negative relationship between the FFT-model predicted right minus left expected values and actual proportions of right-hand choices within each experimental condition. Further, a between subjects t-test also confirmed a statistically significant difference between the coefficients of the two groups $t(30)=3.64$, $p=.001$.

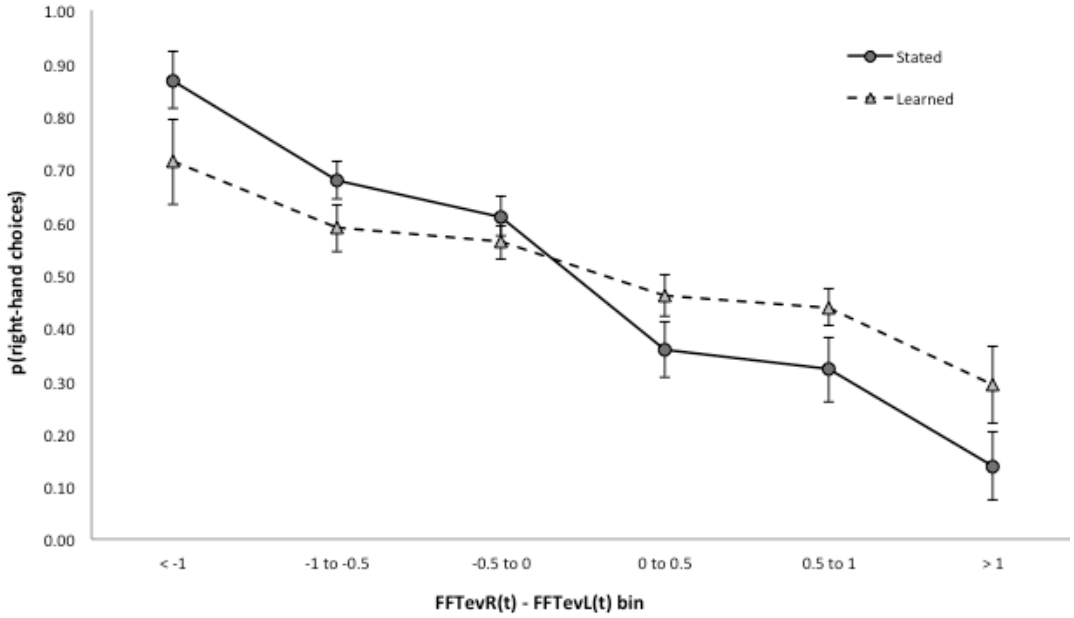


Figure 31: Illustrating the relationship between putative FFT-model derived illusory expected value differences and actual proportions of right hand choices. Error bars show 95% confidence intervals.

6.c.v. Mediation Analyses

To maintain consistency with preceding studies in this thesis, the relationships of both self-reported Intuition vs. Strategy use and Apophenia to participants' choice behaviour was assessed through two independent mediation analyses. These again used the PROCESS macro for SPSS (Hayes, 2012, 2013). Mediation model No. 6 was used, with a single independent variable and two mediator variables. Effects were calculated for each of 10,000 bootstrapped samples. Model results are presented in Figures 32 and 33 (Variable coding is as follows: Stated = 0 Learned = 1, Intuition 1 – 5 Strategy, Belief no Pattern 1 – 5 Belief was Pattern, Proportion of Maximizing Choices/Proportion of FFT-model consistent choices 0 – 1).

The first of these was conducted to determine the relationship to proportions of maximizing choices (see Figure 32). The direct effect of Probability condition on strategy usage was found to be non-significant (t_{30}= .000, R^2= .000). The direct effects on self-

reported apophenia of both Probability condition (t_{29}= 1.556, p = .131), and strategy use (t_{29}= -.491, p = .627) were found to be non-significant, with these two predictors accounting together for 8% of the sample variance (R^2= .084). The direct effects of Probability condition (t_{28}= -4.284, p < .001), and strategy use (t_{28}= 2.481, p = .019) on proportion of maximizing choices were each found to be significant, whereas the direct effect of apophenia was not (t_{28}= -1.327, p = .195). These three predictors accounted together for 53% of the sample variance (R^2= .534).

The unstandardized indirect effects of Probability condition on proportion of maximizing choices were as follows: Through intuition versus strategy use (.000)(.038)= .000, with bias corrected 95% confidence intervals ranging from -.021 to .031. Through apophenia (.000)(-.016)= .000, with bias corrected 95% confidence intervals ranging from -.048 to .003. Through first strategy use then apophenia (.000)(-.116)(-.016)= .000, with bias corrected 95% confidence intervals ranging from -.005 to .003. All indirect effects were non-significant at α =.05.

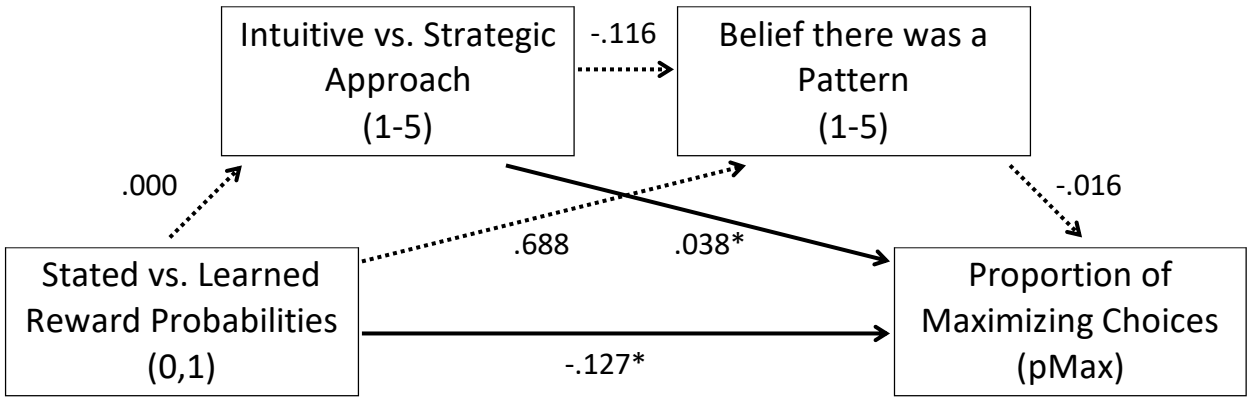


Figure 32: Mediation model for all participants, showing effects on proportions of maximizing choices. Unstandardized regression coefficients shown with asterisk if significant. Solid and dotted lines indicate significant and non-significant effects. α = .05.

The second mediation analysis focused on relationships with proportions of FFT-model consistent choices (see Figure 33). The direct effects of Probability condition (t_{28}= 4.393, p < .001), and strategy use (t_{28}= -3.642, p = .001) on proportion of FFT-model consistent choices were each found to be significant, whereas the direct effect of apophenia was not (t_{28}= .877, p = .388). These three predictors accounted together for 58% of the sample variance (R^2= .576).

The unstandardized indirect effects of Probability condition on proportion of maximizing choices were as follows: Through intuition versus strategy use (.000)(-.064)= .000, with bias corrected 95% confidence intervals ranging from -.050 to .038. Through apophenia (.000)(.012)= .000, with bias corrected 95% confidence intervals ranging from -.009 to .052. Through first strategy use then apophenia (.000)(-.116)(.012)= .000, with bias corrected 95% confidence intervals ranging from -.004 to .004. All indirect effects were again non-significant at α =.05.

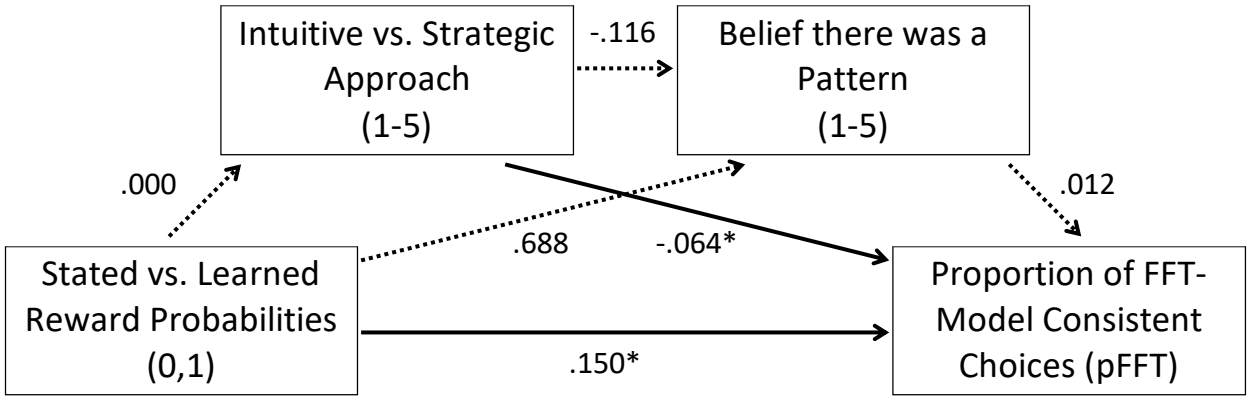


Figure 33: Mediation model for all participants, showing effects on proportions of FFT-model consistent choices. Unstandardized regression coefficients shown with asterisk if significant. Solid and dotted lines indicate significant and non-significant effects. α = .05.

6.c.vi. Behavioural Results Summary

This collection of behavioural analyses has highlighted a number of interesting, and in some cases surprising results. Findings were broadly consistent with those recorded in the behavioural only version of the same standard task, as described in Chapter 2. Higher rates of learning were again evident in the Learned probabilities condition. Significantly higher proportions of maximizing choices were also again observed in the Stated rather than Learned probabilities condition. This was true both over the entire task and, when applying a one-tailed test, also when considering steady-state responses. Subjective estimates of reward probabilities were again found not to differ between groups or from the actual values. No significant between group difference was found in self-reported Apophenia, as in Chapter 2. Surprisingly however, unlike all preceding behavioural studies, the present experiment indicated no link between participants being explicitly informed of the outcome contingencies before the task and subsequent strategy use. As a result of this, no mediated effect of Stated reward probabilities via Intuition vs. Strategy use was evident either. As a general rule for follow-up studies with substantially lower sample sizes than those that they aim to replicate, such a failure could be explained by a combination of lost statistical power and regression to the mean. However, it must be noted that in the present case these effects were found to not only be non-significant, but also effectively non-existent.

Nevertheless, the pattern of independent direct effects on proportions of maximizing choices was once again consistent with the results in Chapter 2. Significant positive effects of stated over learned reward probabilities and strategy over intuition use upon proportions of maximizing choices were both replicated, while there was again no reliable effect of apophenia on choice behaviour. Proportions of choices consistent with a WSLS strategy were once again found to be higher in the Learned condition, and significantly above the individual participant baselines in each condition. These again

remained below the group-level baseline of .7, to be expected from a maximizing strategy, suggesting that WSLS was not a predominant response.

The trial-by-trial predictions of an FFT-based model proved to be an overall poor predictor of choice behaviour. In fact, a negative relationship was found between FFT-model derived illusory right minus left expected value difference and actual proportions of right hand choices. This was true within both experimental groups. This would naturally be expected for participants that maximize, as a maximizing strategy will in most cases predict the opposite choice to the FFT-model. In practice, although absolute maximizing over the entire task is rarely seen, no participants in the present study selected the maximizing option for all of the 420 trials. Nevertheless, this finding indicates that, in the present study, there was in general a substantially stronger correspondence of actual choices to maximizing than to the FFT-model. In fact, the average percentage of maximizing choices was over 90% in the stated probabilities condition, and over 83% across all participants. This is substantially higher than the pure FFT-model expected 70% (strict matching). This high discrepancy between observed and FFT-model predicted proportions of maximizing choices will inevitably lead to an apparently negative group-level relationship between FFT-predictions and actual choice proportions.

Intriguingly, the between group difference does indicate a significantly lesser negative relationship between actual and FFT-model predicted choices within the learned probabilities condition. It is this condition in which it would be theoretically predicted to see a closer accordance with the FFT-based model. This difference then could potentially offer some indication of higher FFT-model consistent choices within a subset of participants, which may be to some extent counteracting the general opposite trend that results from the overmatching and near-maximizing behaviour expressed by others. However, an alternative explanation is that the between group difference may simply be

a side effect of the generally lower group-level proportions of maximizing choices within the learned condition, and unrelated to increased levels of genuine FFT-model related choices. Moreover, the very reason that it would be theoretically expected to observe higher levels of FFT-model consistent choices in the Learned condition is the corresponding expectation of higher reliance on intuition within this condition. Although this difference has been proven robust across two previous studies (Experiment 1 and 2), it was surprisingly absent within the sample obtained for the present study. This suggests that the second, more prosaic, interpretation of the between groups difference might be the more likely explanation. On the other hand, a significant negative effect of strategy use on proportions of FFT-model consistent choices was found, exactly as hypothesised. Moreover, this effect is even greater than the effect of strategy use on proportions of maximizing choices, and so is not explicable solely as a side-effect of this.

The most appropriate interpretation of this pattern of results is not entirely clear. Although multiple effects were consistent with the a priori predictions of Faulty Foraging Theory, the overall relationship between FFT-model choices and group-level behaviour was found to be negative. Moreover, while participants did make fewer maximizing and more FFT-consistent choices within the Learned relative to the Stated condition, group-level behaviour still more closely approximated maximizing than the predictions of the FFT-model even within the Learned condition. These results suggest that the illusory accumulation of expected values in the unchosen option, which forms the foundation of the FFT-model, is not likely to be a dominant factor in determining participants' choices within either experimental condition.

6.d. Electroencephalographic Results

6.d.i. Software

EEG data were preprocessed and analysed in Matlab (Mathworks, MA, USA) using a combination of custom code and built-in functions of the 'Fieldtrip' software toolbox (Oostenveld et al., 2011).

6.d.ii. Pre-processing

Individual trials were extracted from the continuous data stream with reference to trigger values relating to specific task-related events of interest, i.e. the onset of stimulus, or feedback presentation. For the stimulus-locked analysis, the time window of interest was defined as running from 500ms before stimulus presentation to 1500ms following stimulus presentation. For response-locked analysis, the time window of interest was defined as 1500ms before feedback presentation to 1000ms following feedback presentation. The timing of feedback presentation was locked to 500ms post-response, so this time window corresponds equivalently to 1000ms before each choice response to 1500ms after each choice response.

All data were passed through the same three-stage noise removal process as in the preceding study (see Chapter 5). Following this, data were again downsampled to 100Hz, referenced to the average of all scalp electrodes, and then passed through a 10Hz low-pass filter. A manual baseline correction was applied, at the participant level, through subtracting the average activity across all trials for each participant within a specified reference period. This period ran from 500ms pre-stimulus to the onset of stimulus presentation for the stimulus-locked analysis, and from 1500 to 1000ms pre-feedback (1000 to 500ms pre-choice) for the feedback-locked analysis.

6.d.iii. Stimulus-locked Analysis

Two sets of trial-by-trial model predictions, concerning right vs. left choices, were determined for each participant. The first of these was based on a simple maximizing model, which predicts the exclusive selection of the higher probability option over all trials. The second set was determined by the illusory expected value predictions of the FFT-based probability matching model. This model was adapted from, and closely replicates, the local expected value calculation that was used for analysis of the SSFT task in the previous chapter. More precisely, this is calculated through multiplying a given option's resource generation rate by the number of trials since it was last subject to a failed selection, minus the number of times in which it had been subject to a successful selection since this time.

The model predictions for the simple fixed probability model were binary, coded as 0 for a predicted left choice and 1 for a predicted right choice. For the FFT-based model, the predicted right minus left expected value difference, FFTevR(t)-FFTevL(t), for each trial was categorised into one of six bins. These bins were the same as in the behavioural analysis and the previous experiment: <-1, -1 to -0.5, -0.5 to 0, 0 to 0.5, 0.5 to 1 and >1. Each of these bins was coded for incrementally in the predictor model, with values that give an equivalent overall range as the binary model predictors of the fixed probability model: 0, 0.2, 0.4, 0.6, 0.8 and 1. In this sense a 0 represented the lowest right minus left expected value difference of the FFT-model, and therefore the strongest predicted left choice, and a 1 for the opposite.

A series of mass multiple regression analyses were then conducted, for each individual participant, and at each of 64 scalp electrodes and 200 timepoints. Two independent variables were entered into each regression, being the respective predictions of each model for each of the 420 trials. The dependent variable was measured scalp activity. This analysis was conducted through the matlab built-in function 'regstats'. This

analysis outputs a total of 12800 statistics for each participant (64x200 electrodes*timepoints). To control for alpha inflation resulting from multiple comparisons, *t*-statistics of the regression coefficients of each model, for each participant, were then entered into a monte-carlo based cluster analysis (1000 bootstrap samples), which compared them against a dummy variable of all zeros.

6.d.iv. A Priori Between Groups Prefrontal Contrasts - Stimulus-Locked

In addition to the whole scalp analyses described above, an initial a priori planned contrast was conducted focusing on differences in activity exclusively at prefrontal electrodes only (Fpz, Fp1, Fp2, AF7, AF3, AFz, AF4, AF8, F7, F5, F3, F1, Fz, F2, F4, F6, F8) between the Stated and Learned probabilities groups. This contrast was made both separately for high probability left trials and high probability right trials, so as to potentially identify any lateralized differences, and then again following concatenation across all trials.

A single cluster of higher average prefrontal activity was found in those participants who had received Stated probability information before the task began relative to those that had not. This was evident at central prefrontal electrodes for high probability right trials, and with a left lateralized bias across a combination of all trials. There were no significant clusters found for high probability left trials only, though an increase in power may have revealed a similar result to the other contrasts, as one non-significantly related cluster did show a positive trend (p=.060). These results are shown below. Cluster statistics are plotted with a binary mask applied to all data where p>.05. The topography, and evolution of topography of each cluster over time is also shown, with each scalp plot representing a 100ms bin of model-related activity, these begin at 200-300ms and end at 700-800ms.

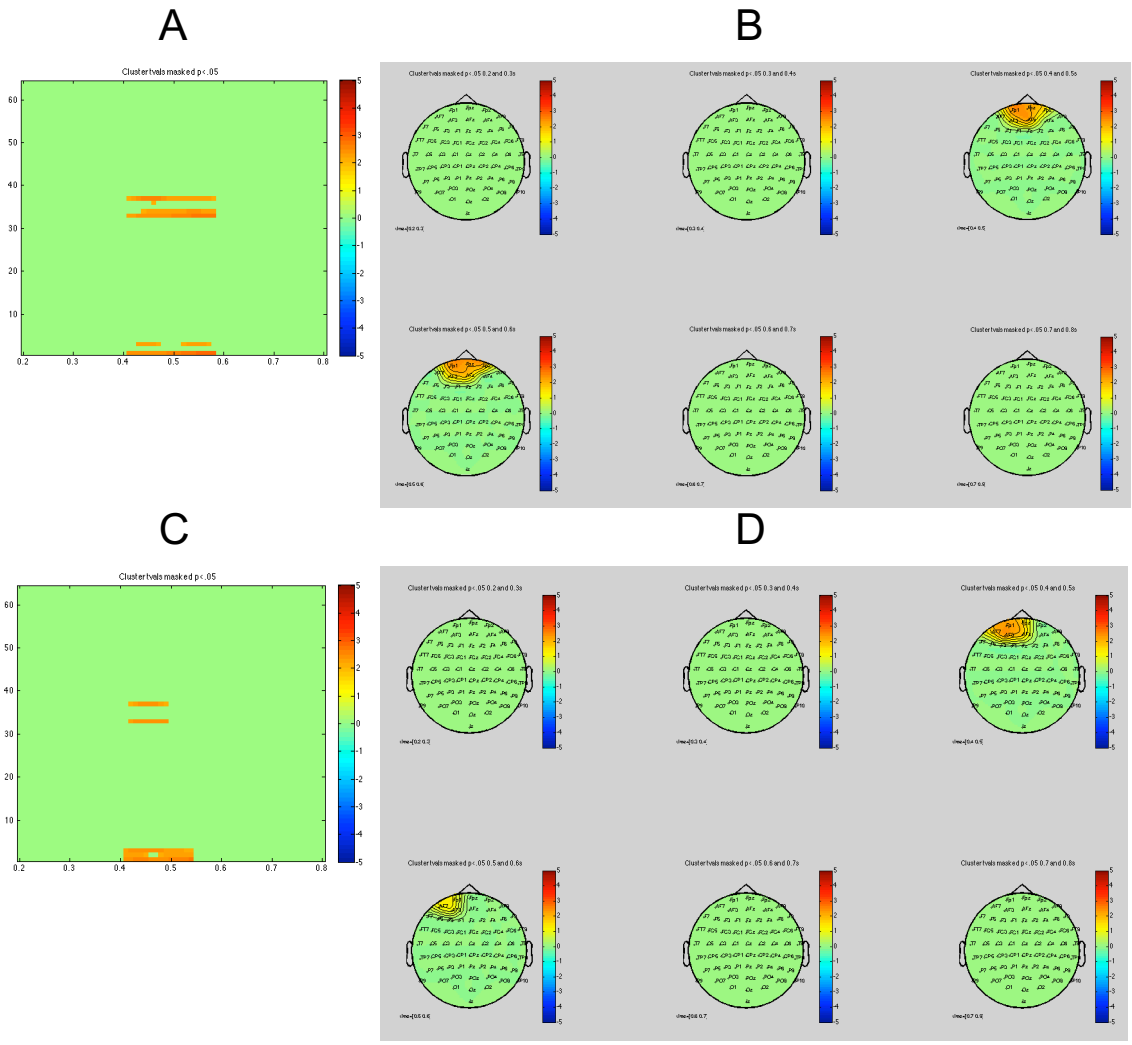


Figure 34: Plots showing significant prefrontal clusters, at which average activity significantly diverged between Stated and Learned probability conditions (p<.05). Panels A and B show differences on high probability right trials only. C and D show differences across all trials. Panels A and C are time-channel charts. B and D are topographic plots. These illustrate cluster position over each of six 100ms time-bins, starting from 200-300ms, and ending at 700-800ms post-stimulus presentation.

6.d.v. A Posteriori Between Group Contrasts – Stimulus-Locked

A number of further exploratory contrasts were then conducted to assess differences in average activity between Stated and Learned probability conditions, now assessing differences in ERP activity over the entire scalp. Differences in strength of the

relationships of sensor activity to each predictive model (*t*-statistics) between Stated and Learned conditions were also investigated. No significant clusters were found for any of these contrasts.

6.d.vi. Assessment of Model Interrelationship

Moving onto the primary analysis it is important to consider that, in theory, any choice pattern that leads to overmatching will entail to a negative relationship between the predictions of the two predictive models. As discussed already, in the most extreme case, of strict maximizing, the illusory FFT-based accumulation in expected value of the unchosen option will cause the lower probability option to apparently supersede the originally higher expected value of the genuine maximizing option by the onset of the third trial. Beyond this point the two predictive models will make entirely opposite trial-by-trial predictions, as the FFT-based illusory EV of the unchosen option will only continue to grow throughout the remainder of the task. Given the ubiquity of at least some degree of overmatching, for many participants a relatively high degree of inverse collinearity between model predictions is therefore expected.

In light of this, a series of correlation analyses were conducted between the two sets of model predictions. This was done in order to assess whether any observed collinearity was within the acceptable range, or was instead too high to justify the entry of both variables together into a multiple regression analysis. Mean *r*-values and corresponding variance inflation factors (VIFs) are shown for each condition in Table 22.

Table 22: Correlation coefficients and corresponding variance inflation factors of the two sets of model predictions, within each experimental group.

	Mean *r* values of correlation of two predictor models		
	m	se	VIF
Overall	-.434	.078	1.260
Stated	-.718	.054	2.064
Learned	-.191	.115	1.038

The maximum level of VIF, at which an unacceptable level of collinearity is indicated, has widely been considered to be 10 (Hair, Anderson, Tatham, & Black, 1995; Kennedy, 1992; Marquard, 1970; Neter, Wasserman, & Kutner, 1989). Stricter thresholds have also been suggested, as low as only four or five (Pan & Jackson, 2008; Ringle, Wende, & Becker, 2015; Rogerson, 2001). Regardless, all calculated Variance Inflation Factors for the present study are below the even more conservative criterion of 2.5. This suggests that the observed levels of collinearity are not so high as to cause concern for the validity of the multiple regression analyses conducted. Results of these analyses are illustrated below.

