Computational Dissection of Schizotypy:

Differential Contingency Weighting

Santiago Castiello de Obeso

St Anne's College

Department of Experimental Psychology

University of Oxford



Dissertation submitted for the degree of

Doctor of Philosophy (D.Phil.)

Michaelmas Term, 2022

Abstract

Schizotypy is a personality dimension related to psychotic experience, *i.e.*, alteration of beliefs and sensory experiences, also is a liability for schizophrenia. In this dissertation I focus on schizotypy and its relationship with perceptual, learning, and agency. I employed computational models, and Contingency Learning Theory as a general framework, to develop novel behavioural tasks used to dissect cognitive features of schizotypy measured in the general population.

I studied the Jumping-to-Conclusions bias with a novel perceptual contingency task (Experiment 1). This task also tests the tendency to perceive associations between events. I found that people scoring high on schizotypy does not perceive associations easily, but they gather more sensory evidence to make correct decisions. In addition, I investigated whether schizotypy moderates how people use absent information to make judgments (Experiments 2 and 3). In an exploratory analysis, I found that high unusual experiences scores were negatively correlated with how people use the absence of events to judge event relatedness. Moreover, I developed a framework −behavioural paradigm and computational model− to study how two agents (one human and one artificial) interact and make judgments of control (Experiments 4 and 5). I used this task to study illusion of control and how people weight external and internal events as sources of control. I found that people scoring high in positive schizotypy have an enhanced illusion of control when estimating their control. Although, this result is dependent on participants not 'interacting' with another agent. I also found that unusual experiences in schizotypy negatively correlate with sensitivity to external controllers (or agents).

In summary, I found that weighting sensory evidence varies in relation to schizotypy. This was supported with my results showing a bias to gather more evidence, a weight reduction for absent information, an enhanced illusion of control when there is nothing else competing for the control, and a reduced sensitivity to perceived control from another agent. Together, my results indicate that high positive schizotypy is related to downweighting sensory evidence with a low saliency baseline (absence of target information, and other agents). Finally, I explored the differences in the concept of weight in computational modelling and I conclude that mathematically different models refer to a similar mechanism, a value multiplying a discrepancy learning signal.

Contents

Abstract ii

Contents iii

Figures viii

Tables ix

Acknowledgments x

Declaration xi

Abbreviations and Glossary xii

Overview 1

Chapter One Dissection of Schizotypy and Contingency Learning Theory 3

1.1. Schizotypy, Psychosis, Cluster A Personality Disorder, and Schizophrenia 3

1.1.1. Psychosis: Disorders, Symptoms, and Experiences 3

1.1.1.1. Conceptual Distinctions of Schizotypy 5

1.1.1.2. Schizotypy 6

1.1.2. Perception, Learning, Agency, and Schizotypy 7

1.1.2.1. Perception 9

1.1.2.2. Associative Learning 10

1.1.2.3. Instrumental Learning 17

1.1.2.4. Sense of Control and Agency 20

1.2. Contingency Learning Theory 21

1.2.1. When Pavlov ‘met’ Skinner: It is about the contingency! 22

1.2.2. A Formal Theory 23

1.2.2.1. The General Rule 23

1.2.2.2. Contingency Computations 25

1.3. Hypotheses and Summary 26

Chapter Two Perceptual Contingency Learning in Schizotypy 28

2.1. Introduction 28

2.1.1. Hyperassociation and Schizotypy 28

2.1.2. The Psychophysics of Contingency 30

2.2. Methods 32

2.2.1. Participants 32

2.2.2. Design and Procedure 32

2.2.3. Data analysis plan 34

2.2.3.1. Model Strong Positive Responses 35

2.2.3.2. Model Decision Time 35

2.2.3.3. Psychophysical and Computational Modelling 35

2.2.3.4. Schizotypy Subscales and Behavioural Parameters 37

2.2.3.5. Information Type Weights Analysis and Recency Analysis 38
2.2.3.6. Code and open access 38
2.3. Results 39
2.3.1. Differences in Samples 39
2.3.2. Confirmatory Mixed Models 39
2.3.2.1. Modelling Responses 39
2.3.2.2. Modelling Decision Time 40
2.3.2.3. Confirmatory Analysis: Summary 42
2.3.3. Drift Diffusion Model and Perception of Contingency 43
2.3.3.1. DDM: Summary 44
2.3.4. Exploratory: Behaviour and Questionnaires 45
2.3.4.1. Exploratory Analysis: Summary 46
2.3.5. Weighting Differences and Information Type 47
2.4. Discussion 48
2.4.1. Contingency and Psychophysics 49
2.4.2. Contrast with previous literature 50
2.4.3. Limitations 50
2.4.4. Conclusions 50
2.5. Appendix 2 52
2.5.1. Included versus excluded samples 52
2.5.2. Subsamples differences 52
2.5.3. Task instructions 54
2.5.4. Trait Specificity: Schizotypy subscales and Depression 55
2.5.5. Sensitivity Analysis: Generalized Mixed Model (with Gamma link function) 56
Chapter Three Sensitivity to Mutual Absent Cues in Schizotypy 58
3.1. Introduction 58
3.1.1. Contingency Judgments, Conditional Reasoning, and Absent Information 59
3.1.2. Disconfirmatory and Absent Evidence in Schizotypy 60
3.1.3. Contingency Learning Theory: Frequency vs Duration 61
3.2. Experiment 2 64
3.2.1. Methods 65
3.2.1.1. Participants 65
3.2.1.2. Psychometric measurements 65
3.2.1.3. Design 65
3.2.1.4. Procedure 67
3.2.1.5. Analysis plan 68
3.2.2. Results & Discussion 69

3.2.2.1. Full model 69
3.2.2.2. Model comparison 70
3.3. Experiment 3 72
3.3.1. Methods 72
3.3.1.1. Participants 72
3.3.1.2. Psychometric measurements 72
3.3.1.3. Design 72
3.3.1.4. Procedure 75
3.3.1.5. Analysis plan 75
3.3.2. Results & Discussion 75
3.3.2.1. Full model 76
3.3.2.2. Model comparison 77
3.4. Pooled Experiments: Exploratory Analysis 77
3.5. Discussion 79
3.5.1. Clinical relevance 80
3.6. Appendix 3 83
3.6.1. Presence and Absence Information Symmetry 83
3.6.2. The Task 84
3.6.2.1. Screen: instructions 84
3.6.2.2. Screen: rating 84
3.6.2.3. Screen: debriefing 85
3.6.2.4. Visual references 86
3.6.3. Sensitivity Analysis Experiment 3 87
Chapter Four The Agency Paradox: Illusion of Control and Schizotypy 88
4.1. Introduction 88
4.1.1. Behavioural Control and Contingency Learning Theory 88
4.1.2. Learning Control, Sense of Agency, and Schizotypy 89
4.1.3. Contingency Judgments with Two Agents 91
4.2. Experiment 4 93
4.2.1. Methods 93
4.2.1.1. Participants 93
4.2.1.2. Questionnaires 94
4.2.1.3. Design 94
4.2.1.4. Analysis 97
4.2.1.5. Software used and open access declaration 100
4.2.2. Results 101
4.2.2.1. Time series 101

4.2.2.2. Delta P and Ratings 102
4.2.2.3. Correlations: Behaviour, Models, and Questionnaires 104
4.2.2.4. Moderation analysis: Positive Schizotypy, Responses and Perceived Control 105
4.2.3. Discussion and Summary 106
4.3. Experiment 5 108
4.3.1. Methods 108
4.3.1.1. Participants 108
4.3.1.2. Questionnaires 109
4.3.1.3. Design 109
4.3.1.4. Analysis 109
4.3.1.5. Software used and open access declaration 110
4.3.2. Results and discussion 110
4.3.2.1. Time series 110
4.3.2.2. Delta P and Ratings 111
4.3.2.3. Correlations: Behaviour, Models, and Questionnaires 114
4.3.2.4. Moderation analysis: Schizotypy, Responses and Perceived Control 117
4.3.3. Discussion and Summary 117
4.4. Discussion 117
4.4.1. Illusion of Control, Unusual Experiences, and Referentiality? 118
4.4.2. Agency Paradox as a Contrast Effect 119
4.4.3. Behavioural Control Theory for two agents: simulations 120
4.4.4. Closing remarks and Limitations 123
4.4.5. Conclusions 124
4.5. Appendix 4 125
4.5.1. Information Theory and Contingency Learning Theory 125
4.5.2. Samples Characteristic 127
4.5.3. Sensitivity Analysis of Moderation analyses effects 128
4.5.4. Participants instructions 128
Experiment 4 (E4) 128
Experiment 5 (E5) 130
Chapter Five Contingency Weighting in Schizotypy: General Discussion 132
5.1. What do Computational Models mean by 'Weighting'? 134
5.1.1 Differential Weighting of Sensory Evidence 135
5.2. Overview of Computational Models 136
5.2.1. Learning rule (or perceptual model) 137
5.2.2. Decision Rule (or response model) 138
5.3. Two Families of Computational Models 138

5.3.1. Bayesian Learning (BL) 138
5.3.2. Reinforcement and Associative Learning (RAL) 140
5.3.2.1. Fixed weights 141
5.3.2.2. Variable weights 143
5.3.3. The intersection of Bayesian and Associative theories 144
5.4. Summary of Learning in SZ and Schizotypy 145
5.4.1. Clinical Population 146
5.4.2. Subclinical Population 147
5.4.3. Summary: Clinical and Subclinical Schizotypy and Learning 148
5.5. Schizotypy and Weighting 148
5.5.1. Interpretation of my Findings 148
5.5.2. Jumping-to-Conclusion and Low Saliency Types 150
5.5.3. Consistency with the Two-Stage Cognitive Theory 151
5.5.4. Summary: What Schizotypy Weights? 152
5.6. Limitations 153
5.7. Future directions 154
5.8. Conclusions 154
5.9. Appendix 5 156
Table 5.A1. Computational Models (Cognitive, Bayesian, Associative and Reinforcement) 157
Table 5.A2. Schizotypy and Characteristics of Learning Studies 161
6. **References** 172

Figures

Figure 2.1.	Paradigm used in Experiment 1	34
Figure 2.2.	Main Results of Experiment 1	42
Figure 2.3.	Drift-Diffusion Mode (DDM) l for Experiment 1	44
Figure 2.4.	Schizotypy Subscales: Exploratory Analysis for Experiment 1	46
Figure 2.5.	Information Weights Before the Response in Experiment 1	48
Figure 2.A1.	Psychometric curves, ΔP, and subsamples for Experiment 1	53
Figure 2.A2.	Schizotypy Subscales: Main Model Exploration for Experiment 1	56
Figure 3.1.	Paradigm used in Experiments 2 and 3	63
Figure 3.2.	Confirmatory analysis Experiments 2 and 3	71
Figure 3.3.	Exploratory analysis for Experiments 2 and 3	79
Figure 3.A1.	Visual Reference for Experiment 2 and 3	86
Figure 4.1.	Paradigm Description for Experiment 4 and 5 (E4 and E5, respectively)	96
Figure 4.2.	Time Series for Experiment 4	102
Figure 4.3.	Main Results for Experiment 4	104
Figure 4.4.	Moderation Analysis Results for Experiments 4 and 5	106
Figure 4.5.	Time Series for Experiment 5	111
Figure 4.6.	Main Results for Experiment 5	114
Figure 4.7.	Behavioural Control Theory for Two Agents	120
Figure 4.8.	Simulation of Two Agents in a Free Operant Task	123
Figure 4.A1.	Information Theory and Contingency Learning Theory	126
Figure 5.1.	Skinner's Three-Term Contingency	141

Tables

Table 1.1.	Contingency Learning Theory applies for both Pavlovian and instrumental learning	24
Table 2.1.	Sociodemographic and General Characteristics Between Samples from Experiment 1	32
Table 2.A1.	Comparison between Included and Excluded participants	52
Table 2.A2.	O-LIFE internal correlation current study vs normative data	54
Table 3.1.	Participants' General Characteristics for Experiments 2 and 3	62
Table 3.2.	Experiment 2: Conditions and ΔP [Frequency (4), and baseline Contingency (3)]	65
Table 3.3.	Experiment 2: Full model fixed effect estimates	70
Table 3.4.	Experiment 3: Conditions and ΔP [Contingency (3), Frequency (4), and Duration (4)]	74
Table 3.5.	Experiment 3: Full model fixed effect estimates	76
Table 3.A1.	Contingency and Mutual Information between S1 and S2	83
Table 3.A2.	Experiment 3 Sensitivity Analysis: Full Model Fixed Effect Estimates	87
Table 4.1.	Two Agents' Contingency Table and Formulas for Experiments 4 (E4) and 5 (E5)	93
Table 4.2.	Conditions for Experiments 4 and 5	95
Table 4.3.	Experiment 4: correlations between behavioural task outputs and questionnaires	107
Table 4.4.	Experiment 5: correlations between behavioural task outputs and questionnaires	106
Table 4.A1.	Sociodemographic and psychological characteristics from Experiments 4 and 5	127
Table 5.A1.	Computational Models (Cognitive, Bayesian, Associative and Reinforcement)	157
Table 5.A2.	Schizotypy and Characteristics of Learning Studies	161

Acknowledgments

Writing this dissertation has been one of the most challenging and gratifying adventures of my life. This work is not only the product of effort and luck, but it is also the achievement of multiple people who have provided me with invaluable support. I am grateful to all who have inspired me and helped me along this path (including Wikipedia, Stackoverflow and Stackexchange).

This personal achievement would not be a reality without the patience and encouraging conversations with my supervisors. I want to thank Robin A. Murphy, my supervisor, who taught me the value of explain complex concepts in multiple ways, and who provided theoretical and emotional provide along the whole Ph.D. adventure. I would also like to thank Michael Browning, my co-supervisor, who showed me the value of simplifying questions and fitting any kind of model.

My experience in Oxford would have not been possible without my research training in Guadalajara. I want to thank José Burgos who introduced me to computational modelling, Mario Treviño who introduced me to the art of data visualization, and my professors at the Centro de Estudios e Investigaciones en Comportamiento and ITESO University. Special thanks also to my North America learning family, Andy Baker, Ralph R. Miller, Jim Witnauer, and Andy Delamater, who inspired my work and I had motivating discussions. I thank my friends, Javier Ortiz Orendain for the vast support with psychosis and related topics, and Manuel Aguilar Moreno for inspiring me to pursue a Ph.D. and the dream to becoming a professor.

Quiero agradecer a mi amada familia, Papá, Mamá, Andrés y Jimena. Gracias por enseñarme a esforzarme en lo que amo y a siempre procurar el bien del otro y para la sociedad. Me siento afortunado de tenerlos como familia. También gracias a mis abuelos, Mamá Tere quien siempre procuró que me sienta en casa en cualquier sitio que me encuentre, y Berna quien desde pequeño me orientó para desarrollar la creatividad. Mi trabajo es para ustedes, así como para mis abuelos Jaime, Eduardo, Milagros, y Gaby. Gracias a todos por darle sentido, significado y diversión a mi vida.

This dissertation would have not been possible without the support of the Beca Institucional from Universidad de Guadalajara that funded most of my studies; the Experimental Psychology Society that supported my short-term scholar visit at Yale; St Anne’s College, which provided a great community, and supported my research and presentations at conferences; Huxter Fund who assisted my research; and Corpus Christi College for my Stipendiary Lectureship (2022-2023).

Gracias.

Declaration

Some of the work described in this dissertation has previously been published or presented in a conference.

Journal Publication

Chapter Three

Castiello, S., Miller, R. R., Witnauer, J. E., Alcaide, D. M., Fung, E., Pitliya, R. J., Morrissey, D. K. C., & Murphy, R. A. (2022). Benefiting from trial spacing without the cost of prolonged training: Frequency, not duration, of trials with absent stimuli enhances perceived contingency. *Journal of Experimental Psychology: General*, 151(8), 1772–1792. https://doi.org/10.1037/xge0001166

Conferences

Chapter Two

Castiello, S., Baker, A., Browning, M., & Murphy, R.A. (2019, April). Recency vs Primacy in Contingency Learning: a pilot experiment evaluating Schizotypy. *Associative Learning Symposium 2019*.

Castiello, S., Baker, A., & Murphy, R.A., (2020, October). Psychophysical bias to perceived positive associations in Schizotypy. *Expectation, Perception, & Cognition Virtual Workshop*. 10.13140/RG.2.2.27883.77608

Chapter Three

Castiello, S., Browning, M., Miller, R.R., & Murphy, R.A., (2021, January). Lower sensitivity to absent cues in schizotypy. Experimental Psychology Society online meeting January 2021. 10.13140/RG.2.2.30851.73766

Chapter Four

Castiello, S., Baker, A. G., Pitliya, R. J., Murphy, R. A. (2022, April). The Agency Paradox: Control Attribution in Multiagent Scenarios. *Associative Learning Symposium 2022.*

Abbreviations and Glossary

ANOVA = Analysis of Variance (statistical method to analyse multiple groups)

BADE = Bias Against Disconfirmatory Evidence (cognitive bias)

BDI-2 = Beck Depression Inventory 2 (depression questionnaire)

BCE = Behavioural Conditional Entropy

BCT = Behavioural Control Theory

CD = Cognitive Disorganization (O-LIFE subscale)

CI = Confidence Interval (range of uncertainty of a statistical parameter)

CLT = Contingency Learning Theory (theory describing the statistical dependency between events)

DSM = Diagnostic and Statistical Manual of Mental Disorders (American diagnostic handbook)

DDM = Drift-Diffusion Model (computational perceptual decision-making model)

DT = Decision Time (dependent variable used in Chapter Two)

ETD = Eye Tracking Disfunction (impairment when visual tracking an object on a screen)

fMRI = functional Magnetic Resonance Imaging (neuroimaging method)

IA = Introvertive Anhedonia (O-LIFE subscale)

IN = Impulsive Nonconformity (O-LIFE subscale)

ITI = Inter-Trial Interval (time between experimental trials)

JTC = Jumping-to-Conclusions (cognitive bias)

NS = Non Significant (null result from a statistical hypothesis testing)

KB = Kamin Blocking (learning phenomenon discovered by Leon J. Kamin)

LI = Latent Inhibition (learning phenomenon discovered by R. E. Lubow & A. U. Moore)

LIRR = Learned Irrelevance (learning phenomenon discovered by Andy G. Baker)

LMM = Linear Mixed Model (statistical method)

LR = Learning Rate (the rate in the value of an association is updated)

O-LIFE = Oxford-Liverpool Inventory of Feelings and Experiences (schizotypy questionnaire)

PC = Perceived Control (used as dependent variable in Chapter Four)

PE = Prediction Error (mechanisms for learning)

rPFC = right prefrontal cortex (usually related with memory retrieval)

RW = Rescorla-Wagner model (computational learning model)

SD = Standard Deviation (describes the variation in a continuous variable)

SDT = Signal Detection Theory (analytical or theoretical method to study decision-making)

SoA = Sense of Agency (subjective experience of controlling outcomes)

SZ = Schizophrenia (psychiatric condition)

UE = Unusual Experiences (O-LIFE subscale)

Overview

Animals learn from senses and perceive the world through the lens of experience, discriminating complex stimuli and extracting regular patterns (Murphy et al., 2004, 2008), allowing adapted interaction with the environment. Research suggests that disruptions to typical processing of experience can be related to psychological variables, such as mood and personality, and may be diagnostic features of psychopathology (Byrom & Murphy, 2018). My dissertation aim is to study how learning patterns vary between different degrees of schizotypy traits.

Across five chapters and five experiments I develop a characterization of the learning phenotype underpinning schizotypy, a latent personality organization associated with psychotic experiences (unusual beliefs and percepts) that has been related to proneness to psychotic diagnoses, such as schizophrenia (SZ; Lenzenweger, 2018a). In addition, evidence suggests that high schizotypy scorers may learn and think in ways that may support creativity and/or superstitious thinking, providing insights into the cognitive mechanisms of unusual and delusional beliefs (Hergovich et al., 2008; Mohr & Claridge, 2015) as well as how typical beliefs are formed. However, the underlying nature of schizotypy has not been fully clarified. Two schools of thought on schizotypy exist and have opposite views[1]: *i)* the first is *categorical* (Lenzenweger, 2010, 2018b), assumes that schizotypy is categorical in nature and a liability for SZ; and *ii)* while the second is *dimensional* (Claridge, 1997; Claridge et al., 1996; Grant et al., 2018), assumes that schizotypy is a typical and varying personality trait in the general population. In my dissertation, I examine how learning about contingent relations involving observed events and those caused by an agent may capture some fundamental features of schizotypy. To this end I employed Contingency Learning Theory (CLT) to account for how people learn to extract associative relationships (Byrom & Murphy, 2019), use this input to construct judgements of causation (Baker et al., 1996, 2005), and how it is related to schizotypy.

[1] I had the opportunity to talk with both Prof. Gordon Claridge in Oxford 2019 and Prof. Mark Lenzenweger in Binghamton 2022.

I developed behavioural paradigms that aim to dissect cognitive processes: perceptual, learning, sense of control, and agency. The first chapter is a general literature review on psychosis, perception, contingency learning, and agency in schizotypy. Next, I present the hypotheses for my studies. In the second chapter, my hypothesis is that higher schizotypy scorers exhibit an hyperassociation between cues, within the millisecond to second timeframe an effect that has not been shown before. In the third chapter, my hypothesis was that high schizotypy scorers would display a reduced sensitivity to absent events when making contingency assessments, in the second to minutes timeframe, an inversion of the first hypothesis. These two effects demonstrate how weighting of observed information impacts decisions. In the fourth chapter I extend this analysis to an instrumental learning procedure. Where I developed a multi-agent free-operant paradigm and investigate how people develop internal and external perceived control (PC) on the basis of the types of information presented. The hypotheses were that high schizotypy scorers would exhibit an overall bias towards perceiving control (the illusion of control) as well as display referential behaviours.

Overall, I found that schizotypy is related to slower decision time (DT) in perception of contingencies, maybe due to reduced learnt confidence in sensory evidence (Chapter Two). In addition, I found a reduced sensitivity to absent information as a function of unusual experience (UE; a subfactor of schzotypy) scores (Chapter Three). Moreover, I found a positive relationship between illusion of control and UE but only when no competing cue is present. This may be due to a reduced weight of the responses' consequences (Chapter Four). Finally, I described a two-agent computational model that captures a contrast effect between control experiences. My data supports that perception, learning, and agency are related to positive schizotypy.

In the fifth chapter I explored the meaning of 'weight' in computational models. Next, I compare this concept between Bayesian and reinforcement/associative learning models. I showed that Bayesian models are alike the associative Mackintosh model, where larger weights are given to predictive information. Finally, I integrate the learning literature of schizotypy and SZ with my experimental findings, then discus limitations, future directions, and conclusions.

Chapter One

Dissection of Schizotypy and Contingency Learning Theory

1.1. Schizotypy, Psychosis, Cluster A Personality Disorder, and Schizophrenia

1.1.1. Psychosis: Disorders, Symptoms, and Experiences

A currently definition of psychosis revolves around alterations in the ability to maintain adequate reality testing either in the form of delusions (alterations in belief formation) or hallucinations (alteration in sensory experience). Although, it is worth noting the difficulties in defining the concept of psychosis, and its mutability over the course of history (Bürgy, 2008). The prototypical psychotic diagnosis is SZ, psychotic experiences occur in other psychiatric disorders, and are also found in non-clinical samples (van Os et al., 2009). Psychotic experiences are found in 8% of the population, psychotic experiences are implied in psychotic symptoms, and these symptoms have a prevalence of 4%. Nevertheless, both experiences and symptoms are implied in psychotic disorders, which are prevalent in 3% of the population (van Os et al., 2009). Psychosis experiences in subclinical population are also found in people scoring high schizotypy personality. However, there are conceptual disagreement in the literature that contribute to the difficulty in definition that I will return to at the end of this section (Claridge, 1997; Lenzenweger, 2010).

When maladaptive personality traits provoke distress or functional impairment, the result can be considered a personality disorder: “An enduring pattern of inner experience and behaviour that deviates markedly from the expectations of the individual’s culture” (American Psychiatric Association, 2013). The DSM fifth edition, group personality disorders in three clusters: A (odd,

eccentric), B (dramatic, emotional), and C (anxious, fearful). The Cluster A of personality disorders is associated with high levels of schizotypy (Barrantes-Vidal et al., 2013), and is divided into: *i)* paranoid personality disorder, a pattern of suspiciousness and distrust in others, and malevolent interpretation of others' intensions; *ii)* schizoid personality disorder, a pattern of alienation from the social life and flat affect; and *iii)* schizotypal personality disorder, a pattern of discomfort in relationships, odd beliefs and percepts, and eccentric behaviour (American Psychiatric Association, 2013). Likewise, schizotypy is also related with psychotic disorders such as: *i)* delusional disorder, a persistent delusion for at least one month; and *ii)* SZ, schizophreniform disorder, or schizoaffective disorder, two or more of the main psychotic symptoms (delusions, hallucinations, disorganized speech, catatonic behaviour, or negative symptoms) for at determined period. To diagnose SZ the presence of delusions or hallucinations is not sufficient, a SZ diagnosis requires evidence of chronicity and another cluster of symptoms that affect thought, behaviour, and emotions.

The difference between psychosis and SZ is that, the former is a broader state of symptoms, and in the latter, psychosis exists in at least one moment of the condition. Thus, psychosis does not imply SZ. Psychosis may happen for multiple reasons, *e.g.*, drug induced psychosis, depressive psychosis, and psychosis in SZ. Alternatively, SZ a condition where psychosis is latent and, in the best case, controlled by antipsychotic medications. However, someone with a diagnosis of SZ and a remission of positive symptoms can still display psychotic experiences, such as, odd beliefs, blunted affect, and conceptual disorganization among other symptoms (Schennach et al., 2015). These symptoms are also sometimes referred as attenuated psychosis or residual SZ.

In terms of treatments for SZ, cognitive-behavioural therapy (CBT) −the main psychotherapeutic technology developed in psychology− as a treatment for SZ has very limited evidence to improve clinical outcomes in patients, based on evidence from Cochrane Reviews (Jones, Hacker, Meaden, et al., 2018; Jones, Hacker, Xia, et al., 2018). Alternatively, psychopharmacological treatments usually target the dopamine system (Kapur & Remington, 2001) and alleviate positive symptoms (Ortiz-Orendain et al., 2017). In contrast, psychological treatments focus on negative symptoms (Klingberg et al., 2011), restructuring thoughts, and supporting behaviour plans (Hofmann

et al., 2012). Yet, SZ leads to poor long-term outcomes, and very low (4%) recovery rates (Novick et al., 2009), but in short psychotic episodes usually prognosis is better. This highlights the importance of understanding the psychological mechanisms underlying the psychotic experience. I start with a historical overview of the scholarly work on SZ and underlying psychotic experiences.

Psychotic experiences while usually interpreted in negative terms, can be found as expressions of both both health and psychopathology. For instance, healthy individuals with psychotic experiences have been used to describe new forms of human perception. Such individuals have been assigned special religious and cultural significance in some societies (*e.g.*, shamanism, Mohr & Claridge, 2015; Silverman, 1967). Similar perceptions have been implicated in artistic expression (*e.g.*, Vincent Van Gogh Blumer, 2002). In the general population, the psychotic experience may be partially captured by the construct of schizotypy (Grant et al., 2018). Nonetheless, forms of psychosis, such as psychotic depression (Dubovsky et al., 2021), may be independent of schizotypy. In summary, schizotypy and psychotic experiences have been displayed in multiple facets of humans, and not all of them imply psychiatric disorders.

1.1.1.1. Conceptual Distinctions of Schizotypy

Schizotypy has been conceptualized by two different schools. On the one hand, a *dimensional* approach, Eysenck (Eysenck, 1947, 1952) suggests that schizotypy is a normal personality and continuous dimension, this approach supports the idea that schizotypy is not necessarily psychopathological (Claridge, 1997, 2018; Grant et al., 2018; Mohr & Claridge, 2015). On the other hand, a *categorical* or *disease-based* approach, Meehl (Meehl, 1962, 1990) sees schizotypy as a latent personality organization, discontinuous in nature, and as a liability for SZ (Barrantes-Vidal et al., 2013; Gooding et al., 2005; Kwapil & Barrantes-Vidal, 2015; Lenzenweger, 2010). The specific relationship between schizotypy and psychosis is unknown. Although, schizotypy as personality, as well as other factors such as genetic (*i.e.*, schizotaxia; Avramopoulos et al., 2002), social (van Os et al., 2009), and drug consumption (Hall & Degenhardt, 2008; Patel et al., 2020; van Os et al., 2009; *cf.*, Daedelow et al., 2021) increase the risk of clinical psychosis −in people with SZ many drugs have been shown to worsen symptoms (D'Souza et al., 2005).

A more recent perspective to detect psychosis vulnerability, comes from phenomenology which aims to assess the anomalous self-experiences. Here, the main core of psychosis vulnerability is a function known as *ipseity*- or self-disorders (L. A. Sass & Parnas, 2003), which are characterized by a lack of sense of minimal-self (*i.e.*, the basic unit of experience), loss of common sense (Hipólito et al., 2018), and attenuated sense of self-presence or agency. This type of phenomenological exploration is also related with non-psychotic manifestation of schizotypy (Raballo & Parnas, 2011). Overall, given psychotic experiences found in schizotypy, the latter has been used as a model to study psychotic symptoms and psychotic diagnoses in general population (Kwapil & Barrantes-Vidal, 2015). This perspective is of special interest in Chapter Four where I test SoA and control.

1.1.1.2. Schizotypy

The nature of schizotypy is still an ongoing debate and no definitive evidence support either posture (Linscott & van Os, 2010). It could be that each posture has their own benefits, *e.g.*, *categorical* reliability and *dimensional* validity (Esterberg & Compton, 2009), or even that schizotypy is both perspectives (Mason, 2014). The aim of my dissertation is not to understand the nature of schizotypy, thus given some experimental and epidemiological evidence supporting the *dimensional* view (Corlett & Fletcher, 2012; Oestreich et al., 2015a, 2015b; Ross et al., 2015b; van Os et al., 2009; Whitford et al., 2017a) in my dissertation I used the *dimensional* approach of schizotypy.

I measured schizotypy with the O-LIFE in its brief (43 items) form (Mason et al., 2005). The O-LIFE has four factors (Claridge, 1994), Unusual Experiences (UE), Introvertive Anhedonia (IA), Impulsive Nonconformity (IN), and Cognitive Disorganization (CD). UE is related to traits for positive symptoms of psychosis (also called positive schizotypy) and contains items related to perceptual aberrations, magical thinking, and hallucinations. IA is the subscale related to negative symptoms of psychosis (also called negative schizotypy), including lack of sources of pleasure and avoidant behaviour. IN contains the items related to impulsive, anti-social, and eccentric behaviour. Finally, CD measures the dimension related to cognitive symptoms of psychosis as well as disorganization; these include poor decision-making, attention, and concentration.

1.1.2. Perception, Learning, Agency, and Schizotypy

In this section I characterise the disrupted cognition, learning and memory in schizotypy and SZ. Historically, Eugen Bleuer coined the term schizophrenia (SZ), from greek *schizen* 'to split' and *phren* 'soul, spirit, or mind', which was proposed as an alternative term to the *dementia precox* (Fusar-Poli & Politi, 2008), popularized by Emil Kraepelin coined by Arnol Pick. In Bleuer's opinion, the main symptom of the group (*i.e.*, schizophrenias) was a splitting of psychic functions (Moskowitz & Heim, 2011). For Bleuer, one of the central features of SZ was the disruption of thoughts, also known as *loosening associations* (*schizen* as split of associations), which was accompanied by signs and symptoms such as ambivalence, avolution, autism, affective blunting, and attentional impairment (Andreasen & Carpenter, 1993). Attention, memory, and executive functioning present at the core of the disorder, *dementia praecox*, reflects the conceptualisation of SZ as a neurocognitive disorder with early onset (Kraepelin, 1919). Cognitive symptoms are thought to be fundamentally distinct from positive and negative symptoms because they do not respond to antipsychotic medication, usually are stable over time, and seem to be independent of psychosis onset (Gold, 2004). Yet, some of the cognitive symptoms are also found in schizotypy (Sahakyan & Kwapil, 2018). The current operational diagnosis of SZ has been criticised for an over-reliance on the positive symptoms while simultaneously overlooking cognitive aspects of the illness (Kahn & Keefe, 2013; Keefe et al., 2005). Studying cognitive symptoms usually requires neuropsychological assessments with a clinician or computer-based tasks.

From the perspective of Mortiz et al.'s (2017) two-stage cognitive theory of psychotic symptoms, delusion formation and delusion maintenance are distinct and may involve separate processes at different times, *i.e.*, stages. On the one hand, Stage 1, delusions are formed by reduced decisional threshold, *i.e.*, liberal acceptance, while on the other hand, Stage 2, delusions are first maintained by an augmented confidence in beliefs (confirmation biases), and second, as a reduced strength of alternative explanations/beliefs to what they already believe (Moritz et al., 2017). The latter is an example of the BADE; this can be measured using a task in which participants, who are exposed to multiple scenarios, are asked to rate the likelihood of multiple interpretations. Then, in

consecutive trials, the scenario reveals more information which makes it easier to assess the likelihood of the interpretations. For instance, people adjust their likelihood ratings as a function of gathering more evidence. Yet, people with SZ change their likelihood ratings following disconfirmatory evidence to a lesser extent than people without SZ (Juárez-Ramos et al., 2014). BADE has been observed also in people with no psychotic diagnosis but high schizotypy scores (Buchy et al., 2007; Woodward et al., 2007), where their interpretations do not change as much as people with low schizotypy scores.

Another neurocognitive account of delusion formation builds on standard models of selective attention that describes how automatic processes of salience detection are used to direct attentional resources (see Flagel et al., 2011). These models suggest that selective attention in SZ and schizotypy is maintained to stimuli that have little relevance to a particular task (Haselgrove et al., 2016; Le Pelley et al., 2010; Morris et al., 2013). This disrupted attentional phenomenon may be due to aberrant salience attribution, such that perceived salience is maintained. This effect has been related to elevated dopaminergic signalling to irrelevant cues (Griffiths et al., 2014; Kapur, 2003). Aberrant salience attribution can be captured by computational models involving learning PEs, which can represent the surprise generated by a high salience cue. For instance, reward PE attenuation is anticorrelated with unusual though content (Corlett et al., 2007). Also, people with SZ show both an attenuated PE response to rewarding cues and an augmented PE response to irrelevant cues in the midbrain (Murray, Corlett, et al., 2008). In people with no SZ diagnosis, PE in striatum also show aberrant activity that has been argued reflect stimulus encoding and negatively correlates with magical and distressing beliefs (Corlett & Fletcher, 2012). Altogether, these findings suggest that the use of PEs in learning may be relevant for how people develop (unusual) beliefs. Nevertheless, given the parallelism between SZ and the general population only varying in the magnitude, the continuous nature of schizotypy has been supported (Claridge, 1997, 2018), rather than the categorical or disease-based model (Lenzenweger, 2010).

In parallel, a body of experimental work has built on pre-clinical studies on the interaction between attention and learning to propose that learning from otherwise irrelevant stimuli is a key

cognitive feature of psychosis (*e.g.*, Griffiths et al., 2014; Kapur, 2003; Le Pelley et al., 2010) and relates to current thinking about how to structure research efforts in psychiatry in general (*e.g.*, Kotov et al., 2017). In the next section I shall examine evidence relating SZ and schizotypy with perception, learning, and agency.

1.1.2.1. Perception

Perceptual impairments in schizotypy and SZ are found at different timeframes. In this section I describe evidence relating schizotypy and SZ with contrast sensitivity, sensory attenuation, perceived causality and apophenia al potentially markers of perceptual differences.

A visual low-level impairment is found in schizotypy which reduces the ability to see shades and patterns. For instance, using a visual psychophysics paradigm, people with high UE show a lower contrast sensitivity, possibly due to an increase in internal sensory noise (Harper et al., 2020). In addition, low-level perceptual impairments are a potential link between symptoms in SZ and the underlying neurobiology, it states that an increase in precision of prior beliefs compensate a failure in sensory attenuation (Adams et al., 2013, 2016; Friston et al., 2016). In a perceptual decision-making paradigm, it has been revealed that the precision of prior beliefs (*i.e.*, starting point from a DDM) rather than increased sensory precision (*i.e.*, drift rate) drive premature responses and impaired response inhibition in people with SZ (Limongi et al., 2018). In summary, it is possible to claim that UE in schizotypy and positive symptoms in SZ may involve a combination of disruptions to both top-down and bottom-up processing. In the former, aberrant percepts may be due to an internal noise in input signals, and the latter an enhanced PE to surprising or novel events.

Evidence of difference in sensory attenuation −understood as a reduction in sensing due to own actions (Pareés et al., 2014)− has been found in non-clinical high schizotypy scorers. Here, people with high schizotypy scores showed a reduced sensory attenuation when they were in control of an experimental machine that produced somatosensorial stimulation in their arm (Lemaitre et al., 2016). In addition, high schizotypy scorers produce similar sensory ratings between external and self-sources of sensorial stimulation (Whitford et al., 2017b).

Perceiving causality differs as a function of schizotypy. Positive symptoms of SZ were related with an increased perception of causality, and cognitive symptoms were negatively correlated (Tschacher & Kupper, 2006). Here, participants were asked to detect when a 'collision' between circles happened. The collision was meant to coincide with a sound, which was delayed or anticipated in relation to the visual collision in different conditions. They found that positive symptoms were related with an increase in perceived causality and cognitive symptoms with a reduction in perceived causality in SZ.

Finally, apophenia is the tendency to relate unrelated events, or to attribute meaning to meaningless events. For example, believing that because a sports commentator on the TV said my name, then it is my lucky day, therefore I will bet all my money to the horse which name has similar characters as the commentators' name. Operationally, apophenia can be thought and measured as a false alarm, *i.e.*, perceive something where nothing was there (Blain et al., 2020). People prone to delusions, and high schizotypy and psychoticism scorers, tend to exhibit higher degrees of apophenia than the low scorers (Blain et al., 2020; Fyfe et al., 2008). Overall, in this section I presented evidence that relates schizotypy with impairment at different perceptual levels, from low-level visual contrast sensitivity and sensory attenuation to high-level apophenia beliefs and perception of causality.

1.1.2.2. Associative Learning

Associative learning is an analytic technique for testing how experience translates into behaviour. The specific term 'associative' refers to the hypothesised links formed during learning between internal representations of experience. The type of tasks involves presenting events with varying contingent relations and test participants sensitivity to those statistical relationships. Although not strictly speaking a causal task, participants perceive causality based on contingency or correlation (Goldman, 1967, 1979) especially if one cue precedes the occurrence of the second event or outcome, a form of predictive learning (Baker et al., 2005). For instance, an event (A) causes another event (B) if the presence of A increases the likelihood of B (Allan, 1980; Baker et al., 1996). In this section I describe some associative learning phenomena such as Kamin Blocking (KB), Latent Inhibition (LI),

and Learned Irrelevance (LIRR), then I describe neuroimaging and eye tracking evidence converging in the same idea, increased irrelevant learning in people scoring high in schizotypy.

Kamin Blocking (*KB*). One special set of phenomena of particular interest for associative learning are the cue competition effects, such as KB (Kamin, 1967). For example, KB is the reduced sensitivity to the associative value of a cue to predict the presence of an outcome given that a simultaneously presented cue is already a good predictor of the same outcome. In phase one agents learnt about the meaning of A+ (cue A rewarded), whereas in phase two they learnt about a new cue B also presented in compound with A, *i.e.*, AB+. Agents showed interference in learning about B in the second phase given that A accompanies learning about B, thus A blocks B, *i.e.*, KB effect. Simple models of learning (Bush & Mosteller, 1951; and model free or Q-learning, Sutton & Barto, 2018) cannot account for KB because they do not have some comparator or contrast process that permits interaction between cues (*e.g.*, Delamater, 2012; Donahoe et al., 1993; Mackintosh, 1975; Pearce & Hall, 1980; Rescorla. & Wagner, 1972). Even though KB is a central leaning phenomenon there are conditions that interfere with its emergence (Maes et al., 2016) and some discussion whether the contrast process occurs at time of acquisition or performance.

Schizotypy is related with a reduction in the KB effect (Haselgrove & Evans, 2010; Moran et al., 2012; *cf.*, Humpston et al., 2017) and self-reported delusional scores showed differences in neural activity in fronto-striatal PE in a KB task (Corlett & Fletcher, 2012). One explanation for the KB effect occurrence is that in the preexposure phase A acquires all the associative strength, and in the compound phase the associative strength left is split between A and B, leaving B with little left to gain (Rescorla & Wagner, 1972). In other words, B is being blocked by A, because A was already a good predictor of the reward (+). An explanation of the KB effect impairment in schizotypy may be due to a non-selective learning process (Haselgrove & Evans, 2010), this is that high schizotypy scorers may process cues independently and do not compute the interaction between cues, thus B will be learned as well as A. There is another variation of KB called backward blocking, which is also disrupted in people with proneness to delusions (Corlett et al., 2007). In this version, AB+ is presented in the first stage, and both cues acquire association with the outcome; and in the second phase A is presented and

paired with the outcome. If tests of B following this second phase show a reduction in the association with the outcome in the absence of any specific B experience Backward Blocking is said to have occurred (Kruschke & Blair, 2000; Shanks, 1985). Even though the early demonstrations of KB and Backward Blocking both require two phases of learning, relative validity learning designs can find similar effects in which both type of trial are presented in the same learning phase (Baetu et al., 2005; Baker et al., 2000; Baker & Mackintosh, 1979; Murphy et al., 2001).

Thus is KB is agents learn less about cues if more salient or predictable cues have been acquired. In KB (both forward and backward) B is less valid than A and this difference enhances attention to A and reduces attention to B (Mackintosh, 1975). Predictive value can be the relevant feature producing saliency from exteroceptive inputs. Therefore, the KB paradigm is a method of studying salience attribution (Corlett et al., 2007; Corlett & Fletcher, 2012). The salience or relevance of a cue is linked to dopamine (D2) activity, and SZ has been conceptualized as an aberrant state of salience (Kapur, 2003). Rats exposed to amphetamines (dopamine D2 agonist) show abolished KB (Crider et al., 1982). Also, a KB reduction is reported in general population under amphetamine (Gray et al., 1997; S. H. Jones et al., 1997). Dopamine is related with reward learning (Flagel et al., 2011; Pessiglione et al., 2006), and SZ seems to be impaired in prefrontal areas −causing negative symptoms− which at the same time leads to excessive dopamine activity in mesolimbic areas −causing positive symptoms− (Davis et al., 2006).

Jones, et al. (1992) found that people with acute but not chronic SZ showed enhanced learning of an irrelevant cue, using a KB learning procedure. In their study, two groups one without any diagnosis and the second with SZ were randomized in two groups: preexposure (P) and non-preexposure (NP) groups. In the first stage, participants were exposed to sequences of cues and were asked to predict an outcome, for group P, cue A was always followed by the outcome. However, for NP the outcome was not reliably predicted by any cue. In a second stage, participants were exposed to a novel auditory cue B, then A and B presented in compound were followed by the outcome. They found that participants without diagnosis in group P needed more trials to learn the cue B-outcome association in the last test stage in comparison with group NP, *i.e.*, KB effect. Yet, chronic patients in

P and NP showed no difference; and more surprisingly acute patients showed a faster learning to the redundant cue B (S. H. Jones et al., 1992b). However, this KB effect was not found in the general population with high psychoticism (S. H. Jones et al., 1990), but when participants from the general population received amphetamines, KB was disrupted (S. H. Jones et al., 1997). In addition, other studies suggest that the impairment in KB is found only in no-paranoid patients but remains intact in paranoid patients (Moran et al., 2003; Oades et al., 2000). Finally, the reduction of the KB effect seems to be extended to first order relatives of people with SZ (Serra et al., 2001). Together these findings support a link between the schizotypy spectrum −in both clinical and non-clinical population− and an increase learning to novel and non-predictive cues, which also may be linked with the dopamine hypothesis of SZ.

To confirm the previous impression, there are at least two additional learning paradigms that can provide meaningful data to support the impression. These are learned irrelevance (LIRR) and latent inhibition (LI), which have also shown impairments in high schizotypy scorers (Haselgrove et al., 2016; Le Pelley et al., 2010; Schmidt-Hansen et al., 2009).

Latent Inhibition (*LI*). LI is the reduced expression of behaviour towards a previously extinguished cue (Lubow & Moore, 1959). For example, a cue P that in a previous phase has been expose to an extinction protocol (P-); then in a subsequent acquisition phase P is paired with reward (P+). What is observed is that in the second phase, agents slowly learn from P in comparison of another cue N that was not previously exposed to the extinction protocol only reward learning in the second phase (N+). Thus, it is said that agents display LI to P when comparing to N (Lubow, 1973). One of the possible mechanisms of LI is that given the negative predictive value of P, *i.e.*, P predicts no reward, P's attention is reduced (Baker & Mackintosh, 1979). LI is a critical learning phenomenon because it raises theoretical question with respect what is being learn when there is no reward. LI provides evidence that agents learn statistical relationships in the absence of a reward.

Interestingly, like KB a reduced LI effect has been related to people with high scores of schizotypy and with a diagnosis of SZ (Evans et al., 2007; Granger et al., 2012, 2016; Schmidt-

Hansen et al., 2009; Serra et al., 2001; Vaitl et al., 2002; *cf.* Swerdlow et al., 1996). Evidence for an effect of LI is observed at both the behavioural and physiological levels. For instance, in unmedicated patients with a diagnosis of SZ, reduced LI was shown by a reduced electrodermal conditional response to the novel stimulus (Vaitl et al., 2002). These results suggest an impairment in LI due to a non-differential conditional response to the preexposed and novel stimulus. Thus, this support the claim that people with SZ are less sensitive to novel stimulus, perhaps to surprising stimulus. However, other studies have failed to replicate the LI impairment in people with SZ argued that the core issue is an overall learning impairment (Swerdlow et al., 1996).

In addition to these effects with clinical populations, LI impairments are found in groups with high schizotypy scores in the general population (Evans et al., 2007; Granger et al., 2012, 2016; Schmidt-Hansen et al., 2009). For example, Granger et al. (2016) conducted an experiment with a general population sample where: *i)* no target stimulus was presented in the preexposed phase, *ii)* participants were required to read out loud the stimulus letters presented in Phase 1 (*i.e.*, controlling for possible LIRR and conditioned inhibition), and *iii)* were asked to press a key anticipating the occurrence of the target stimulus in Phase 2. They found a LI effect, the reaction times (RT) for the preexposed stimulus weas larger than the RT for the non-preexposed stimulus, *i.e.*, slower learning to the stimulus that was seen in Phase 1. In addition, they found a positive relationship between positive schizotypy (UE) and RT for the preexposed cue. This suggests that LI as KB are a transdimensional learning phenomena in clinical and non-clinical forms of schizotypy.

Learned Irrelevance (*LIRR*). Research on the effect of LI on behaviour has grappled with whether the effect involves learning about the predictive value of a cue or its attentional capture. The LIRR procedure involves a preexposed cue that is presented without an outcome, but for which the outcome does occur during the same session but is presented with a null correlation (non-contingent) (Baker, 1976; Baker & Mackintosh, 1979). Thus, the first phase in LI, the cue P is non-rewarded (P-), and in the case of LIRR, P is presented uncorrelated the reward (P +/-). This non-contingent presentation of P and the reward makes impossible that P became informative of the reward.