6.d.vii. Simple Maximizing Model Related Activity – Stimulus-Locked

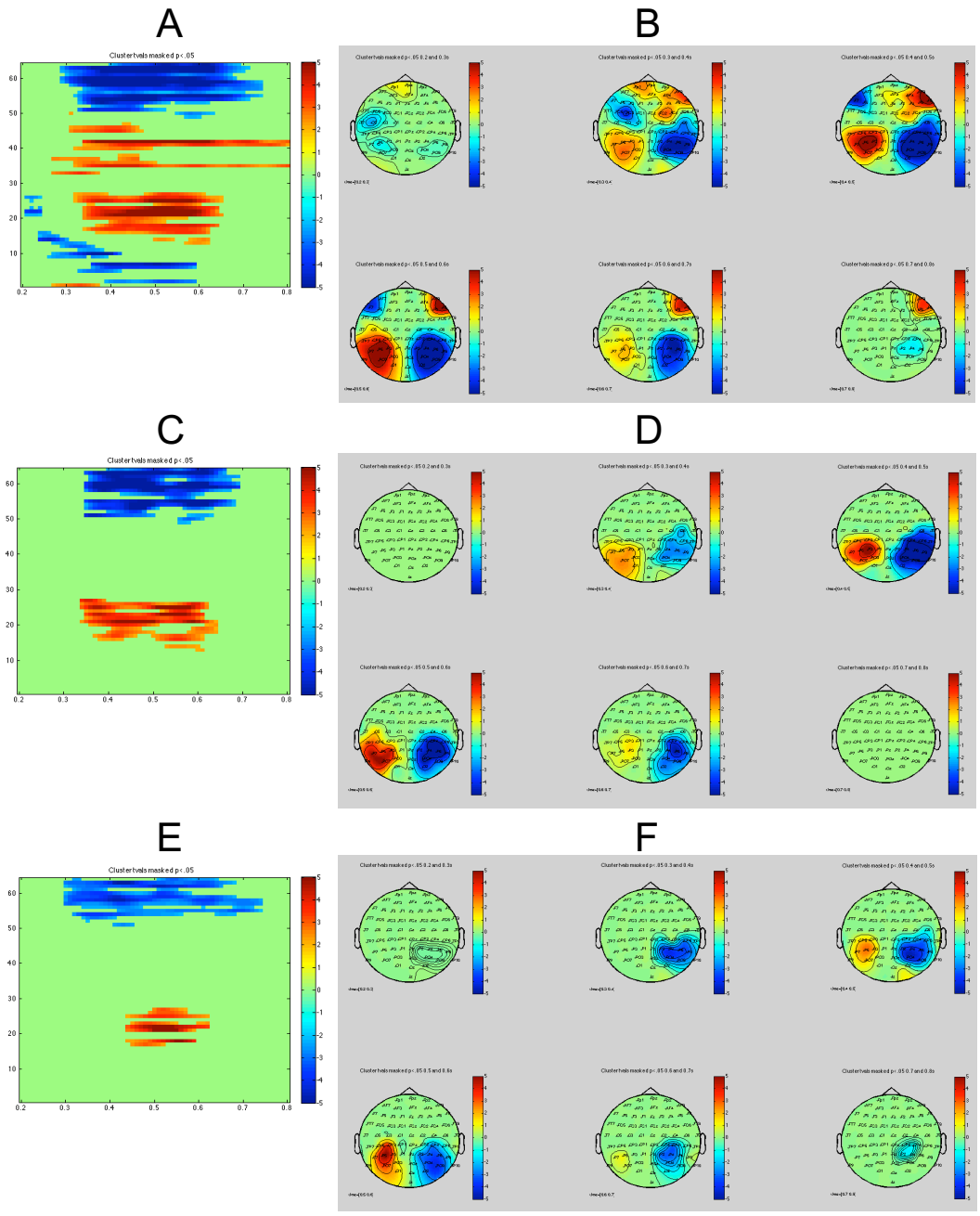


Figure 35: Showing significant clusters of evR(t)-evL(t) related activity (p<.05). Panels A and B show clusters across all 32 participants. C and D show clusters for the Stated condition. E and F show clusters within the Learned condition. Panels A, C and E are time-channel charts. B, D and F are topographic plots. These illustrate cluster position over each of six 100ms time-bins, starting from 200-300ms, and ending at 700-800ms post-stimulus presentation.

6.d.viii. Faulty Foraging Model Related Activity – Stimulus-Locked

No significant clusters of scalp activity relating to the predictions of the FFT-based illusory local expected value model were found, at an α criterion of .05. This was true both independently within each Stated and Learned probability groups alone, and also when concatenating data from all 32 participants.

6.d.ix Response and Feedback-locked Analysis

An adapted version of the stimulus-locked analysis was subsequently applied to feedback-locked data. In fact, as feedback was always presented exactly 500ms following the choice-response this analysis can be considered to be time-locked to both response and feedback. In light of this, a longer segment was extracted for each trial, running from 1000ms pre-choice to 1000ms post-feedback, for a total window of 2500ms. For this analysis the baseline reference period ran from 1500ms to 1000ms pre-feedback, which equates to 1000ms to 500ms pre-response.

For the feedback-locked analysis there were a number of variables of interest. These will be considered in two pairs, due to theoretical interrelatedness. The first of these pairs was whether the higher probability option was presented to the right or left of the screen, coupled with whether the participant actually chose the right or left option. The second was whether participants made a maximizing or non-maximizing choice on a given trial, coupled with whether or not they predicted correctly and were therefore rewarded, on the same trial. The majority of these variables of interest are theoretically informative both before and after feedback presentation. However, excluding the prospect of precognition, one cannot expect there to be any meaningful difference in activity relating uniquely to the valence of future feedback prior to its presentation. In this case, any variance that appears to relate to the nature of so far unseen feedback must instead truly result from a relation to maximizing vs. non-maximising, which is the only

knowable factor that is associated with feedback valence. In light of this consideration, for the response-locked data, analyses were split into two separate pipelines, which focus on pre and post-feedback related activity respectively.

For the first of these pipelines, a mass multiple regression analysis was conducted with the first pair of predictor variables, comprising whether the high probability option was presented to the right or left hand side of the screen and whether the participant actually selected the option to the right or left of the screen. Although the interaction of these two variables should relate whether or not participants made a maximizing choice, a second mass simple linear regression analysis was conducted to directly assess this factor. In this second analysis the sole predictor variable was whether participants made a maximizing or non-maximizing choice. This analysis would more clearly elucidate any maximizing related effects and, in particular, should indicate the direction of any such effects.

The two analyses that were conducted for the second pipeline were largely equivalent to the first. As before, the first was a mass multiple regression analysis with the predictor variables of to which side of the screen the maximizing option was presented, along with the position of the option that the participant actually chose. By contrast with the pre-feedback analysis however, the second mass regression analysis was another multiple regression, with the predictor variables of not only whether or not participants had selected the higher or lower probability option, but also whether they subsequently received positive or negative feedback, having correctly or incorrectly predicted which colour light would switch on in that trial.

Each of these mass regression analyses was conducted through the matlab built-in function 'regstats', and included all 64 scalp electrodes and each of the 100 timepoints. For the pre-feedback pipeline these ran from 1000ms pre-feedback to the onset of feedback presentation. As feedback was locked to 500ms following the response, this

equates to 500ms pre-choice to 500ms post-choice. For the post-feedback pipeline these ran from the onset of feedback presentation to 1000ms post-feedback. Each mass regression analysis outputs a total of 6400 statistics for each participant (64x100 electrodes*timepoints). In order to control for alpha inflation, resulting from multiple comparisons, *t*-statistics of the regression coefficients of each model, for each participant, were subsequently entered into a monte-carlo based cluster analysis (1000 bootstrap samples), in which they were compared against a dummy variable of all zeros.

6.d.x. A Priori Prefrontal Contrasts – Feedback Locked

As in the stimulus-locked analysis, prior to the exploratory analysis described above, a series of more spatially targeted analyses were applied. These were designed to directly assess several a priori experimental hypotheses. In these analyses contrasts were limited to activity measured at exclusively prefrontal electrodes (Fpz, Fp1, Fp2, AF7, AF3, AFz, AF4, AF8, F7, F5, F3, F1, Fz, F2, F4, F6, F8). Findings are described below.

6.d.xi. Stated vs. Learned condition – Feedback Locked

No significant clusters were found for overall differences in response-locked ERP activity between Stated and Learned probability conditions, in either pre or post-feedback periods.

6.d.xii. Max vs. Min and Win vs. Lose – Feedback Locked

In the stimulus-locked analysis the prediction was made that Stating reward probabilities before the task would lead to generally higher prefrontal activity. In the response-locked analysis more specific response triggers are available, which allow for the discrimination of maximizing and non-maximizing choices at an individual trial level. This allows for the assessment of a number of further a priori predictions. First, that maximizing responses will also relate to a higher pre-choice level of prefrontal activity,

reflecting cognitive top-down control. Second, that an inverse pattern may be evident in the post-choice period, reflecting higher compensatory top-down control following a previous non-maximizing response. Third, that there will similarly be higher prefrontal activity after unsuccessful than successful trials, resulting from a greater need for reflection on potential Strategy update. Here the shorthand terms Max and Min are introduced to describe Maximizing and non-maximizing responses, and Win and Lose to describe successfully and unsuccessfully predicted trials. Results are illustrated in Figures 36, 37 and 38.

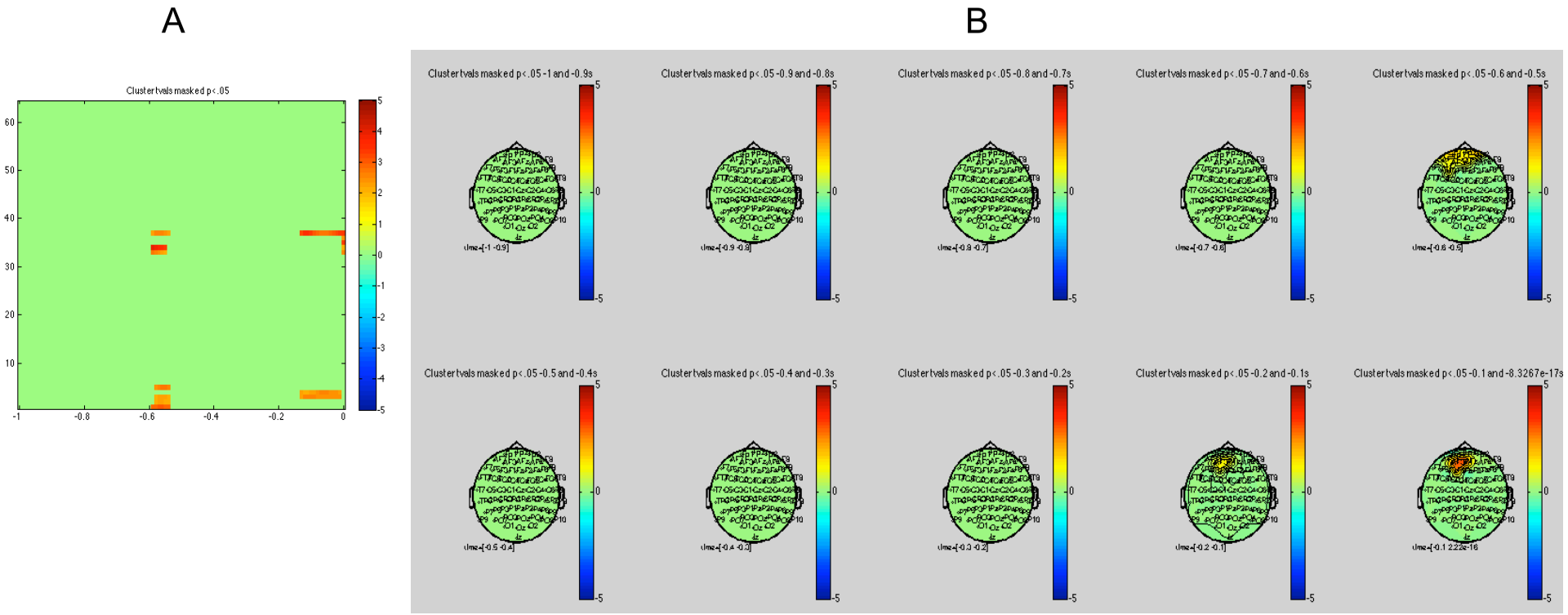


Figure 36: Plots showing significant pre-feedback Max vs. Min related prefrontal clusters, at which average activity significantly diverged between maximizing and non-maximizing choices (p<.05). Panel A is a time-channel chart. Panel B is a topographic plot, illustrating cluster position over each of ten 100ms time-bins, starting from 500-400ms pre-choice, and ending at 400-500ms post-choice.

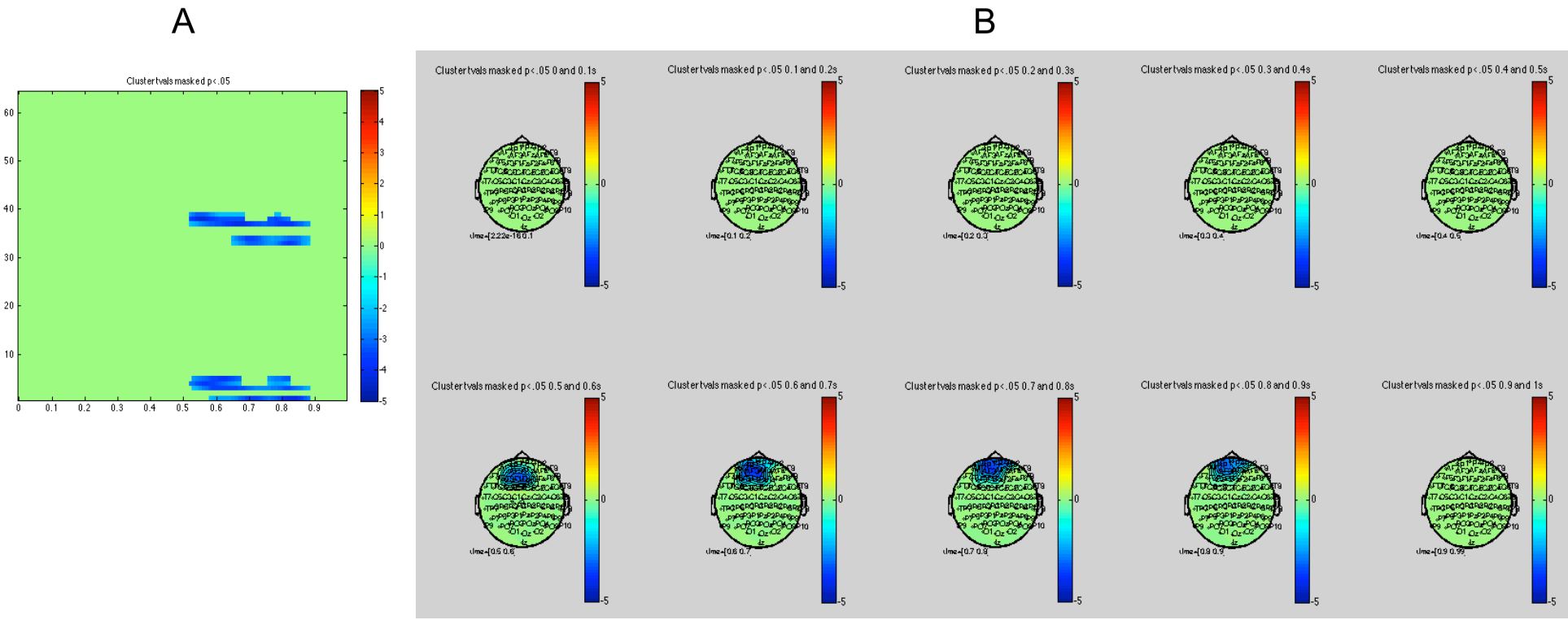


Figure 37: Plots showing significant post-feedback Max vs. Min related prefrontal clusters, at which average activity significantly diverged between maximizing and non-maximizing choices (p<.05). Panel A is a time-channel chart. Panel B is a topographic plot, illustrating cluster position over each of ten 100ms time-bins, starting from 0-100ms post-feedback, and ending at 900-1000ms post-feedback.

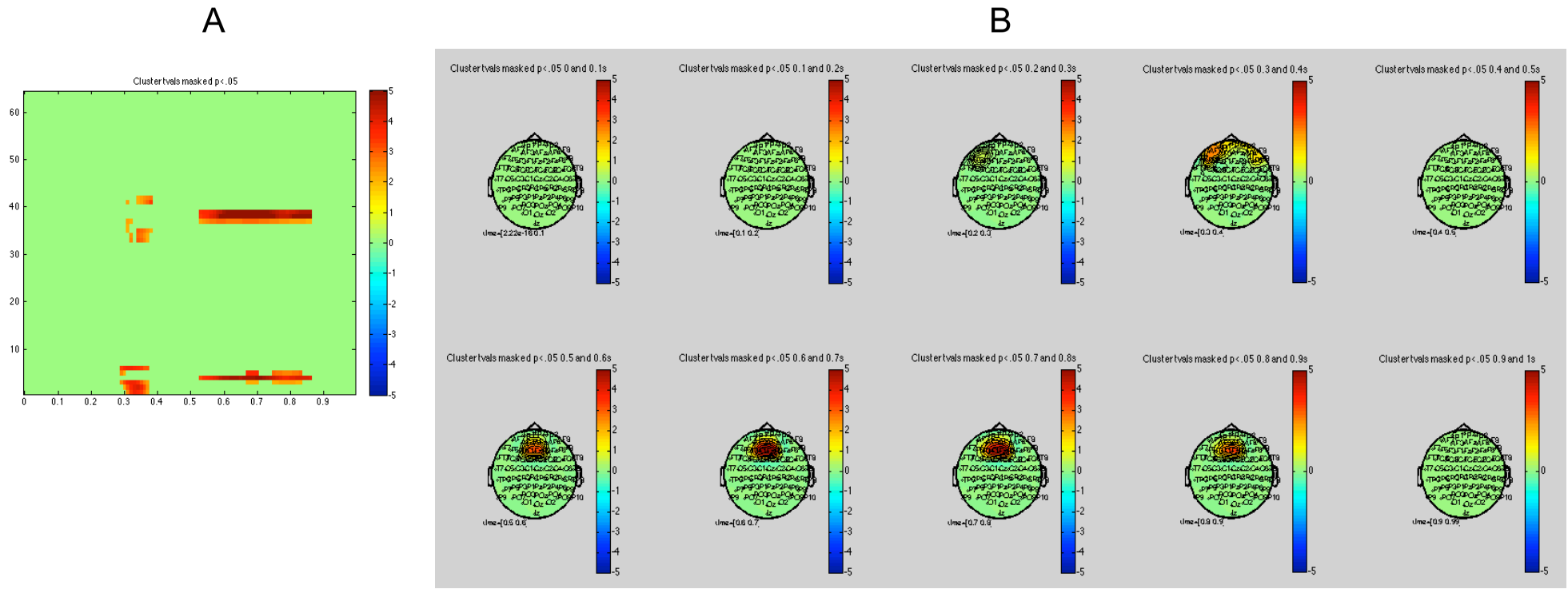


Figure 38: Plots showing significant post-feedback Win vs. Lose related prefrontal clusters, at which average activity significantly diverged between positive and negative feedback (p<.05). Panel A is a time-channel chart. Panel B is a topographic plot, illustrating cluster position over each of ten 100ms time-bins, starting from 0-100ms post-feedback, and ending at 900-1000ms post-feedback.

6.d.xiii. Whole-scalp Analysis – Feedback Locked

Further exploratory contrasts were then conducted to assess differences in average activity, across all 64 scalp electrodes, that related to the three pre-feedback factors: HighProbR vs. HighProbL; ChoseR vs. ChoseL; Max vs. Min; and the four post-feedback factors: HighProbR vs. HighProbL; ChoseR vs. ChoseL; Max vs. Min; Win vs. Lose. Results of these contrasts are described and illustrated below, in Figures 39-43.

No significant clusters were evident in relation to the right vs. left position of the high probability option (HighProbR vs. HighProbL), over either the pre- or post-feedback periods.

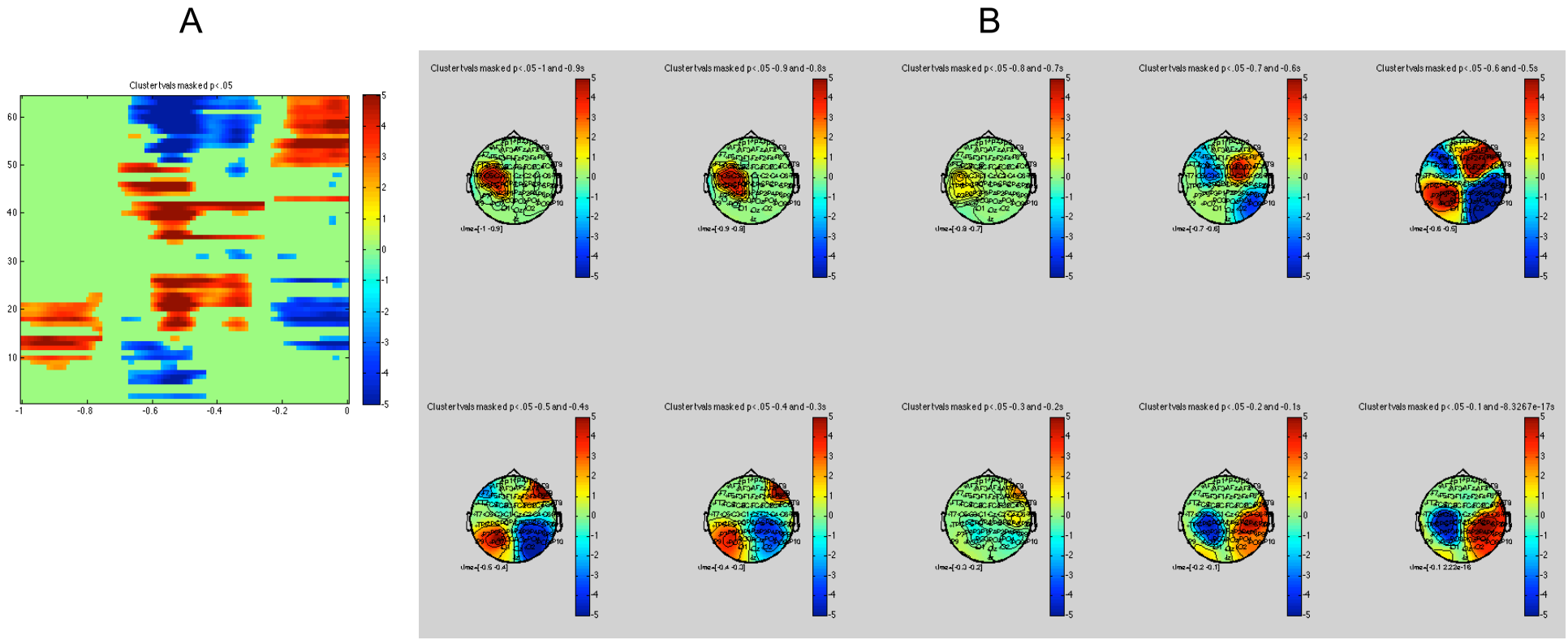


Figure 39: Plots showing significant pre-feedback ChoseR vs. ChoseL related clusters, at which average activity, across all 32 participants, significantly diverged between choices of the right and left hand option (p<.05). Panel A is a time-channel chart. Panel B is a topographic plot, illustrating cluster position over each of ten 100ms time-bins, starting from 500-400ms pre-choice, and ending at 400-500ms post-choice.

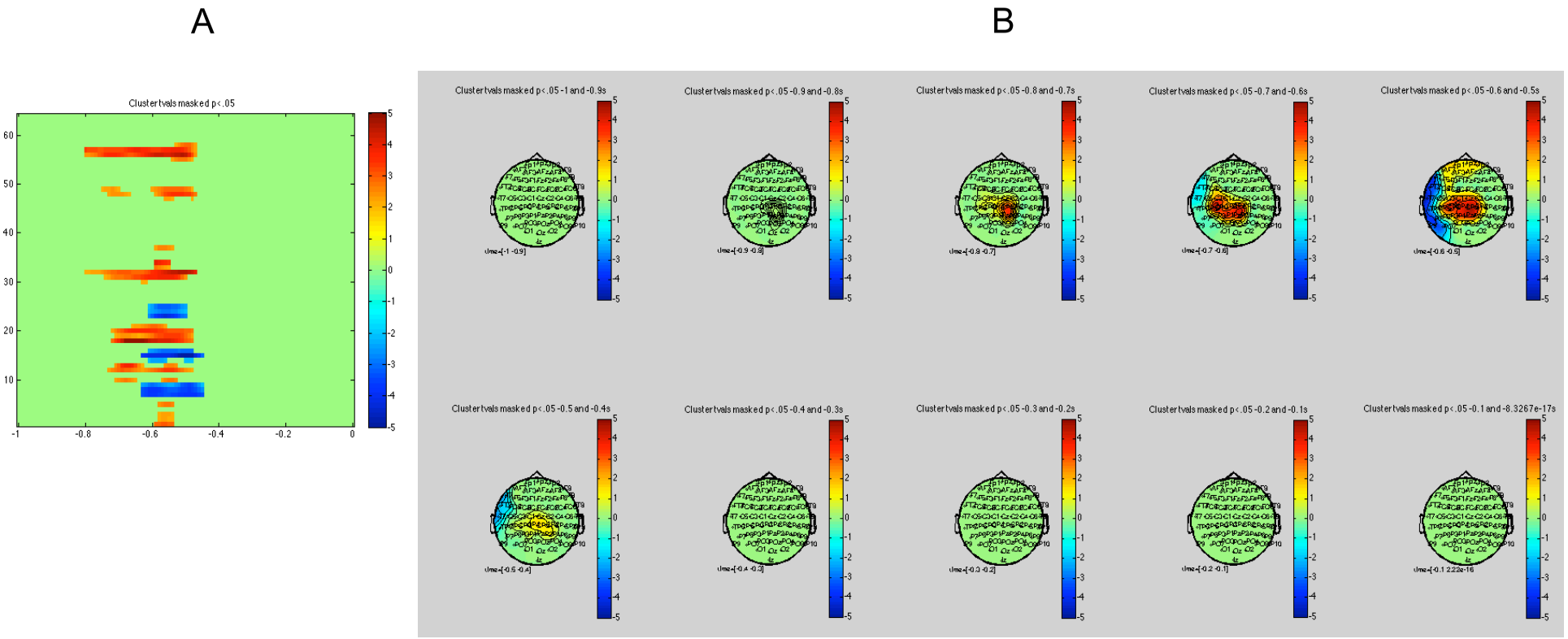


Figure 40: Plots showing significant pre-feedback Max vs. Min related clusters, at which average activity, across all 32 participants, significantly diverged between maximizing and non-maximizing choices ($p<.05$). Panel A is a time-channel chart. Panel B is a topographic plot, illustrating cluster position over each of ten 100ms time-bins, starting from 500-400ms pre-choice, and ending at 400-500ms post-choice.