In the case of humans' contingency judgments or causal learning, an LIRR experiment can involve a more complex experimental design because the original design may be too easy to learn. Thus, normally with humans in the first phase participants are preexposed to cue compounds (AX, AZ, BX, and BZ) with different outcomes (+ and *) with the rule that only two of the four cues are completely predictive of each of the two outcomes, *i.e.*, AX+, AZ+, BX*, and BZ* (Le Pelley & McLaren, 2004). Therefore, A always predicts + and B always predicts *, but X and Z does not reliably predictors of + nor *. The LIRR effect is exposed in a subsequent phase where learning from X and Z is reduced in comparison of A and B. This effect may be due to a change in the predictive value of the cues in the first stage (Mackintosh, 1975). The predictive value has been demonstrated experimentally where responses are given faster to a target location that is predicted by a cue (Le Pelley et al., 2013).

In the case of people with SZ, similarly to KB and LI, LIRR has been shown to be impaired (Gal et al., 2005; Young et al., 2005). This is observed by an increase learning to the non-predictive cues. Participants observed a sequence of characters and were asked to anticipate the occurrence of a target "X". When comparing the second phase between two conditions, preexposed and non-preexposed, the RT of participants with no diagnosis of SZ are slower for the preexposed condition and faster to the non-preexposed, *i.e.*, the LIRR effect. However, patients with a diagnosis of SZ do not display the effect and their RT are similar suggesting that they attend similarly to the cue in the preexposed and the non-preexposed conditions. Moreover, only people with chronic SZ seem to struggle to learn overall (Gal et al., 2005; Young et al., 2005).

In addition to clinical population, people with high schizotypy scores also show an impairment in LIRR (Haselgrove et al., 2016; Le Pelley et al., 2010; Morris et al., 2013; Schmidt-Hansen et al., 2009). Using variations of Le Pelley and McLaren (2004)'s task, where in a first phase participants learnt AX+, AZ+, BX* and BZ*, and in the second phase participants learnt AB (unique outcome predictive cues) and XZ (non-unique outcome predictive cues), Le Pelley et al. (2010) showed that high positive schizotypy scorers learned AB and XZ. This suggests an enhanced learning to irrelevant (non-predictive) cues in high positive schizotypy scorers.

Moreover, to complement the behavioural evidence, eye tracking provides a way to measure cues' predictive values (Beesley, Nguyen, et al., 2015). Supporting the selective attention impairments, Eye Tracker Disfunction (ETD) has been documented in people with SZ (Levy et al., 2010) and high schizotypy scorers (Lenzenweger & O'Driscoll, 2006). ETD refers to smooth pursuit of slow-moving objects, and results showed that SZ tend to fixate on non-informative areas in comparison to general population (Beedie et al., 2011). In a general population sample, evidence suggests that either predictiveness and/or uncertainty influences learning and gaze to cues, also found independently to participants awareness, *i.e.*, implicit learning (Beesley, Pearson, et al., 2015). When participants are presented with certain and uncertain cues, participants tend to look more at the uncertain ones, as an exploratory gaze behaviour (Beesley, Nguyen, et al., 2015), supporting a variant of associative theory that relies on attentional tuning, the Pearce-Hall model (Pearce & Hall, 1980). Although, this model contrasts with other, attentional theories that predict the opposite result, that more informative cues should attract attention (Le Pelley, 2004; Mackintosh, 1975).

With respect LIRR, eye tracking, and schizotypy, Haselgrove et al. (2016) used eye tracking to assess if learning irrelevance in schizotypy was related with the proportion of time that participants saw predictive versus non-predictive cues. They found that participants with lower scores of schizotypy focus on relevant cues, whereas high schizotypy scores divide their attention between relevant and irrelevant cues (Haselgrove et al., 2016).

Overall, KB, LI, and LIRR studies suggest a difference in learning about normally irrelevant cues, and it may be due to a fundamental impairment in selective attention processes, for high schizotypy scorers, people with SZ, and relatives of people with SZ. In addition to the behavioural results, in neuroimaging studies there is some evidence showing that the PE in a KB paradigm is disrupted by schizotypy and found relation with an increased striatal activation to irrelevant cues (Corlett & Fletcher, 2012). Corlett & Fletcher (2012) argued that the striatum maintains 'irrelevant' cue-outcome associations, which could be interpreted as if KB occurs due to a decay in the associations formed by the blocked cue and the outcome, so a lack of KB is because of maintaining

those associations. In addition, an fMRI study suggested that the medial frontal gyrus, previously related with computing uncertainty, was related with a reduced KB (Moran et al., 2012).

1.1.2.3. Instrumental Learning

People exhibit their learning when their actions are used as 'instruments' to obtain feedback (appetitive or aversive consequences) from the environment. Here, learning implies the action-reward contingencies (*i.e.*, instrumental learning); it has been suggested that people scoring high in schizotypy or diagnosed with SZ display a disruption in the meaning of actions (Morris et al., 2015, 2018).

Meaning of action (or judged feedback) can be defined and tested using instrumental learning. Outcome devaluation and degradation involve changing the value of a learnt outcome and expecting behaviour to alter if the outcome is represented as a goal. Morris et al (2018) used a 2-action free-operant with degradation and devaluation conditions and showed that people with SZ learnt the relation between their actions and rewards (food delivered at the end of the experiment) in the same way as people with no SZ diagnosis. However, SZ showed a reduced sensitivity to contingency degradation, *i.e.*, same response rates in conditions where their actions produced the outcome and conditions where did not. This indicates that people with SZ may have a disruption in action-outcome contingency learning (Morris et al., 2018), which is mapped onto hypoactivity in the Caudate (Morris et al., 2015). Furthermore, in a one-action free-operant task, high UE scorers did not show differences when responding to two reinforcement programs: Random Ratio (RR) and Random Interval (RI; Randell et al., 2009, 2012). RR programs are rules of reward delivery dependent on a random number of responses, in contrast to RI are programmes dependent on a random time. A standard response pattern is to find a higher response rate to RR than to RI (Ferster & Skinner, 1957). The lack of differentiation may be an indicator of a lack of contingency action-outcome sensitivity, where high schizotypy scorers keep responding even when they could get the same outcomes with fewer responses (*i.e.*, RI). Finally, the lack of sensitivity between schedules was evident at the response level but also at the conscious level, participants were not able to verbalize differences in schedules (Randell et al., 2012).

The weight of actions-produced outcomes seems to be affected in schizotypy and SZ. Consistent with the previous results a psychopharmacological study using ketamine −a non-competitive N-methyl-D-aspartate (NMDA) receptor agonist that induces dissociative symptoms and is used as a model of psychosis−, a study found a reduce PE weighting (Vinckier et al., 2016). Healthy participants on ketamine played a Reinforcement Learning (RL) task, where they choose between different cues, one associated with high win and high loss and another with low win and low loss. They fit a RL model and found that ketamine reduced the impact of confidence, a meta-cognitive construct, which in this case modulates choice (exploration/exploitation) and learning at a higher order level. In sum, ketamine reduced the exploitation and the impact of PE for learning in SZ (Murray, Corlett, et al., 2008; Vinckier et al., 2016). A reduced confidence in PE can be reframed as a reduced rate in the belief update mechanism which also is found in SZ (Bansal et al., 2022). Hence, all the previous studies (Morris et al., 2015, 2018; Murray, Cheng, et al., 2008; Murray, Corlett, et al., 2008; Vinckier et al., 2016) agree with a similar fundamental mechanism in SZ, *i.e.*, a PE unbalance in different levels of the processing cognitive hierarchy (Adams et al., 2013; Corlett et al., 2019; Sterzer et al., 2018).

In addition to contingency-based learning, impairments in goal-directed behaviour have been observed in SZ, this is how rewards guide their choices. For example, in young people in risk of psychosis, Waltz et al (2015) showed that CD and social disfunction were related with a reduced ability to use positive feedback to guide their choices (Waltz et al., 2015). This result contradicts Morris et al. (2018) who found that feedback related goal-directed behaviour seemed to be intact in SZ. Although these studies have completely distinct population (people with SZ vs young at risk of psychosis). In a similar two-bandit arm task, Murray, Corlett, et al (2008) found no differences comparing LRs nor the exploration/exploitation parameter (decision temperature) between healthy participants and people with SZ. However, they showed abnormal reward PE in dopaminergic midbrain, striatum, and limbic system, in people with positive psychosis symptoms (Murray, Corlett, et al., 2008; Schlagenhauf et al., 2014). In summary, the relation between feedback and learning in people with SZ and risk of psychosis may depend on the population.

To model the subcomponents of decision-making with actions and DT it is possible to use computational models. For example, the DDM dissects decision into multiple and distinct components such as how much sensory evidence an agent requires make a choice (threshold), if there is a starting decisional bias toward one of two options (starting point), or how quick sensory evidence is accumulated (drift rate). In one study, people with SZ exhibited more cautious choice selection (threshold parameter), punishment learning deficits (reduced drift rate for punishment trials), and slow motor/encoding time (Moustafa et al., 2015). However, no difference was found in the starting point, thus no decisional bias was found, this contrasts with a previous study which argue that the decisional bias was the main difference in SZ (Limongi et al., 2018). Moustafa et al. (2015)'s results in deficit in punishment feedback is supported by another study that found a reduced reward sensitivity (Schlagenhauf et al., 2014). Altogether, this indicates an outcome downweighing computation in SZ. Although, this seems different to the intact goal-directed behaviour found by Morris et al. (2018). Yet, evidence suggests that positive and negative outcomes are computed distinctively (Pulcu & Browning, 2017). Thus, the difference in reward valences may explain the distinct results.

Finally, something fundamental to consider for instrumental learning in SZ is the relevancy of memory (Collins & Frank, 2012). Recent findings suggest that there is an important working memory (WM) component in RL and that this may explain reward-based learning impairments in people with SZ (Collins et al., 2014, 2017). Collins and collaborators (2017) found that people with SZ had difficulties learning conditions with high memory load but did not find overall differences in RL. These results indicates that WM and not RL acquisition explained reward-based learning impairments in SZ. Nonetheless, Morris et al (2018) and Collins et al (2017) are not mutually exclusive. RL impairments in people with SZ are due to causal understanding and WM but not to goal-directed acquisition.

Overall, it seems that people with SZ and high schizotypy scorers display at least three fundamental impairments regarding learning the relation between their actions and their consequences. First, the feedback sensitivity is reduced. Second, a WM impairment affect reward-based learning. Third, the causal value of their action, contingency sensitivity, or meaning of action.

Learning the relations between action-reward and no-action-rewards have separable mechanisms at the behavioural and neural levels (Jocham et al., 2016). Goal-directed actions and sensitivity to contingent events allows individuals to act in their environment to achieve outcomes. Next section incorporates describe evidence between sense of control and people with SZ or high schizotypy scorers.

1.1.2.4. Sense of Control and Agency

Agency is the feeling of being the one who make choices and therefore is responsible of the consequences occurring to the self (Moore, 2016). But also, this feeling of experiencing the consequences due to oneself is closely linked with the agent being able to produce certain consequences rather than other, therefore, controlling consequences or outcomes. In this section, I will briefly examine the Sense of Control and SoA and its relationship with SZ and schizotypy.

In philosophy of mind, the SoA as well as the sense of ownership are part of the basic unit of experience called *minimal self*, a pre-reflexive point of actions, thoughts, and experience (Gallagher, 2000). Phenomenological research in psychopathology suggests that the main core problem in SZ may be related this basic form of self, *i.e.*, *ipseit* disorder (Henriksen & Parnas, 2014; Laing, 1960; Raballo & Parnas, 2011; L. A. Sass & Parnas, 2003). Therefore, agency may be another window to study psychotic experiences linking experimental and clinical disciplines.

Sense of control and SoA have been studied using experimental paradigms where participants are asked if a seen hand movement belongs to them. Daprati and colleagues (1997) asked participants to see videos with hand movements −the videos were distorted through a camera and two mirrors− and identify the source of the movement: self or external. In some conditions, participants saw their own hand, in others the experimenter's hand. When asked to rate how much they felt that the seen hand was their own people with SZ, in comparison of general population, generated more errors when they saw the experimenter's hand, they were insensitive to the change and more willing to perceive their agency on the basis of visual input but with no proprioceptive feedback. Although, when they

saw their own hand no errors of attribution to external agency were made. Finally, the errors in agency attribution were more likely in positive symptoms of SZ.

In another study, people pressed a key (self-paced) and heard a delayed tone in two conditions: SoA and perceived delay. They were asked to rate their instrumentality over the tone, and the perceived delay, respectively for each condition. Overall, high schizotypy scorers exhibited reduced SoA. In addition, their SoA was strongly reduced in the perceived delay condition (Asai & Tanno, 2008). Furthermore, it has been found in schizotypy that less SoA is related to less surprise, *i.e.*, less PE, hence less beliefs updating (Moore, Dickinson, et al., 2011). Nonetheless, ketamine administrated to people with no diagnosis of SZ produce a reduced SoA, similar to what is found in SZ (Moore, Turner, et al., 2011). SoA can be changed by ketamine, which is in part why it is used as a model of psychosis in general population and animal research.

Overall, people with SZ and high schizotypy scorers tend to exhibit a reduced SoA which maps onto theories of phenomenological psychopathology. If SoA is a subcomponent of the self, then agency tasks may be mapping a fundamental core in SZ. In Chapter Four, I develop this idea using an illusion of control task in a multiagent scenario.

1.2. Contingency Learning Theory

Throughout this introduction the elements of a learnt contingency seem fundamental for understanding how schizotypy modulates perception of the environment. Contingency learning more generally is the study of how people learn about associations between stimuli or/and events. This is considered a model of predictive or causal learning (Baker et al., 1996, 2005; Goldman, 1967; Shanks, 1985, 1995).

In a predictive learning task, experimenters use experimental designs and abstract cues, or novel correlational scenarios, participants match their contingency judgments to cue-outcome (O) associations to the true experimental probabilistic relations between them, *e.g.*, a sample fertiliser is related with a crop field (Baetu & Baker, 2009). This requires both memory for the events and some

integration of those memories. An example is the *tank and plane task* (Baker et al., 1993; Shanks & Dickinson, 1987), in this study, participants played a computer game, and learnt to actively prevent an outcome (*i.e.*, the destruction of a tank) based on a participant-dependent cue (*i.e.*, participants decided when to camouflage or not) and a participant-independent cue (*i.e.*, in some trials participants saw a spotter plane that could change the likelihood of an O). Participants learnt the relation between cues and outcome. However, they also showed that when the independent cue (*i.e.*, plane) was more valid as a predictor of the O and camouflage was a partial predictor of it, participants judgments for the effectiveness of their actions was considerably reduced. This is an example of relative validity learning (Wagner et al., 1968), where the credit assignment or the predictive value of the action (camouflage) depends on an incontrollable cue (plane); and illustrates one of the core learning phenomena under consideration in my dissertation —that people's contingency judgments interact when two predictive cues provide generative or preventative information of the O; *i.e.*, cue competition (Baker et al., 1993, 2000). This interaction between self and external generated cues is one of the central ideas that I explore in Chapter Four.

1.2.1. When Pavlov 'met' Skinner: It is about the contingency!

The experiments presented in my dissertation are framed within the CLT. This theory applies for both Pavlovian and instrumental learning and can be used as an approximation to study causal learning in humans (Baker et al., 2005; Dickinson, 2001). For my dissertation contingency refers to a statistical relationship relation between two events, see **Table 1.1**, which is described formally by the ΔP value (Allan, 1980).

Learning the contingency between two stimuli, a stimulus and a response, or response and its consequence (*e.g.*, instrumental) is a fundamental form of learning (Baker et al., 1996), involving perceiving, remembering, and learning the correlation between events. Such statistical relations can be positive, in which people learn that the presence of one event predicts another (bacon is a useful predictor of eggs as they are positively correlated in some people's breakfast experience), but predictive relations can be negative, where the presence of one event signals the absence of another

(toothpaste and wine, blessedly, tend to be experiences that are negatively correlated). While the simultaneous presence of events is clearly relevant for learning a positive relation, and the occurrence of either event alone is relevant for negative correlations, the mutual co-absence of both events is less intuitively relevant for either type of correlation. However, the absence of events is not the absence of information (the experiments described in Chapter Three are reported in Castiello et al., 2022). For example, in determining the cause of stress-related headaches, the absence of stress and headaches might be informative for inferring the relation between stress and headaches (Murphy et al., 2017).

To consider how contingent relations are formed, it is worth considering the two well-established empirical factors for establishing the perception of a relationship between two events (*e.g.*, E1 and E2) which are: (1) the temporal and spatial contiguity, and (2) the statistical dependency or contingency which defines a correlative relationship. The first factor has been identified to be necessary, if not sufficient, for associative learning and memory co-activation (Miller & Barnet, 1993; for a review, see Savastano & Miller, 1998); a stimulus will become associated with other stimuli and be able to retrieve that stimulus if presented in the same moment in time and space (Molet et al., 2010). The second factor is contingency, a descriptor of the correlation between the two events. When two stimuli are positively contingent, there is a greater likelihood of observing E1 and E2 together (*i.e.*, conjunction or co-occurrence) and absent together (*i.e.*, co-absence). Seeing only one stimulus without the other characterizes negative contingencies (*i.e.*, disjunction). The contingency metric has been formalised by ΔP (delta p; Allan, 1980), the one-way relation between two binary (present/absent) variables. ΔP has been widely used as an empirical model of learning and as a proxy for causal sensitivity (Baker et al., 1993, 2005; Dickinson et al., 1984b; Shanks & Dickinson, 1987; Vallée-Tourangeau et al., 2005a).

1.2.2. A Formal Theory

1.2.2.1. The General Rule

The ΔP contingency between two events, see **Table 1.1**, is a function of the number of co-occurrences of E1 and E2 (A trials), the number of occurrences of E1 alone (B trials), the number of

occurrences of E2 alone (C trials), and the number of co-absences of E1 and E2 (D trials). First described by Allan (1980) as a normative metric for human learning defined as the difference in the conditional probability of the presence of E2 given the presence of E1 [p(E2|E1)], and the presence of E2 given the absence (~) of E1 [p(E2|~E1)]. Hence:

$$\Delta P = p(E2|E1) - p(E2|{\sim}E1) = \frac{A}{A+B} - \frac{C}{C+D}, \qquad \text{(eq. 1.1)}$$

When there is no relation between the two events, that is, ΔP = 0, E1 provides no information about the occurrence of E2. However, if ΔP > 0, the presentation of E1 signals an increase in the likelihood of E2 relative to the baseline occurrence of E2, and if ΔP < 0, the presence of E1 indicates a decrease in the likelihood of E2 relative to the baseline occurrence of E2.

Table 1.1. Contingency Learning Theory applies for both Pavlovian and instrumental learning

Pavlovian Learning			**Instrumental Learning**		
	Outcome	~Outcome		Reward	~Reward
Cue	A	B	Response	A	B
~Cue	C	D	~Response	C	D

Generalization: Events

	E2	~E2
E1	A	B
~E1	C	D

Note: where ~ means absence of. E1 and E2 are two distinct events (*e.g.*, cues, outcomes, or responses).

Perhaps not surprisingly, increasing the number of times the two events are paired together (A trials), everything else being equal, increases judgments of a positive relation (*e.g.*, Murphy et al., 2011). Perhaps less intuitively, according to ΔP, augmenting the number of times that both stimuli are absent together (D trials) should also heighten their contingency and by extension perceived association. In contrast, presenting either event alone (B or C trials), is consistent with a negative contingency, and a decrease in the perceived association (Murphy et al., 2021).

Previous experiments have shown that the four types of events or trials are not equally effective in changing people's judgments (*e.g.*, Murphy et al., 2021). Particularly, a form of D trial

neglect has been found, such that participants tend to give weight to the events in the following manner: A > (B ≈ C) > D (Kao & Wasserman, 1993; Wasserman et al., 1993). This selectivity for positive evidence has also been found in other behavioural and cognitive tasks in which participants show sensitivity to the relevance of all types of information but regularly demonstrate a bias towards confirmatory evidence and away from so called negative evidence (Wason, 1968). I explored this form of neglect in Chapter Three.

1.2.2.2. Contingency Computations

In addition to a normative model like ΔP, CTL can also be described from a process-based perspective, where learning the contingency is described in a trial-by-trial basis. The RW model (Rescorla & Wagner, 1972) provides associative strength calculations that converge on the value of ΔP (Baker et al., 1989). The RW model calculates the associative strength $\mathbf{v}_t$[2] at every next trial $t+1$ between the cues $\mathbf{c}_t$ (present/absent) and a particular outcome λ_t at trial t. The model is as follows:

$$\mathbf{v}_{t+1} = \mathbf{v}_t + \Big(\theta\big(\lambda_t - (\mathbf{c}_t \bullet \mathbf{v}_t)\big)\Big)\,\mathbf{c}_t, \qquad \text{(eq. 1.2)}$$

where the θ represents a LR parameter(s). Usually, RW model assumes that $\theta = \alpha * \beta_0 * \beta_1$, which is the product between LR (α), non-reward (β_0), and reward sensitivity (β_1). Although, the RW model has some limitations in Chapter Five I describe in detail other associative and Bayesian learning models.

Finally, CTL has been used to study individual differences for example to understand individual differences, including depression and schizotypy (Byrom, 2013; Haselgrove & Evans, 2010; Murphy & Msetfi, 2014). Additionally, Haselgrove & Evans (2010) proposed that high schizotypy scorers showed a KB reduction due to a lack of a cue interaction component (or cue competition) as described by the RW model. From eq. 1.2, high schizotypy scorers may not update $\mathbf{v}_t$ using the component $\mathbf{c}_t \bullet \mathbf{v}_t$, which makes the cue competition possible, *i.e.*, sum of all present cues' associative strength. In computational models, this may be more like a Bush-Mosteller or Q-learning

[2] Lowercase and bold letter represent column vectors.

model, *i.e.*, which does not compute cue competition (Bush & Mosteller, 1951; Sutton & Barto, 2018).

1.3. Hypotheses and Summary

People with high schizotypy scores exhibit impairments at various perceptual levels −from low-level visual contrast sensitivity and sensory attenuation to high-level apophenia beliefs and perception of causality−, thus in the first experiment I tested whether perceptual differences emerged in the milliseconds timeframe when assessing contingencies. Therefore, the first hypothesis (Chapter Two) was whether participants scoring high in schizotypy would exhibit a higher detection of positive contingencies between stimuli. Also, I previous evidence suggested that SZ require more sensory evidence for decision-making, for instance larger DTs. On one hand, if people scoring high in schizotypy exhibit larger DT, this would support previous results suggesting sensory evidence is trusted less in this population and that JTC does not apply in short time scales. On the other hand, if schizotypy is related with shorter DTs, this increases the generalizability of JTC to perceptual contingency paradigms.

In addition, studies have found associative learning impairments in schizotypy and SZ. These studies mainly focus on cue competition, attention to irrelevant cues, and aberrant salience. Although, with specific relevance for judgments and beliefs, evidence also showed cognitive biases against disconfirmatory or confirmatory evidence. Thus, I used an associative learning paradigm with contingency assessment to study information relevant for cognitive biases and belief formation as the absence of events and disconfirmatory logical evidence. I had two hypotheses (Chapter Three), first I hypothesized that people scoring high in schizotypy would be less sensitive to trials with absent stimuli. Second, I hypothesized that frequency, not duration, of trials with absent stimuli enhances perceived contingency.

Contingency learning in instrumental paradigms have been linked with people with diagnosis of SZ and scoring high on schizotypy. For example, feedback sensitivity and awareness of consequences of actions. Thus, these impairments could be related with how people experience agency. Given that

psychotic symptoms are related with external forcers controlling one's mind and referential ideation. In last two experiments I tested PC in a free-operant task and hypothesized that illusion of control would be correlated with positive schizotypy scores. Also, I tested whether illusion of control is related with referential behaviour, for example, an upweight of responses (internally generated) or downweight of an external source of control. Finally, as a complement to the differential weighting possibility, I tested weather schizotypy moderates the relationship between responses and PC.

In summary, differences in perception, learning, and agency have been associated with schizotypy. In the next chapters, I present results relating schizotypy with perceptual learning (Chapter Two), associative learning (Chapter Three), instrumental learning, agency, and sense of control (Chapter Four). For the last chapter I conducted a narrative review and systematically extract the results learning papers on schizotypy, and SZ. Also, I review the concept of weighting and compare two modelling approaches (Bayesian and Reinforcement/Associative Learning). Lastly, I conclude by incorporating my results to the current evidence in schizotypy and learning.

Chapter Two

Perceptual Contingency Learning in Schizotypy

2.1. Introduction

In this chapter I investigated the perception of contingency using a psychophysical approach, *i.e.*, two-alternative forced task. Where trials were contingencies between two cues in rapid sequence of events −in the millisecond to second domain−. The research focus was to study hyperassociability in schizotypy.

2.1.1. Hyperassociation and Schizotypy

The perception of a contingency is much like any other perceptual task except that the events that define the presence of such a relation are not predictable on the basis of a single event. In some signal detection tasks the goal is to acknowledge the perception of the presence or absence of an event, and the two types of error that can be performed (misses and false alarms) in combination with the success (hits and correct rejection) are an indication of sensitivity and detection. Contingencies have shown similar properties as other physical dimensions (Allan et al., 2008; Perales et al., 2005), but the difference is that detection involves a macrotimescale to include all the trials of the experiment. Each trial presents a single piece of evidence and over time the participant increases their accuracy in judging the presence or absence of the relation. Research suggests that some of the properties of the task are isomorphic to instance the underlying gaussian noise using SDT (Maia et al., 2018). Thus, in this chapter I studied contingency learning from a perceptual learning perspective. This is important, because understanding the perceptual bases of contingencies may inform how people use event types conforming contingencies, as well as causation (Baker et al., 2005; Dickinson, 2001).

There are several possible areas of research with differing response requirements where the effects of schizotypy have been observed to involve perceiving contingencies. For instance, apophenia is the tendency to perceiving a connection when there is none which has been described as a form of type I errors (false positive; Blain et al., 2020). For example, someone with SZ may believe that a commentator on the TV is communicating with the patient's parent because they both look at their mobile phones at the same time. Pareidolia is similar to apophenia but exclusively with the visual system, *e.g.*, we all see more than clouds when we look at them. Interestingly, apophenia is related to psychoticism personality traits (Blain et al., 2020) as well as in schizotypy and delusional proneness (Fyfe et al., 2008). My first hypothesis is that a similar effect will emerge with schizotypy, an enhanced perception for detecting positive contingencies (*i.e.*, $\Delta P > 0$), *i.e.*, hyperassociability, a predisposition to perceived positive associations between stimuli. If this result was found, it would provide a method to study this tendency in a lab setting. Further an apophenic-like precepts occur at a very low level, *i.e.*, perception in the millisecond range, would generalise the effect from naturalistic experiences and suggest that the processes are fundamental.

Enhanced learning to irrelevant stimuli in schizotypy (Haselgrove et al., 2016; Haselgrove & Evans, 2010; Le Pelley et al., 2010) indicates an easiness for creating associations. This result could occur due to an upweighting of perceived cues as hyper salient (Kapur, 2003), *i.e.*, A trials (**Figure 2.1**, both $S1$ and $S2$). Although is learning impaired due to saliency or is it memory what is impaired? Schizotypy has shown impairments in episodic memory (Sahakyan & Kwapil, 2018), and in SZ working memory seems to be the critical factor impaired in learning tasks (Collins et al., 2014, 2017). The task used in this chapter is immune to deficits in memory due to its low temporal resolution. So, participants were asked to detect contingencies in the millisecond timeframe.

Another interesting phenomenon found in both clinical and subclinical schizotypy is the Jumping-to-Conclusion (JTC) bias (Dudley et al., 2016; Juárez-Ramos et al., 2014; Ross et al., 2015a; Woodward et al., 2007). JTC is when people decide or conclude something with little or not enough data about the circumstance. It seems that this bias is not due to an overestimation of the gathering information cost, but a high noise in decision-making (Moutoussis et al., 2011). The task that I used

for this chapter has a component that could be used for understanding JTC, described in the next section. The current task can be seen as a JTC-like paradigm where participants saw streams of information relevant for the contingency. Participants are asked to stop the stream when they perceive the contingency, so they control the amount of information they receive.

2.1.2. The Psychophysics of Contingency

Contingencies (ΔP; Allan, 1980) as with physical dimensions, *e.g.*, brightness or motion, exhibit psychophysical properties such as the psychometric function (Allan et al., 2008). Classical psychophysical methods and models can be used to the study of contingencies, such as SDT (Maia et al., 2018; Perales et al., 2005) and DDM. On one hand, the use of SDT permits the research to dissect discriminability (*d'*) from the response bias (*C*); similar to the steepness and inflexion point from a psychometric function. On the other hand, using DT in a DDM makes it possible to differentiate between multiple components in perceptual decision-making process (Ratcliff, 1978; Voss & Voss, 2007, 2008), such as: sensory evidence accumulation threshold (*a*), sensory evidence accumulation rate (*v*), non-decision time (*t0*), and the perception bias or sensory evidence accumulation starting point (*zr*).

To understand how contingencies are perceived over a very short temporal scale (*i.e.*, milliseconds to seconds), I used both analytic models described above. Interestingly, the DDM has been tested previously in people with SZ. Results suggested an increase in prior beliefs toward detecting a signal, *zr* (Limongi et al., 2018). Also, SZ exhibit an increase in sensory evidence thresgold, *a* (Moustafa et al., 2015). Therefore, my second hypothesis is that high schizotypy scorers would gather more evidence to make correct choices. The rational is, as argued by Moustafa et al (2015), people with SZ learn to reduce certainty on their sensory evidence. Therefore, to correctly solve tasks high schizotypy scorers need more sensory evidence. Also, this finding would support a continuum dimensional perspective between SZ and schizotypy. Although, this is a counterintuitive result given JTC, which I discuss in depth Chapter Five.

I developed the rapid-trial streaming procedure (Crump et al., 2007) consisting of a quick presentation of one of four types of information or event types (A: $S1$ & $S2$; B: $S1$ & $\sim S2$; C: $\sim S1$ & $S2$; and D: $\sim S1$ & $\sim S2$; **Figure 2.1**). I modified the task, so participants do not need to wait until the end of each stream to make their choice. This is an attempt to easily detect differences in the amount of information participant needed to classify contingencies. Participants were asked to correctly classify contingencies (trials) as Strong Positive or Strong Negative. Research suggests that people weight the A event more than B and C, and those more than D (Kao & Wasserman, 1993; Murphy et al., 2021; Wasserman et al., 1993). Thus, with the current chapter's data I presented a novel analytical approach to estimate the weights that previously has been hypothesised to be A > B = C > D.

The perceived weights of the four event types may differ between individual contributing to the cognitive symptoms of SZ. For example, high schizotypy is related with the biases against disconfirmatory evidence (BADE; Juárez-Ramos et al., 2014; Woodward et al., 2007), where the disconfirmatory evidence of an association between two stimuli may be understood as the using the absence of information. Hence, high schizotypy scorers may downweight the absence of events, *e.g.*, D events. Furthermore, in a different subclinical population, depressive traits are related with a reduce weight of contextual information (Castiello, Senan, et al., 2020; Msetfi et al., 2005, 2017).

In summary, first, I hypothesized that schizotypy would be related with an increase in perceiving positive contingencies, *i.e.*, hyperassociation. This would be manifested with more detections of a positive contingencies, as well as with a large response bias *C* (SDT) and a larger starting point *sr* (DDM). Second, I hypothesized that DT would be longer in schizotypy. Although shorter DT would be congruent with a JTC explanation in a very short timeframe scale. Third, I predicted a positive correlation between schizotypy and DDM threshold. Fourth, I conducted exploratory analysis between subscales of schizotypy and computational and behavioural parameters to assess if the main hypothesis are exclusive only of total schizotypy scores. Finally, I was interested in addressing the effect of specific event types (A, B, C, or D) in perceptual decision-making. Thus, I presented five stages of analysis described in detail below.

2.2. Methods

2.2.1. Participants

A total of 126 participants from four samples, (**Table 2.1**) were collected. Mean age of 25.5 (SD 7.88) and 69 F (55%). However, ~23 % were excluded (F 17 and M 12) from the analysis. This was due to invalid trials, which I used as part of the exclusion criteria. The inclusion criteria were: *i)* having more than 60% of correct trials overall, *ii)* at least 100 valid trials, defined as trials with a DT between 125 ms and 8000 ms, which correspond to the minimal time for seeing an event (**Figure 2.1**) and the duration of a whole trial stream video (*i.e.*, 8 sec), respectively. The exclusion criteria were critical to assure that participants were exposed to the contingencies. A final sample of ninety-seven participants were included in the analysis. I compared the included versus excluded participants and there were no important differences between them (**Table 2.A1**). The ethics committee of the University of Oxford (R60840/RE001) approved the research.

Table 2.1. Sociodemographic and General Characteristics Between Samples from Experiment 1

	Experiments subsample comparisons				
Sample name:	Students	Online	Multiple	Students	Total
N, # (%)	68	25	18	15	126
Age, M (SD)	23.4 (8.32)	33.28 (5.37)	23.25 (2.14)	23.53 (2.26)	*25.46 (7.88)
Sex F, # (%)	46 (67.65)	5 (20)	10 (55.56)	8 (53.33)	*69 (54.76)
Included participants, # (%)	55 (80.88)	14 (56)	17 (94.44)	11 (73.33)	*97 (76.98)
Depression, M (SD)	14.85 (11.65)	8.12 (11.07)	-	-	*13.04 (11.83)
Schizotypy, M (SD)					
Total	15 (7.9)	13.08 (8.63)	13.89 (5.21)	13.07 (6.45)	14.23 (7.54)
Cognitive Disorganization	5.63 (2.86)	1.72 (2.57)	5.33 (2.54)	4.53 (2.23)	*4.68 (3.07)
Introvertive Anhedonia	2.46 (2.17)	5.12 (2.4)	2.72 (2.08)	2.13 (1.73)	*2.98 (2.39)
Impulsive Nonconformity	3.72 (2.44)	3.28 (2.84)	2.39 (1.24)	2.87 (2.5)	3.34 (2.43)
Unusual Experiences	3.19 (2.72)	2.96 (2.85)	3.44 (2.28)	3.53 (1.96)	3.22 (2.59)

Note: # = frequency; % = percentage; M = mean; SD = standard deviation; F = female. * Represent a p value < 0.05 for a test between samples (ANOVA for continuous and chi-squared for categorical variables).

2.2.2. Design and Procedure

After responding on the short O-LIFE version (O. Mason et al., 2005) and the BDI-2 (Beck et al., 1961; only for two samples see **Table 2.1**) questionnaires, participants performed the experimental

task (see **Figure 2.1**). See the comparison of the O-LIFE scores of our samples against the normative data in **Table 2.A2**.

The perceptual contingency decision-making task that I used was inspired by the rapid-trial streaming procedure from Crump et al. (2007). Each stream consisted of 8000 ms presentation of four events (A, B, C, or D; see **Figure 2.1**) which were randomized without replacement within every 125 ms bin. Therefore, each video has 64 bins and the whole task consisted of 180 such trials in which two stimuli held a contingency established by their ΔP, *i.e.*, p($S2|S1$) - p($S2|{\sim}S1$) (**Figure 2.1**). I used nine different contingency conditions, varying from -1 to 1 in 0.25 increments ΔPs. The trials were quasi-randomized presented in 20 blocks of nine trials, each block contained one trial per condition. The nine conditions' ΔP where fully randomized within a block/.

Participants were asked "*to decide whether you think the shapes and drawings are POSITIVELY or NEGATIVELY associated with each other. Use the keyboard to respond and press "c" when you think they are NEGATIVELY associated and "m" when you think they are POSITIVELY associated.*" For more details see **2.5.3. Appendix**. Each trial stopped immediately after participants made a response. Feedback was provided when the stream ended, *i.e.*, after participants responses.

Participants received feedback, correspondingly "correct" when they produced a Strong Positive response for positive contingencies and Strong Negative response for negative contingencies, otherwise they read "incorrect". In the case of $\Delta P = 0$, half correct were Strong Positive and the other half Strong Negative, *i.e.*, false feedback given that 0 ΔP is not either positive or negative. If participants did not make a response before the end of each trial, then ended with a blank screen and until participant responded, next trial was presented. If a participant responded at this point, the trial was excluded, and next trial started. The time from the trial onset to the response is called DT. Given the right skewed distribution of DT, these analysis were conducted using the natural logarithm of the DT [*ln*(DT)].

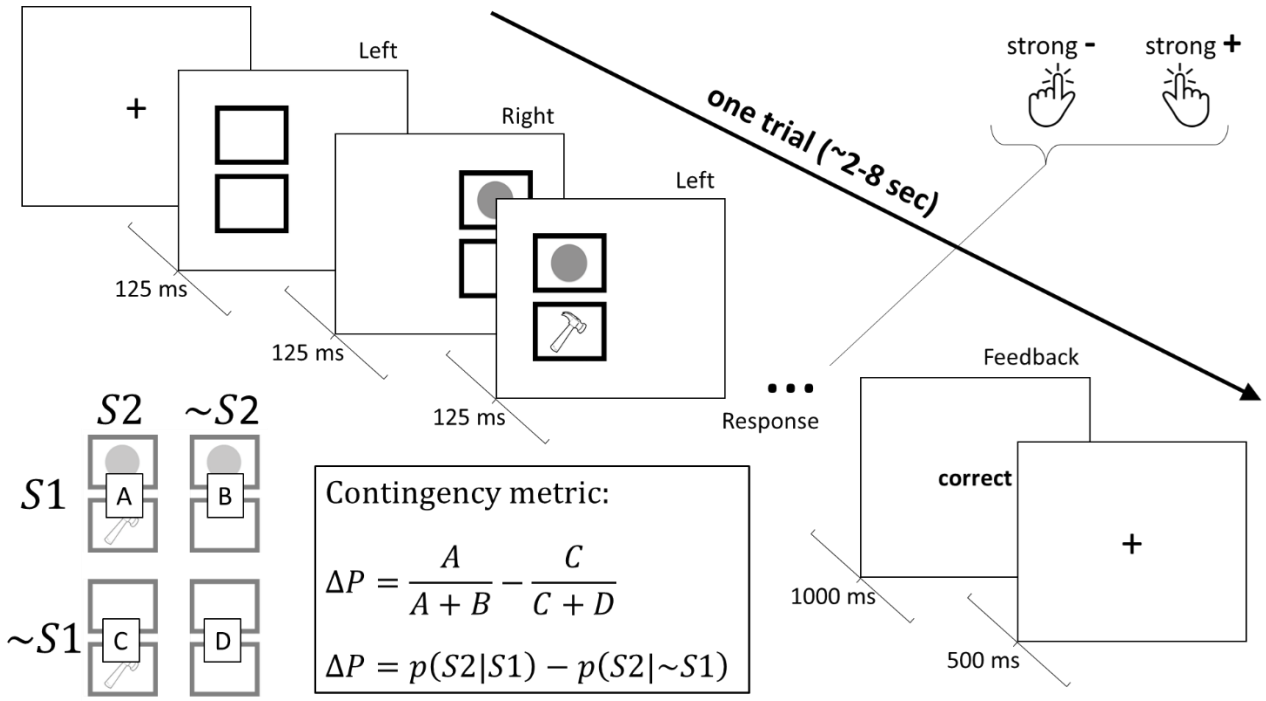


Figure 2.1. Paradigm used in Experiment 1. Schematic of the psychophysics contingency task, or rapid-streaming procedure (Crump et al., 2007). Bottom left there is a contingency table between stimulus 1 ($S1$) and stimulus 2 ($S2$), where ~ means absent. Bottom middle in a black square we showed how to calculate the contingency (ΔP) between $S1$ and $S2$.

2.2.3. Data analysis plan

I conducted analysis in five stages. First, modelling responses with logistic mixed model. Second, a LMM predicting DT logarithmically transformed [ln(DT)] and sensitivity analysis with gamma link function. Third, a computational dissection using DDM. Fourth, exploring all the computational and behavioural parameters and their relationship with schizotypy subscales and questionnaires. Finally, for the fifth stage I explored the effect of individual cells in the milliseconds immediately before the response.

I used questionnaires scores as continuous variables and I mean-cantered them by subtracting the mean and dividing them their SD. However, DT was an exception because I transformed it to the natural logarithmic space.

All models, logistic, linear, and gamma, had the same random effect structure where responses were nested by participants and participants nested by samples. Finally, to assess the regressors' statistical significance I used the 95% CI of the coefficients (Coef.).

2.2.3.1. Model Strong Positive Responses

The first stage of the analysis was to assess the main hypothesis regarding whether schizotypy exhibited an enhanced perception of positive contingencies. I used a logistic mixed model to predict Strong Positive responses. This model included conditions (ΔP, as continuous), scaled schizotypy (continuous), and natural logarithm of DT [*ln*(DT); continuous], as well as their respective interactions. The main effect of schizotypy would corroborate first hypothesis. Furthermore, the interaction between ΔP and schizotypy would suggest a differential sensitivity to classify contingencies in schizotypes. Finally, the third order interaction or schizotypy and *ln*(DT), would suggests a further examination of the *ln*(DT). See next section Model Decision time.

2.2.3.2. Model Decision Time

For the second stage of analysis, I examined DT by fitted a LMM predicting the natural logarithm of DT [*ln*(DT)]. The regressors were, feedback (correct = 1), response type (Strong Positive = 1), and scaled schizotypy. Then, I ran a sensitivity analysis for the DT model, by fitting a generalized LMM, with a Gamma link function as is appropriate for untransformed data (Lo & Andrews, 2015). No difference in the direction of the effect was found between modelling linearly *ln*(DT) and directly DT with Gamma link function.

2.2.3.3. Psychophysical and Computational Modelling

For the third stage of analysis, I extracted parameters from psychometric curve, SDT, and DDM (see below). Even though, SDT and psychometric parameters are practically the same, I decided to do this from an atheoretical perspective because all the parameters would be used for the exploratory analysis of the fourth stage.

2.2.3.3.1. Psychometric Curve

To calculate the bias towards perceiving a positive or a negative association (inflexion point: i) and the contingency conditions' discriminability (function steepness: s), I fit a psychometric or sigmoidal function to each participant:

$$p(Strong\ Positive\ response) = \frac{1}{1+exp\langle -s * (\Delta P - i)\rangle}, \quad \text{(eq. 2.1)}$$

where s is the sigmoidal slope which represent the function slope (near to 0 values mean no ΔP sensitivity, so participants respond 0.5 for negative and positive ΔP), and i is the inflexion point and represents an overall bias towards pressing one button over the other.

2.2.3.3.2. Signal Detection Theory (SDT)

I quantified how well participants discriminated positive from negative contingencies and how biased their responses were by using SDT. I calculated detection sensitivity (*d'*) and response criterion (*C*) as follows:

$$\text{Detection Sensitivity} = d' = z[p(h)] - z[p(fa)]; \quad \text{(eq. 2.2)}$$

$$\text{Response criterion} = C = -1 * \{z[p(h)] + z[p(fa)]\}/2; \quad \text{(eq. 2.3)}$$

where $p(h)$ and $p(fa)$ are called hit rate and false alarm rate, respectively, and are calculated as follows:

$$\text{Hit Rate} = p(h) = \frac{hit}{hit + miss} \quad \text{(eq. 2.4)}$$

$$\text{False Alarm Rate} = p(fa) = \frac{false\ alarm}{false\ alarm + correct\ rejection}; \quad \text{(eq. 2.5)}$$

where *hit* is detecting a $\Delta P > 0$ when there is actually $\Delta P > 0$; *miss* is detecting $\Delta P < 0$ when the trial presented a $\Delta P > 0$; *false alarm* is detecting a $\Delta P > 0$ when there is a $\Delta P < 0$; and correct rejection is detecting a $\Delta P < 0$ when there is actually a $\Delta P < 0$ (Castiello et al., 2022; Perales et al., 2005).

2.2.3.3.3. Drift Diffusion Model (DDM)

This perceptual computational model assumes that sensory evidence is accumulated as a stochastic process (Brownian motion) and the agent decides when to respond when the process reaches one of the two −upper and lower− boundaries (Ratcliff, 1978). The calculated parameters

were the sensory evidence accumulation threshold (a); the sensory decision bias or evidence starting point (z); the non-decision time ($t0$); and the stochastic process tendency or drift rate (v). To fit the DDM, I used the *fast-dm.exe* ('fast method for diffusion model analysis; Voss & Voss, 2007) and a, z, $t0$, and v were extracted for each participant.

2.2.3.4. Schizotypy Subscales and Behavioural Parameters

This fourth stage of analysis was exploratory. I ran two models to predict all the possible dependent variables (DV) individually. The first model includes schizotypy subscales as a regressor, the second includes total schizotypy and total depression scores. The dependent variables are clustered in five: DDM, SDT, Psychometric curve, general task parameters, and general task parameter split by correctness. I estimated a total of twenty DV individually with the two models described above. See below:

1. DDM (4 dependent variables): *a*, *zr*, *t0*, and *v*.
2. SDT (2 dependent variables): *C*, hit rate, false alarm rate, and *d'*.
3. Psychometric or sigmoidal function (2 dependent variables): steepness, and inflexion point.
4. General Task (4 dependent variables): probability of strong positive response [p(positive response)], probability of correct [p(correct)], mean of ln(DT) [mean(ln(DT))], and mean of DT [mean((DT)], −these last two are not the same (Lasserre, 1991)−.
5. Given the interaction schizotypy and correct responses, the last cluster incorporates the previous (4) general task, but split it by correct and incorrect trials, therefore (6 dependent variables): probability of strong positive response for correct [p(pos. resp. | correct)] and for incorrect trials [p(pos. resp. | incorrect)]; the mean of ln(DT) for correct [mean(ln(DT) | correct)] and incorrect [mean(ln(DT) | incorrect)] trials; and finally, the DT average for correct [mean(DT | correct)] and incorrect [mean(DT | incorrect)] trials.

For each DV within each cluster, I ran a linear model with the schizotypy subscales as regressors. Using the four O-LIFE subscales in a model has been used (Evans et al., 2007; Granger et al., 2012, 2016). I did not find any significant relation with Bonferroni correction, 0.05/20 = 0.0025.

So, as a purely exploratory exercise I used α of 0.05. Thus, in the former case significant results should be not taken as strong evidence.

2.2.3.5. Information Type Weights Analysis and Recency Analysis

In addition, I was interested in evaluating the influence of each event type on the responses. To achieve this, I conducted a Recency analysis, where I fitted logistic regressions to predict responses (Strong positive = 1, Strong negative = 0) with event type as factor regressor (A, B, C, and D) for every 125 ms before the response for the last 1000 ms of each trial. The regressors were fitted for every participant and for every one of the last 12 (125 ms each bin t) events seen just before participants responses. I calculate the event type weights as followed.

$$log\left(\frac{p(strong\ resp.)}{1-p(strong\ resp.)}\right) = W_A + \beta_B \mathrm{B}_t + \beta_C \mathrm{C}_t + \beta_D \mathrm{D}_t\ , \qquad \text{(eq. 2.6)}$$

where B_t, C_t, and D_t are dummy variables and encode the information event identity (A, B, C, or D) in a specific bin t, before the response. The reference event type was A, therefore, the weight for A (W_A) is the intercept. Moreover, the weights for B, C, and D were calculated by adding the dummy variable to the intercept; $W_B = W_A + \beta_B$, $W_C = W_A + \beta_C$, $W_D = W_A + \beta_D$, respectively.