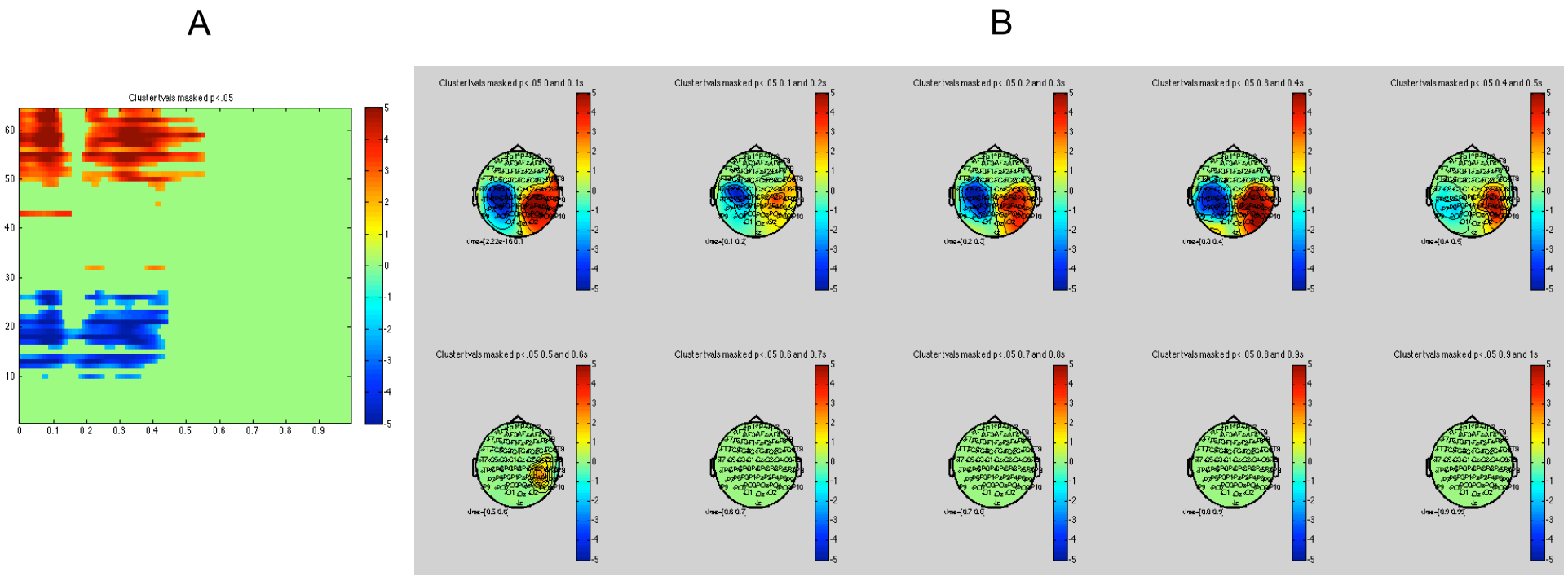


Figure 41: Plots showing significant post-feedback ChoseR vs. ChoseL related clusters, at which average activity, across all 32 participants, significantly diverged between choices of the right and left hand option (p<.05). Panel A is a time-channel chart. Panel B is a topographic plot, illustrating cluster position over each of ten 100ms time-bins, starting from 0-100ms post-feedback, and ending at 900-1000ms post-feedback.

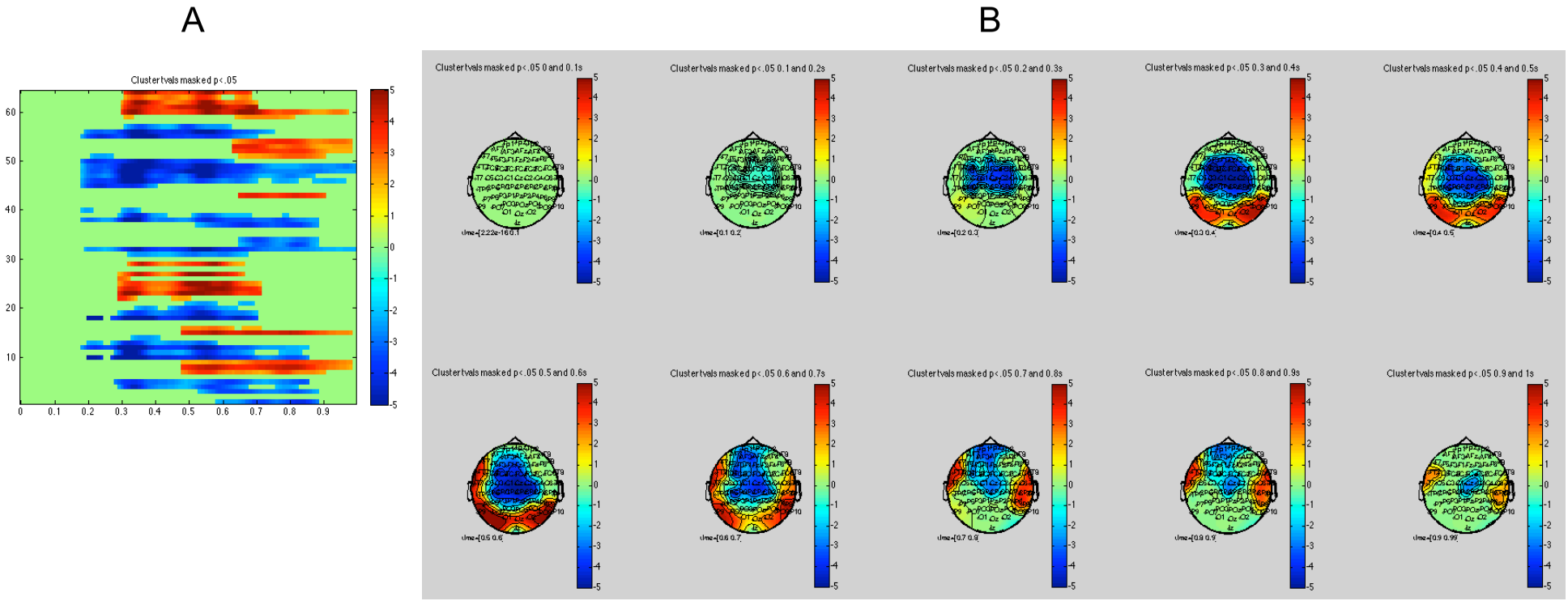


Figure 42: Plots showing significant post-feedback Max vs. Min related clusters, at which average activity, across all 32 participants, significantly diverged between maximizing and non-maximizing choices (p<.05). Panel A is a time-channel chart. Panel B is a topographic plot, illustrating cluster position over each of ten 100ms time-bins, starting from 0-100ms post-feedback, and ending at 900-1000ms post-feedback.

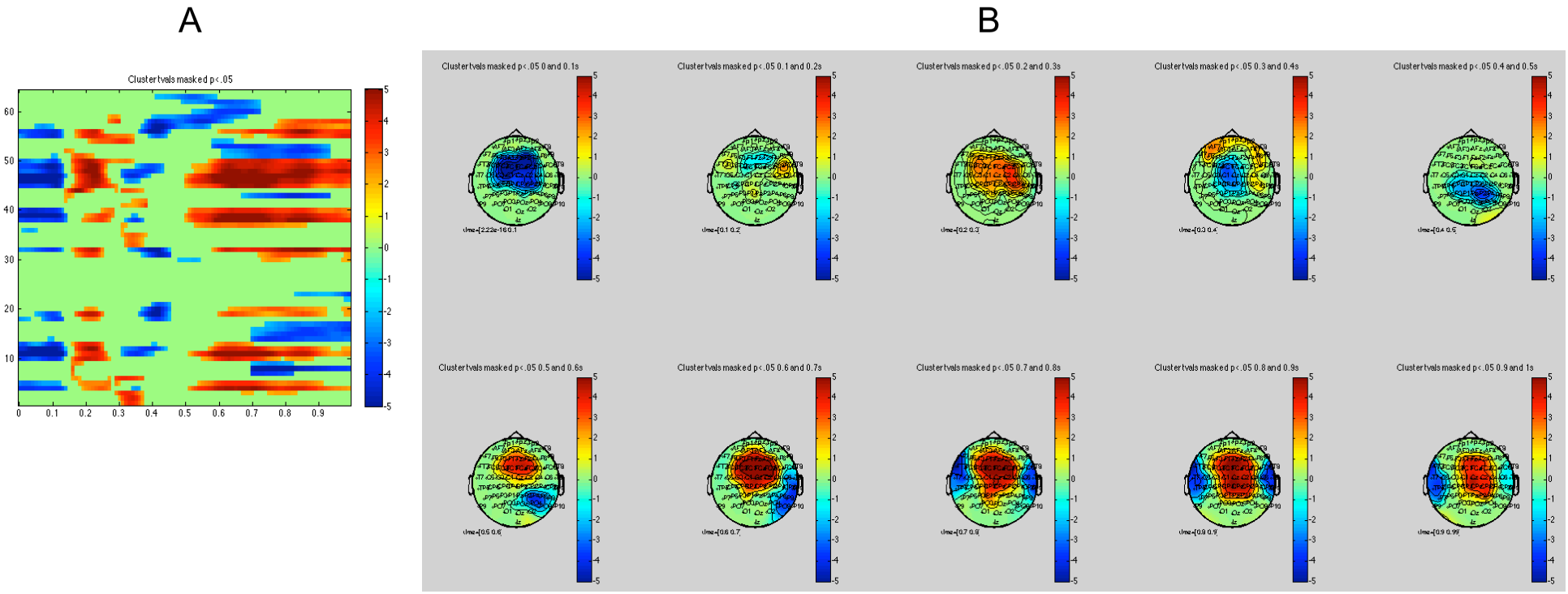


Figure 43: Plots showing significant post-feedback Win vs. Lose related clusters, at which average activity, across all 32 participants, significantly diverged between positive and negative feedback (p<.05). Panel A is a time-channel chart. Panel B is a topographic plot, illustrating cluster position over each of ten 100ms time-bins, starting from 0-100ms post-feedback, and ending at 900-1000ms post-feedback.

6.d.xiv. Prediction Error Coding – Feedback Locked

The prediction error will inherently correlate negatively with maximizing choices and positively with positive feedback. As a result, wherever observed activity simultaneously correlates, in an inverse manner, with the variables Max vs. Min and Win vs. Lose, this could potentially be better attributed to prediction error coding. Prediction error was not entered directly into the above analysis as a factor, due to the intrinsically high level of interrelationship with both Win vs. Lose and Max vs. Min. It was also not explicitly modelled as an interaction between these two variables, as while it would be possible to code for an interaction term that covaries with the prediction error, it is not possible to do so precisely. Although prediction error relates systematically with both Max vs. Min and Win vs. Lose, the simple product of the levels of these two variables cannot itself form an accurate representation of it. Instead of explicitly including a prediction error term in this analysis, it is possible to infer its presence in the manner described above. This potential factor will also be further illustrated by a series of more descriptive and easily interpretable ERP plots that will be presented later in the chapter.

6.d.xv. Assessment of Lateralization of Observed Effects – Feedback Locked

While activity relating to right vs. left hand choices showed a distinct pattern of lateralization, several clusters of maximizing and feedback-valence related activity with central foci were identified. These components may truly reflect homogenous central activity. Alternatively this apparent centrality may reflect a subtler lateralization, which depends upon whether the participant was about to choose, or had just chosen, the right or left hand option. In order to assess this potential lateralization of choice (Max vs. Min) and feedback (Win vs. Lose) related effects, a series of independent analyses were conducted for trials in which participants chose the option presented to the left or right hand-side of the screen. Results of these analyses are illustrated below in Figures 44-46.

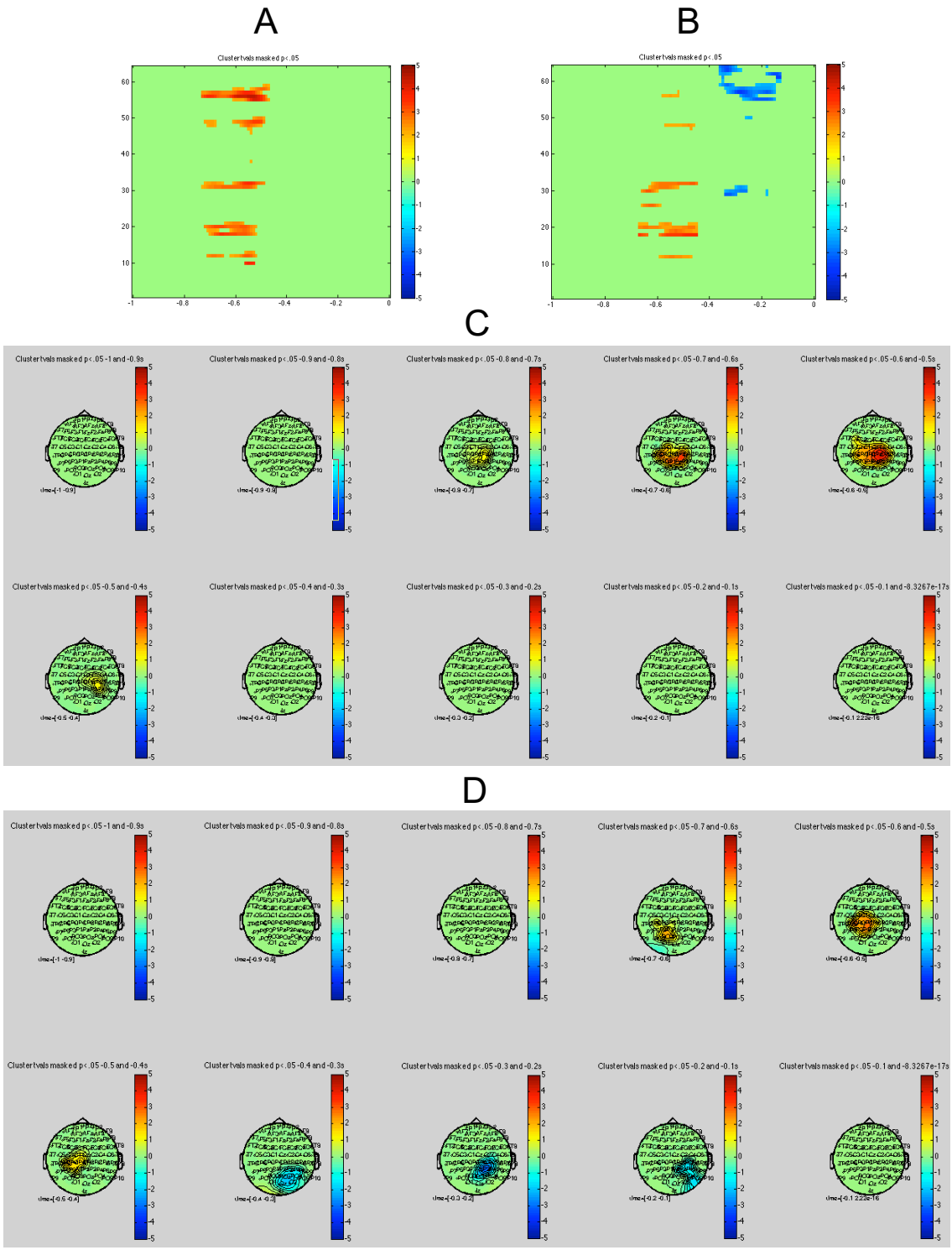


Figure 44: Plots showing significant clusters of pre-feedback Max vs. Min related activity (p<.05). Panels A and B are time-channel charts. Panels C and D are topographic plots, illustrating cluster position over each of ten 100ms time-bins, running from 500-400ms pre-choice to 400-500ms post-choice. Panels A and C show left-hand choice trials only. Panels B and D show right-hand choice trials only. Colorbars represent t-values.

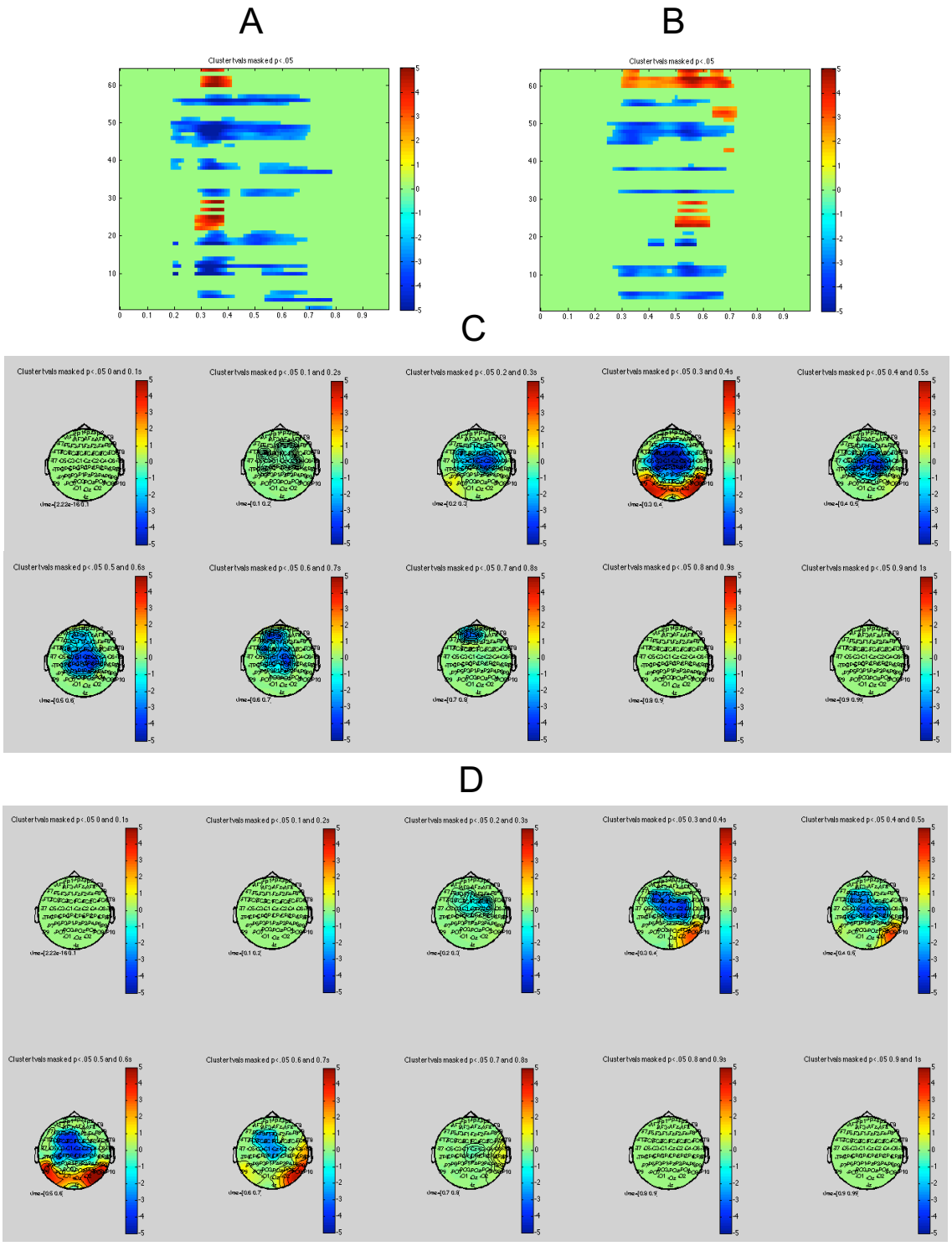


Figure 45: Plots showing significant clusters of post-feedback Max vs. Min related activity (p<.05). Panels A and B are time-channel charts. Panels C and D are topographic plots, illustrating cluster position over each of ten 100ms time-bins, running from 0-100ms to 900-1000ms post-feedback. Panels A and C show left-hand choice trials only. Panels B and D show right-hand choice trials only. Colorbars represent t-values.

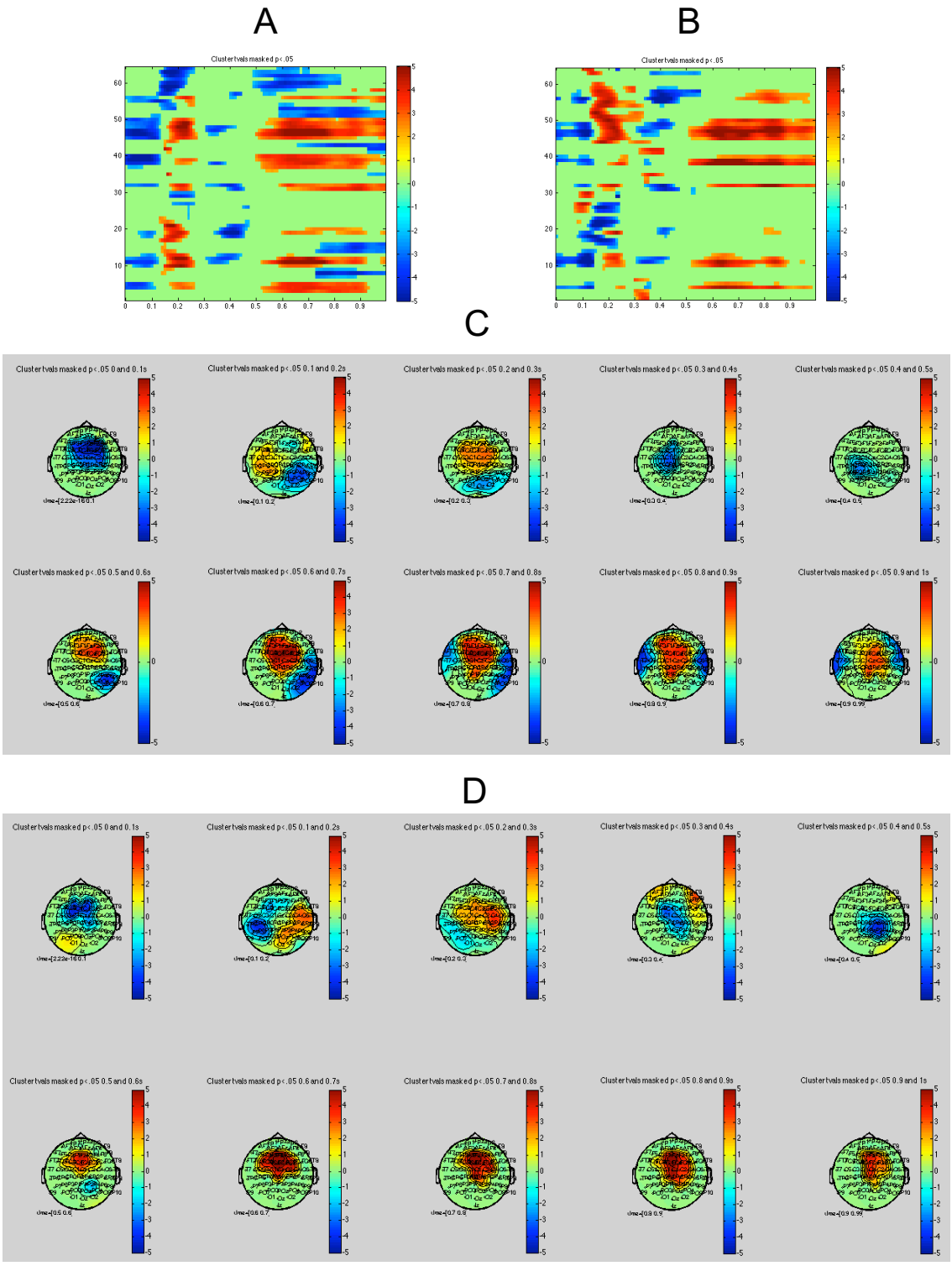


Figure 46: Plots showing significant clusters of post-feedback Win vs. Lose related activity (p<.05). Panels A and B are time-channel charts. Panels C and D are topographic plots, illustrating cluster position over each of ten 100ms time-bins, running from 0-100ms to 900-1000ms post-feedback. Panels A and C show left-hand choice trials only. Panels B and D show right-hand choice trials only. Colorbars represent t-values.