2.2.3.6. Code and open access

All analysis and visualizations were made using R (R Core Team, 2021). The scripts used for visualizations and statistical analysis, as well as the data files are available by request to the authors. To fit the mixed models, I used the *lme4* (Bates et al., 2015) and complemented with *lmerTest* (Kuznetsova et al., 2017). Additional packages used for data processing were *reshape2*, *hmisc*, *devtools*, and *dplyr* (Harrell Jr, 2021; Wickham, 2007; Wickham et al., 2022; Wickham H et al., 2021); for exporting models to csv tables, I used *report* (Makowski et al., 2021), and data visualization, *ggplot2* (Wickham, 2016), *viridis* (Garnier et al., 2021), and *ggpubr* (Kassambara, 2020).

2.3. Results

2.3.1. Differences in Samples

I compare the four subsamples (**Table 2.1**), and there was a difference in age, where the Online sample (n = 25) had a larger mean 33.28 versus 23.40 from the other samples (t = 7.6, p < 0.05). Also, the online sample contained fewer woman (F 5) and fewer participants from this sample were included (n = 14). Also, from the two samples with depression scores, the Online sample scored lower (t = -2.56, p < 0.05). Finally, with respect to the schizotypy scores, there was no difference in total O-LIFE scores (t = -0.76, p = NS), nor IN (t = -0.12, p = NS) and UE (t = -0.52, p = NS). Although, the Online sample showed reduced scores in CD (t = -6.35, p < 0.05) but larger in IA (t = 5.0, p < 0.05).

With respect the behavioural differences, I ran a logistic mixed model per sample (**Figure 2.A1**). I did not find any main effect on schizotypy but in the Multiple sample (multiple sources, online, lab, and personal invitations) I found a significant interaction between contingency and schizotypy. This suggests a reduced sensitivity to classify contingencies only in that sample. Even with differences between samples, I pooled them to conduct a main analysis with a larger sample size (n = 97).

2.3.2. Confirmatory Mixed Models

2.3.2.1. Modelling Responses

I pooled the four samples and present results from 97 participants. For visualization purposes I split the group in terms of schizotypy terciles (**Figure 2.2A**). For the mixed models I scaled schizotypy scores and used it as a continuous variable. Overall, participants classified the two contingencies with accuracy, *i.e.*, linear relation between ΔP and probability of Strong Positive responses. Moreover, I found that high schizotypy scorers required more time to correctly perceive contingencies.

Figure 2.2A shows the model estimates. This confirmed the previous assessment given the main effect of conditions (ΔP), Coef. = 2.35 [95% CI from 2.27 to 2.42]. No main effect of schizotypy was found, Coef. = -0.01 [-0.09 to 0.07]. However, *ln*(DT) was negatively related with Strong Positive responses, Coef. = -0.24 [-0.31 to -0.17]. The significant interactions were ΔP x DT, Coef. = 0.35 [0.24 to 0.47]; and the triple interaction with schizotypy, Coef. = 0.13 [0.02 to 0.25] (**Figure 2.2B**), no other interactions were significant. Furthermore, we compared the previous model (deviance = 19180; **Figure 2.2B**) against a reduced model with no schizotypy as a regressor (deviance = 19186). The test suggests that schizotypy does not increase the goodness of fit, $X^2(4) = 5.38, p = 0.25$.

2.3.2.2. Modelling Decision Time

To test the second hypothesis, concerning whether schizotypy is related to requiring more sensory evidence to make decisions and to explore the triple order interaction found in **Figure 2.2B**. I modelled *ln*(DT). First, I visualized *ln*(DT) as a function of schizotypy (**Figure 2.2C**), and split Strong Negative (left panel) from Strong Positive (right panel) responses, then separated using colours to display accuracy for either correct (green) or incorrect (red) trials. **Figure 2.2C inset** shows the same response type (Strong Negative and Strong Positive) and accuracy (correct and incorrect), but no schizotypy.

The results reveal that schizotypy is related with an increase in *ln*(DT) only for correct trials. To test this hypothesis, I fitted a LMM to predict *ln*(DT), which revealed the positive interaction between correct and schizotypy, which means that on correct trials schizotypy is related with a larger *ln*(DT), Coef. = 0.03 [95% CI from 0.00 to 0.05]. The model also suggests that overall correct trials are faster, Coef. = 0.03 [95% CI from 0.01 to 0.06], and a negative interaction between correct and Strong Positive response, Coef. = -0.12 [95% CI from -0.16 to -0.09]. No other significant main effect nor interaction was found. Nonetheless, when I compared the previous model (deviance = 28770; **Figure 2.2D**) against a reduced model with no schizotypy as a regressor (deviance = 28783), the likelihood ration test suggests that schizotypy increase the goodness of fit, $X^2(4) = 12.51, p < 0.05$.

Finally, I conduced a sensitivity analysis modelling the raw DT with generalized mixed model with a gamma link function (see **2.5.5. Appendix 2**), the results point to the same direction. The model also exhibits the positive interaction between schizotypy and correct trials. Therefore, transforming DT does not affect the main result.

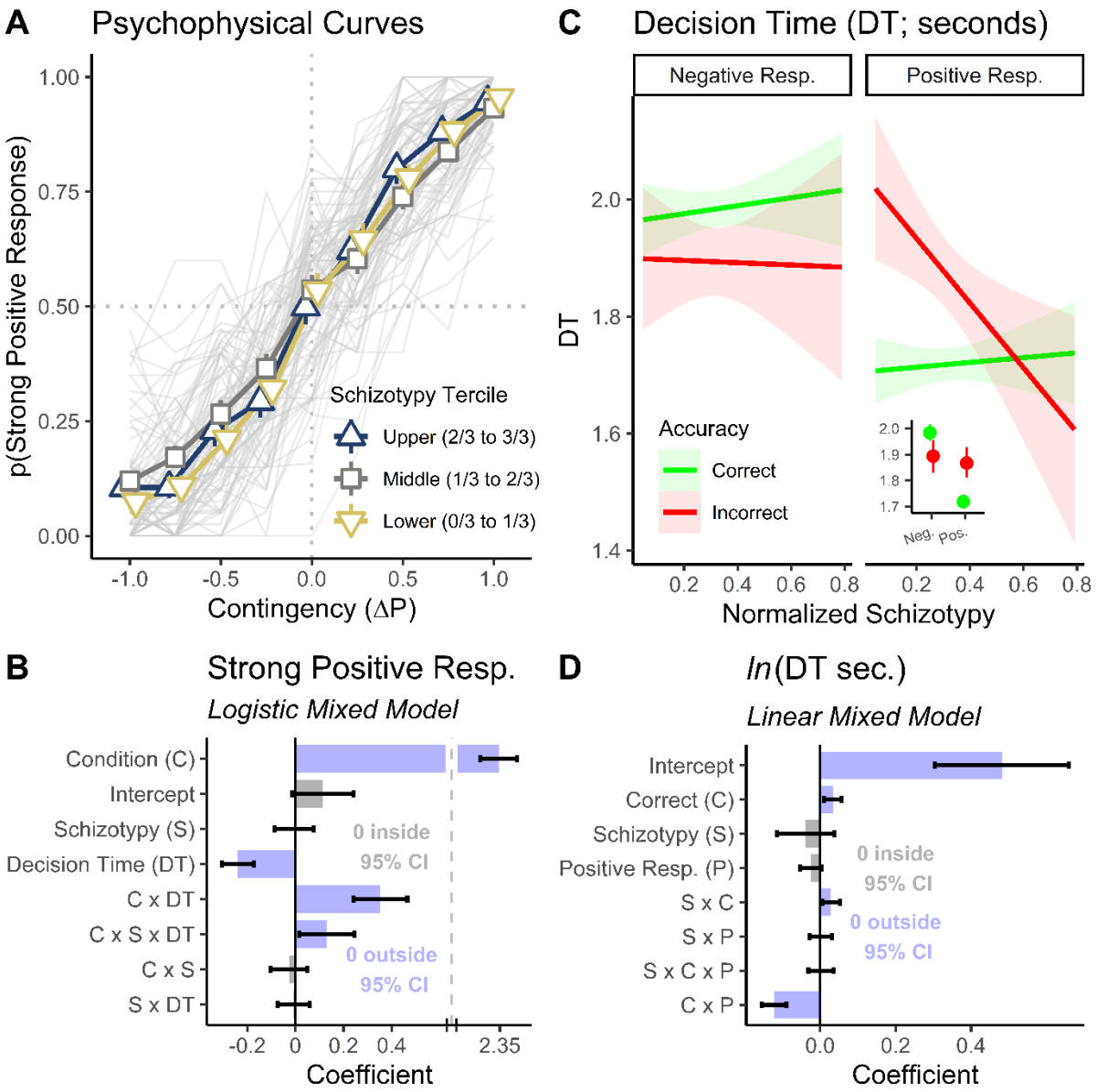


Figure 2.2. Main Results of Experiment 1. (**A**) Probability of responding Strong Positive as a function of contingency condition (ΔP), each grey line represents a participant (n = 97). (**B**) Coefficients (with the 95% CI) of a logistic mixed model used to predict Strong Positive responses with condition, schizotypy, *ln*(DT) and the respective interactions. (**C**) DT as a function normalized schizotypy (score divided by the maximum possible score). Left panel negative responses and right panel positive responses; green colour is the effect for correct trials, and red for incorrect trials. Inset panel shows the grand mean for positive/negative responses and correct/incorrect trials. (**D**) Coefficients (with the 95% CI) of a LMM used to predict the *ln*(DT) with Strong Positive responses, correct trials, schizotypy, and their interactions.

2.3.2.3. Confirmatory Analysis: Summary

Participants correctly classify positive and negative contingencies. Positive contingencies were easier to perceive, as reflected in a reduced DT in positive correct against negative correct trials (**Figure 2.2C inset** and **Figure 2.2D**). Moreover, the results from **Figure 2.2B**, suggest that the task

accuracy does not depend on schizotypy. But the data also suggests that this may be because they spend more time gathering information to get correct trials (**Figure 2.2D**). This support Moustafa et al (2015) evidence and extend those results to general population with high schizotypy scores and counters previous suggestions related to JTC that decisions are speeded up.

2.3.3. Drift Diffusion Model and Perception of Contingency

I fit four DDM: Threshold (*a*) model, Starting Point (*zr*) model, Drift Rate (*v*) model, and non-Decision time (*t0*) model. Each model estimated parameters per condition and per participant. For example, the Threshold (*a*) model estimate one *a* per condition per participant (*i.e.*, nine *a* one for each Δ*P*), and only one *zr*, *v*, and *t0* per participant. The visualization of these four models is displayed in **Figure 2.3**.

I ran four LMM, one by model subpanel from **Figure 2.3**. The aim was to assess schizotypy main effect or the interaction with any of the parameters. Each model has participant as random intercept, and the regressors were condition (Δ*P*), scaled schizotypy, and the interaction. **Figure 2.3A** exhibit a negative relation between *a* and Δ*P* [*t* score = -3.8, $p < 0.001$], no other regressor was significant. A similar result occurred in **Figure 2.3B**, only a positive relation between Δ*P* and *zr* [*t* score = 30.58, $p < 0.001$], no schizotypy nor its interaction. **Figure 2.3C** displayed the same monotonical relation between Δ*P* and *v* [*t* score = 58.22, $p < 0.001$], and no other significant effects. Finally, **Figure 2.3D** showed the same Δ*P* and *v* effect [*t* score = -8.25, $p < 0.01$], no schizotypy effect but a trend to significance in the interaction Δ*P* and schizotypy [*t* score = 1.73.2, $p = 0.09$].

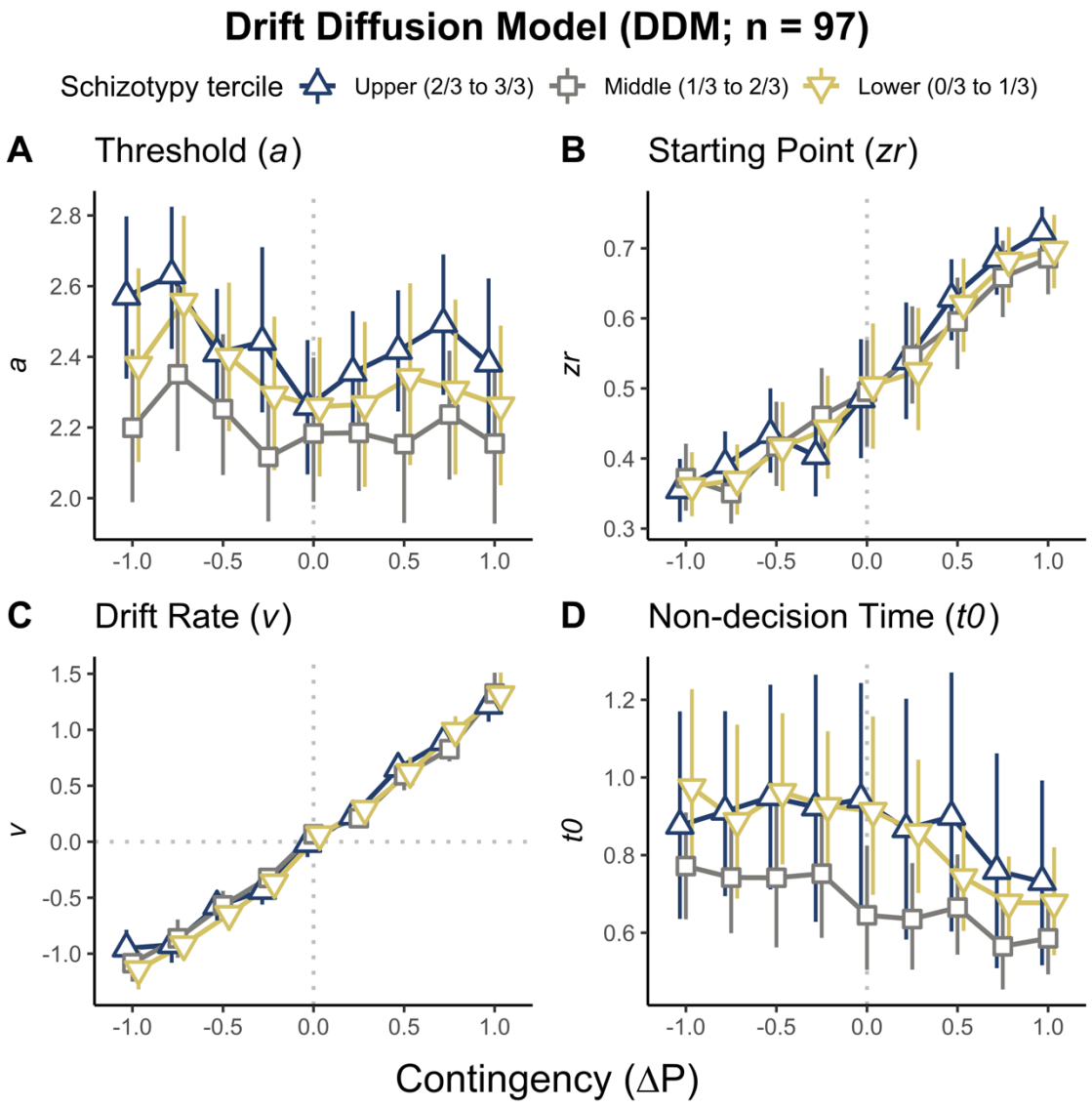


Figure 2.3. Drift-Diffusion Mode (DDM) l for Experiment 1. DDM fitted parameters as a function of contingency. X-axes shows the nine conditions (ΔP) levels. Y-axes shows the DDM parameters. **A**: threshold (*a*), amount of information needed for a choice to be taken. **B**: starting point (*zr*), position where the Brownian motion process starts at each trial. **C**: drift rate (*v*), the evidence accumulation process bias towards one of the two thresholds (*a* or 0). **D**: Non-decision time (*t0*), a time interval where the visual system is not processing the trials.

2.3.3.1. DDM: Summary

Overall, I did not find any specific finding regarding schizotypy and DDM parameters. Although, there are interesting results regarding the differences in perceiving Strong Positive and Strong Negative contingencies. For example, for positive contingencies participants need a lower threshold of evidence accumulation and lower *t0*. Overall, this suggest that positive contingencies are

easier to perceive, and match previous literature (Vallée-Tourangeau et al., 2005b), for instance negative contingencies are less stable and elicit less strong judgments (Msetfi et al., 2007).

2.3.4. Exploratory: Behaviour and Questionnaires

In this section I show how questionnaires relating to behavioural and computational parameters, which I called DV, they are displayed in the y-axis (**Figure 2.4**). I ran two models per DV, model 1 included four schizotypy subscales as regressors, and model 2 included schizotypy and depression total scores as regressors (forty were the total number of models ran).

Using Bonferroni correction, which is appropriate due to the exploratory nature of this section, no significant results were found. Although, I also ran the analysis without using the correction to detect possible, but week, relations. With that in mind, **Figure 2.4**, hinted some interesting effects. Mostly the panels Task and Task Corr. (subdivided by correct and incorrect) unveil the negative relation between IA and the mean of DT (**Figure 2.4** Task), but specifically this seems to be related exclusively with incorrect trials (**Figure 2.4** Task Corr.). Also, it seems that IN −which relates with impulsive, anti-social and eccentric behaviour− is related with a larger steepness of the psychometric function, *i.e.*, better contingency discriminability.

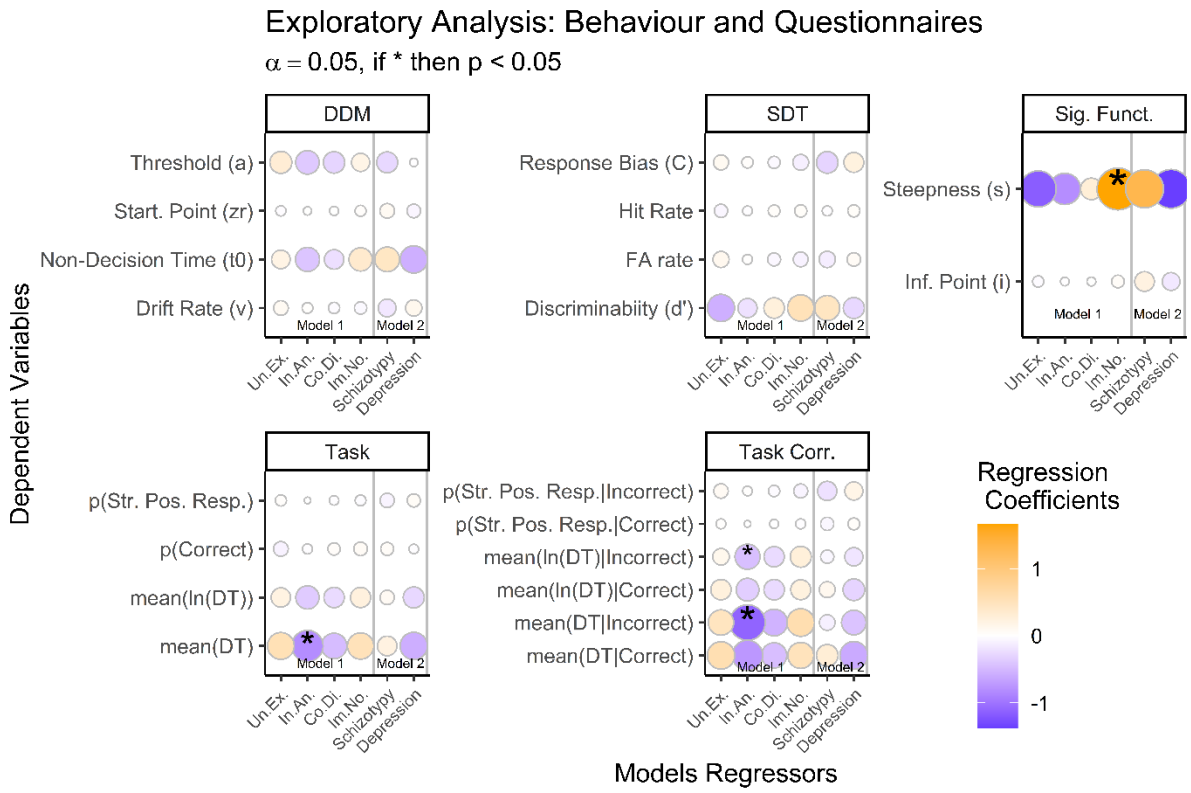


Figure 2.4. Schizotypy Subscales: Exploratory Analysis for Experiment 1. No Bonferroni correction is used. X-axis: Un.Ex. = Unusual Experiences; In.An. = Introvertive Anhedonia; Co.Di. = Cognitive Disorganization; Im.No = Impulsive Nonconformity. Results of an exploratory analysis predicting all (4 + 4 + 2 + 4 + 6) 20 dependent variables (DV) with two sets of models: using schizotypy subscales (Model 1) and using schizotypy and depression total scores (Model 2). The panels represent the method used to calculate the DV (see **2.2.3.4. Methods**). In the y-axis, each panel has DV (behavioural or computational). The regression coefficient' size and direction are shown by the size of the circle (magnitude) and the colour (direction). Overall, given a participant-row data base type, it seems to be little evidence relating schizotypy and behaviour in this task.

2.3.4.1. Exploratory Analysis: Summary

Overall, this may suggest that IA is related with a reduced DT for incorrect trials. This is linked to negative schizotypy (negative symptoms of psychosis), which includes lack of sources of pleasure and avoidant behaviour. For instance, high IA people may have found the task boring, so they just responded quickly to finish earlier. However, this also would imply a negative relation with the task accuracy. However, this was not the case. Finally, IN is related with better contingency discrimination. This result is counterintuitive because IN is related with impulsive, but I do not see how impulsive behaviour an advantage is to classify contingencies better. Given the exploratory nature of this stage of analysis, the results should be taken with caution.

2.3.5. Weighting Differences and Information Type

The goal of this analysis was to assess the different impact that each event type may have contributed to the responses. I modelled the effect of each event type in the 1500 ms prior to the response. I used logistic regressions (Strong Positive response = 1, Strong Negative response = 0) to estimate the weights for each event type (see **2.2.3.5. Methods**). The weights for the stimuli conjunction in both presence and absence events, A and D (W_A and W_D), predicted Strong Positive responses and the inverse effect for stimuli disjunction events, B and C (W_B and W_C; **Figure 2.5A**). Interestingly, I observed that participants were more influenced by information perceived in the -1000 to -500 ms window before their response in comparison of information perceived around -250 ms.

In addition, W_B and W_C events behave similarly (overlap function in **Figure 2.5A**). However, W_A seems to be a stronger predictor of Strong Positive response in comparison of W_D which has a flatter function. To compare both conjunction (W_A and W_D) and disjunction (W_B and W_C) events I obtained the zero-complement (*i.e.*, multiplied disjunction weights by -1) to get all the weights in positive and comparable scale (**Figure 2.5A inset**). Interestingly, B and C had similar weights, but A and D no, where A events showed a larger effect around 500 ms. This corroborates previous findings of reduced weights to D events (Kao & Wasserman, 1993; Murphy et al., 2021; Wasserman et al., 1993). However, this does not imply that D trials are not used at all (see Castiello et al., 2022). This co-absent event type will be the focus of next chapter.

I tested the effect on schizotypy, event types, and its interaction in zero-complement weights (**Figure 2.5B**) by running an LMM-based ANOVA. There was no main effect on schizotypy in weights [$F(1, 93.6) = 0.01$, $p = 0.92$]. Nor an interaction effect between schizotypy and event type [$F(3, 4495.1) = 0.58$, $p = 0.63$]. So schizotypy did not affect how participants used the event types. However, I found a main effect on event types [$F(3, 4495.2) = 51.68$, $p < 0.001$], where D event type showed an overall reduce weight. Finally, given low-level perceptual features have been impaired in UE (Harper et al., 2020), as an exploratory analysis I used UE and information type (positive: A and D coded as 1; negative: B and C, coded as 0) to predict zero complement weights. I found a negative

main effect on information type [Estimate = -0.16, p < 0.001], suggesting that people overall use less positive evidence (A and D)1.5 seconds previous the response. But no main effect on UE [Estimate = -0.04, p = 0.15]. Interestingly, the interaction UE and information type was positive and significant. Suggesting that UE is associated with reduced weight for negative information, *i.e.*, the disjunction of cues.

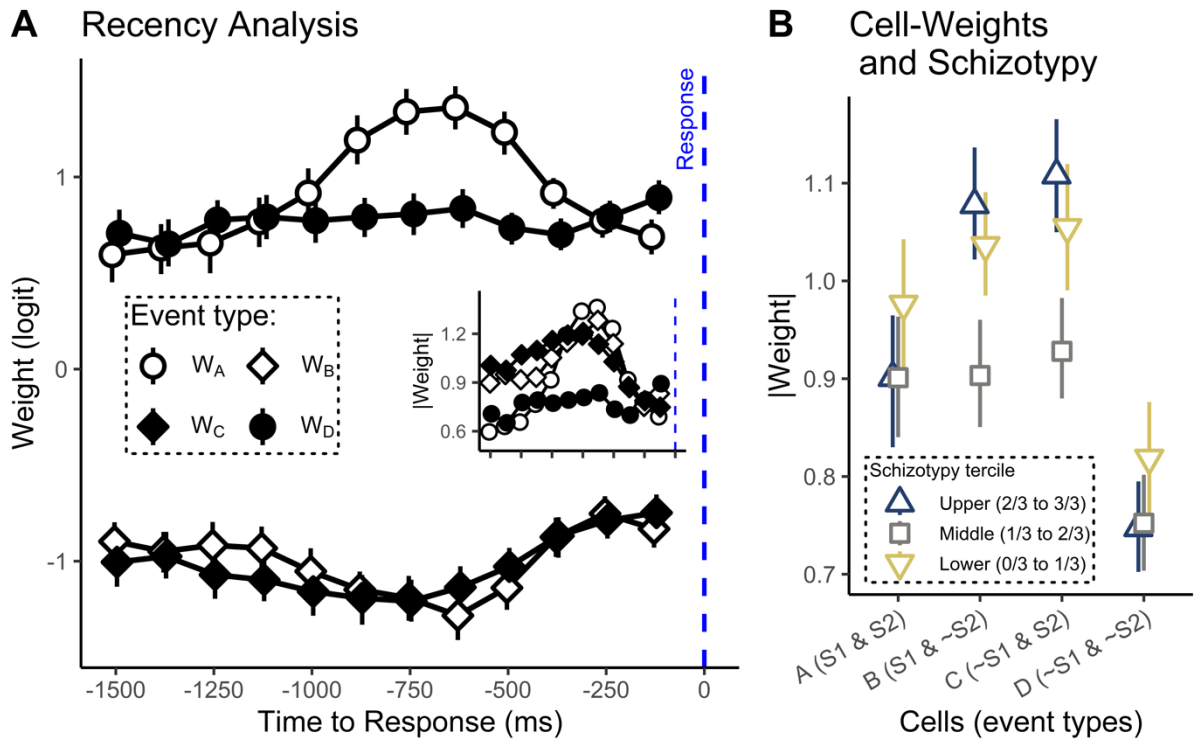


Figure 2.5. Information Weights Before the Response in Experiment 1. **A**. Events type weights for the last 12 events (*i.e.*, 1.5 secs) just before the response was made (blue dotted vertical line at time 0). Inset is the absolute weight means. **B**. zero-complement, *i.e.*, W_B and W_C multiplied by -1.

2.4. Discussion

In this chapter the data failed to find evidence to support the hypothesis that schizotypy would exhibit hyperassociability, a response bias to perceive positive contingencies−. However, I corroborated the second hypothesis, where schizotypy was related with an increased DT for correct trials. This means gathering more sensory evidence to perform adequately the contingency task. Moreover, the exploratory analysis suggests that IA is related with a reduced DT. If I consider the exploratory result, then it may be related to why IA also showed a reduced sensitivity to contingencies

(**Figure 2.A2**). It is possible that high IA scorers lacked motivation or did not attend properly to the task, so when trials seemed difficult, they responded more quickly and less accurately. Therefore, they displayed of reduced DT which made participants less sensitive to discriminate contingencies.

Interestingly, the exploratory analysis (**Figure 2.A2**) also suggests that UE is related with a lower sensitivity to detect contingencies. This result agrees with findings showing that people scoring high in UE were unable to discriminate between two reinforcement schedules contingencies (Randell et al., 2009, 2012). This may suggest an overall deficit in assessing statistical regularities in positive schizotypy. Thus, partially supporting the intuition of the hyperassociability hypothesis.

Furthermore, the positive relation between schizotypy and DT for correct trials, confirm previous findings where people with SZ exhibited slower DT (Strube et al., 2022). Also, people with SZ may learn to reduce confidence in sensory input during their life, therefore they need to gather more sensory information to make correct choices (Moustafa et al., 2015), again contrary to JTC.

2.4.1. Contingency and Psychophysics

Independent to schizotypy, I found that positive contingencies are easier to perceive, this was supported by the negative relation between *a* and *t0* with contingency ΔP. That is, the more positive ΔP the less *a* and *t0* were required to choices. In addition, I found that the mutual absent stimuli in the immediately previous time to the response, did not impact the response as much as the mutual present stimuli.

This is interesting with respect to CLT. People do not use the co-absent information (D events) as much as they use the mutual presence of stimuli nor the disjunction events (A, B, and C events). Therefore, I corroborate the contingency weighting rule: $W_A > (W_B = W_C) > W_D$ (Kao & Wasserman, 1993; Murphy et al., 2021; Wasserman et al., 1993). This suggest a differential event weighting where co-absent events (D trials) are perceived as less relevant −perhaps, because they are less salient? However, my results expand the contingency event weighting rule to short timeframe, *i.e.*, seconds to milliseconds. These findings raise the question about how people use the mutual

absent information in contingency assessment (see Castiello et al., 2022). Hence, in the next chapter I focused on the relevance of co-absent events.

2.4.2. Contrast with previous literature

The present schizotypy results with general population using DDM did not find similar results to previous experiments in people with SZ (Limongi et al., 2018). Limongi et al., found an increased effect of priors (larger starting point, *zr*) in people with SZ. However, my results did not find any relation between schizotypy and *zr*. I only find a larger *zr* for positive contingencies in contrast to negative ones. Finally, I found a similar result as Moustafa et al (2015). They found that people with SZ had a larger sensory threshold accumulation (*a*) for punishment trials, this means larger DT. Even though in my experiment I did not use punishment trials, I found that larger DT for correct trials in schizotypy, *i.e.*, in order to get the same accuracy more sensory evidence was needed.

2.4.3. Limitations

Psychophysical vision research has shown that UE subscale showed a reduced contrast sensitivity (Harper et al., 2020). This may suggest a potential confound in my results, where the lack of contrast sensitivity may reduce the accuracy of contingency perception, therefore larger DT. However, I believe that this is not the case because incorrect trials did not exhibit an increase in DT. In addition, people diagnosed with SZ perceive long trials as shorter and have less discriminability to colour and time (Ciullo et al., 2018). So given the cognitive overlap with SZ and schizotypy, it may be that the increase of DT is due to a difference in perception of time.

2.4.4. Conclusions

Overall, this study showed three main results. First, schizotypy did not change the overall probability of perceiving positive contingencies in the millisecond timeframe; this supports the idea that individuals can discriminate the contingency task without a response bias, *e.g.*, lateralization of responses. Second, overall positive contingencies are easier to perceive. Third, in correct trials individuals scoring high in schizotypy tended to respond slowly. They need more sensory evidence to make a correct choice. This is the opposite of what JTC bias would predict, *i.e.*, people with high

schizotypy scores and with SZ tend to use less information to decide (Hua et al., 2020; Le et al., 2019).

2.5. Appendix 2

2.5.1. Included versus excluded samples

Table 2.A1. Comparison between Included and Excluded participants

	Included	Excluded	Statistical test
N, # (%)	97 (77)	29 (23)	
Age, M (SD)	25.67 (8.49)	24.79 (5.59)	W = 1303.5, p = 0.92
Sex F, # (%)	52 (54)	17 (58)	$X^2(1) < 0.001$, p = 0.99
Depression, M (SD)	12.99 (11.52)	13.21 (12.92)	W = 858, p = 0.8
Schizotypy, M (SD)			
Total	15.68 (6.53)	17.31 (8.88)	t(37.5) = -0.92, p = 0.37
Cognitive Disorganization	4.72 (2.99)	4.55 (3.36)	W = 1454.5, p = 0.78
Introvertive Anhedonia	4.12 (1.87)	4.83 (2.41)	W = 1148, p = 0.13
Impulsive Nonconformity	3.85 (2.18)	3.93 (2.56)	W = 1373, p = 0.85
Unusual Experiences	2.99 (2.45)	4 (2.92)	W = 1119, p = 0.09

Note: the statistical tests were selected as a function of the normality (Shapiro Wilk test) and homoscedasticity (Levene test) of the samples. # = frequency; % = percentage; M = mean; SD = standard deviation; t = t statistic (from t-test); W = w statistic (from Wilcoxon test); F = female.

2.5.2. Subsamples differences

I ran a logistic mixed model for each sample, then I visualised the included participants per sample (**Figure 2.A1)**. All the samples showed a main effect on condition, *i.e.*, ΔP (Coef. > 2.1, and p < 0.001). However, no main effect on schizotypy was found (coefficients between -0.06 and 0.10, and p > 0.18). Finally, only in the Multiple sample (n = 17) I found a reliable interaction between ΔP and schizotypy, Coef. = -0.16 [95% CI from -0.31 to -0.01, p < 0.05], no other interactions were significant (Coef. Between -0.11 and 0.04, and p > 0.3).

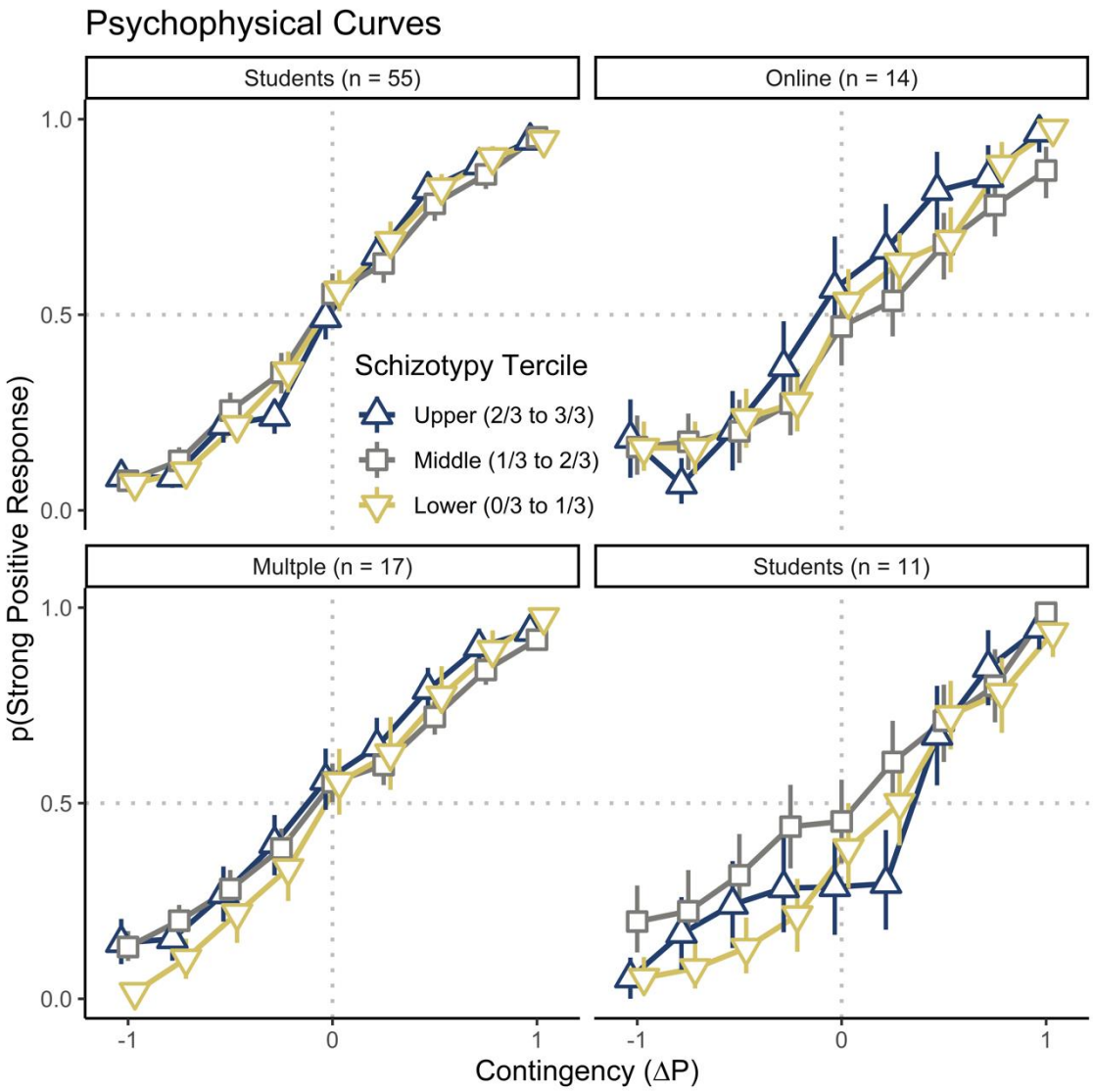


Figure 2.A1. Psychometric curves, ΔP, and subsamples for Experiment 1. Probability of responding Strong Positive as a function of contingency conditions (ΔP). Each subpanel is one of the four subsamples. The participants were split in schizotypy terciles only for visualization purposes, the analyses add schizotypy as a continuous variable.

In addition to comparing the sociodemographic between samples (**Table 2.1**) and general performance in the task (**Figure 2.A1**). In **Table 2.A2** I presented the O-LIFE internal subscales correlations each sample and all the included participants. Compare against the normative data (Mason et al., 1995).

Table 2.A2. O-LIFE internal correlation current study vs normative data

	Current study (n = 97)			Normative data (n = 508)		
	UE	CD	IA	UE	CD	IA
CD	0.47***			0.42**		
IA	0.26*	0.03		-0.08	0.25**	
IN	0.47***	0.41***	0.17	-0.46**	0.33**	-0.10*
	Sub Samples of the current study					
	Students (n = 55)			Online (n = 14)		
	UE	CD	IA	UE	CD	IA
CD	0.65***			0.29		
IA	0.34*	0.36**		0.26	0.53	
IN	0.54***	0.5***	0.12	0.29	0.61*	0.65*
	Multiple (n = 17)			Students (n = 11)		
	UE	CD	IA	UE	CD	IA
CD	0.11			0.66*		
IA	0.4	-0.06		0.37	0.32	
IN	0.62**	0.03	0.36	0.59	0.22	0.38

Note: *** = p < 0.001; ** = p < 0.01; * = p < 0.05. all correlations were conducted with Pearson classic method.

2.5.3. Task instructions

Instructions:

In this part of the experiment, you will be watching a series of rapidly presented shape (circle) and drawing (hammer). Your task is to decide whether you think the shapes and drawings are POSITIVE or NEGATIVE associated with each other. Use the keyboard to respond and press "c" when you think they are NEGATIVE associated and "m" when you think they are POSITIVE associated. For every choice that you made you will be given feedback (correct/incorrect) after you press one of the keys.

** A POSITIVE association should be given when the shape and drawing are regularly presented together and absent together.*

** A NEGATIVE association should be given when the shape is regularly presented without the drawing and the drawing is regularly presented without the shape.*

Press the appropriate keys (NEGATIVE: key "c"; or POSITIVE: key "m") once you think to have enough information to pick the correct answer. Try to focus so that you can get the most correct answers as possible.

Press {space} to continue

2.5.4. Trait Specificity: Schizotypy subscales and Depression

In this section I explored how specific is the main null effect found in the main text (3.2.; **Figure 2.2B**). I ran extra five models, one for each schizotypy subscales and one for depression. See **Figure 2.A2** to see the results, where I plotted significant regressors with colour signalling the direction of the effect.

I did not find any overall response biases towards perceiving positive contingencies (*i.e.*, no questionnaire main effect). Although UE, IA, and CD showed a significant interaction with ΔP. The results found here suggests that UE and IA have lower sensitivity to classify between positive and negative contingencies. However, CD has a higher sensitivity to detect contingencies. Nonetheless, given that here I used the whole trial by trial data set, to see the results of the analysis using the participant-based data base see (**Figure 2.4**).

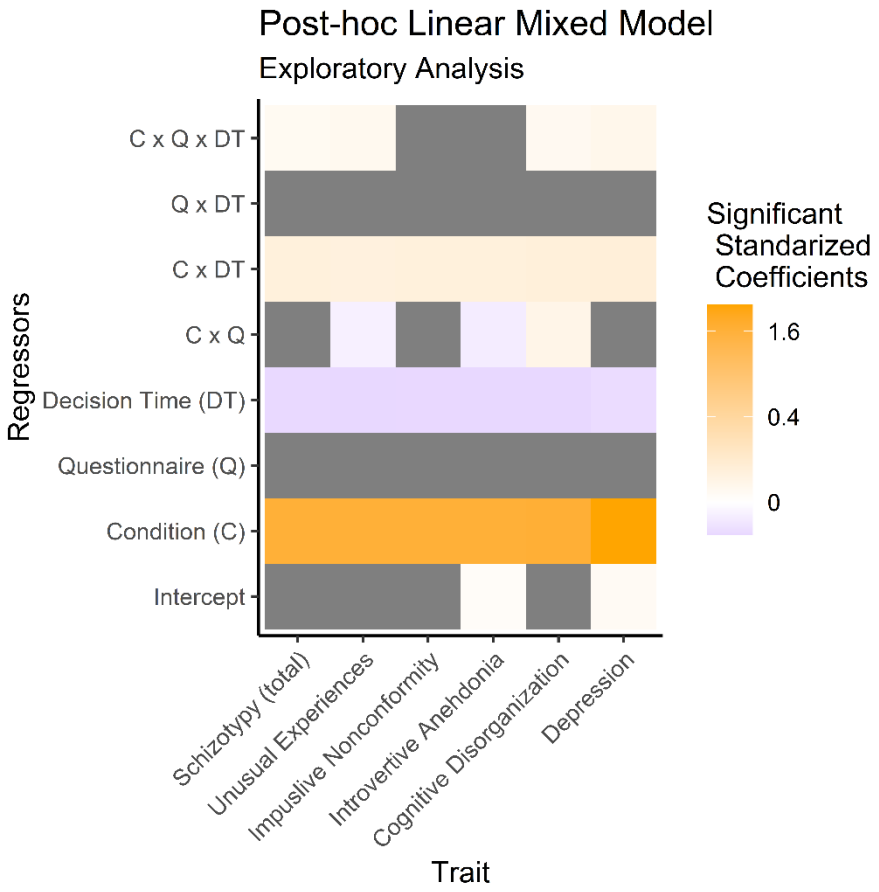


Figure 2.A2. Schizotypy Subscales: Main Model Exploration for Experiment 1. X-axes are the psychological traits, total schizotypy and its subscales (UE, IN, IA, and CD) and depression. Y-axes represent the regressors for each model, it was the same as is in **Figure 2.2B**. The grey squares represent a null effect (i.e., p value > 0.05).

2.5.5. Sensitivity Analysis: Generalized Mixed Model (with Gamma link function)

Some issues emerged due to the right skewed nature of reaction times or DT distributions. Some solutions are: *i)* transforming the raw scores, or *ii)* using a specific link function in a generalized mixed model. Although is not the same running a linear model to a log-transformed right skewed data than a generalized model with a specific link function (Lasserre, 1991). Sometimes may be better to run the generalized model with the adequate link function for the raw data.

Therefore, I conducted a sensitivity analysis examining whether fitting a generalized mixed model with a gamma link function modify the results presented in the main text. After fitting the model, the only difference is that the current model does not show the significance of correct as main effect. Although, the relevant estimate for our hypothesis was the interaction between schizotypy and

correct trials, which also remained significant in this model. In summary, both transformed or raw data models point to the same direction, *i.e.*, a positive interaction between correct and schizotypy scores. Which means that to achieve correct trials people with high schizotypy scores must gather more sensory evidence.

Chapter Three

Sensitivity to Mutual Absent Cues in Schizotypy

Note on the publication from this chapter

This chapter is a secondary analysis of our already published paper (Castiello et al., 2022). In Castiello et al (2022), we used a rapid-trial streaming procedure in three experiments and found that the perceived association between events is enhanced with increasing frequency, but not duration, of mutually absent events (like an ITI). These findings suggest ways in which the benefits of trial spacing, during which both events are absent, could be obtained without increasing total training time. Centrally, this can be done by frequent repeating of shortened co-absent events, each marked by a trial contextual cue. We found this "free lunch" effect also in extinction (Witnauer et al., 2022).

Even though the published paper did not include any schizotypy analysis. I wanted to pursue this line of research because I was aware that the BADE is found in people with SZ and poeple scoring high in schizotypy and delusional proneness (Buchy et al., 2007; Juárez-Ramos et al., 2014; Woodward et al., 2007). I believed that disconformatory evidence is similar to co-absent events. Therefore, in this chapter I used two of the published experiments in which I collected schizotypy scores.

3.1. Introduction

In the previous chapter I found an asymmetry between weighting mutual presence of stimuli and co-absence of them. In informational terms (Shannon, 1948), there should be no difference (see **3.6.1. Appendix**), *i.e.*, presence and absent of an information bit are symmetric. Although intuitively it feels self-evident that animals weight the presence of something more than its absence. Evidence backed this intuition (Murphy et al., 2021; Wasserman et al., 1993). However, research suggests that

people still use absent information, but with a lower impact (Murphy et al., 2021; Wasserman et al., 1993). The absence of events in not the absence of information. For example, in determining the cause of stress-related headaches, the absence of stress and headaches is informative for calculating the relation between stress and headaches (Murphy et al., 2017). In this chapter, I hypothesized that schizotypy would display a reduced sensitivity to perceiving the absence of information.

3.1.1. Contingency Judgments, Conditional Reasoning, and Absent Information

In the previous chapter, I used ΔP in a psychophysical approach to contingencies, using a two-alternative choice task. In the current chapter, I used contingency judgments, which has been used widely for the study (Baetu & Baker, 2009; Baker et al., 1993, 2005; Dickinson, 2001; Dickinson et al., 1984b; Shanks & Dickinson, 1987; Wasserman et al., 1993). While the simultaneous presence of stimuli is clearly relevant for learning a positive relation, and the occurrence of either event alone is relevant for negative correlations, the co-absence of both events (*e.g.*, ITI) is less intuitively relevant for either type of correlation, for example to increase ratings in positive contingencies and reduce ratings in negative contingencies (Mercier & Parr, 1996). Although, research suggests the opposite for negative contingencies using the rapid-trial streaming procedure (Castiello et al., 2022; Murphy et al., 2021; Witnauer et al., 2022). But how do absences become relevant and inform particular predictive relations? Moreover, tasks that assess human reasoning find that information about counterfactual (a type of absence) is often used less by participants (*e.g.*,Wason, 1968).

Like the contingency table used in previous chapter (see trial types in **Figure 3.1**), the Wason task used four faced-up (p, $\sim p$, q, $\sim q$) two-sided cards; on one side is either p or $\sim p$, and on the other q or $\sim q$. Participants are asked to test statement "if p then q" ($p \rightarrow q$) by selecting a card (Cummins et al., 1991). There are only two logically valid (p and $\sim q$) selections to accept or reject $p \rightarrow q$ and two invalid ones ($\sim p$ and q). The invalid choices provide ambiguous information. On the one hand, trying to reject $p \rightarrow q$ by selecting $\sim p$ independent of the other side of the card it is an example of the "Denying the Antecedent" fallacy. One the other, hand rejecting $p \rightarrow q$ with q is an "Affirming the

Consequent" fallacy. Both, selecting *p* (*Modus Ponens*) and ~*q* (*Modus Tollens*) are valid, and Wason (1968) observed that people are more likely to select *p* than ~*q*. This evidence reveales people's biases against using counterexamples or disconfirming (~*q*) evidence. In the next section I described the evidence that relates conditional reasoning and contingency learning judgments with schizotypy and SZ.