6.d.xvi. Descriptive ERP Plots

In order facilitate the interpretation of these feedback-locked results, which include effects and interactions pertaining to multiple factors, a simpler and more conventional ERP analysis was conducted. The primary aim here was, through a more descriptive presentation of the data, to disentangle the often temporally and spatially overlapping effects relating to both maximizing vs. non-maximizing choices and positive vs. negative feedback. This analysis focused on groups of midline electrodes selected a priori, due to their common use in feedback-related negativity (FRN) based research.

Data plotted here was passed through the same pre-processing stages and baseline correction as in the previous analyses. This was a participant level average, taken across all trials, over the first 500ms of the segment: from -1500ms to -1000ms pre-feedback and -1000ms to -500ms pre choice. The processed data for each trial was then combined, through averaging, into a single 'meta-participant', so as to increase the number of trials that comprise the ERP for each condition, and therefore maximise reliability. It is important to bear in mind that the ERP plotted for each trial type will comprise an average of differing numbers of individual trials. In fact, the low number of trials in which participants had successfully predicted an incidence of the lower probability outcome was a particular consideration in concatenating data across subjects. Some participants had only a single digit number of trials in this category over the entire task. In light of this, averaging within subjects prior to combining into a between subject grand average would disproportionately weight the measured activity from those with a small number of these trials. This could distort the average and in turn lower the signal to noise ratio.

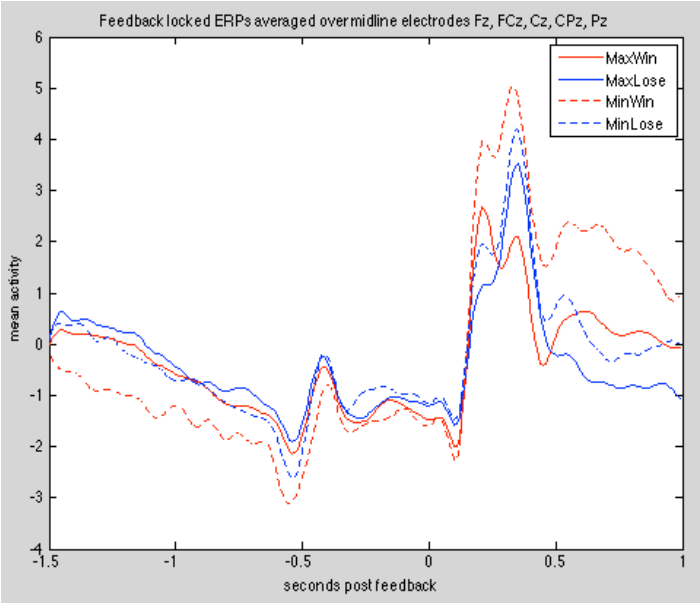
Averages are plotted for each of four conditions, these are termed 'Max-Win', 'Max-Lose', 'Min-Win' and 'Min-Lose', and refer respectively to trials in which participants either successfully predicted the higher probability option, unsuccessfully

predicted the higher probability option, successfully predicted the lower probability option or unsuccessfully predicted the lower probability option. Averages across both maximizing and non-maximizing trial types and both win and lose trial types are also shown on separate plots. The contributions of each trial type to these averages are not weighted by number of trials of which they comprise; rather equal weight is given to each. For instance, the 'Max' ERP plotted here is the average of the 'Max-Win' and 'Max-Lose' averages, even though the 'Max-Win' average itself comprises over twice as many trials as the 'Max-Lose' average. This is done to ensure that any differences evident between maximizing and non-maximizing, or between win and lose trials are independent of the other factor. As the factors are to some extent collinear, this would present a clear confound if averages were not calculated in this manner. It is important to bear in mind however that this does lead to the effective overweighting of trials in some conditions; most notably the 'Min-Win' trials, as there were significantly fewer of these overall, as discussed above. The number of individual trials within each trial type, across all participants, is shown in Table 23:

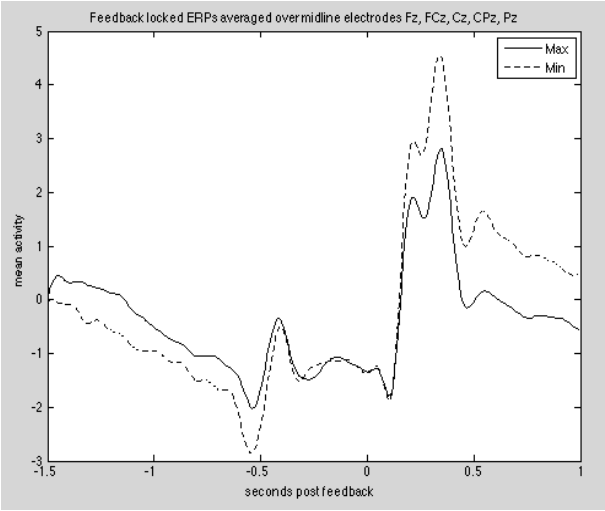
Table 23: Showing number or trials from which the ERP for each of the four trial-types is calculated, across both experimental conditions.

Trial type	Number of trials		
	All participants	Stated only	Learned only
Max-Win	6181	3401	2730
Max-Lose	2583	1424	1159
Min-Win	518	141	377
Min-Lose	1210	358	852

A



B



C

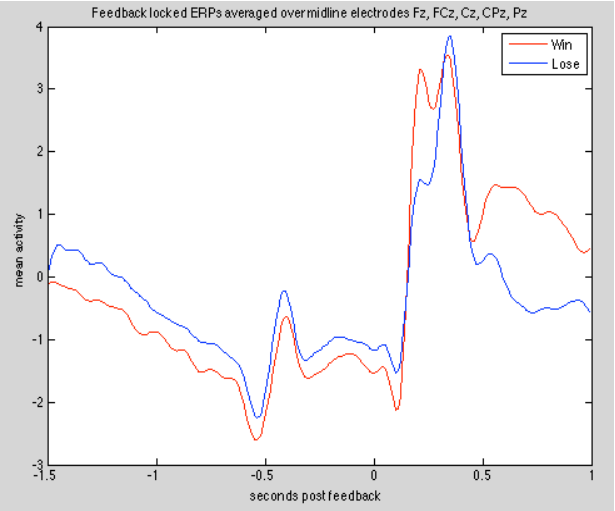
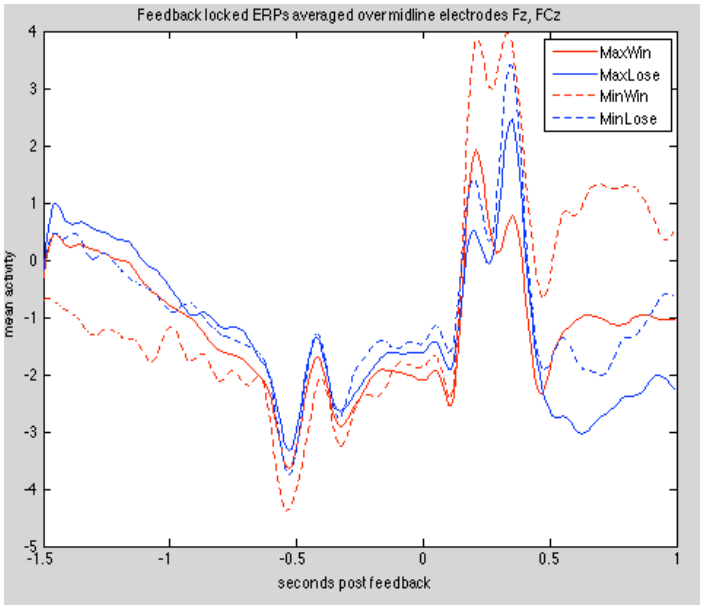
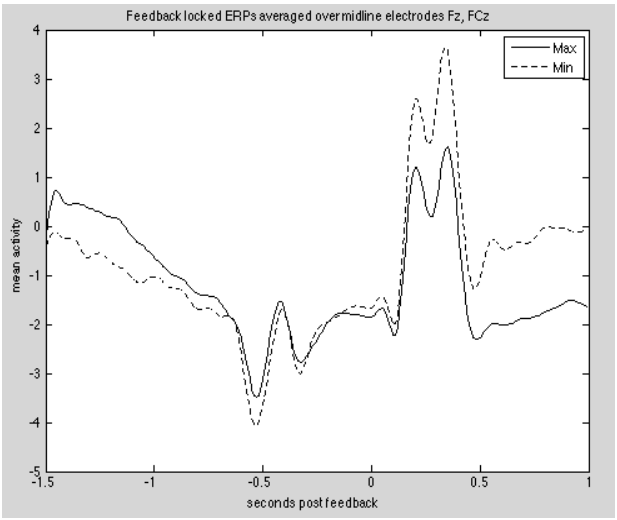


Figure 47: Showing average activity across the four Max vs. Min and Win vs. Lose trial-type combinations, calculated across the five midline electrodes Fz, FCz, Cz, CPz and Pz, across all participants. Panel A shows separate ERP traces for each of the four trial-types. Panel B shows average ERPs for all maximizing vs. all non-maximizing trials. Panel C shows average ERPs for all rewarded vs. unrewarded trials.

A



B



C

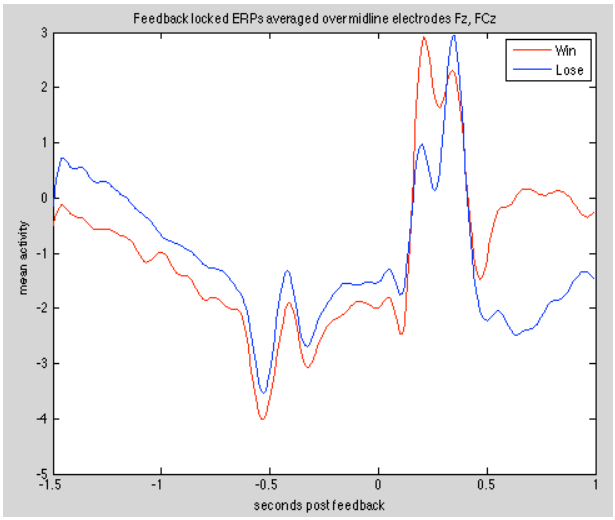
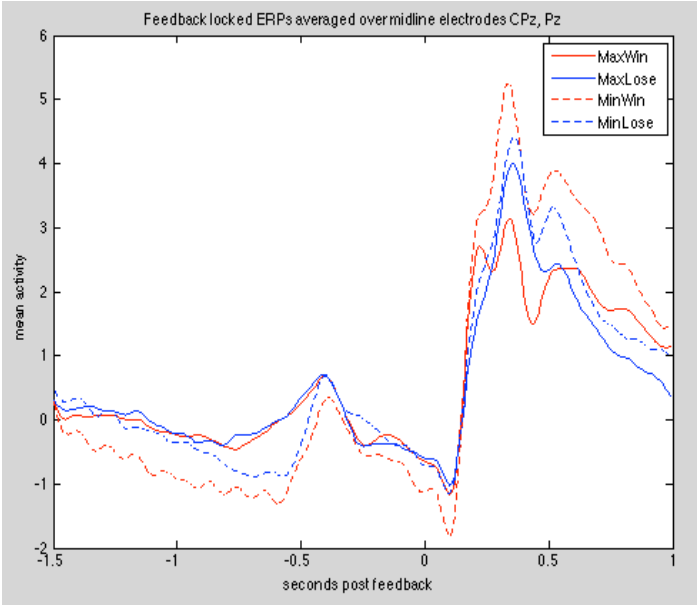
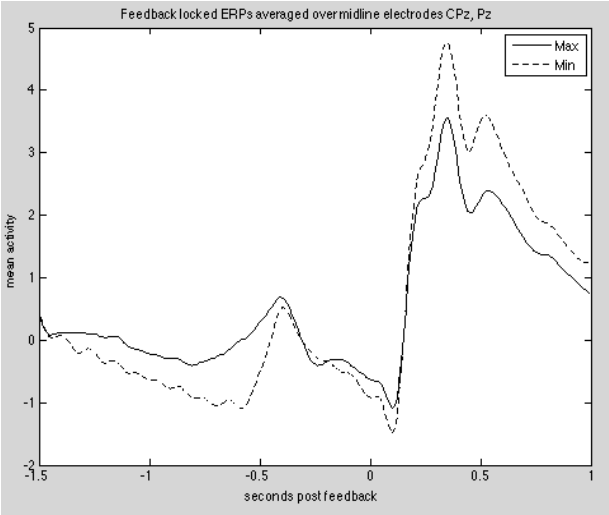


Figure 48: Showing average activity across the four Max vs. Min and Win vs. Lose trial-type combinations, calculated across the two fronto-central electrodes Fz, and FCz, across all participants. Panel A shows separate ERP traces for each of the four trial-types. Panel B shows average ERPs for all maximizing vs. all non-maximizing trials. Panel C shows average ERPs for all rewarded vs. unrewarded trials.

A



B



C

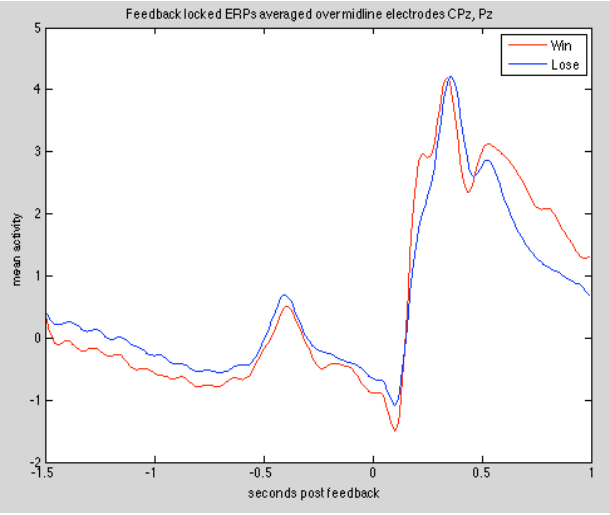
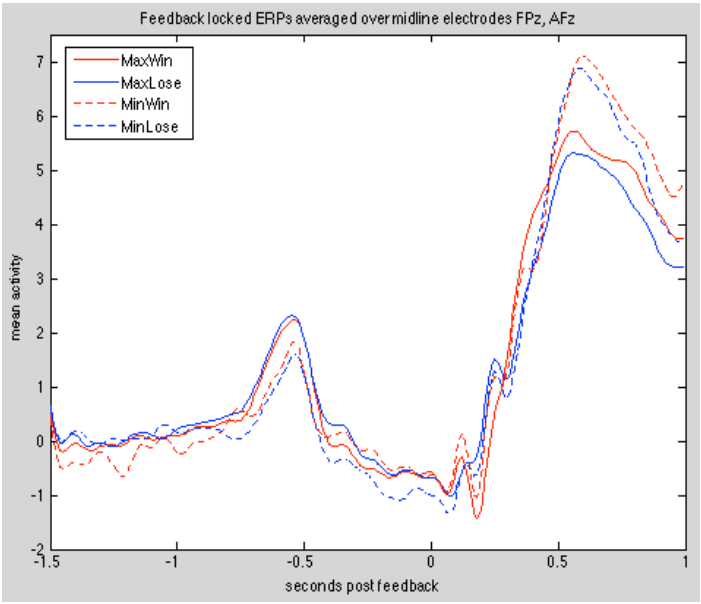
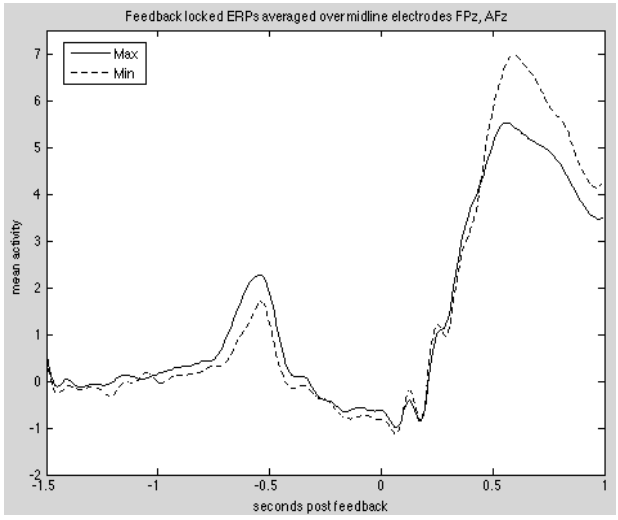


Figure 49: Showing average activity across the four Max vs. Min and Win vs. Lose trial-type combinations, calculated across the two parieto-central electrodes CPz, and Pz, across all participants. Panel A shows separate ERP traces for each of the four trial-types. Panel B shows average ERPs for all maximizing vs. all non-maximizing trials. Panel C shows average ERPs for all rewarded vs. unrewarded trials.

A



B



C

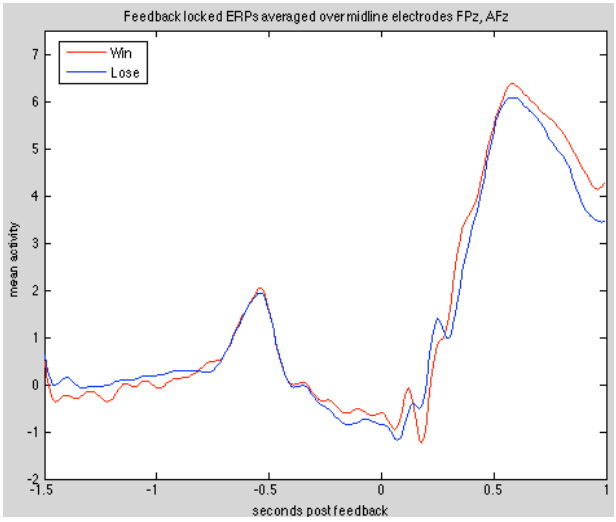
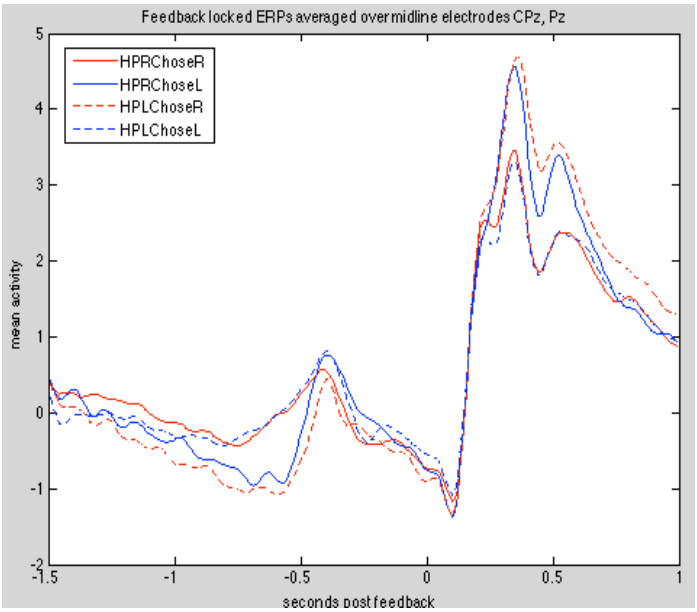


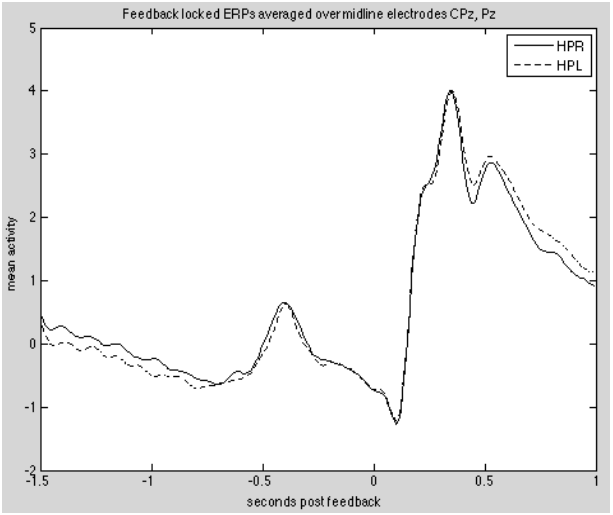
Figure 50: Showing average activity across the four Max vs. Min and Win vs. Lose trial-type combinations, calculated across the two prefrontal midline electrodes FPz, and AFz, across all participants. Panel A shows separate ERP traces for each of the four trial-types. Panel B shows average ERPs for all maximizing vs. all non-maximizing trials. Panel C shows average ERPs for all rewarded vs. unrewarded trials.

To provide another perspective on the difference in activity at parieto-central sensors between maximizing and non-maximizing choice trials the next group of charts splits ERP traces by whether or not the high probability option was presented to the right or left of the screen (HPR vs. HPL) and whether or not participants selected the left or right-hand option (ChoseR vs. ChoseL). The interaction of these two factors determines whether or not a maximizing choice was made.

A



B



C

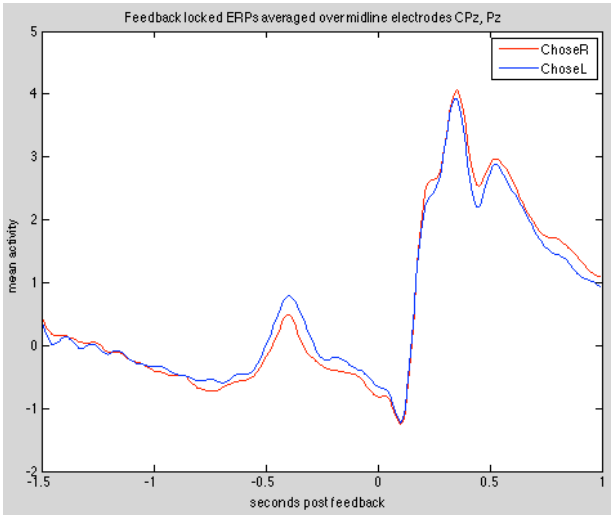


Figure 51: Showing average activity across the four HPR vs. HPL and ChoseR vs. ChoseL trial-type combinations, calculated across the two parieto-central electrodes CPz and Pz, across all participants. Panel A shows separate ERP traces for each of the four trial-types. Panel B shows average ERPs for all HPR vs. all HPL trials. Panel C shows average ERPs for all ChoseR vs. ChoseL trials.

6.d.xvii. Within Group Descriptive ERP Plots

To facilitate between group comparisons, the series of charts that follow (Figures 52-55) plot average activity separately for participants within each the Stated and Learned probabilities conditions, for each of the four Max vs. Min and Win vs. Lose trial type combinations.

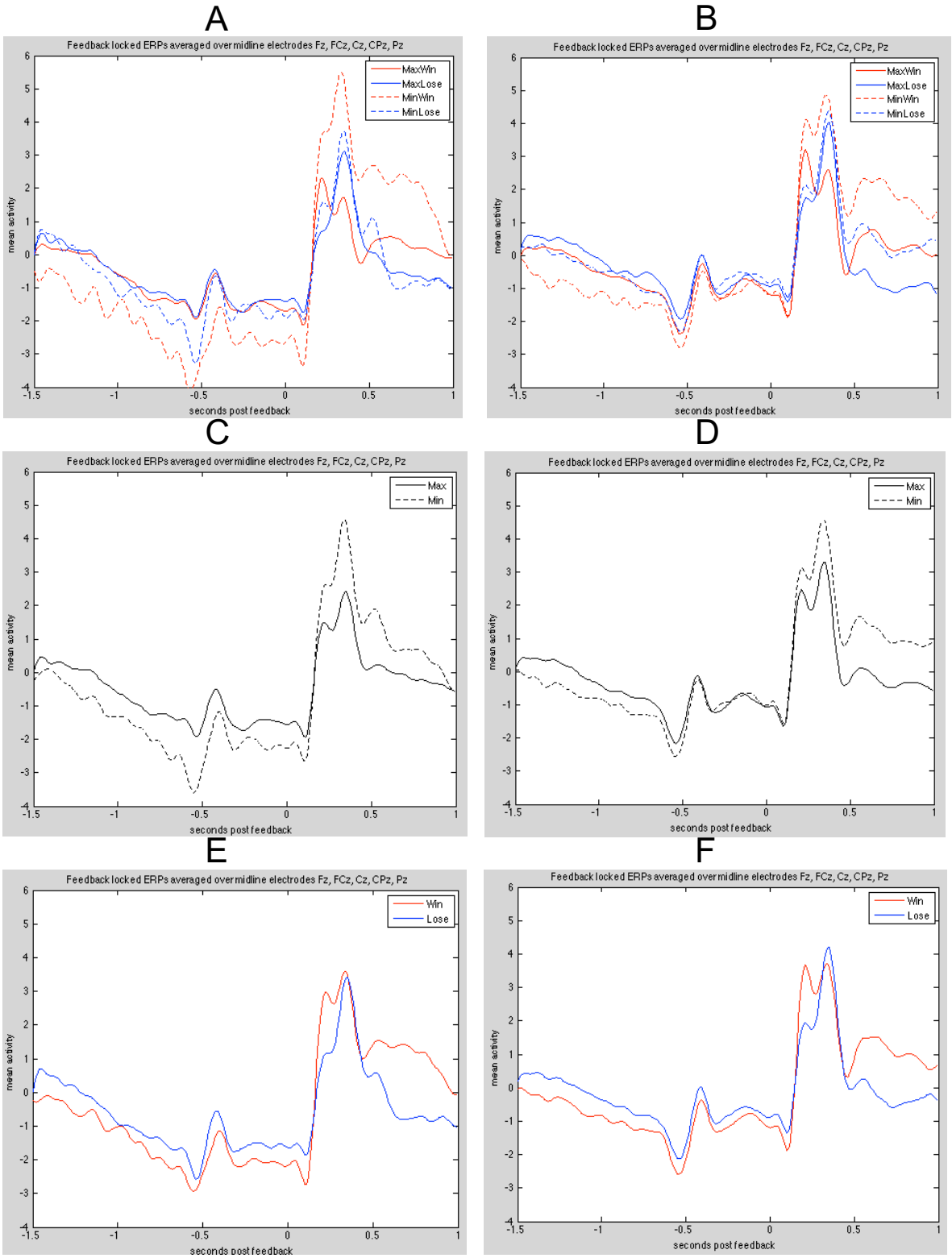


Figure 52: Showing average activity across the four Max vs. Min and Win vs. Lose trial-type combinations, calculated across the five midline electrodes Fz, FCz, Cz, CPz and Pz. Panels A, C and E show average activity within the Stated condition, B, D and F for the Learned condition. Panels A and B show separate ERP traces for each of the four trial-types. Panels C and D show average ERPs for all maximizing vs. all non-maximizing trials. Panels E and F show average ERPs for all rewarded vs. unrewarded trials.