3.1.2. Disconfirmatory and Absent Evidence in Schizotypy

Understanding humans' contingency judgments is relevant for translational psychiatry research (Byrom & Murphy, 2018). High schizotypy scorers are more prone to form odd or paranormal beliefs; hence a tendency for superstitions (Hergovich et al., 2008a). For example, superstitious beliefs, such as homeopathy may be caused by a reduced weight of absent (or disconfirmatory) evidence. Thus, people may only remember when they took homeopathy and felt better but they do not when they did not take homeopathy (counterexample) and felt better. Interestingly, in the Wason task, schizotypy was related with ignoring counterexamples, such as ~*q* (Sellen et al., 2005), similar effects were found in people with SZ (Kornreich et al., 2017).

Furthermore, in a free-operant procedure, people with paranormal beliefs are more likely to press a button to produce an outcome; hence less likely to expose themselves to disconfirmatory evidence, *i.e.*, not pressing the button (Blanco et al., 2015). This suggests a differential event weighting between confirmatory (pressing) and disconfirmatory (not pressing). In addition, illusion of control, which is judging to control an outcome when there is no control ($\Delta P = 0$), has been related with superstitious beliefs (Griffiths et al., 2018). Finally, schizotypy is related with BADE (Juárez-Ramos et al., 2014; McLean et al., 2016; Moritz & Woodward, 2006). This bias is shown when high schizotypy scorers do not change their judgments even when disconfirmatory evidence is presented.

Evidence suggests that general population weighs stimuli co-presence more than stimuli co-absence (Kao & Wasserman, 1993; Murphy et al., 2021; Wasserman et al., 1993). I thought the absent evidence is similar to counterexamples. Therefore, people high in schizotypy might be predicted to

use less co-absent information (disconfirmatory) while learning a contingency between any two stimuli.

3.1.3. Contingency Learning Theory: Frequency vs Duration

In addition to the relevance of disconfirmatory and absent evidence in human contingency judgments. The absent evidence is studied in animals, for instance through the Intertrial-Interval (ITI). Although, methodological differences make difficult direct results' comparison, evidence from animal studies suggests that the ITI (which could be conceptualized as co-absent events, *i.e.*, no CS nor US) is critical for learning. Furthermore, the evidence supports duration (time) instead of frequency of the ITI during the acquisition process (*e.g.*, Balsam et al., 2010; Escobar et al., 2002; Gallistel & Papachristos, 2020; Harris et al., 2015; Harris & Bouton, 2020; Urcelay et al., 2009). However, during extinction, the presentation of the CS without the US, CS frequency seems to be more relevant for predicting decreases in conditioned responding (Chan & Harris, 2019; Harris & Andrew, 2017).

The paradigm that I used in this chapter is the rapid-streaming procedure (Allan et al., 2008; Crump et al., 2007; Murphy et al., 2021). The contingency was calculated with ΔP from the trial types (A, B, D, and D trials) described in **Figure 3.1**. Where an A trial is the mutual presence of stimuli, a B trial is the presence of S1 only, a C trial is the presence of S2 only, and a D trial is the mutual absence of stimuli. In the presented two experiments D trials represent the co-absence of stimuli, which are the trials that I manipulate parametrically in the frequency and duration dimension.

In this chapter I presented the analysis of the effect on schizotypy in perceiving co-absent stimuli events (D trials). This reanalysis is based on two of the three experiments already published: Experiments 2 and Experiment 3 from Castiello et al (2022). There we found that presenting trials without stimuli (co-absent event) enhanced the perceived association between two target stimuli, and that the frequency but not the duration of the co-absent events was the critical factor of the enhanced perceived association. Experiment 1 found an effect of D trial (co-absent stimuli) frequency, but not D trial duration, on judgments of a single type of contingency (zero contingency: no correlation between the two stimuli). There was no evidence that the increase of D trial duration nor the introduction of a

3-min delay (retention interval) between learning and judgments resulted in any change in the expression of the learned relationship. Experiment 2 extended the findings to positive and negative baseline contingencies, where the trial frequency effect was replicated. Experiment 3 provided a control for overall D trial duration that corrected for increases in D trial frequency potentially having a corresponding effect on overall duration. Here, the D trial frequency effect was found to be independent of changes in overall D trial duration. The consistent result of the three experiments is that increasing the frequency of co-absent stimuli (D trials) enhances the perceived association between the two stimuli.

In summary, in two experiments I manipulated, the frequency and duration of co-absent stimuli events and assessed their impacted contingency judgements. I used those experiments to investigate the effect of schizotypy in perceiving D-trials (co-absence of stimuli). Based on the relevance of absent information for both humans and animals, as well as the importance of disconfirmatory evidence in contingency judgments, my hypothesis was that high schizotypy scorers would be hyposensitive to D-trials.

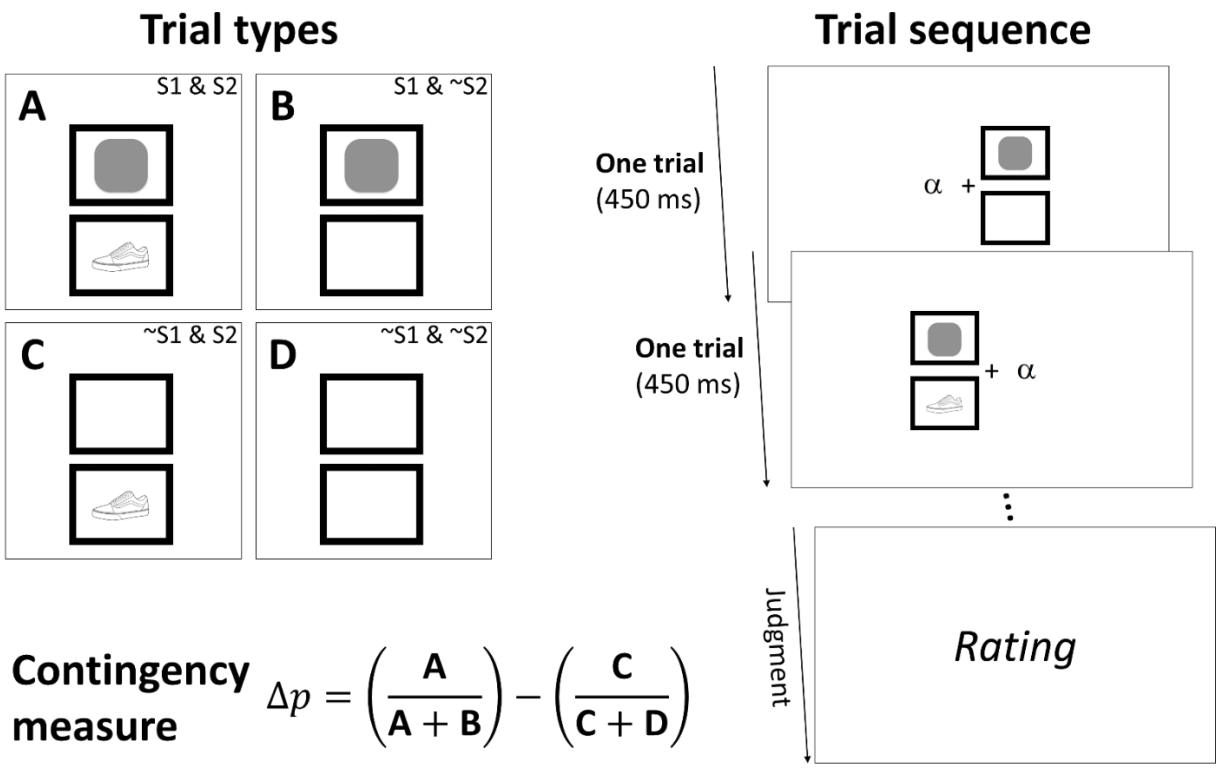


Figure 3.1. Paradigm used in Experiments 2 and 3. Diagram of the task used in both experiments. **Trial types**. There are four possible combinations between shapes (S1) and drawings (S2) that were used as stimuli in experiments 2 and 3; ~ indicates *not* (or absence of). For details see 3.6.2. Appendix. The black squares are trial markers (TM) and their texture change in every condition. **Trial sequence**. This is an example of the first two trials in one condition. The stimuli in each trial were presented on the left and right of the fixation cross, alternating in consecutive trials to explicitly mark different trials. Black squares or trial markers (TM) were used to mark trials and their context. **Contingency measure**. Delta p (ΔP) is a measure of continency (Allan, 1980) defined by the difference in conditional probabilities representing the four event types.

Table 3.1. Participants' General Characteristics for Experiments 2 and 3

	Experiment 2		Experiment 3	
	Mean (SD)	range (from - to)	Mean (SD)	range (from - to)
Age	20.40 (2.82)	19 - 38	27.38 (8.85)	18 - 55
Sex, F|M (%)	24|21 (53|47)		27|13 (67.5|32.5)	
Schizotypy	16.47 (8.7)	4 - 43	12.57 (7.26)	2 - 28
CD	6.31 (3.52)	0 - 11	4.27 (3.34)	0 - 11
IA	2.42 (2.11)	0 - 10	2.89 (2.27)	0 - 10
IN	3.82 (2.36)	0 - 10	2.68 (2.14)	0 - 8
UE	3.91 (3.06)	0 – 12	2.73 (2.52)	0 - 8

Note: SD = Standard Deviation; CD = Cognitive Disorganization; IA = Introvertive Anhedonia; IN = Impulsive Nonconformity; UE = Unusual Experiences.

3.2. Experiment 2

This experiment aimed to replicate the D trial frequency effect (sensitivity to co-absent events) obtained in Experiment 1 from Castiello, et al (2022), not presented here due to lack of questionnaire data. In addition, for the matter of the current chapter, I assessed the effect of schizotypy in co-absent events sensitivity. I have two hypotheses. The first is relevant for studying schizotypy, I hypothesised that high schizotypy scorers are less sensitive to co-absent events, hence use less D trials to rate contingency between two stimuli ($S1$ and $S2$). Second, relevant for learning theory, I tested the relevant D trials factors for contingency judgments: frequency or duration.

Schizotypy may modulate how people use counterfactuals (a possible instance of D trials). This modulation may be due to four underlying accounts: *i)* lower weight to D trials (Cheng, 1997; White, 2004); *ii)* testing effect, where performance during a test of association is enhanced if information acquired on earlier A, B, and C training trials is retrieved on later trials (Roediger & Butler, 2011); *iii)* reduced associative interference by the context, more 'empty' trials promotes more context extinction (Bouton, 1993; Miller & Escobar, 2002); or *iv)* reduced encoding interference, here it is assumed that memory for the pairing between any two events takes time to process, and increasing the spacing between each A trial permits more (Engle, 2016; Waugh & Norman, 1965). However, I have found that the more likely explanation was a reduced D trials weighting (Castiello et al., 2022).

3.2.1. Methods

3.2.1.1. Participants

The volunteer participants characteristics are presented in **Table 3.1**. Forty-five participants, 24 females and 21 males (age between 19 and 38, M = 20.4, SD = 2.82). The experiments were conducted in an open laboratory space, where multiple participants were tested at the same time. Our participant exclusion criteria were participants who provided the same rating on all conditions (i.e., were not sensitive to the manipulation) or those who did not finish their session for whatever reason; based on these criteria, no participants were excluded. This research was approved by the Medical Science Interdivisional Research Ethics Committee of University of Oxford; reference number: R60840/RE001.

3.2.1.2. Psychometric measurements

As in the previous chapter, schizotypy was measured with the short O-LIFE (Claridge et al., 1996; Mason et al., 2005; Mason & Claridge, 2006).

3.2.1.3. Design

The experimental design consisted of twelve conditions (plus one practice condition; total 13 conditions), four negative, four zero, and four positive contingencies (see **Table 3.2**). At the end of each condition participants provided a contingency rating between -10 and +10. The conditions were presented by stages of contingencies. The decision not to randomise the contingencies across training was based on experience with some within subject design experiments where complete randomization is related to increased error variance. The order of contingency stages (negative, zero, and positive) was randomly assigned, and within a contingency stage the four frequency conditions were randomized.

In each condition, the trial marker (TM; different border textures) and a further discrete stimulus (a Greek letter that was unique for each condition) on the opposite side of the screen from the stimuli on each trial. See **Figure 3.A1** for a visual reference.

For each of the three baseline contingencies (negative, zero, and positive), I performed the same manipulation of the four D frequencies: 0 (None), 24 (Baseline), 72 (More), and 216 (Many). The baseline D trial frequency for every contingency was 24. Negative contingencies included 4 A trials and 44 B trials; zero contingencies included 24 A and 24 B trials, and positive contingencies included 44 A trials and 4 B trials, with all three contingency types including 24 C trials. These numbers resulted in the negative baseline contingency having a ΔP = [4/(4+44)]-[24/(24+24)] ≈ -0.42; zero baseline contingency having a ΔP = [24/(24+24)]-[24/(24+24)] = 0, and positive baseline contingency having a ΔP = [44/(44+4)]-[24/(24+24)] ≈ +0.42 (see ΔP values for each condition in **Table 3.2**). For example, the Negative More condition (ΔP = -0.17; negative baseline contingency with 72 D trials) was presented in four blocks, each one with 1 A trial, 11 B trials, 6 C trials, and 18 D trials. Therefore, the calculated ΔP for this block was the same as the overall contingency (i.e., 1/(1+11) – 6/(6+18) ≈ -0.17). To minimize potential trial order effects within conditions, the trials within all conditions were presented in three blocks, with the ΔP for that condition being met in each block. The four trial types were always 450 ms in duration. Hence, there was no ITI other than the D trials. The practice condition was a replication of the baseline condition with a zero contingency, with 24 A, 24 B, 24 C, and 24 D trials. This experiment has not been preregistered.

Table 3.2. Experiment 2: Conditions and Δ*P* [Frequency (4), and baseline Contingency (3)]

Contingency and A, B, and C trials		**D trial Frequency**			
		None 0	Baseline 24	More 72	Many 216
Positive	(44 A, 4 B, 24 C)	-0.1	0.42	0.67	0.82
Zero	(24 A, 24 B, 24 C)	-0.5	0	0.25	0.4
Negative	(4 A, 44 B, 24 C)	-0.9	-0.42	-0.17	-0.02

Note: Each cell within the grid represents a condition, the number represents Δ*P*.

3.2.1.4. Procedure

I used ePrime 2.0 (Psychology Software Tools, Pittsburgh, PA) to programme the task. Participants provided informed consent and were asked to turn off their mobile phones. The participants completed the computer task alone in a computer cubicle and followed on-screen instructions. When they finished the behavioural task, participants completed the short O-LIFE version (Mason et al., 2005).

The task involved observing the appearance of two figures, a geometric shape (S1) and a drawing (S2; see **Figure 3.1**) selected from two independent lists. Shapes and drawings were consistently yoked for all participants, but assignment of yoked pairs to conditions was completely randomized without replacement, independently across participants. As presented in the right-hand panel of **Figure 3.1**, every trial was displayed in a sequentially alternating position on the screen respective to the fixation cross (*e.g.*, Right-Left-Right-Left, etc.). I did this to distinguish trials and avoid ambiguities arising when two consecutive trials of the same type are presented, which would have made it difficult to differentiate whether there were two of the same type of trial or one longer duration trial. Additionally, using the trial markers (TM; black squares around the stimuli as depicted in **Figure 3.1**) provided another feature to mark individual trials.

Each condition was programmed into four training blocks, with the Δ*P* for that condition being met within each block; participants were not informed of the transitions between blocks. Within

each of the four blocks, the trials were randomized without replacement. Presenting trials by blocks avoided random selection biasing the overall relationship, either at the beginning or end of the condition. Note that manipulating the frequency of D trials influences the ΔP as given by the formula in **Figure 3.1**. The nominal ΔP for every condition is given in **Table 3.2**. Before starting the 12 experimental conditions, participants experienced a practice condition, which was not used for the analysis. In the practice condition, each trial type was presented 36 times at 450 ms, and the $\Delta P = 0$; thus, it was identical to the baseline condition and post hoc analysis found that the practice condition received similar ratings to the randomly presented baseline condition.

3.2.1.5. Analysis plan

I modelled ratings using LMM. I divided the analysis into two parts: *i)* inferences based on the Full model's standardized estimates and the 95% CI, and *ii)* model comparison approach, where comparing the Full model and all possible reduced models with the Bayesian Information Criterion (BIC; Schwarz, 1978) and the Akaike Information Criterion (AIC; Akaike, 1973).

The Full model included *i)* D trials frequency, *ii)* contingency stage (negative, zero, and positive), *iii)* and total schizotypy score, as well as the triple order interactions. In addition, all models included the stage order as a covariate and the same random effect structure. Where participants were random intercept effects, and frequency and contingency stage were random slope effects; resulting in a maximal random effects structure (Barr et al., 2013).

I fitted the models by substituting the regressors levels by contrasts (Schad et al., 2020). For frequency (0, 24, 72, 216) the contrasts were: -13, -9, -1, and 23[3]; for contingency stage (negative, zero, positive) the contrasts were: -1, 0, and 1; and for total schizotypy I scaled the score by subtracting the mean and dividing the scores by the SD.

For the model comparison approach, I fitted all possible reduced models. Then obtained the ΔBIC by subtracting the lowest BIC model to all the others. Same procedure with ΔAIC. With this, I

[3] Frequency contrasts: f = [0,24,72,216]; contrasts: (f - mean(f))/6 = [-13, -9, -1, 23]

obtained the most parsimonious model, and I could appreciate how much each factor affect to the overall model parsimony.

All analysis and visualizations were made using R (R Core Team, 2021). I used the following packages. LMMs were run with *lme4* (Bates et al., 2015)and *lmerTest* (Kuznetsova et al., 2017). Additional packages used for data processing were *reshape2* (Wickham, 2007) and *dplyr* (Wickham et al., 2022); models' standardized estimates and export csv tables were obtained by *report* (Makowski et al., 2021), calculation of Bayes factors used *bayestestR* (Makowski et al., 2019), and data visualization used *ggplot2* (Wickham, 2016), *viridis* (Garnier et al., 2021), and *ggpubr* (Kassambara, 2020).

3.2.2. Results & Discussion

In **Figure 3.2A** I present the ratings as a function of D trials frequency for each contingency stage, from negative in the left to positive in the right panels. This figure shows the total schizotypy in terciles and a linear fitted for each tercile. Next, **Figure 3.2B** shows visually the Full model standardized estimates (Std. Est.) with the 95% CI, this model is also shown numerically in **Table 3.3**. Finally, **Figure 3.2C** presents the model comparisons.

3.2.2.1. Full model

There was strong evidence that the Frequency of trials contributed to increased ratings of contingency, this was supported by the frequency effect, Std. Est. = 0.30 [95% CI from 0.21 to 0.40], and Contingency stage, Std. Est. = 0.69 [95% CI from 0.63 to 0.75]. Schizotypy was not significant as main effect, not any interaction which include it. The only significant interaction was Frequency and Contingency, Std. Est. = -0.12 [95% CI from -0.16 to -0.08].

To explore the F x C interaction, I ran three extra models, one for each contingency stage. Visually (**Figure 3.2A**) negative contingencies seem to have a larger frequency effect, this was corroborated with the size of the frequency standardized estimates from the three post-hoc models. Where the largest Std. Est. was the negative-contingency frequency, Std. Est. = 0.52 [95% CI from

0.35 to 0.68]; then the zero-contingency frequency, Std. Est. = 0.38 [95% CI from 0.23 to 0.52]; and the positive contingency frequency, Std. Est. = 0.34 [95% CI from 0.20 to 0.48]. Finally, in these models neither schizotypy nor F x S were significant.

Table 3.3. Experiment 2: Full model fixed effect estimates

Variable	Estimate	*t* value	df	*p* value	Std. Est.	lower CI	higher CI
(Intercept)	1.27	3.12	523	<.01**	0.00	-0.09	0.09
Frequency (F)	0.12	6.34	523	<.001***	0.30	0.21	0.40
Contingency (C)	4.78	21.79	523	<.001***	0.69	0.63	0.75
Schizotypy (S)	-0.01	-0.03	523	ns	0.00	-0.09	0.09
Stage Order	-0.29	-1.85	523	ns	-0.04	-0.09	0.00
F x C	-0.06	-5.82	523	<.001***	-0.12	-0.16	-0.08
F x S	0.00	0.03	523	ns	0.00	-0.09	0.10
C x S	-0.21	-0.95	523	ns	-0.03	-0.09	0.03
F x C x S	0.00	0.10	523	ns	0.00	-0.04	0.04

Note: df = degree of freedom; Std. Est. = Standardized Estimate; lower CI and higher CI = 95% confidence interval; *ns* = non-significant.

3.2.2.2. Model comparison

In the second part of the analysis, I fit all possible reduced models (see **Figure 3.2C**). The most parsimonious model, *i.e.*, lower BIC and AIC, was the one including only frequency and contingency stage. However, this model was followed by a model including schizotypy as a main effect. Although, even there, the schizotypy regressor was not significant.

In summary, this analysis suggests that participants increase their perceived contingency as a function of adding more frequent D trials in the contingencies. Also, that this effect applies for all three baseline contingencies. However, schizotypy does not seems to impact how people used the co-absent trials to guide their contingency judgments.

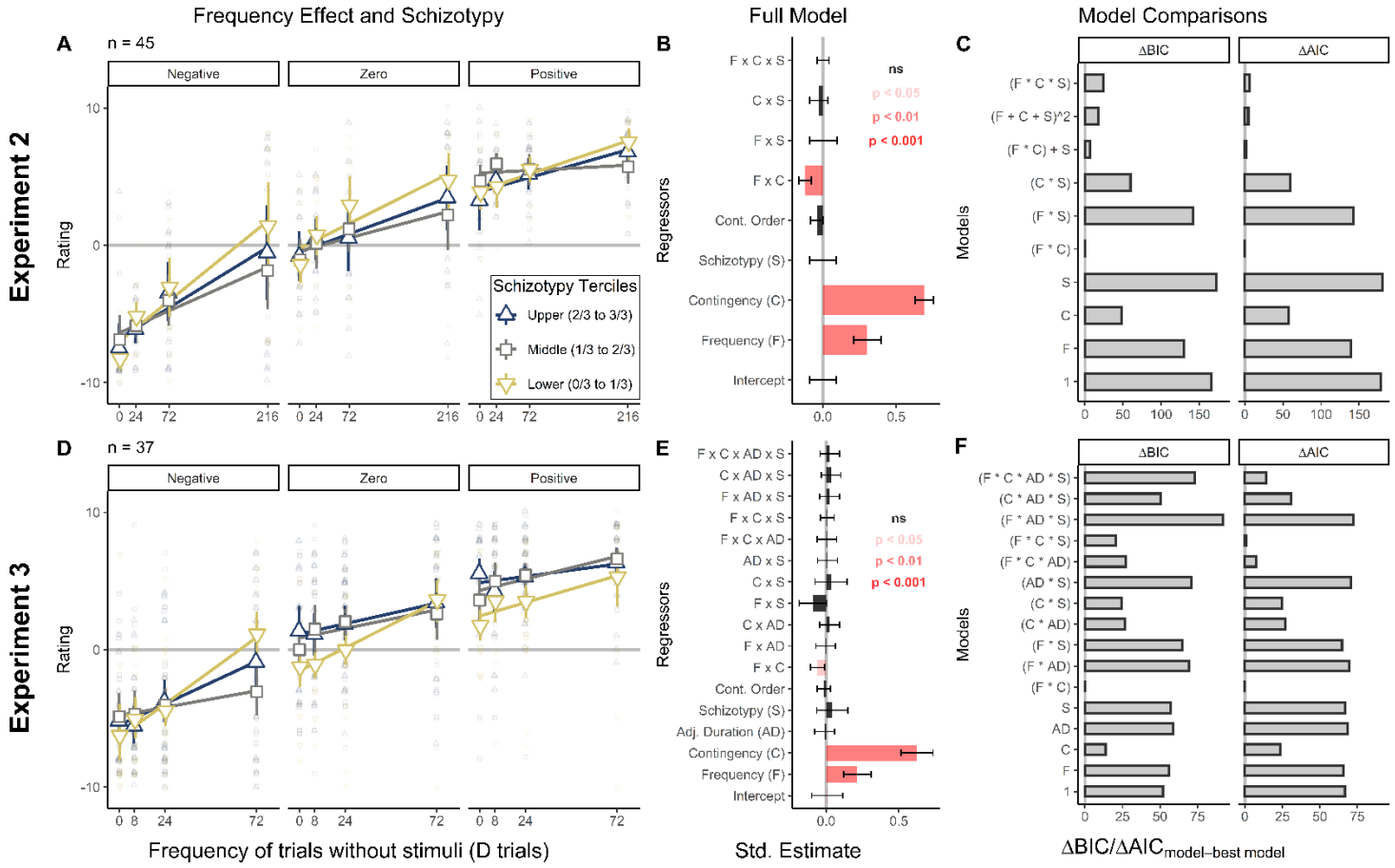


Figure 3.2. Confirmatory analysis. Panels **A**, **B**, and **C** correspond to Experiment 2; and **D**, **E**, and **F** to Experiment 3. Panels **A** and **D** show contingency ratings at the end of conditions as a function of co-absent (D trials) trial frequency and contingency type (Negative, Zero, and Positive) in columns panels. Colours represents total schizotypy terciles. Panels **B** and **E** showed the standardized estimates (with the 95% CI) from the Full Model, i.e., including all design factors. Panels **C** and **F** show the BIC and AIC difference (Δ) between all possible alternative models minus the lower BIC model (i.e., the best model).