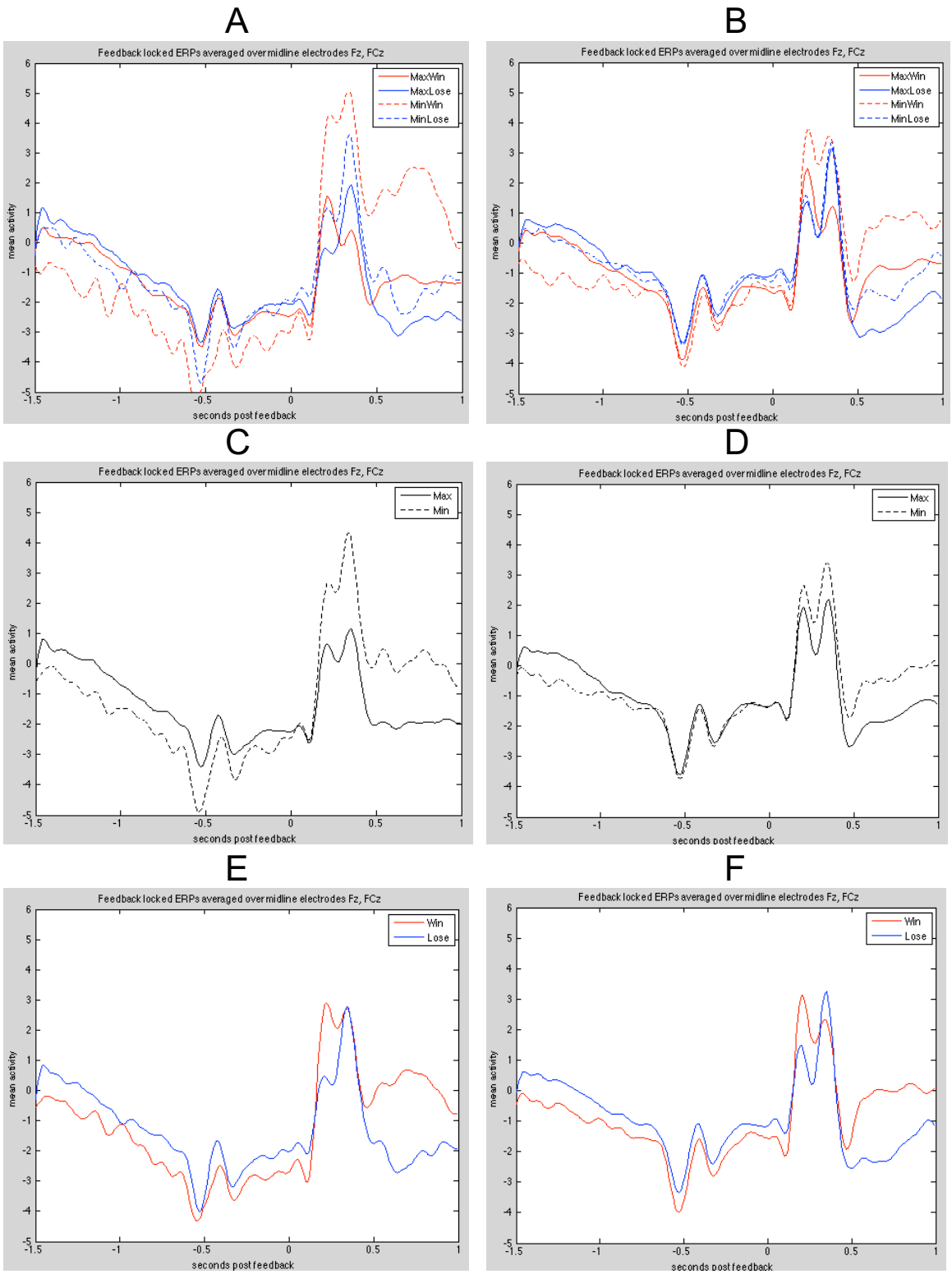


Figure 53: Showing average activity across the four Max vs. Min and Win vs. Lose trial-type combinations, calculated across the two fronto-central electrodes Fz and FCz. Panels A, C and E show average activity within the Stated condition, B, D and F for the Learned condition. Panels A and B show separate ERP traces for each of the four trial-types. Panels C and D show average ERPs for all maximizing vs. all non-maximizing trials. Panels E and F show average ERPs for all rewarded vs. unrewarded trials.

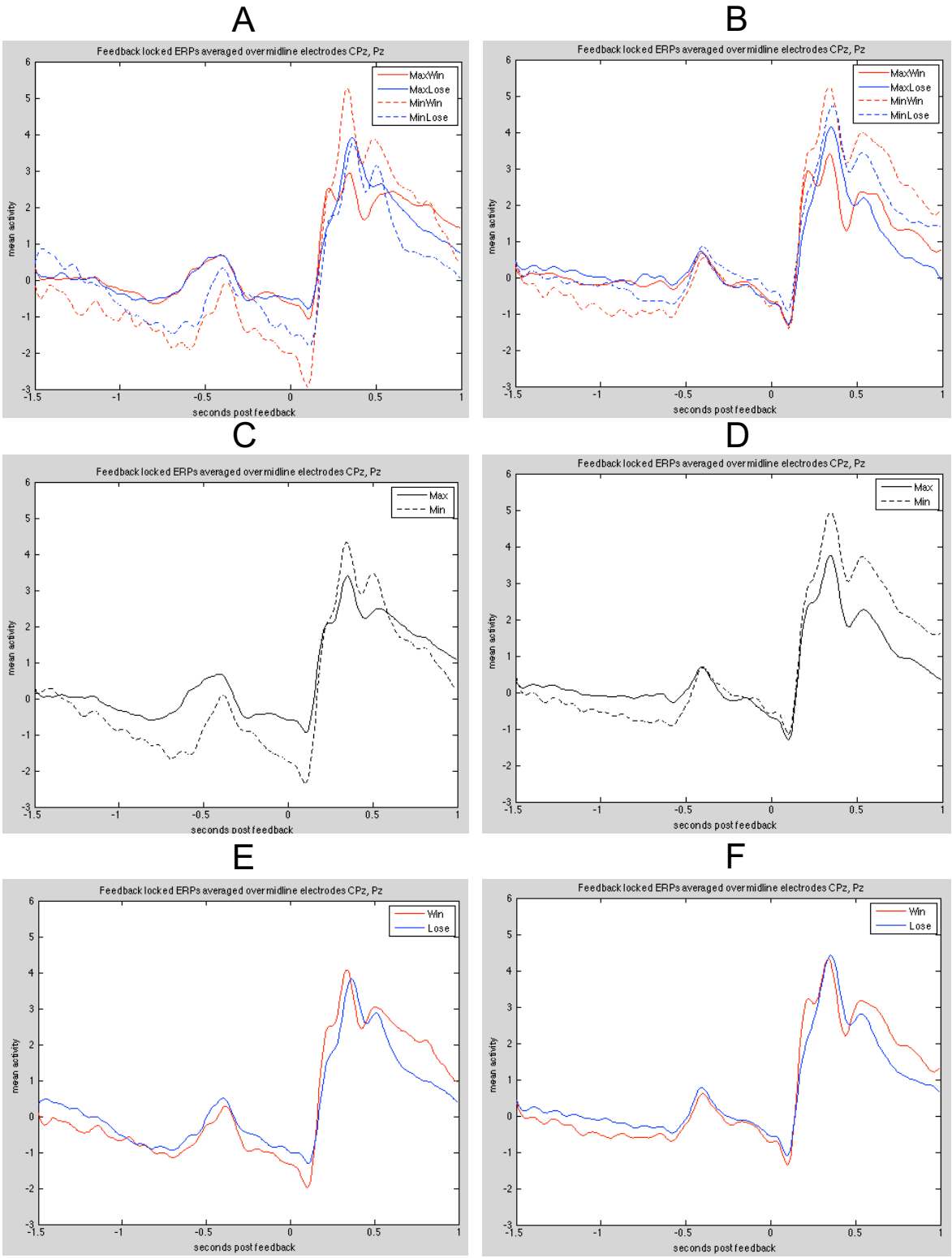


Figure 54: Showing average activity across the four Max vs. Min and Win vs. Lose trial-type combinations, calculated across the two parieto-central electrodes CPz and Pz. Panels A, C and E show average activity within the Stated condition, B, D and F for the Learned condition. Panels A and B show separate ERP traces for each of the four trial-types. Panels C and D show average ERPs for all maximizing vs. all non-maximizing trials. Panels E and F show average ERPs for all rewarded vs. unrewarded trials.

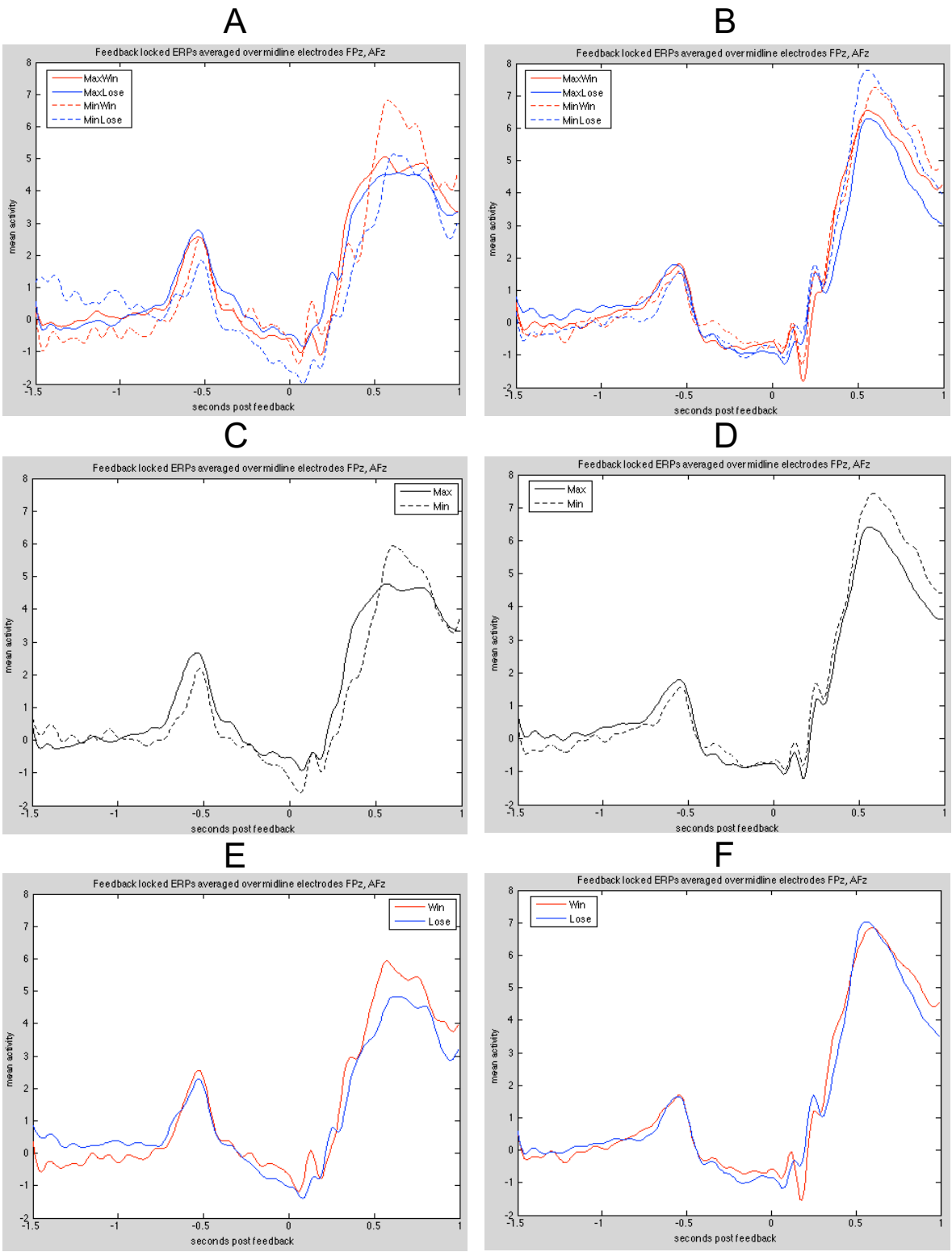


Figure 55: Showing average activity across the four Max vs. Min and Win vs. Lose trial-type combinations, calculated across the two prefrontal midline electrodes FPz and AFz. Panels A, C and E show average activity within the Stated condition, B, D and F for the Learned condition. Panels A and B show separate ERP traces for each of the four trial-types. Panels C and D show average ERPs for all maximizing vs. all non-maximizing trials. Panels E and F show average ERPs for all rewarded vs. unrewarded trials.

6.d.xviii. Electrophysiological Results Summary

Cluster-based Results Summary

The primary electrophysiological analysis applied non-parametric cluster permutation testing to correct for multiple comparisons driven alpha inflation, while determining the statistical significance of the relationships of measured scalp activity to a number of factors, which were in turn output from a series of mass multiple regression analyses. This revealed a number of interesting results, some of which support and some of which do not support various experimental hypotheses.

Stimulus-locked Results Summary

First, an a priori stimulus-locked contrast, which focused on only prefrontal electrodes, revealed a cluster of significantly higher pre-choice prefrontal activity within the Stated than the Learned probabilities condition. This confirms one experimental hypothesis. This contrast was more reliable within trials in which the high probability option had been presented to the right than the left hand side of the screen, and when averaged across all participants the cluster appears to be predominantly left lateralized.

Whole-scalp analysis focused on activity relating to the two predictive models of expected value. The first of these was a simple maximizing model, which always predicted the actual fixed expected values, and therefore the choice of the higher fixed probability option. A large amount of activity was found that significantly related to this model, across multiple clusters. The predominant pattern of value-related activity was evident both immediately preceding and at around the average time of choice response. Spatially, this pattern was broadly similar to that observed in the SSFT-based EEG study documented in the preceding chapter, comprising a dipolar pattern of frontal and parietal positive negative clusters. Here, this activity again seemed to show a right hemispheric dominance, but in this case was quite clearly evident in both hemispheres in an inverse

pattern. These pairs of lateralized clusters most likely reflect dipoles of a single source within each hemisphere, which are each tangential to the scalp and located in motor or premotor areas. This activity is therefore highly consistent-with action-value coding and likely reflects motor preparation and execution of the motor response.

The second model calculated expected values on a trial-by-trial basis, taking account of a putative illusory accumulation of value in the unchosen option over time. This occurs in a manner that derives theoretically from Faulty Foraging Theory, and is functionally equivalent to the genuine local expected values model of the preceding SSFT study. Somewhat anticlimactically, following the cluster-permutation based correction for multiple comparisons, the present analysis revealed no clusters of activity that reliably related to this FFT-based model.

Feedback-locked Results Summary

Again first, a series of a priori contrasts were conducted, limiting sensors of interest to those positioned over prefrontal regions only. A total of six contrasts were made, between Stated and Learned probability conditions, between maximizing and non-maximizing choice trials, and between rewarded and unrewarded trials, each of these independently for both the pre and post-feedback periods. First, no clusters of response or feedback-locked activity were found that differed reliably between Stated and Learned probability conditions, either pre- or post-feedback.

However, a number of clusters were identified reflecting activity that differed significantly depending upon whether or not a maximizing choice was made. These findings supported two experimental hypotheses. First one cluster of higher prefrontal activity was evident over the 100ms that immediately preceded maximizing relative to non-maximizing choice responses; a second cluster of increased prefrontal activity was also evident in the 200ms preceding feedback, following maximizing choices. Second, in the post-choice period an inverse effect was found, with a cluster of reliably higher

prefrontal activity following non-maximizing choices being evident from 500ms to 900ms post-feedback. Two independent clusters of prefrontal activity were also identified that positively related with positive over negative feedback. This is an opposite pattern to a different experimental hypothesis. The first of these onset from 200 until 400ms post-feedback presentation and was evident at more anterior electrodes (F7, F5, AF7, AF3, Fp1, Fpz, FP2, AF8) with a slight left dominance, though evident at electrodes in both hemispheres. The second of these appeared later, from around 500ms to 900ms post-feedback, both centrally and also at the most posterior sensors included in this analysis (F1, Fz, F2).

Whole-scalp analysis also looked independently at pre- and post-feedback activity. Contrasts were made again between maximizing and non-maximizing choices as well as between positive and negative feedback, additional factors were whether the maximizing option was presented to the right of left of the screen along with whether participants chose the option selected to the right or left of the screen. First, no clusters of activity were found that related to the side to which the high probability option was presented, in either pre- or post-feedback periods. Clusters of activity relating to the selection of the right or left hand option, in the 100-200ms immediately before and after the choice response, were effectively the same as those described in the stimulus-locked analysis in relation to the simple maximizing model. This is exactly as would be expected, given that these stimulus-locked clusters were already assumed to reflect preparation and execution of the motor response. Interestingly, the feedback-locked analysis also reveals a subsequent shift into an inverse pattern of lateralized activity, over solely parietal regions. This is evident over the 200ms preceding feedback presentation, which suggests that it may constitute some form of anticipatory activity. Alternatively, it might reflect a physiological recovery effect, resulting from the immediately preceding levels of greatly elevated local activity. However, post-feedback analysis reveals the same clusters to

remain evident up to 500ms following feedback presentation, this persistence in the post-feedback period suggests that it may provide some form functional coding. This could possibly relate to the chosen value, which can be combined with feedback information to form a prediction error. Intriguingly, one other temporally independent cluster was revealed, of left lateralized activity at sensorimotor sensors. This positively related to choices of the right-hand option from as early as 500ms prior to the enaction of the choice response (1000ms pre-feedback).

With respect maximizing vs. non-maximizing activity, two clusters of activity were apparent. The first positively relates with maximizing choices, and is evident predominantly over centro-parietal regions (CP1, Pz, CPz, CP2) from as early as 400ms pre-choice, persisting until around 100ms post-choice. In the 100ms immediately preceding the choice response this cluster also extends to incorporate activity at a number of prefrontal electrodes. This is congruent with findings of the a priori analysis, which focused exclusively on prefrontal sensors. The second cluster negatively relates with maximizing choices, and comprises measured activity, running from 200ms pre-choice to 100ms post-choice across a range of highly left lateralized electrodes. Interestingly, when considering participants in the Stated probabilities group only, while the positively maximizing related centro-parietal and prefrontal cluster remains effectively the same, the left lateralized negatively related cluster is replaced by a right lateralized negative cluster. When considering the Learned group alone no significant clusters are evident, though this is likely down principally to the lower power of these within group analyses.

In the post-feedback period, a markedly different pattern of maximizing related activity emerges. A large central cluster is evident that negatively relates to maximizing; this encompasses electrodes over both right and left frontal, sensorimotor and to some extent parietal regions. This cluster begins to materialize from as early as 100ms post-

feedback presentation and remains, to some extent, all the way until 1000ms post-feedback. However, the broad central component shrinks dramatically from 600-700ms post-feedback and a prefrontal element emerges from 600-900ms post-feedback, this element will constitute the same maximizing related activity identified in the a priori prefrontal-focused analysis. A similarly temporally changeable cluster of positively maximizing related activity is also evident from 300ms to 1000ms post-feedback. This cluster emerges at both right and left lateralized parieto-occipital sensors, before extending to first left, then also right lateralized more anterior electrodes from 500ms and 600ms respectively. From 700ms post-feedback the original parieto-occipital activity disappears, leaving only the two now spatially unconnected clusters of positively maximizing related activity at the left-most (F7, FT7, T7) and right-most electrodes (FT8, C6, T8, CP6, TP8) respectively.

Positive vs. negative feedback related analysis was limited to the post-feedback period for obvious theoretical reasons. This analysis revealed an interesting temporally evolving pattern of feedback related activity. First, a large and puzzling cluster of negatively reward related activity was present over a broad range of central sensors from 0 to 100ms post-feedback, a period too early for this to credibly reflect genuine feedback related activation. The question of what exactly this cluster may represent remains intransigent, though it seems most likely to have emerged due to some extraneous factor. From 200ms to 300ms post-feedback, this broad grouping of fronto-central sensors briefly measures positively reward related activity, then from 300-400ms there is a fronto-central cluster of negatively reward related activity once again. This is now also coupled with a cluster of prefrontal positively reward related activity, which mirrors the first cluster identified in the earlier discussed prefrontal only analysis. From 400-500ms post-feedback this prefrontal cluster has disappeared, leaving only the central negatively reward related cluster, though this seems to have now shifted from fronto-central

sensors to predominantly parieto-central ones, though with a slight right-lateralized bias. Over time, from 500 to 900ms this negative cluster shifts to become more and more clearly right-lateralized, while a fronto-central group of now positively reward related activity becomes evident from 500-600ms, before expanding to encompass more posterior parieto-central regions as well, from 700ms all the way until 1000ms post-feedback. From 700-1000ms the right-lateralized negatively reward related cluster is joined by a similar contralateral cluster, on the left hand side of the scalp. In fact, from 700-900ms these three reward related clusters of activity together mirror the opposite pattern observed for maximizing related activity over the same time period. This activity, which relates to both maximizing and reward in such a clearly inverse manner, is a prime candidate for prediction error coding.

Descriptive ERP-based Results Summary

The ERP charts provide a more descriptive and intuitively understandable presentation of the same data upon which the mass regression and cluster-based analyses were conducted. This allows for a more clear and specific interpretation for a number of the findings that have already been substantiated by these prior inferential analyses.

All ERP charts were plotted for activity at pairs or groupings of midline electrodes. For all but one set of charts ERPs were grouped according to four conditions depending upon a combination of whether or not the participants had made a maximizing choice and whether or not their prediction turned out to actually be correct. Across all midline electrodes there appear to be two main task-related ERPs. The first of these occurs immediately around the choice response, the second following feedback presentation, peaking from around 200-400ms post-feedback. Theoretically informative findings will be described below across the different groupings of sensors according to their anterior-posterior position.

I. Prefrontal Electrodes FPz and AFz

A clear rise in activity is evident at the two prefrontal sensors peaking immediately before the choice-response (-550 to -500ms pre-feedback). This activity is higher, on average, when preceding a maximizing than a non-maximizing choice. Moreover, both prefrontal only and whole-scalp cluster analyses indicated that this between group difference was statistically significant. This activity is also higher, overall, within the Stated rather than Learned condition. A priori cluster analysis, which was limited to prefrontal regions only, did also indicate that this between group difference was significant, although this contrast was made on stimulus-locked data. When interpreting this difference it is also worth bearing in mind that analysis of the questionnaire data did not find higher self-reported strategy use in Stated than Learned groups, though a clearly greater number of maximizing choices were made in the former condition. Interestingly the ERP charts seem to indicate a Stated vs. Learned*Max vs. non-Max interaction effect, with the relative increase in prefrontal activity preceding a maximizing choice being approximately twice as large within the Stated than the Learned probabilities condition.

In the post-feedback period a large positive ERP is evident rising from around 200ms to a peak at 500-550ms post-feedback. This shows some signs of discriminating between rewarded and unrewarded trials in the early period, as it rises fluctuating between higher activity for Win trials from 100-150ms, higher activity for Lose trials from 200-270ms, then once again for Win trials from 300-400ms. The final of these three differences, between 300-400ms, is shown to be statistically significant from results of the cluster analysis, though this is not the case for the two preceding short inverse relationships evident from 100-270ms. Both at, and for a period following, the peak of this post-feedback ERP (500-900ms), prefrontal sensor activity is clearly higher following non-maximizing choices. Moreover, both prefrontal focused and whole scalp cluster

analyses reveal this difference to be significant. This post-feedback peak is also markedly higher overall in the Learned rather than Stated conditions.

II. Frontal Electrodes Fz and FCz

Next, the pattern of activity at fronto-central electrodes Fz and FCz will be discussed. Notably, it is sensors over this area at which feedback related negativity and prediction error coding has been most commonly previously reported. Here, instead of a simple increase to boundary, the ERP centred over the choice response is instead an N-shaped negative then positive inflection, this sharply changing potential follows an extended period of gradually decreasing activity, beginning over 800ms pre-response. The choice-response immediately follows the trough of the negative inflection. Here, the depth of the trough that immediately precedes the choice response is greater on maximizing than non-maximizing trials. This also corresponds to a higher trough-to-peak difference, as the subsequent peaks are roughly equal across both maximizing and non-maximizing choices. Interestingly, this apparent difference is seemingly driven entirely by participants within the Stated probabilities condition as, within the Learned condition, the ERP traces for maximizing and non-maximizing choices almost exactly overlay for the entire period from 1000ms pre-feedback until feedback onset. In any case, none of the apparent maximizing related differences at these electrodes were identified as statistically significant in the cluster analysis.

Further, in the post-feedback period a large and sharply defined positive ERP is evident at these electrodes. This likely represents P3a (Polich, 2007). Here, this potential displays two temporally distinct sub-components, the first peaking at around 200ms, and the second at around 400ms. These peaks are strongly reminiscent of the dual post-feedback components identified by (Philiastides et al., 2010), and show a number of theoretically interesting characteristics. When contrasting maximizing vs. non-

maximizing choices and rewarded vs. unrewarded trials independently, the pattern of results across these two components seem to further bear out the interpretations drawn by Philiastides et al. The 200ms fronto-central peak shows a clear discrepancy according to reward valence (in this task received vs. withheld rewards), while the later 400ms peak no longer shows any relationship with reward valence, but does show a relationship with prediction error magnitude.

However, from assessing relationships across the four conditions I propose that a different set-of relationships can better account for the patterns of activity observed here. First, while the early peak, at around 200ms, does code for prediction error valence, the present data quite strongly indicates that it also codes for prediction error magnitude. It does this through coding for prediction error in a largely linear manner, with large negative prediction error showing the lowest activity, and large positive prediction error showing the highest. This entails that averaging across positive and negative valence conditions will lead to the apparent loss of most of the effect of prediction error magnitude; high magnitude positive and negative prediction error activity average to a similar level as low magnitude positive and negative prediction errors do. However, prediction errors for the four conditions were as follows: Min-Win .7, Max-Win .3, Min-Lose -0.3, Max-Lose -0.7. This is also the precise order in which the ERP peaks of each of these conditions are found. This interpretation is also supported by results from the cluster analysis, which demonstrated that at 200ms, activity at these electrodes (along with a broader cluster of frontal electrodes to both the right and left of the midline), significantly relates both positively to reward and negatively to maximizing. Together, these are the defining characteristics of the prediction error. Intriguingly, during this early 200ms peak the 'Min-Win' condition seems to be somewhat of an outlier. This condition represents trials in which participants had successfully predicted the lower probability option, and the corresponding ERP seems disproportionately high. In fact, this ties in

theoretically well with a previously hypothesised disproportionate level of subjective utility for the successful prediction of the low probability option (Brackbill & Bravos, 1962; Brackbill et al., 1962; Rubinstein, 1959). Moreover, this difference appears to be driven principally by participants within the Stated probabilities condition, suggesting that this disproportionate utility may result in particular from the successful selection of an option that is explicitly known to be of lower probability.

During the short period between the peaks of the two post-feedback sub-components (from 300-350ms post-feedback) activity briefly drops for all four conditions, down to a similar trough for each except the Min-Win condition, which remains at a markedly higher level. At the later sub-peak (400ms post-feedback), the significance of the pattern of activation across the four conditions is less simple to interpret. The cluster analysis results indicate that while the negative relationship with maximizing at fronto-central sensors is maintained through this period, the concurrent relationship with reward at the same sensors switches briefly (from 300-400ms) from a positive one, indicative of prediction error coding, to a negative one. This suggests a qualitatively different aspect to activity at this time. Even more interestingly, the pattern of ERP activity also appears to differ at this point depending upon experimental condition. Within the Stated probabilities condition the order, from highest to lowest ERP peak is as follows: 'MinWin' > 'MinLose' > 'MaxLose' > 'MaxWin'. Within the Learned probabilities condition however, although the precise order is the same, the size of the difference between the 'MinWin', 'MinLose' and 'MaxLose' conditions is drastically reduced, particularly so in the case of the 'MinLose' and 'MaxLose' conditions so as to make them almost equivalent. However, all three remain substantially higher than the 'MaxWin' condition. The most convincing hypothesis for what this late component pattern represents seems to be salience of the outcome, in proportion to each outcome's overall frequency across the task. In this case non-maximizing choices

result in generally higher activity because they are less frequent than maximizing ones. Rewarded non-maximizing trials lead to a higher peak than unrewarded non-maximizing trials, as these are the less common choice-outcome combination, whereas the opposite it true for maximizing trials, for the same reason. Moreover, the compression of the difference between 'MinWin', 'MinLose' and 'MaxLose' trials, particularly between 'MinLose' and 'MaxLose' trials in the Learned condition is reflected in a similarly smaller difference in the number of trials between the four trial-types in this condition. To demonstrate, while ratio between Win and Lose trials will necessarily remain at approximately 2.33 between the less and more likely outcome, the ratio of 'MinLose' to 'MaxLose' trials drops from 3.98 in the Stated probabilities condition to only 1.36 in the Learned condition, mirroring the drop in differences between the ERP peaks of these choice-outcome combinations between the same experimental conditions.

Immediately following this large positive potential (450ms onwards) activity drops again in a sharp negative inflection, in which again the Min-Win condition remains markedly higher than the other three trial conditions. The whole scalp cluster analysis results reveal that fronto-central activity returns to a positive relationship with reward from 500ms, and then continues to significantly differentiate between Win and Lose trials in this manner up until the end of the time-window of interest at 1000ms post-feedback presentation. The same analysis also indicates that the negative relationship with maximizing choices at fronto-central electrode sites remains evident until around 900ms post-feedback.

III. Parietal Electrodes Cpz and Pz

At the parieto-central sensors CPz and Pz the pattern of pre-and post choice activation is broadly similar to that observed at the other midline electrode sites. In a manner similar to what is seen at prefrontal sensors Fz and FCz, a gradual and extended period (over 800ms) of decreasing activity is evident leading up to the choice response.

Here however, the patterns are clearly different between maximizing and non-maximizing choices, this pre-choice decrease in activity is substantially greater on non-maximizing than maximizing trials. Furthermore, whereas on maximizing trials activity shifts into a slow increase from around 250ms pre-response, to a peak at around 80ms post-response, on non-maximizing trials the decrease in activity continues until it reaches a trough immediately preceding enaction of the choice response (50ms pre-response). This is then followed by a steep upturn in activity until it reaches a similar level as is reached on maximizing trials at 80ms post-choice. Cluster analysis reveals that this divergence in activity between maximizing and non-maximizing trial ERPs is statistically significant at these sensors from 300ms before, and until 50ms after, the choice-response. Interestingly, the increase to bound activity seen at these sensors appears similar to that measured at prefrontal sites. However, as this apparent bound is reached here at around 80ms following enaction of the choice-response it logically cannot represent actual accumulation of activity necessary to make the decision itself. Here, it is instead the preceding decrease to trough activity that is contemporaneous with the increase to bound activity seen at prefrontal sites and predictive of the choice-response.

At these sensors the pattern of activity between conditions over the post-choice but pre-feedback period is also notable. Within the Stated condition activity seems markedly lower following non-maximizing choices, whereas in the Learned probabilities condition no such difference is evident. However, although this pattern is of high theoretical interest, potentially indicating anticipatory or chosen-value related activity, cluster analysis did not reveal these differences to be statistically significant. Interpretation is therefore highly speculative and to propose any firm conclusion would be inappropriate.

In the post-feedback period, the pattern of activity once again bears a resemblance to that recorded at fronto-central sites, with some key differences. Again

there is a sharp rise in activity building from around 150ms after feedback presentation to a peak soon after 300ms post-feedback. There is some sign of an earlier sub-peak at around 150-200ms post-feedback, at which time the cluster analysis indicates that activity is significantly higher for both non-maximizing and rewarded trials, characteristic of prediction error coding. At these parietal sensors however this earlier peak is severely depressed relative to the later component, at 300-350ms post-feedback. At this later peak, activity negatively relates with maximizing but not with feedback valence. The pattern across the four conditions is equivalent overall to that of the later fronto-central peak, in inverse proportion to the frequency of occurrence of each choice-feedback combination, although the discrepancies here between Stated and Learned probabilities conditions are not the same as at Fz and FCz. For the 150ms period immediately following this main peak, as activity drops back to a relative trough, ERPs both significantly negatively relate to maximizing and reward. A third sub-peak follows this at around 500ms post-feedback, from which time activity reverts to a positive relationship with reward and an apparent return to prediction error coding, a pattern which remains evident to some extent until the end of the time-window of interest at 1000ms post-feedback.

6.e. Discussion

In this experiment EEG techniques were applied to record electrophysiological activity while participants completed a standard probability matching task, in which reward contingencies were fixed and i.i.d. As in the previous chapter, both behaviour and electrophysiological activity were compared with two predictive models of expected value. The first of these was a simple maximizing model. The fixed trial-by-trial predictions of this model were wasteful on the SSFT but are optimal on the fixed probability schedule used in the present study. The second model is equivalent to the

local expected value model used on the previous experiment, and derives its expected value predictions in line with Faulty Foraging Theory, assuming the illusory accumulation of subjective reward value in the unchosen option. The trial-by-trial predictions of this model were optimal on the SSFT, but are wasteful on the fixed probability task used in the present study.

Behavioural results were generally consistent with previous studies. As before, group-level overmatching was found, while participants were substantially more likely to maximise having been explicitly told the reward probabilities before the task. Furthermore, strategy use once again proved a significant predictor of maximizing choices, although the present study did not replicate the relationship between stating probabilities and strategy use, or therefore the mediated effect of stated probabilities via strategy use upon maximizing choices. As hypothesized, an inverse pattern of significant effects was evident in relation to the Faulty Foraging based model, with higher proportions of FFT-model consistent choices being associated with the use of intuition and Learned reward probabilities. Despite this expected pattern of experimental effects, overall FFT-model derived expected values proved a poor predictor of participants' trial-by-trial choices. A negative relationship between FFT-model predictions and actual choices indicated a substantially better fit of the simple maximizing model to actual choice-data. Surprisingly, this was true even within the Learned reward probabilities condition, in which FFT-related choices were expected to be most common. Other behavioural findings included a replication of the lack of effect of apophenia on proportions of maximizing choices. Self-reported apophenia were also found to be unrelated to proportions of FFT-model consistent choices. WSLS-consistent choices were once again higher in the Learned probabilities condition and, although significantly higher than individual participant baselines, not high enough to suggest that this was a common approach to the task.

Stimulus-locked electrophysiological analysis primarily explored relationships between measured scalp activity and the predicted expected values of each model. This revealed multiple clusters of electrophysiological activity that strongly related to the simple fixed probability model, which were highly consistent with action-value coding of the actual reward probabilities. However, no activity was found that significantly related to the more complex FFT-based model. The core experimental hypothesis in relation to Faulty Foraging Theory was therefore not supported in the present study.

Further analyses did reveal a number of theoretically informative findings, several of which directly support other experimental hypotheses. First, a stimulus-locked contrast between Stated and Learned probability groups, which focused for a priori reasons on prefrontal electrodes only, indicated higher prefrontal activity in the Stated than the Learned probabilities condition in the immediately pre-choice period. A congruent response-locked difference in prefrontal activity was also evident, with significantly higher measured pre-choice activity at prefrontal electrodes being associated with maximizing over non-maximizing choices. Moreover, this relationship was reversed in the post-feedback period, with a cluster of prefrontal electrodes measuring significantly lower activity following maximizing than non-maximizing choices, from 500-900ms post-feedback. This finding was also expected, under the hypothesis that it may represent post-choice compensatory top-down activity for the following trial. However, a similar contrast between rewarded and unrewarded trials revealed a cluster of prefrontal activity from 300-370ms that positively related with successful predictions. This activity seems inconsistent with predictions, as Lose trials should require greater error correction.

Whole scalp response-locked analyses also indicated significantly higher parieto-central activity on maximizing than non-maximizing trials, from 300ms prior to until 80ms following enaction of the motor response. In the post-feedback period, a broad

and temporally extended central cluster of activity was evident that related negatively with maximizing choices, and which was paired with activity at an overlapping cluster of sensors that fluctuated between positive and negative relationships with reward. A series of ERP plots of activity at groups of midline electrodes positioned along the anterior-posterior axis were used to illustrate this activity from another perspective. These show a large positive potential with an onset from around 150ms post-feedback, and peaks ranging from 200-500ms post-feedback at several electrode sites on the anterior-posterior axis. Activity within this overall positive component seems at various locations and time-points points to code for reward prediction error. This is consistent with previous feedback related negativity studies, which have indicated a relative suppression in activity following negative feedback (Gehring & Willoughby, 2002; Nieuwenhuis et al., 2004), which may be greater for unexpected outcomes (Cohen et al., 2007; Holroyd et al., 2003; Yasuda et al., 2004), and may constitute coding of the prediction error overlaid onto the P3a (Chase et al., 2011; Foti et al., 2011) for review see Walsh and Anderson (2012).

Furthermore, at the fronto-central sensors Fz and FCz, this ERP comprised dual-components, as reported by (Philiastides et al., 2010). Consistent with the findings of this study, the earlier component, peaking at around 200ms, coded more for prediction error valence rather than magnitude, and the later component, peaking at around 400ms, coded more for prediction error magnitude than valence. However, further assessment of the ERP traces, combined with the findings of the cluster analysis led to the proposal of a different conclusion. The earlier peak seems not to exclusively code for prediction error valence, but to code for prediction error in a linear manner, from which information can be determined about both valance and magnitude. In this case high magnitude positive and high magnitude negative prediction errors are associated with the lowest and highest measured activity respectively. The later

component relates negatively to maximizing and positively to prediction error magnitude, but the most convincing account of the pattern of activation across choice-feedback combinations relates instead to the frequency of occurrence of each of these combinations. In this manner, this later component may represent some form of weighting to be attached to the immediately preceding value-update, with the purpose of compensating for the vastly different number of instances that information regarding each of the different choice-feedback combinations is encountered over the task.

To conclude, behavioural hypotheses were broadly supported. Electrophysiological results were more mixed. Multiple hypotheses were supported with respect to prefrontal activity, which was considered here as a proxy for top-down cognitive control and found to relate to maximizing choices in the pre-choice period, and non-maximizing choices in the post-choice period. However, the central theoretical hypothesis of the present study remains unsupported; results indicated no clusters of activity that were significantly modulated in accordance with the putative trial-by-trial subjective expected values derived from the Faulty Foraging model. A number of further interesting electrophysiological findings were also made, particularly in relation to the nature of post-feedback FRN and prediction error coding. Nevertheless, in the context of this thesis, the core findings of this study were found to further support the case for probability matching as an intuitive heuristic-driven response, but not to support the further hypothesis that this response is rooted in a non-normative overextension of an optimal foraging strategy.

Chapter 7: General Discussion

7.a. Summary of Key Findings

This thesis had two core objectives. The first of these was to ascertain the relative contributions of several existing accounts of probability matching in determining choice behaviour, particularly regarding the maximization versus diversification of choices. This aim formed the basis of the two experiments conducted in Chapter 2, which were designed to assess the putative effects of both apophenia and strategy use in a single study, through the combined application of experimental manipulation and questionnaire based self-report measures, in conjunction with mediation analysis. These two experiments also established the basic task paradigm; from which the experimental design of the remaining studies of the thesis would be adapted. This was an iterated two-alternative forced choice, or two-armed bandit task. In Chapter 2, both experiments applied what is termed the standard probability matching preparation. In this version reward contingencies are fixed and i.i.d. Here, the exclusive selection of the high probability option is the optimal response, but an abundance of previous literature (Vulkan, 2000) has indicated that a majority of participants will instead diversify choices, with many demonstrating a tendency to match their ratios of choice allocations to the ratio of reward probabilities.

Results of these first two studies indicated a reliable positive effect of explicitly stating reward probabilities upon subsequent proportions of maximizing choices. This was true both directly and, as hypothesised, in a mediated effect via engendering increased levels of strategy use, which was in turn also found to robustly increase proportions of maximizing choices. More surprisingly, despite attempting two distinct manipulations of apophenia: pattern information in Experiment 1 and aggregate feedback in Experiment 2, no effect was found of either upon self-reported levels of belief there was a pattern in the reward sequence or, in turn, proportions of maximizing

choices. Moreover, no direct effect was found of self-reported apophenia on choices within any experimental condition. This evidence complements the findings of Koehler and James (2009), who also failed to identify any relationship between pattern search and probability matching. The potential contributions of two further putative factors of probability matching were also assessed in these two experiments. These were that people might select the low probability option because they believe it is necessary to do so to achieve 100% success, and that they may do so due to a higher subjective utility being obtained from what may be perceived as the skilful prediction of the less likely option. However, when including these into a multiple regression model alongside the two primary factors discussed above, neither were found to provide any significant further contribution to choice behaviour.

In Chapter 3, an experiment was conducted to assess behaviour on a modified version of the standard task, which was designed to more closely model one core factor of natural foraging environments. More specifically, the persistence and accumulation of unharvested rewards over time was introduced as an additional factor to the reward schedule. Crucially, the introduction of this factor entails that the exclusive selection of the high resource generation rate option is no longer the optimal approach; probability matching will now provide maximum pay-offs. This was termed the simple simulated foraging task, or SSFT. Choice behaviour on this task, along with its relation to other key factors, was compared directly with equivalent measurements from the standard task preparation. Results were found to be in close accordance with multiple experimental hypotheses deriving from the predictions of Faulty Foraging Theory, which proposes that probability matching results from the instinctive overextension of a naturally optimal foraging strategy. This is discussed in more detail within the general introduction.

The second core objective of this thesis was to then explicitly assess the potential mechanisms underlying whichever factor was found to make the most substantive

contribution to choice behaviour, through the combined application of behavioural and electrophysiological methods. The evidence of the first two experiments, reported in Chapter 2, indicated that the predominant factor is the use of strategy over intuition, in accordance with the heuristic account of probability matching. Further, the findings reported in Chapter 3 were consistent with the predictions of Faulty Foraging Theory, as a potential underlying model of why probability matching may be the more intuitive and readily available response pattern.

The plan for the remaining studies of the thesis was therefore to bring to bear electrophysiological methods to further investigate these two findings. For the former of these, the approach taken was to examine relative levels of prefrontal activity, on the standard task, between both Stated and Learned probability conditions and between choices in which a maximizing vs. non-maximizing choice was made. The findings of an extensive body of pre-existing empirical literature suggest that this measure can be considered an appropriate proxy for top-down cognitive control, as discussed in Chapter 1. In order to further assess the validity of the Faulty Foraging model, as a putative underlying explanation for the intuitive dominance of choice diversification, it was planned to investigate potential internal representations of the increase in expected reward value over time within underexploited options. This proved a more complex proposition.

First it was necessary to establish the practical feasibility of discriminating such incremental increases in internal representations of expected value through electroencephalographic recording techniques. Chapter 4 documents a re-analysis of an existing but unpublished data set, which was collected the year prior to work beginning on this thesis. This analysis was able to successfully identify a number of sensors and time-points at which stimulus-locked activity was significantly modulated according to the actual lateralized difference in expected value. Moreover, further analysis indicated

this to be a linear monotonic relationship. This finding was consistent with previous studies that have observed action-value coding in motor effector regions (Wunderlich et al., 2009). Following this, Chapter 5 documents a study in which electroencephalographic activity was examined from participants completing the SSFT, in which unharvested rewards did genuinely persist and stochastically accrue over time. Here, stimulus-locked activity was once again identified at sensors and timepoints that were consistent with action-value coding. Scalp activity relating to the fixed resource generation probabilities of the two options was most prevalent. However, evidence was also found for activity that significantly related to local expected values, which were calculated on a trial-by-trial basis depending upon previous choice allocations and outcomes.

Chapter 6 describes the final experiment conducted. This experiment was designed to extend the findings of Chapter 5 from the SSFT to the standard probability matching preparation. On this task rewards do not veridically accumulate in the unchosen option, but a substantial proportion of participants have nevertheless been shown to diversify choices in a manner consistent with if they did. The ultimate aim here regarding Faulty Foraging Theory was to identify the illusory internal representation of imagined reward accrual over time, even on a task in which rewards were not objectively altered in any manner by previous choice allocations or outcomes. For this study, a stricter approach was applied to the EEG analysis, based upon non-parametric cluster-permutation testing. This provided a comprehensive correction for potential inflation of the Type 1 error rate that results from the multiple comparisons inherent to a preceding mass regression based stage of the analysis. Results revealed clusters of significantly greater stimulus-locked prefrontal activity in the Stated relative to the Learned probabilities condition. Moreover, feedback-locked analysis revealed clusters of reliably higher prefrontal activity immediately preceding maximizing rather than non-maximizing choices, paired with a significant inverse pattern in the post-feedback period. Together

these findings add coherent and complementary electrophysiological evidence to the already broad body of behavioural support for the heuristic account of probability matching. With respect to the model-based analysis, a clear pattern of stimulus-locked dipolar activity was evident. This was broadly consistent with the findings of the preceding two EEG studies and again congruent with action-value coding. A range of stimulus-locked activity was prevalent in relation to the simple fixed probability model, in which activity was compared to the genuine objective expected value difference. However, when considering the trial-by-trial subjective expected value predictions derived from the Faulty Foraging Theory based model, no significant clusters of activity were found.

To summarise the core elements and conclusions of the thesis, both behavioural an electrophysiological results were in consistent support of the view that probability maximizing relates to the exertion of top-down of cognitive control, to first identify that the maximizing strategy is optimal, and then to override what is a more readily available and otherwise prepotent tendency towards diversification of choices. However, attempts to address the question of precisely why probability matching appears to be the more intuitive and readily available choice response were less equivocal. This research evaluated one possible and theoretically appealing account of this in particular: Faulty Foraging Theory. This proposes that probability matching represents the overextension of an evolutionarily stable foraging strategy, which has become established over countless generations of natural selection. A comparative assessment of the patterns of behaviour and effects between a simple simulated foraging task and the standard probability matching preparation were generally congruent with the predictions of Faulty Foraging Theory. However, an electrophysiological analysis was unable to identify any clusters of activity that significantly related to the illusory accumulation of subjective

reward value within the unchosen option, despite uncovering copious evidence for internal representations of the actual fixed reward contingencies.

7.b. Potential Alternatives to Faulty Foraging Theory

It is important to bear in mind that a null result does not eliminate the Faulty Foraging model from consideration; an absence of evidence can have a variety of causes. Nonetheless, it would be inappropriate to claim that the findings of this thesis together constitute robust evidence in support of Faulty Foraging Theory. It is therefore apt to also consider alternative explanations for why choice diversification may be an intuitive response on the standard probability matching task, of which there are several.

First, the presence of competitors provides just as valid a driving factor in engendering matching as an optimal response as the potential for resource accumulation and depletion. In fact, while the possibility for overextension of assumed resource accumulation was modelled in the present thesis, the majority of previous studies, both in relation to both the IFD and its possible application to probability matching, have instead focused on competition as the primary basis (Schulze et al., 2015; Weber, 1998). Moreover, competition for resources has provided an equally long-standing and consistently important selection pressure as the standalone potential for resources to accumulate and deplete. This factor could therefore equally have become instantiated at a biological level through natural selection, leading to the instinctive overextension of matching behaviour to non-competitive environments when acting under intuition.

Alternatively, rather than probability matching being manifest as a default response at any genetic level, the tendency to diversify choices may be an over-learned behaviour, developing from common environmental principles that are encountered in everyday life, and which can establish heuristic responses. If for instance outcomes are not assumed to be fixed, a characteristic that is relatively uncommon in real-life

situations, then the periodic exploration of alternatives is an important behaviour for the purposes of maintaining an accurate and up to date representation of the reward value of different options. This is crucial in permitting the behavioural flexibility to more optimally track potential shifts in reward contingencies over time. A study conducted by Green, Benson, Kersten, and Schrater (2010) offers an overt empirical illustration of this point. This study revealed that diversification naturally results from model-based learning when there is uncertainty regarding the fixedness of reward schedules. In fact, the authors found that an optimal Bayesian algorithm converged upon probability matching as its output when it was initialized with an incorrect world-model, which was based upon the ecologically plausible assumptions that reward probabilities were neither i.i.d. nor coupled to those of their alternatives. The representativeness heuristic offers an additional example of how probability matching may result from the mistaken application of seemingly simple common sense logic. Irrespective of whether or not someone assumes that reward contingencies are fixed and i.i.d, they are likely to believe that the best response pattern will generally approximate salient characteristics of the actual outcome sequence.

In addition to these possibilities, another factor that may merit consideration is the potential element of satisfaction that may be gained from successfully predicting the low probability option in a manner that some may perceive as skilful. Previous proponents of this account were discussed in Chapter 1 (Brackbill & Bravos, 1962; Rubinstein, 1959) and this provides a competing explanation to heuristic accounts for why paying more generally leads to less matching (Brackbill et al., 1962; Shanks et al., 2002; Siegel & Goldstein, 1959). In fact, evidence reported in Chapter 2 suggests that within the two non-aggregate feedback conditions participants did, on average, endorse the view that greater satisfaction obtained from the successful prediction of the lower probability option was an adequate motivation to justify choice diversification.

Moreover, electrophysiological evidence reported in Chapter 6 suggested that a disproportionately higher subjective utility is obtained through successful prediction of the less likely option than can be accounted for by the prediction error alone. Nevertheless, returning to an inferential analysis conducted in Chapter 2, the tangible effect of endorsing this view was found to be relatively trivial, changing choice-behaviour on only a very few trials. On balance, this factor does not therefore seem well placed to explain individual-level probability matching. Still, it may engender some level of choice diversification across a range of participants. This might in turn contribute to probability matching that is apparent at the group-level.

One further potential basis for suboptimal diversification is that of momentary maximizing. As discussed in Chapter 1, this account does not strictly predict probability matching. It is primarily for this reason that the present thesis did not investigate this explanation in greater depth, but this is not necessarily a good reason to dismiss it. In fact, when considering behavioural results from the studies that were conducted, strict individual-level probability matching was not commonly evident; apart from on the SSFT, in which case matching behaviour was the optimal response and so would be expected to become widely adopted regardless. The empirical data from this thesis actually suggest that probability matching may not be a monolithic behaviour expressed within a large and distinct plurality of participants. Rather, it may be better characterised as a broadly group-level effect, which relies upon averaging over many participants, including some that severely undermatch reward probabilities and many that consistently overmatch. This presently observed distribution of behaviour is much more consistent with the local maximizing account than the more traditional view of probability matching as a ubiquitous individual-level response pattern.

In reality it is likely that choice diversification is driven by the combined effects of a range of factors. When reward contingencies are not presented before the task,

participants need time to learn through experience. Considering behaviour over short task durations can therefore cause an illusory inflation of the apparent degree of probability matching. Even when probabilities are explicitly stated from the outset some participants may not believe them, due to some expected level of deception in Psychology studies. Some participants may also naturally assume that probabilities vary over time, which may engender both exploration of alternative options and momentary maximizing, in accordance with apparent local reward probabilities. Stochastic fluctuations in these will naturally lead to some degree of choice diversification. Additionally, many participants clearly do mistakenly perceive an ostensibly predictable pattern in the reward sequence, although existing evidence advocating that this may substantially alter choices was not replicated. Further, some participants do seem to experience a higher subjective utility from selecting what they know to be the lower probability option, in particular when they are successful in obtaining a reward by doing so.

Irrespective of the contribution of other factors, the evidence of this thesis consistently indicates that the most reliable, influential and substantial cause of probability matching is reliance upon intuition. Precisely why this is the case is not so clear, despite the efforts documented here. The examination of Faulty Foraging Theory as one putative explanatory account was ultimately inconclusive. In light of this, the tendency to diversify may still be the result of long-standing selection pressure applied by natural foraging dynamics, such as resource accumulation and depletion, or competition for limited resources. Alternatively it may instead be an over-learned response, resulting from various possible environmental factors. Nevertheless, when considering both behavioural and electrophysiological findings of the present thesis together with existing empirical literature, it now seems beyond reasonable doubt that choice diversification is the most readily available and intuitive strategy. Probability maximizing therefore requires

the effortful application of top-down executive control. This is necessary first to determine the less readily available maximizing response, and then to deliberately overrule the otherwise prepotent tendency for diversification of choices.

7.c. Practical Implications and Potential Avenues of Future Research

This thesis documents predominantly 'blue-sky' academic research. The phenomenon of suboptimal probability matching is not frequently observed in natural environments precisely because of the scarcity of i.i.d. random ratio schedules in nature. Nevertheless, there are a number of examples where the behaviours and effects documented here might directly apply to specific modern real-world situations.

Perhaps the most clear-cut of these cases is that of gambling. This is particularly true for slot-machines, fixed odds betting terminals and virtual or online equivalents, such as the recent rise of computer games which, although otherwise unrelated to gambling, encourage the purchase of 'loot-boxes'. Each of these superficially varied forms of gambling takes advantage of the relatively addictive nature of variable or random ratio reward schedules to keep people engaged in the gambling process (cf. James, O'Malley & Tunney, 2016; Madden, Ewan & Lagorio, 2007).

However, although the random ratio reward schedule closely matches that of the standard probability matching preparation, it is important to remember that the present thesis only investigated mechanisms behind the suboptimal diversification of choices in 2AFC situations. Here the task parameters include the assumption that some prediction must be made. Although quite possibly pertinent in some manner to gambling tasks with similar reward schedules, the findings made here may not directly apply to a situation in which the only good decision is often not to play at all. For instance, one core finding of this thesis was that probability maximizing can be facilitated through the provision of reward probability information. Also, that this takes effect both directly and also

indirectly through engendering strategy use, in a manner analogous to greater engagement of System 2. When applied to the case of gambling this indicates that providing explicit information about the expected pay-off of each gamble might directly help people make more optimal decisions. This is already required by UK law for fixed-odds betting terminals, in the form of % RTP. Further, it also suggests that other methods of facilitating greater deliberation could enhance this effect. However, the question of whether these effects generalise to the optimal gambling response of not playing requires more specific research before any firm conclusion can be drawn.

In principle, the findings of this research could also be applied to investment decisions. Here there may be an inherent tendency to over-diversify investments, between options where some are known to be more likely to yield a higher or lower return. However, in this situation it is again difficult to apply findings from the present thesis directly, due to additional factors. One key issue here is the unpredictability of the actual reward probabilities, which are in this case not only not explicitly known, but also neither fixed nor i.i.d. Moreover, diversification of investments is explicitly acknowledged as a valuable strategy for reducing overall risk. While some investments may offer higher average returns, to put everything behind a single option, when you have the chance not only to gain but to lose, is in many cases not the most sensible choice.

These considerations do raise some interesting research questions. For instance, how might we expect people's choices, even in simplified situations such as the standard 2AFC probability matching task, to be affected through factors such as the potential to lose as well as gain money? Similarly, how might choices be affected by the possibility to diversify predictions not only across different trials, but also within trials, such that a certain allocation can be given to each outcome as a method of mitigating risk at the expense of overall expected value?

One further intriguing example of a real-world circumstance in which probability matching could be applied is that of moral, or charitable decisions. For instance, when choosing to allocate money between charities do people give exclusively to whichever they believe to be the most deserving? Or, alternatively, are donations diversified, and if so is this done in proportion to the perceived worthiness of the cause? The really interesting question here is that of what should be done in this case. In the standard probability matching task the aim is to maximize winnings, so a single choice strategy is objectively optimal. A utilitarian argument might be that the same applies in a charitable situation; if a given unit of money will do more good through one charity than the same amount through another then it could be reasoned that the best course of action is to give all available funds to the former.

However, there are examples that seem intuitively to counteract this idea. Consider two hypothetical charities, one provides food and shelter for vulnerable homeless individuals in a country with a lower cost of living, and the other does the same, but in a country with a slightly higher cost of living. Due to discrepancies in the cost of providing services between the two countries the former charity can provide a night's food and shelter for five individuals for the same cost as the second can for only four individuals. Here it seems optimal to allocate funds to the charity which can do most good, but it feels intuitively wrong to do so exclusively, ignoring the plight of all those in the second country simply due to where they live. These are difficult questions, and perhaps ones that cannot be satisfactorily answered even with research directly designed to address them. However, achieving some level of understanding of just how individuals perceive these situations, along with how charitable behaviour might be biased in relation to the potentially inherent tendency to diversify choices, seems a worthy and valuable goal.

References

Abrahams, M. V. (1989). Foraging guppies and the ideal free distribution - the influence of information on patch choice. *Ethology, 82*(2), 116-126. Retrieved from <Go to ISI>://WOS:A1989AA59100003

Amiez, C., Joseph, J. P., & Procyk, E. (2006). Reward encoding in the monkey anterior cingulate cortex. *Cerebral Cortex, 16*(7), 1040-1055. doi:10.1093/cercor/bhj046

Arkes, H. R., Dawes, R. M., & Christensen, C. (1986). Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes, 37*(1), 93-110. doi:10.1016/0749-5978(86)90046-4

Ashby, F. G. (1983). A biased random-walk model for 2 choice reaction-times. *Journal of Mathematical Psychology, 27*(3), 277-297. doi:10.1016/0022-2496(83)90011-1

Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & Cognition, 32*(8), 1369-1378. doi:10.3758/bf03206327

Basten, U., Biele, G., Heekeren, H. R., & Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences of the United States of America, 107*(50), 21767-21772. doi:10.1073/pnas.0908104107

Baum, W. M., & Kraft, J. R. (1998). Group choice: Competition, travel, and the ideal free distribution. *Journal of the Experimental Analysis of Behavior, 69*(3), 227-245. doi:10.1901/jeab.1998.69-227

Bautista, L. M., Alonso, J. C., & Alonso, J. A. (1995). A field-test of ideal free distribution in flock-feeding common cranes. *Journal of Animal Ecology, 64*(6), 747-757. doi:10.2307/5853

Bell, K. E., & Baum, W. M. (2002). Group foraging sensitivity to predictable and unpredictable changes in food distribution: Past experience or present circumstances? *Journal of the Experimental Analysis of Behavior, 78*(2), 179-194. doi:10.1901/jeab.2002.78-179

Bennett, D., Murawski, C., & Bode, S. (2015). Single-trial event-related potential correlates of belief updating. *eNeuro 2, :ENEURO.0076-15.2015. 10.1523/ENEURO.0076-15.2015*

Blanckenhorn, W. U., Morf, C., & Reuter, M. (2000). Are dung flies ideal-free distributed at their oviposition and mating site? *Behaviour, 137*, 233-248. doi:10.1163/156853900502051

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review, 113*(4), 700-765. doi:10.1037/0033-295x.113.4.700

Box, G. E. P. (1953). Non-normality and tests on variances. *Biometrika*, 40, 318-335.

Box, G. E. P. (1954a). Some theorems on quadratic forms applied in the study of analysis of variance problems: I. Effect of inequality of variances in the one-way classification. *Annals of Mathematical Statistics*, 25, 290-302.

Box, G. E. P. (1954b). Some theorems on quadratic forms applied in the study of analysis of variance problems: II. Effect of inequality of variances and of correlation of errors in the two-way classification. *Annals of Mathematical Statistics*, 25, 484-498.

Box, G. E. P., & Anderson, S. L. (1955). Permutation theory in the derivation of robust criteria and the study of departures from assumptions. *Journal of the Royal Statistical Society*, 17, 1-34.

Brackbill, Y., & Bravos, A. (1962). Utility of correctly predicting infrequent events. *Journal of Experimental Psychology, 64*(6), 648-&. doi:10.1037/h0046489

Brackbill, Y., Starr, R. H., & Kappy, M. S. (1962). Magnitude of reward and probability-learning. *Journal of Experimental Psychology, 63*(1), 32-&. doi:10.1037/h0045051

Bray, S., & O'Doherty, J. (2007). Neural coding of reward-prediction error signals during classical conditioning with attractive faces. *Journal of Neurophysiology, 97*(4), 3036-3045. doi:10.1152/jn.01211.2006

Brunia, C. H. M. (1999). Neural aspects of anticipatory behavior. *Acta Psychologica, 101*(2-3), 213-242. doi:10.1016/s0001-6918(99)00006-2

Brunswik, E. (1939). Probability as a determiner of rat behavior. *Journal of Experimental Psychology, 25*(2), 175-197. doi:10.1037/h0061204

Chase, H. W., Swainson, R., Durham, L., Benham, L., & Cools, R. (2011). Feedback-related Negativity Codes Prediction Error but Not Behavioral Adjustment during Probabilistic Reversal Learning. *Journal of Cognitive Neuroscience, 23*(4), 936-946. doi:10.1162/jocn.2010.21456

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. *Journal of Neuroscience, 29*(39), 12315-12320. doi:10.1523/jneurosci.2575-09.2009

Clinch, J. J., & Keselman, H. J. (1982). Parametric alternatives to the analysis of variance. *Journal of Educational Statistics*, 7, 207-214.

Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage, 35*(2), 968-978. doi:10.1016/j.neuroimage.2006.11.056

Conrad, K. (1958). *Die beginnende Schizophrenie. Versuch einer Gestaltanalyse des Wahns (in German)*. Stuttgart: Georg Thieme Verlag.

Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1993). A PET study of visuospatial attention. *Journal of Neuroscience, 13*(3), 1202-1226. Retrieved from <Go to ISI>://WOS:A1993KP71100029

Cotton, J. W., & Rechtschaffen, A. (1958). 2-choice and 3-choice verbal-conditioning phenomena - replication report. *Journal of Experimental Psychology, 56*(1), 96-96. doi:10.1037/h0038693

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. doi:10.1152/jn.01014.2002

Croxson, P. L., Walton, M. E., O'Reilly, J. X., Behrens, T. E. J., & Rushworth, M. F. S. (2009). Effort-Based Cost-Benefit Valuation and the Human Brain. *Journal of Neuroscience, 29*(14), 4531-4541. doi:10.1523/jneurosci.4515-08.2009

David, F. N., Johnson, N. L. (1951) the effect of non-normality on the power function of the f-test in the analysis of variance, *Biometrika*, 38, 43–57.

Davison, M., & McCarthy, D. (1988). *The matching law : a research review*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Derks, P. L. (1962). Generality of conditioning axiom in human binary prediction. *Journal of Experimental Psychology, 63*(6), 538-&. doi:10.1037/h0042796

Derks, P. L., & Paclisanu. (1967). Simple strategies in binary prediction by children and adults. *Journal of Experimental Psychology, 73*(2), 278-&. doi:10.1037/h0024137

Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Primate analogue of the Wisconsin Card Sorting Test: Effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behavioral Neuroscience, 110*(5), 872-886. doi:10.1037/0735-7044.110.5.872

Dreisig, H. (1995). Ideal free distributions of nectar foraging bumblebees. *Oikos, 72*(2), 161-172. doi:10.2307/3546218

Edwards, W. (1961). Probability-learning in 1000 trials. *Journal of Experimental Psychology, 62*(4), 385-&. doi:10.1037/h0041970

Elliott, R., & Dolan, R. J. (1998). Activation of different anterior cingulate foci in association with hypothesis testing and response selection. *Neuroimage, 8*(1), 17-29. doi:10.1006/nimg.1998.0344

Elliott, R., Rees, G., & Dolan, R. J. (1999). Ventromedial prefrontal cortex mediates guessing. *Neuropsychologia, 37*(4), 403-411. doi:10.1016/s0028-3932(98)00107-9

Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review, 104*(2), 301-318. doi:10.1037/0033-295x.104.2.301

Fantino, E., & Esfandiari, A. (2002). Probability matching: Encouraging optimal responding in humans. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 56*(1), 58-63. doi:10.1037/h0087385

Fiorina, M. P. (1971). Critique and comment - note on probability matching and rational choice. *Behavioral Science, 16*(2), 158-&. doi:10.1002/bs.3830160206

Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011). Event-Related Potential Activity in the Basal Ganglia Differentiates Rewards from Nonrewards: Temporospatial Principal Components Analysis and Source Localization of the Feedback Negativity. *Human Brain Mapping, 32*(12), 2207-2216. doi:10.1002/hbm.21182

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25-42. doi:10.1257/089533005775196732

Fretwell, D.S., Lucas, H.L., 1970. On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheoretica* 19, 16-32

Fretwell, S. D. 1972. Populations in a Seasonal Environment. Princeton, NJ: Princeton University Press.

Frey, S. H. (2008). Tool use, communicative gesture and cerebral asymmetries in the modern human brain. *Philosophical Transactions of the Royal Society B-Biological Sciences, 363*(1499), 1951-1957. doi:10.1098/rstb.2008.0008

Fuster, J. M. (2000). Executive frontal functions. *Experimental Brain Research, 133*(1), 66-70. doi:10.1007/s002210000401

Gallistel, C. R. (1990). *The organization of learning*. Cambridge (Mass): MIT Press.

Gaissmaier, W., & Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition, 109*(3), 416-422. doi:10.1016/j.cognition.2008.09.007

Gao, J., & Corter, J. E. (2015). Striving for perfection and falling short: The influence of goals on probability matching. *Memory & Cognition, 43*(5), 748-759. doi:10.3758/s13421-014-0500-4

Gardner, R. A. (1957). Probability-learning with 2 and 3 choices. *American Journal of Psychology, 70*(2), 174-185. doi:10.2307/1419319
Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science, 295*(5563), 2279-2282. doi:10.1126/science.1066893
Gigerenzer, G. (1996). Rationality: Why social context matters. In P. B. Baltes & U. M. Staudinger (Eds.), *Interactive minds: Life-span perspectives on the social foundation of cognition* (pp. 319-346). Cambridge, UK: Cambridge University Press.
Gigerenzer, G. (2000). *Adaptive Thinking: Rationality in the Real World*. New York: Oxford University Press.
Gigerenzer, G., & Fiedler, K. (2007). *Minds in environments: The potential of an ecological approach to cognition.*
Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball - on the misperception of random sequences. *Cognitive Psychology, 17*(3), 295-314. doi:10.1016/0010-0285(85)90010-6
Glascher, J., Hampton, A. N., & O'Doherty, J. P. (2009). Determining a Role for Ventromedial Prefrontal Cortex in Encoding Action-Based Value Signals During Reward-Related Decision Making. *Cerebral Cortex, 19*(2), 483-495. doi:10.1093/cercor/bhn098
Godin, J. G. J., & Keenleyside, M. H. A. (1984). Foraging on patchily distributed prey by a cichlid fish (teleostei, cichlidae) - a test of the ideal free distribution-theory. *Animal Behaviour, 32*(FEB), 120-131. doi:10.1016/s0003-3472(84)80330-9
Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences, 5*(1), 10-16. doi:10.1016/s1364-6613(00)01567-9
Goldstone, R. L., & Ashpole, B. C. (2004). Human foraging behavior in a virtual environment. *Psychonomic Bulletin & Review, 11*(3), 508-514. doi:10.3758/bf03196603
Goodnow, J. J. (1955). Determinants of choice-distribution in two-choice situations. *The American journal of psychology, 68*(1), 106-116. doi:10.2307/1418393
Goodnow, J. J., & Postman, L. (1955). Probability learning in a problem-solving situation. *Journal of Experimental Psychology, 49*(1), 16-22. doi:10.1037/h0047072
Grand, T. C. (1997). Foraging site selection by juvenile coho salmon: Ideal free distributions of unequal competitors. *Animal Behaviour, 53*, 185-196. doi:10.1006/anbe.1996.0287
Grand, T. C., & Grant, J. W. A. (1994). Spatial predictability of resources and the ideal free distribution in convict cichlids, cichlasoma-nigrofasciatum. *Animal Behaviour, 48*(4), 909-919. doi:10.1006/anbe.1994.1316
Grant, D. A., Hake, H. W., & Hornseth, J. P. (1951). Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *Journal of Experimental Psychology, 42*(1), 1-5. doi:10.1037/h0054051
Gray, R. D. (1994). Sparrows, matching and the ideal free distribution - can biological and psychological approaches be synthesized. *Animal Behaviour, 48*(2), 411-423. doi:10.1006/anbe.1994.1255
Green, C. S., Benson, C., Kersten, D., & Schrater, P. (2010). Alterations in choice behavior by manipulations of world model. *Proceedings of the National*

Academy of Sciences of the United States of America, 107(37), 16401-16406. doi:10.1073/pnas.1001709107
Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711-1725. doi:10.1111/j.1469-8986.2011.01273.x
Haaland, K. Y., Harrington, D. L., & Knight, R. T. (2000). Neural representations of skilled movement. *Brain, 123*, 2306-2313. doi:10.1093/brain/123.11.2306
Hair, J. F. J., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate Data Analysis (3rd ed)*. New York: Macmillan.
Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology, 42*(2), 161-170. doi:10.1111/j.1469-8986.2005.00278.x
Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology, 71*(2), 148-154. doi:10.1016/j.biopsycho.2005.04.001
Hake, H. W., & Hyman, R. (1953). Perception of the statistical structure of a random series of binary symbols. *Journal of Experimental Psychology, 45*(1), 64-74. doi:10.1037/h0060873
Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *Journal of Neuroscience, 28*(22), 5623-5630. doi:10.1523/jneurosci.1309-08.2008
Harper, D. G. C. (1982). Competitive foraging in mallards - ideal free ducks. *Animal Behaviour, 30*(MAY), 575-584. doi:10.1016/s0003-3472(82)80071-7
Hauser, T. U., Iannaccone, R., Stampfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *Neuroimage, 84*, 159-168. doi:10.1016/j.neuroimage.2013.08.028
Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White Paper]. Retrieved from http://www.afhayes.com/public/process2012.pdf.
Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis : a regression-based approach*.
Heekeren, H. R., Marrett, S., Bandettini, P. A., & Ungerleider, L. G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature, 431*(7010), 859-862. doi:10.1038/nature02966
Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience, 9*(6), 467-479. doi:10.1038/nrn2374
Heinrich, B. (1979). Majoring and minoring by foraging bumblebees, bombus-vagans - experimental-analysis. *Ecology, 60*(2), 245-255. doi:10.2307/1937652
Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4*(3), 267-272. doi:10.1901/jeab.1961.4-267
Herrnstein, R. J., & Loveland, D. H. (1975). Maximizing and matching on concurrent ratio schedules. *Journal of the Experimental Analysis of Behavior, 24*(1), 107-116. doi:10.1901/jeab.1975.24-107
Hill, T. & Lewicki, P. (2007). *STATISTICS: Methods and Applications.* StatSoft, Tulsa, OK.

Hinson, J. M., & Staddon, J. E. R. (1983). Matching, maximizing, and hill-climbing. *Journal of the Experimental Analysis of Behavior, 40*(3), 321-331. doi:10.1901/jeab.1983.40-321

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis. of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*(4), 679-709. doi:10.1037//0033-295x.109.4.679

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport, 14*(18), 2481-2484. doi:10.1097/01.wnr.0000099601/41403.a5

Horsnell, G. (1953). The effect of unequal group variances on the F-test for the homogeneity of group means. *Biometrika*, 40, 128–136.

Huettel, S. A., Mack, P. B., & McCarthy, G. (2002). Perceiving patterns in random series: dynamic processing of sequence in prefrontal cortex. *Nature Neuroscience, 5*(5), 485-490. doi:10.1038/nn841

Huk, A. C., & Shadlen, M. N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *Journal of Neuroscience, 25*(45), 10420-10436. doi:10.1523/jneurosci.4684-04.2005

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology, 387*(2), 167-178. doi:10.1002/(sici)1096-9861(19971020)387:2<167::aid-cne1>3.0.co;2-z

Hsu, P. L. (1938). Contributions to the theory of Student's t test as applied to the problem of two samples. *Statistical Research Memoirs*, 2, 1-24.

James, R. J. E., O'Malley, C., & Tunney, R. J. (2016). Why are some games more addictive than others: The effects of timing and payoff on perseverance in a slot machine game. *Frontiers in Psychology*. doi:10.3389/fpsyg.2016.00046.

Janssen, L., Meulenbroek, R. G. J., & Steenbergen, B. (2011). Behavioral evidence for left-hemisphere specialization of motor planning. *Experimental Brain Research, 209*(1), 65-72. doi:10.1007/s00221-010-2519-5

Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology, 41*(4), 291-297. doi:10.1037/h0056878

Jensen, J., Smith, A. J., Willeit, M., Crawley, A. P., Mikulis, D. J., Vitcu, I., & Kapur, S. (2007). Separate brain regions code for salience vs. valence during reward prediction in humans. *Human Brain Mapping, 28*(4), 294-302. doi:10.1002/hbm.20274

Jessup, R. K., Busemeyer, J. R., & Brown, J. W. (2010). Error Effects in Anterior Cingulate Cortex Reverse when Error Likelihood Is High. *Journal of Neuroscience, 30*(9), 3467-3472. doi:10.1523/jneurosci.4130-09.2010

Jones, M. H., & Liverant, S. (1960). Effects of age-differences on choice behavior. *Child Development, 31*(4), 673-680. doi:10.1111/j.1467-8624.1960.tb04982.x

Kahneman, D., & Frederick, (2002). In T.Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics of intuitive judgment*.New York:Cambridge University Press.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty : heuristics and biases*. Cambridge ; New York: Cambridge University Press.

Kahneman, D., & Tversky, A. (2000). *Choices, values, and frames*. New York Cambridge, UK: Russell sage Foundation ; Cambridge University Press.

Kelly, S. P., & O'Connell, R. G. (2013). Internal and External Influences on the Rate of Sensory Evidence Accumulation in the Human Brain. *Journal of Neuroscience, 33*(50), 19434-19441. doi:10.1523/jneurosci.3355-13.2013
Kennedy, P. (1992). *A Guide to Econometrics*. Oxford: Blackwell.
Kennerley, S. W., Dahmubed, A. F., Lara, A. H., & Wallis, J. D. (2009). Neurons in the Frontal Lobe Encode the Value of Multiple Decision Variables. *Journal of Cognitive Neuroscience, 21*(6), 1162-1178. doi:10.1162/jocn.2009.21100
Kim, J. N., & Shadlen, M. N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience, 2*(2), 176-185. Retrieved from <Go to ISI>://WOS:000078231600018
Kim, S., Hwang, J., & Lee, D. (2008). Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron, 59*(1), 161-172. doi:10.1016/j.neuron.2008.05.010
Kim, S. G., Ashe, J., Hendrich, K., Ellermann, J. M., Merkle, H., Ugurbil, K., & Georgopoulos, A. P. (1993). Functional magnetic-resonance-imaging of motor cortex - hemispheric-asymmetry and handedness. *Science, 261*(5121), 615-617. doi:10.1126/science.8342027
Koehler, D. J., & James, G. (2009). Probability matching in choice under uncertainty: Intuition versus deliberation. *Cognition, 113*(1), 123-127. doi:10.1016/j.cognition.2009.07.003
Koehler, D. J., & James, G. (2010). Probability matching and strategy availability. *Memory & Cognition, 38*(6), 667-676. doi:10.3758/mc.38.6.667
Kogler, C., & Kuehberger, A. (2007). Dual process theories: A key for understanding the diversification bias? *Journal of Risk and Uncertainty, 34*(2), 145-154. doi:10.1007/s11166-007-9008-7
Kohlmann, S. G., & Risenhoover, K. L. (1997). White-tailed deer in a patchy environment: A test of the ideal-free-distribution theory. *Journal of Mammalogy, 78*(4), 1261-1272. doi:10.2307/1383069
Kraft, J. R., & Baum, W. M. (2001). Group choice: The ideal free distribution of human social behavior. *Journal of the Experimental Analysis of Behavior, 76*(1), 21-42. doi:10.1901/jeab.2001.76-21
Kraft, J. R., Baum, W. M., & Burge, M. J. (2002). Group choice and individual choices: modeling human social behavior with the Ideal Free Distribution. *Behavioural Processes, 57*(2-3), 227-240. doi:10.1016/s0376-6357(02)00016-5
Lamb, A. E., & Ollason, J. G. (1993). Foraging wood-ants formica-aquilonia yarrow (hymenoptera, formicidae) tend to adopt the ideal free distribution. *Behavioural Processes, 28*(3), 189-198. doi:10.1016/0376-6357(93)90092-6
Lau, B., & Glimcher, P. W. (2007). Action and outcome encoding in the primate caudate nucleus. *Journal of Neuroscience, 27*(52), 14502-14514. doi:10.1523/jneurosci.3060-07.2007
Lin, A., Adolphs, R., & Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience, 7*(3), 274-281. doi:10.1093/scan/nsr006
Lindman, H. R. (1974). *Analysis of variance in complex experimental designs*. San Francisco: W. H. Freeman & Co.
Luu, P., Tucker, D. M., Derryberry, D., Reed, M., & Poulsen, C. (2003). Electrophysiological responses to errors and feedback in the process of action regulation. *Psychological Science, 14*(1), 47-53. doi:10.1111/1467-9280.01417

Madden, G. J., Ewan, E. E., and Lagorio, C. H. (2007). Toward an animal model of gambling: delay discounting and the allure of unpredictable outcomes. *J. Gambl. Stud.* 23, 63–83. doi: 10.1007/s10899-006-9041-5.

Madden, G. J., Peden, B. F., & Yamaguchi, T. (2002). Human group choice: Discrete-trial and free-operant tests of the ideal free distribution. *Journal of the Experimental Analysis of Behavior, 78*(1), 1-15. doi:10.1901/jeab.2002.78-1

Marquard, D. W. (1970). Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics, 12*(3), 591-&. doi:10.2307/1267205

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., & Bestmann, S. (2008). Trial-by-Trial Fluctuations in the Event-Related Electroencephalogram Reflect Dynamic Changes in the Degree of Surprise. *Journal of Neuroscience, 28*(47), 12539-12545. doi:10.1523/jneurosci.2925-08.2008

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience, 10*(5), 647-656. doi:10.1038/nn1890

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective. 2nd ed.* Mahwah, NJ: Lawrence Erlbaum associates.

Mesulam, M. M. (1999). Spatial attention and neglect: parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 354*(1387), 1325-1346. doi:10.1098/rstb.1999.0482

Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience, 1*(1), 59-65. doi:10.1038/35036228

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167-202. doi:10.1146/annurev.neuro.24.1.167

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*. 63 (2): 81–97. PMID 13310704. doi:10.1037/h0043158

Miller, M. B., Valsangkar-Smyth, M., Newman, S., Dumont, H., & Wolford, G. (2005). Brain activations associated with probability matching. *Neuropsychologia, 43*(11), 1598-1608. doi:10.1016/j.neuropsychologia.2005.01.021

Milner, B. (1963). Effects of different brain lesions on card sorting - role of frontal lobes. *Archives of Neurology, 9*(1), 90-&. Retrieved from <Go to ISI>://WOS:A19635701A00003

Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience, 9*(6), 788-798. doi:10.1162/jocn.1997.9.6.788

Myer, J. L.; & Well, A. D. (2003). *Research Design and Statistic Analysis*. 2nd ed. New Jersey : Lawrence Erlbaum.

Neimark, E. D. (1956). Effects of type of nonreinforcement and number of alternative responses in 2 verbal-conditioning situations. *Journal of Experimental Psychology, 52*(4), 209-220. doi:10.1037/h0047325

Neimark, E. D., & Shuford, E. H. (1959). Comparison of predictions and estimates in a probability-learning situation. *Journal of Experimental Psychology, 57*(5), 294-298. doi:10.1037/h0043064

Neter, J., Wasserman, W., & Kutner, M. H. (1989). *Applied Linear Regression Models*. Homewood, IL: Irwin.

Newell, B. R., Koehler, D. J., James, G., Rakow, T., & van Ravenzwaaij, D. (2013). Probability matching in risky choice: The interplay of feedback and strategy availability. *Memory & Cognition, 41*(3), 329-338. doi:10.3758/s13421-012-0268-3

Newell, B. R., & Rakow, T. (2007). The role of experience in decisions from description. *Psychonomic Bulletin & Review, 14*(6), 1133-1139. doi:10.3758/bf03193102

Nies, R. C. (1962). Effects of probable outcome information on 2-choice learning. *Journal of Experimental Psychology, 64*(5), 430-&. doi:10.1037/h0048225

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. H. (2004). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neuroscience and Biobehavioral Reviews, 28*(4), 441-448. doi:10.1016/j.neubiorev.2004.05.003

O'Connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nature Neuroscience, 15*(12), 1729-+. doi:10.1038/nn.3248

O'Doherty, J. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current Opinion in Neurobiology, 14*(6), 769-776. doi:10.1016/j.conb.2004.10.016

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. doi:10.1126/science.1094285

Olejniczak, P. (2006). Neurophysiologic basis of EEG. *Journal of Clinical Neurophysiology, 23*(3), 186-189. doi:10.1097/01.wnp.0000220079.61973.6c

Oliveira, F. T., Mcdonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *J. Cogn. Neurosci., 19*, 1994–2004. doi:10.1162/jocn.2007.19.12.1994

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*. doi:10.1155/2011/156869

Otto, A. R., Taylor, E. G., & Markman, A. B. (2011). There are at least two kinds of probability matching: Evidence from a secondary task. *Cognition, 118*(2), 274-279. doi:10.1016/j.cognition.2010.11.009

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature, 441*(7090), 223-226. doi:10.1038/nature04676

Pan, Y., & Jackson, R. T. (2008). Ethnic difference in the relationship between acute inflammation and serum ferritin in US adult males. *Epidemiology and Infection, 136*(3), 421-431. doi:10.1017/s095026880700831x

Passingham, R. E. (1993). *The frontal lobes and voluntary action.* Oxford: Oxford University Press.

Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8-13. doi:10.1016/j.jneumeth.2006.11.017

Peters, J., & Buchel, C. (2010). Neural representations of subjective reward value. *Behavioural Brain Research, 213*(2), 135-141. doi:10.1016/j.bbr.2010.04.031

Peterson, C. R., & Ulehla, Z. J. (1965). Sequential patterns and maximizing. *Journal of Experimental Psychology, 69*(1), 1-4. doi:10.1037/h0021597

Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology, 48*(5), 656-664. doi:10.1111/j.1469-8986.2010.01136.x

Philiastides, M. G., Biele, G., Vavatzanidis, N., Kazzer, P., & Heekeren, H. R. (2010). Temporal dynamics of prediction error processing during reward-based decision making. *Neuroimage, 53*(1), 221-232. doi:10.1016/j.neuroimage.2010.05.052

Polania, R., Krajbich, I., Grueschow, M., & Ruff, C. C. (2014). Neural Oscillations and Synchronization Differentially Support Evidence Accumulation in Perceptual and Value-Based Decision Making. *Neuron, 82*(3), 709-720. doi:10.1016/j.neuron.2014.03.014

Polich, J. (2007). Updating p300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10), 2128-2148. doi:10.1016/j.clinph.2007.04.019

Poling, A., Edwards, T. L., Weeden, M., & Foster, T. M. (2011). The matching law. *Psychological Record, 61*(2), 313-322. Retrieved from <Go to ISI>://WOS:000296105700010

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, 9*(7), 545-556. doi:10.1038/nrn2357

Ratcliff, R. (1978). Theory of memory retrieval. *Psychological Review, 85*(2), 59-108. doi:10.1037//0033-295x.85.2.59

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873-922. doi:10.1162/neco.2008.12-06-420

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9*(5), 347-356. doi:10.1111/1467-9280.00067

Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review, 106*(2), 261-300. doi:10.1037//0033-295x.106.2.261

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current theory and research*, 64-99.

Richmond, B. (1980). A New Look at an Old Friend. *Plexus (Resource Policy Center, Thayer School of Engineering, Dartmouth College, Hanover, NH)*.

Ringle, C. M., Wende, S., & Becker, J. (2015). *SmartPLS 3*. Bönningstedt: SmartPLS. Rogerson, P. A. (2001). London: Sage.

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience, 22*(21), 9475-9489. Retrieved from <Go to ISI>://WOS:000179031600037

Rubinstein, I. (1959). Some factors in probability matching. *Journal of Experimental Psychology, 57*(6), 413-416. doi:10.1037/h0038962

Ruchkin, D. S., Sutton, S., & Tueting, P. (1975). Emitted and evoked p300 potentials and variation in stimulus probability. *Psychophysiology, 12*(5), 591-595. doi:10.1111/j.1469-8986.1975.tb00052.x

Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E., & Behrens, T. E. (2011). Frontal Cortex and Reward-Guided Learning and Decision-Making. *Neuron, 70*(6), 1054-1069. doi:10.1016/j.neuron.2011.05.014

Samejima, K., Ueda, Y., Doya, K., & Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science, 310*(5752), 1337-1340. doi:10.1126/science.1115270

Schall, J. D. (2001). Neural basis of deciding, choosing and acting. *Nature Reviews Neuroscience, 2*(1), 33-42. doi:10.1038/35049054

Schluter, N. D., Rushworth, M. F. S., Passingham, R. E., & Mills, K. R. (1998). Temporary interference in human lateral premotor cortex suggests dominance for the selection of movements - A study using transcranial magnetic stimulation. *Brain, 121*, 785-799. doi:10.1093/brain/121.5.785

Schubotz, R. I., & von Cramon, D. Y. (2002). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fMRI study. *Neuroimage, 15*(4), 787-796. doi:10.1006/nimg.2001.1043

Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992). Neuronal-activity in monkey ventral striatum related to the expectation of reward. *Journal of Neuroscience, 12*(12), 4595-4610. Retrieved from <Go to ISI>://WOS:A1992KC84000003

Schulze, C., van Ravenzwaaij, D., & Newell, B. R. (2015). Of matchers and maximizers: How competition shapes choice under risk and uncertainty. *Cognitive Psychology, 78*, 78-98. doi:10.1016/j.cogpsych.2015.03.002

Schulze, C., van Ravenzwaaij, D., & Newell, B. R. (2017). Hold it! The influence of lingering rewards on choice diversification and persistence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. http://dx.doi.org/10.1037/xlm0000407

Serrien, D. J., Ivry, R. B., & Swinnen, S. P. (2006). Dynamics of hemispheric specialization and integration in the context of motor control. *Nature Reviews Neuroscience, 7*(2), 160-167. doi:10.1038/nrn1849

Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology, 86*(4), 1916-1936. Retrieved from <Go to ISI>://WOS:000171562000035

Shafir, E., & LeBoeuf, R. A. (2002). Rationality. *Annual Review of Psychology, 53*, 491-517. doi:10.1146/annurev.psych.53.100901.135213

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal-lobe damage in man. *Brain, 114*, 727-741. doi:10.1093/brain/114.2.727

Shanks, D. R., Tunney, R. J., & McCarthy, J. D. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making, 15*(3), 233-250. doi:10.1002/bdm.413

Siegel, S., & Goldstein, D. A. (1959). Decision-making behavior in a 2-choice uncertain outcome situation. *Journal of Experimental Psychology, 57*(1), 37-42. doi:10.1037/h0045959

Skyrms, B. (1996) *The evolution of the social contract*. Cambridge University Press. [aKES]

Skyrms, B. (1997). Game Theory, Rationality and Evolution, in M. L. Dalla Chiara *et al*. (eds.), *Structures and Norms in Science*, Dordrecht: Kluwer Academic Publishers, pp. 73–85.

Smith, P. L. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology, 44*(3), 408-463. doi:10.1006/jmps.1999.1260

So, N., & Stuphorn, V. (2010). SEF encodes and monitors action value in a gambling task. *Society for Neuroscience Abstract Viewer and Itinerary Planner, 40*. Retrieved from <Go to ISI>://BCI:BCI201100533404

Sohn, J. W., & Lee, D. (2007). Order-dependent modulation of directional signals in the supplementary and presupplementary motor areas. *Journal of Neuroscience, 27*(50), 13655-13666. doi:10.1523/jneurosci.2982-07.2007

Sokolowski, M. B. C., Tonneau, F., & Baque, E. F. I. (1999). The ideal free distribution in humans: An experimental test. *Psychonomic Bulletin & Review, 6*(1), 157-161. doi:10.3758/bf03210824

Stanovich, K. E. (1999). *Who is rational? : studies of individual differences in reasoning*. Mahwah, N.J.: L. Erlbaum Associates.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*(5), 645-+. doi:10.1017/s0140525x00003435

Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science, 304*(5678), 1782-1787. doi:10.1126/science.1094765

Summerfield, C., & Tsetsos, K. (2012). Building bridges between perceptual and economic decision-making: neural and computational mechanism. *Frontiers in Neuroscience, 6*. doi:10.3389/fnins.2012.00070

Sur, S., & Sinha, V. K. (2009). Event-related potential: An overview. *Industrial Psychiatry Journal, 18*, 70-73.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science, 150*(3700), 1187-&. doi:10.1126/science.150.3700.1187

Sutton, S., Tueting, P., Zubin, J., & John, E. R. (1967). Information delivery and sensory evoked potential. *Science, 155*(3768), 1436-&. doi:10.1126/science.155.3768.1436

Tan, L., Sosa, F., Talbot, E., Berg, D., Eversz, D., & Hackenberg, T. D. (2014). Effects of predictability and competition on group and individual choice in a free-ranging foraging environment. *Journal of the Experimental Analysis of Behavior, 101*(2), 288-302. doi:10.1002/jeab.76

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition Without Control: When a Little Frontal Lobe Goes a Long Way. *Current Directions in Psychological Science, 18*(5), 259-263. Retrieved from <Go to ISI>://WOS:000271465400002

Tomarken, A. J., & Serlin, R. C. (1986). Comparison of ANOVA alternatives under variance heterogeneity and specific noncentrality structures. *Psychological Bulletin*, 99(1), 90-99.

Tregenza, T., & Thompson, D. J. (1998). Unequal competitor ideal free distribution in fish? *Evolutionary Ecology, 12*(6), 655-666. doi:10.1023/a:1006529431044

Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature, 398*(6729), 704-708. Retrieved from <Go to ISI>://WOS:000079920100050

Tueting, P., Sutton, S., & Zubin, J. (1970). Quantitative evoked potential correlates of probability of events. *Psychophysiology, 7*(3), 385-&. doi:10.1111/j.1469-8986.1970.tb01763.x

Unturbe, J., & Corominas, J. (2007). Probability matching involves rule-generating ability: A neuropsychological mechanism dealing with probabilities. *Neuropsychology, 21*(5), 621-630. doi:10.1037/0894-4105.21.5.621

Urbano, A., Babiloni, C., Onorati, P., Carducci, F., Ambrosini, A., Fattorini, L., & Babiloni, F. (1998). Responses of human primary sensorimotor and supplementary motor areas to internally triggered unilateral and simultaneous bilateral one-digit movements. A high-resolution EEG study. *European Journal of Neuroscience, 10*(2), 765-770. doi:10.1046/j.1460-9568.1998.00072.x

Valentin, V. V., & O'Doherty, J. P. (2009). Overlapping Prediction Errors in Dorsal Striatum During Instrumental Learning With Juice and Money Reward in the Human Brain. *Journal of Neurophysiology, 102*(6), 3384-3391. doi:10.1152/jn.91195.2008

Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2003). Predicting events of varying probability: uncertainty investigated by fMRI. *Neuroimage, 19*(2), 271-280. doi:10.1016/s1053-8119(03)00122-8

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys, 14*(1), 101-118. doi:10.1111/1467-6419.00106

Wahlstrom, L. K., & Kjellander, F. (1995). Ideal free distribution and natal dispersal in female roe deer. *Oecologia, 103*(3), 302-308. Retrieved from <Go to ISI>://WOS:A1995RR94800006

Wallis, J. D., & Miller, E. K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *European Journal of Neuroscience, 18*(7), 2069-2081. doi:10.1046/j.1460-9568.2003.02922.x

Walsh, M. M., & Anderson, J. R. (2011). Modulation of the feedback-related negativity by instruction and experience. *Proceedings of the National Academy of Sciences of the United States of America, 108*(47), 19048-19053. doi:10.1073/pnas.1117189108

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews, 36*(8), 1870-1884. doi:10.1016/j.neubiorev.2012.05.008

Weber, T. P. (1998). News from the realm of the ideal free distribution. *Trends in Ecology & Evolution, 13*(3), 89-90. doi:10.1016/s0169-5347(97)01280-9

Wenzlaff, H., Bauer, M., Maess, B., & Heekeren, H. R. (2011). Neural Characterization of the Speed-Accuracy Tradeoff in a Perceptual Decision-Making Task. *Journal of Neuroscience, 31*(4), 1254-1266. doi:10.1523/jneurosci.4000-10.2011

West, R. F., & Stanovich, K. E. (2003). Is probability matching smart? Associations between probabilistic choices and cognitive ability. *Memory & Cognition, 31*(2), 243-251. doi:10.3758/bf03194383

Wise, S. P., Murray, E. A., & Gerfen, C. R. (1996). The frontal cortex basal ganglia system in primates. *Critical Reviews in Neurobiology, 10*(3-4), 317-356. Retrieved from <Go to ISI>://WOS:A1996VZ23900003

Wilcox R. R. (1987). New designs in analysis of variance. *Annu. Rev. Psychol*. 38, 29–6010.1146/annurev.ps.38.020187.000333

Wolford, G., Miller, M. B., & Gazzaniga, M. (2000). The left hemisphere's role in hypothesis formation. *Journal of Neuroscience, 20*(6), art. no.-RC64. Retrieved from <Go to ISI>://WOS:000085724200002

Wolford, G., Newman, S. E., Miller, M. B., & Wig, G. S. (2004). Searching for patterns in random sequences. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 58*(4), 221-228. doi:10.1037/h0087446

Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2009). Neural computations underlying action-based decision making in the human brain. *Proceedings of the National Academy of Sciences of the United States of America, 106*(40), 17199-17204. doi:10.1073/pnas.0901077106

Yasuda, A., Sato, A., Miyawaki, K., Kumano, H., & Kuboki, T. (2004). Error-related negativity reflects detection of negative reward prediction error. *Neuroreport, 15*(16), 2561-2565. doi:10.1097/00001756-200411150-00027

Yellott, J. I. (1969). Probability learning with noncontingent success. *Journal of Mathematical Psychology, 6*(3), 541-&. doi:10.1016/0022-2496(69)90023-6

Yu, A. J., and Huang, H. (2014). Maximizing masquerading as matching in human visual search choice behavior. *Decision* 1, 275. doi: 10.1037/dec0000013

Appendices

i. Appendix A: Wording of Questionnaire Items (with the yellow light as the high reward probability option)

Please estimate on what proportion of trials (%) the yellow and blue lights switched on over the entire task.

Yellow Light: %

Blue Light: %

Overall, to what extent do you feel that you used your intuition to make your choices on the task, as opposed to any explicitly held plan or strategy?

(1 = pure intuition, 5 = pure strategy)

1 2 3 4 5

- Please answer whether you agree or disagree with the following statements (1 = strongly disagree, 2 = slightly disagree, 3 = unsure, 4 = slightly agree, 5 = strongly agree):

There was a pattern in the sequence.

1 2 3 4 5

To have any chance of getting 100% correct, you must select the blue light some of the time.

1 2 3 4 5

It is sometimes worth deliberately picking the light that is less likely to switch on just because it is more satisfying to win after doing so.

1 2 3 4 5

ii. Appendix B: Task Instructions. Slide 1

Welcome to the experiment. You will be presented with a series of choices. On each trial, a yellow and a blue light bulb will be shown to either side of the screen. Later in the trial, one of these will switch on. Your task is to predict the bulb that will subsequently switch on in each trial. For each correct choice, 2p will be added to your total winnings. You can choose between these lights by pressing the 'q' key for the option shown to the left of the screen, or the 'p' key for the option shown to the right of the screen. The side to which each colour bulb is presented will vary randomly between trials. The aim is to win as much money as possible throughout the task, by predicting the correct light colour on as many trials as possible. Once you have read and understood the task instructions, press the space bar to continue. If you have any questions, please ask the experimenter.

iii. Appendix C: Task Instructions. Additional Reward Probability Information (only shown in Stated probabilities group)

Further information: On each trial, the probability of the blue bulb switching on will be 0.7, while the probability of the yellow bulb switching on will be 0.3. This means that the blue bulb can be expected, on average, to switch on in 70% of trials, and the yellow bulb, on average, to switch on in 30% of trials. Once you have read and understood this information, press the space bar to continue. If you have any questions, please ask the experimenter.

iv. Appendix D: Task Instructions. Additional Pattern Information (only shown in 'Instruction' Apophenia manipulation)

Further information: One colour bulb may turn on more often than the other one. Other than this, the bulbs will turn on at random. There will be no pattern or other way to accurately predict which of the two lights will turn on in each trial. Once you have read and understood this information, press the space bar to continue. If you have any questions, please ask the experimenter.

v. Appendix E: Statistical Considerations Regarding Potential Violations of Assumptions of ANOVA

Across these analyses a number of issues were raised regarding assumptions of ANOVA. Mauchly's tests commonly revealed significant violations of the assumption of sphericity. This is likely due to high early variance in choice allocations, partially driven by uncertainty in estimates of reward probabilities, converging towards the maximising response as the task progressed. However, this violation is not a serious concern for the within-subjects contrasts, as these derive from differences between sums of weighted-means with only a single degree of freedom. For within subject effects, degrees of

freedom were adjusted according to Greenhouse-Geisser correction. Kolmogorov-Smirnov tests revealed a non-normal distribution across a number of blocks. This most likely reflects a ceiling effect on maximising responses in the stated probabilities group. However, with the large sample size of 30 participants in each group central limit theorem should render the deviation from normality irrelevant, as the sampling distribution of the mean will approximate normality irrespective of the distribution of the variable in the population (Hill & Lewicki, 2006). Levene's tests for equality of error variances were also found to be significant for several blocks, with heterogeneity of variance up to a maximum ratio of approximately 1:7. This is again most likely due to a ceiling effect in the stated probabilities condition leading to a much lower variance. However, ANOVA is relatively robust to such violations, especially under the conditions of large and equal group sizes (for discussion, see Box, 1953; Box, 1954a; Box, 1954b; Box & Anderson 1955; Clinch & Keselman, 1982; David & Johnson, 1951; Horsnell, 1953; Hsu, 1938; Lindman, 1974, p.33; Myers & Well, 2003; Tomarken & Serlin, 1986). Moreover, Wilcox (1987) reviewed several studies *'showing that in a four-group case with 12 observations in each group when the variances are in the ratio of 1:1:1:16, the probability of Type 1 error is .101 instead of the nominal value of .05. When sample sizes are larger, the effect of unequal variances is reduced.'* (Maxwell and Delaney, 2004). The much greater sample size and significantly lower ratio of smallest to largest variance of the present case will mean that the effect on Type 1 error rate is substantially lower. In practice, any inflation of Type 1 error rates due to these issues should have no bearing on the effective significance of the between-subjects main effects or interactions observed at $p<.01$.