3.3. Experiment 3

This experiment aimed to replicate the previous experiment results as well as controlling the overall time exposure to D trials. In the previous, I found that frequency of D trials contributed to increased ratings; however, the conditions with increased frequency also received a greater overall duration. In this way some D trials frequency effect was confounded with total trial duration, this is because D trials durations were constant at 450ms when I manipulated the frequency. Hence, increasing D frequency also increases the total D trial time experienced by participants. The goal for Experiment 3 was to inversely vary D-duration while altering D-frequency events. Additionally, I sought to replicate the D frequency effect across the three contingencies. Finally, I sought to replicate the lack of schizotypy and frequency interaction observed in Experiment 2.

3.3.1. Methods

3.3.1.1. Participants

Forty participants, 27 females and 13 males (age between 18 and 55, M = 27.37, SD = 8.84), served in the experiment (see **Table 3.1**). The participants were paid £10.00 at the completion of the experiment. No participants were excluded by using the same exclusion criterion of Experiment 2. However, I lost the schizotypy scores of three participants, hence their ratings were excluded from visualization and analysis. This research was approved by the Medical Science Interdivisional Research Ethics Committee of University of Oxford; reference number: R60840/RE001.

3.3.1.2. Psychometric measurements

Same questionnaire as Experiment 2 was used. Short O-LIFE (O. Mason & Claridge, 2006).

3.3.1.3. Design

The experimental design consisted of 24 conditions, plus a practice condition which was not used for the analysis. I tested the D frequency effect with respect to three baseline contingencies, Negative ($\Delta P \approx -0.42$), Zero ($\Delta P \approx 0$), and Positive ($\Delta P \approx 0.42$). **Table 3.4** depicts all 24 conditions. The manipulations were: *i*) four D frequencies (0, 8, 24, and 72) with the constant duration (450 ms),

ii) four D durations (0, 150, 450, and 1350 ms) with the same frequency (24), and *iii*) two D frequency x duration inverse varied conditions: a frequency of 8 trials of 1350 ms, and a frequency of 72 trials of 150 ms; for each of the three baseline contingencies.

I used 24 as the baseline frequency. This was done to avoid a potential ceiling effect that I found with positive contingencies in Experiment 2 as shown in **Figure 3.2A**. This change did not undermine sensitivity to D trial frequency, which was corroborated by fitting a model omitting the 216-frequency condition in Experiment 2; for this analysis I found a reliable effect of frequency, contingency-stage, and their interaction, Fs > 10, ps < 0.01.

Negative contingencies included 4 A trials and 44 B trials: zero contingencies, 24 A and 24 B trials, and finally, positive contingencies, 44 A trials and 4 B trials with all three contingency types including 24 C trials. These numbers resulted in the negative baseline contingency having a ΔP = [4/(4+44)]-[24/(24+24)] ≈ -0.42; zero baseline contingency having a ΔP = [24/(24+24)]-[24/(24+24)] = 0, and positive baseline contingency having a ΔP = [44/(44+4)]-[24/(24+24)] ≈ +0.42 (see ΔP values for each condition in **Table 3.4**). All conditions were presented in four blocks, with the ΔP for that condition being met in each block. A, B, and C trials were always 450 ms in duration, with only D trials being manipulated. The practice condition was a replication of the baseline condition with a zero contingency, with 24 A, 24 B, 24 C, and 24 D trials. This experiment was not preregistered.

Table 3.4. Experiment 3: Conditions and Δ*P* [Contingency (3), Frequency (4), and Duration (4)]

Contingent	D trial Duration	D trial Frequency			
		0	**8**	**24**	**72**
Positive (44 A, 4 B, 24 C)	**0**	None -0.08			
	150			Short 0.17*	More adjusted 0.67; 0.42*
	450		Few 0.17	Baseline 0.42	More 0.67
	1350		Few adjusted 0.17; 0.42*	Long 0.67*	
Zero (24 A, 24 B, 24 C)	**0**	None -0.5			
	150			Short -0.25*	More adjusted 0.25; 0*
	450		Few -0.25	Baseline 0	More 0.25
	1350		Few adjusted -0.25; 0*	Long 0.25*	
Negative (4 A, 44 B, 24 C)	**0**	None -0.92			
	150			Short -0.67*	More adjusted -0.17; -0.42*
	450		Few -0.67	Baseline -0.42	More -0.17
	1350		Few adjusted -0.67; -0.42*	Long -0.17*	

Note: Each rectangle represents a condition, with the name of the condition and the number is the corresponding Δ*P*. * When D trial duration was varied (0 [same as no D trials: None], 150, 450, and 1350 ms), ΔP was calculated based on total duration in each trial type.

3.3.1.4. Procedure

Experiment 3 used the same procedure as in Experiment 2 except where otherwise stated. See Supplementary Materials from Castiello et al. (2022) for more details.

3.3.1.5. Analysis plan

For this statistical analysis, I used the same plan, program, packages, and functions as in Experiments 1. The Full model included *i)* D trials frequency, *ii)* contingency stage (negative, zero, and positive), *iii)* adjusted by duration (coded as 1), *iv)* and total schizotypy score, as well as the fourth order interactions. Hence, the Full model was: *rating ~ (frequency * contingency * adjusted_duration * schizotypy) + contingency_order + (1 + frequency + contingency | participant)*. Participants had a random intercept effect and contingencies, and frequency were the random slope effects; and as a covariate contingency stage order was included. The same contrasts were used for frequency as in Experiments 1.

To have the 'frequency adjusted by duration' as a complete design factor, I duplicated the 0 frequency condition data in the analyse, so adjusted and non-adjusted at 0 frequency were the same. Because repeating the same condition may bias the analysis by undermining the assumptions of independent observations and variability, I conducted a sensitivity analysis with the zero frequency conditions removed. The pattern of results was similar with both analysis strategies (see **Table 3.A2**).

3.3.2. Results & Discussion

Experiment 3 results are presented following that of Experiment 2. **Figure 3.2D** shows ratings as a function of D trials frequency for each contingency stage, from negative to positive. I observed that ratings increased as a function of D trials, and this was the case for the three contingency baselines (stages). Also, total schizotypy is shown in terciles and a linear fitted for each tercile, it seems that schizotypy is associated with reduced frequency effect. Then, **Figure 3.2E** shows the Full model standardized estimates (Std. Est.) with the 95% CI, and the exact model values are shown in **Table 3.5**. Finally, **Figure 3.2F** presents the models comparisons.

3.3.2.1. Full model

Similar to Experiment 2, a significant main effect of Frequency was observed, ratings increased as a function of D trials, Std. Est. = 0.21 [95% CI from 0.12 to 0.31], and Contingency stage, Std. Est. = 0.63 [95% CI from 0.52 to 0.73]. Schizotypy and adjusted by durations were not significant as main effects nor as any interaction which include them (**Figure 3.2E**). The only significant interaction, as Experiment 2, was Frequency and Contingency, Std. Est. = -0.06 [95% CI from -0.11 to -0.01].

Also, I explored the F x C interaction by running three extra models, one for each contingency stage. The largest Std. Est. was the negative-contingency frequency, Std. Est. = 0.33 [95% CI from 0.17 to 0.50]; then the zero-contingency frequency, Std. Est. = 0.26 [95% CI from 0.08 to 0.44]; and finally, the positive contingency frequency, Std. Est. = 0.19 [95% CI from 0.06 to 0.32]. Finally, in these models neither schizotypy nor Adjusted by duration were significant.

Table 3.5. Experiment 3: Full model fixed effect estimates

Variable	Estimate	*t* value	df	*p* value	Std. Est.	lower CI	higher CI
(Intercept)	0.95	2.20	971	<.05*	0.01	-0.10	0.11
Frequency (F)	0.09	4.43	971	<.001***	0.21	0.12	0.31
Contingency (C)	4.24	11.17	971	<.001***	0.63	0.52	0.73
Adj. Duration (AD)	-0.06	-0.32	971	*ns*	-0.01	-0.08	0.06
Schizotypy	0.24	0.77	971	*ns*	0.04	-0.06	0.15
Stage Order	-0.11	-0.72	971	*ns*	-0.02	-0.06	0.03
F x C	-0.03	-2.44	971	<.05*	-0.06	-0.11	-0.01
F x AD	0.00	-0.08	971	*ns*	0.00	-0.07	0.06
C x AD	0.16	0.68	971	*ns*	0.02	-0.04	0.09
F x S	-0.04	-1.86	971	*ns*	-0.09	-0.19	0.00
C x S	0.24	0.62	971	*ns*	0.03	-0.08	0.14
AD x S	0.05	0.27	971	*ns*	0.01	-0.06	0.08
F x C x AD	0.00	0.12	971	*ns*	0.00	-0.06	0.07
F x C x S	0.00	0.32	971	*ns*	0.01	-0.04	0.06
F x AD x S	0.01	0.68	971	*ns*	0.02	-0.04	0.09
C x AD x S	0.23	0.98	971	*ns*	0.03	-0.03	0.10
F x C x AD x S	0.01	0.72	971	*ns*	0.02	-0.04	0.09

Note: df = degree of freedom; Std. Est. = Standardized Estimate; lower CI and higher CI = 95% confidence interval; *ns* = non-significant.

3.3.2.2. Model comparison

Next, I fitted all possible reduced models (**Figure 3.2F**). Here, the most parsimonious model, *i.e.*, lower BIC and AIC, was the one including only frequency and contingency stage. Only in the case of ΔAIC, it seems that the model including schizotypy (F * C * S) may be almost as good as the model without schizotypy (F * C). In addition, in this former model, the triple interaction with schizotypy did not reached significance, Std. Est = 0.04 [95% CI from 0.00 to 0.07].

Overall, this experiment supports the same conclusions as Experiment 2. Participants increased their perceived contingency as a function of increased D trials. However, the statistical evidence does not support the idea that schizotypy influenced people's sensitivity to absence information (D trials).

In the next section I conducted an exploratory analysis by pooling both experiments. I decided to do this because in both experiments, it seems that the schizotypy terciles slopes followed a decreasing order, smaller slope for the upper terciles. Also, the model comparisons suggest that the second-best models are the ones including schizotypy. Therefore, if there is an effect, this may be a small effect. Thus, I calculated a post-hoc power analysis to know which correlation magnitude will be needed for N = 82, $\alpha = 0.05$, and $\beta > 0.80$ (Zhang & Mai, 2022). So, to detected negative correlation between sensitivity (D trials frequency slope) and schizotypy the correlation should be smaller than -0.272.

3.4. Pooled Experiments: Exploratory Analysis

There was no reliable interaction between D trials' frequency and schizotypy in either experiment. However, in the model comparison analysis (**Figure 3.2C** and **3.2F**) it was observed that the second-best model after the one including only frequency and contingency are the ones including schizotypy. Based on this, I decided to explore the individual sensitivity to D trials and correlate it with the total schizotypy and its subscales.

I conducted this exploratory analysis by fitting individual linear models per participant. Therefore, I describe the relationship between contingency ratings and the natural logarithm of D trials frequency with two parameters: an intercept and a slope (see individual linear fit in **Figure 3.3A**). On one hand, a large slope means more sensitivity to D trials frequency; on the other hand, a large intercept means an overall increased perception of positive contingencies. Furthermore, I correlated the linear models' parameters with the five schizotypy scores (total score + 4 subscales). I normalized the scores by dividing each score vector by the largest possible value. Given that the correlations did not pass Shapiro-Wilk normality test, I conducted Spearmen rank-order correlations (rho).

In **Figure 3.3B** I show the scatterplots between slopes and schizotypy scores. No adjustment of p value was used, and the slope is negatively correlated with UE, $rho = -0.219$, $p < 0.05$. This suggests that participants scoring highly on UE were less sensitivity to co-absent information. Although, this is significant is not as big as the calculated in the power analysis, not even using Pearson correlation, $r(80) = -0.235$; $p < 0.05$. No other, correlation with the slopes was found. Finally, in **Figure 3.3C** I presented correlations between the intercepts and the schizotypy scores. Interestingly, the intercept is correlated with UE, $rho = 0.257$, $p < 0.05$ which may indicate an overall bias toward perceiving positive contingencies, in other words, hyperassociativity. This indicates that hyperassociation may be a specific feature of positive schizotypy, perhaps this contributes to why some associative learning studies found only correlations with UE (Evans et al., 2007; Granger et al., 2012; Le Pelley et al., 2010; Schmidt-Hansen et al., 2009).

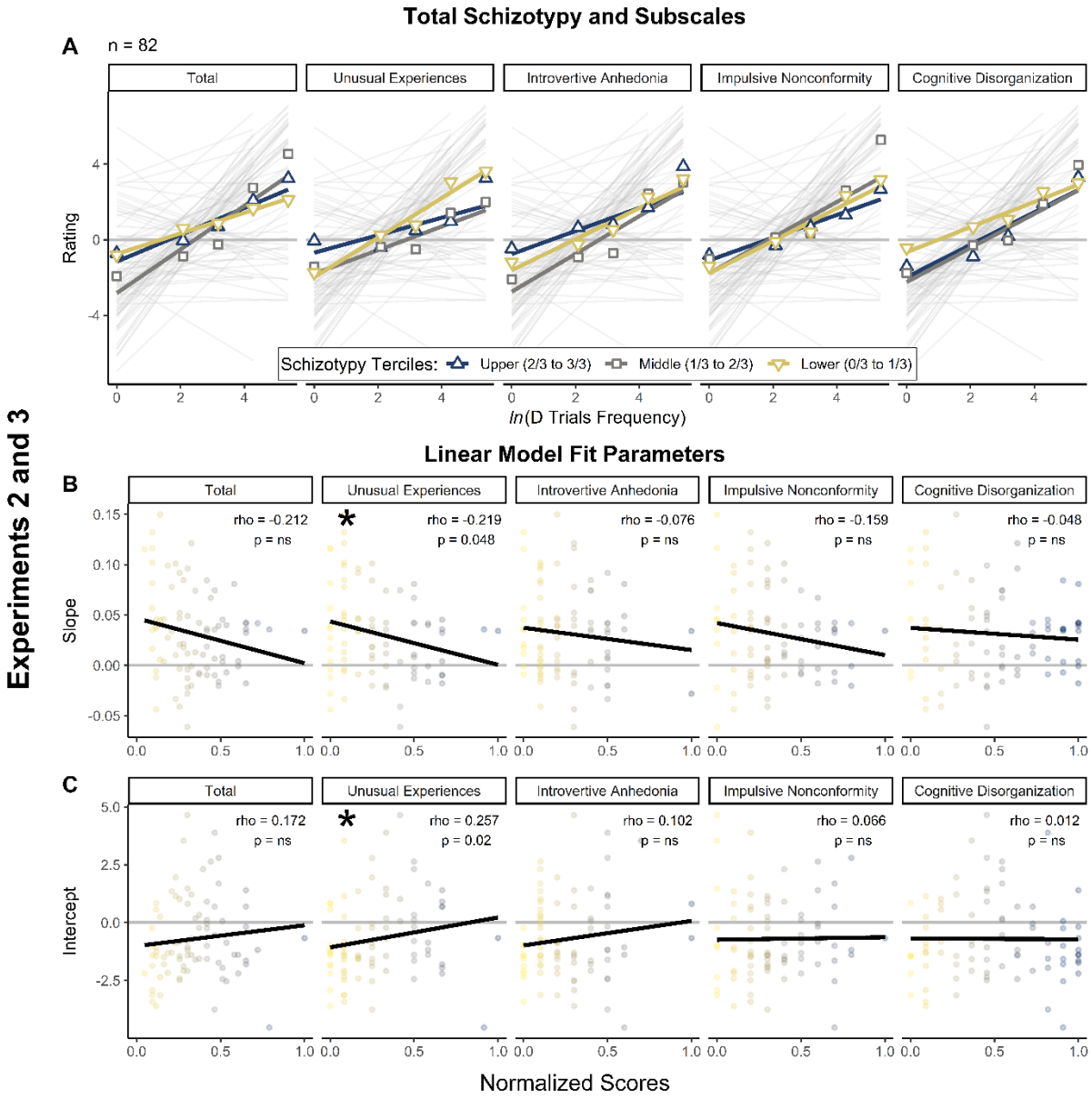


Figure 3.3. Exploratory analysis for Experiments 2 and 3. **A** Ratings as a function of the natural logarithm of D trials frequency (1 was used for frequencies of 0). Colours represents total schizotypy terciles, and grey lines are individuals. Columns' panels for **A**, **B**, and **C** panels are schizotypy total, and the four subscales. Panels **B** and **C** are the results of a linear fit [rating = intercept + slope * ln(D trials frequency)] per participant. **B** slopes and **C** intercepts as a function of normalized schizotypy and subscales scores.

3.5. Discussion

In this associative learning task using the rapid-streaming procedure, I did not find a reduced sensitivity to events with absence of stimuli (D trials) and schizotypy. Previous research suggests that people that are more likely to believe in the paranormal do not expose themselves to the information where there is an absent response (Blanco et al., 2015). This possible difference may be due to the

fundamental difference in tasks: associative vs instrumental. Independent of schizotypy, our results supported the D trials accounts based on *i)* participants tracking the ΔP contingency and *ii)* reduced associative interference with information concerning A trials by the context. This was also found, because in negative contingencies D trials increased ratings (*cf.*, (cf.,Mercier & Parr, 1996).

I presented the analysis of two experiments (Castiello et al., 2022), where I reported that people perceived contingencies between two stimuli is increased by including short trials without stimuli. The hypothesis being tested was that high schizotypy scorers would be less sensitive to using trials without target stimuli (D trials). However, our confirmatory analysis supports the conclusion that with this paradigm schizotypy does not affect how people use D trials in contingency learning.

However, in an exploratory analysis, I pooled the data from both experiments, and obtained an absent information sensitivity metric (the slope of the D trials frequency effect). Then, I found that UE were related with lower sensitivity to co-absent information in contingency learning. Also, UE was positive related with an overall enhanced perception of positive contingencies, *i.e.*, hyperassociation. However, due to the exploratory nature of this analysis this result should be taken with caution. Nonetheless, this may suggest that hyperassociation may be an aspect of positive schizotypy (UE), thus future experiments should focus on this specific relation.

3.5.1. Clinical relevance

I used a rapid-streaming procedure for associative learning and presented evidence that schizotypy −viewed as a normally distributed personality trait (Claridge, 1997; Claridge et al., 1996; Grant et al., 2018) and not only as a liability for SZ (Lenzenweger, 2010, 2018a)− is not associated with a hyposensitivity to absent information. However, in different settings and paradigms, differential information weighting was associated with schizotypy and people with SZ. For example, cognitive biases such as Liberal Acceptance, Bias Against Confirmatory Evidence (BACE), BADE, Jumping to Conclusion (Juárez-Ramos et al., 2014; McLean et al., 2016; Moritz & Woodward, 2006), and insensitivity to denying the consequent reasoning type (Johnson-Laird & Wason, 1970). Nonetheless, from a learning perspective a superstitious belief might be created as the result of

overweighting the outcomes of their actions (Griffiths et al., 2019; Matute, 1996; Matute et al., 2011), which I explored in the next chapter.

My rational for this chapter was that the relation between schizotypy and reduced absence of evidence weighting would emerge in an associative learning context (such as the present tasks). However, our findings point out that for this effect to emerge different learning or perceptual requirements are needed, and not passive viewing of a contingency. For instance, value-based decision making, such as instrumental learning, or even more complex scenarios in the realm of reasoning, such as BADE. Therefore, our results may reach a boundary effect, in the sense that, to find a differential weighing effect −*e.g.*, a reduce weight to absent information or an exuberant weight to present information (aberrant salience; Kapur, 2003)− response-consequent contingencies may be necessary.

The pooled exploratory result indicates that UE may exhibit a reduced sensitivity to absence information. The effect of the absent events has been suggested to play an important role in how depressive traits individuals make judgments using the contextual information (Byrom et al., 2015; Msetfi et al., 2005, 2013, 2017). Furthermore, pharmacological evidence suggests that escitalopram (Selective Serotonin Recapture Inhibitor) increases action-context discrimination in mildly depressed participants (Msetfi, Kumar, et al., 2016) which may be due to a differential configural processing in people with different individual characteristics (Msetfi et al., 2015) as well as how people are sensitive to temporal changes during the contingency (Msetfi et al., 2012). This variety of phenomena are closely linked to how does people process the absence of cues (Castiello, Senan, et al., 2020).

Finally, in the exploratory analysis, I also found a positive correlation between UE and the intercept of a linear model fitted to contingency judgments. This means that overall, UE is related with perceiving positive associations between cues. A possible similar effect to the illusion of control found in people with tendencies to superstitious beliefs (Griffiths et al., 2018), as well as in schizotypy and SZ. Even though the current data does not fully support the hypothesis regarding a reduced sensitivity for processing absent information in schizotypy. I considered that the question is a

good heuristic to understand judgments and cognitive biases (such as the one manifested in superstitions or odd beliefs).

3.6. Appendix 3

3.6.1. Presence and Absence Information Symmetry

In this section I showed that the presence and absence of stimuli (S) provided the same information among them. Assuming that S are random variables −*e.g.*, $S \sim Be(p): S \in \{0,1\}$, *i.e.*, a Bernoulli experiment with probability of success of p−, then from Information Theory (Shannon, 1948), I know that the entropy of S is:

$$H(S) = -\sum_{o \in O} p(s) * log_2[p(s)] \quad \text{(eq. 3.A1),}$$

where higher H implies more bits of information are required to describe S. Then, the joint entropy amount two $S1$ and $S2$ is:

$$H(S1, S2) = -\sum_{s1 \in S1} \sum_{s2 \in S2} p(s1, s2) * log_2[p(s1, s2)] \quad \text{(eq. 3.A2).}$$

Finally, the mutual information between them is:

$$I(S1, S2) = H(S1) + H(S2) - H(S1, S2) \quad \text{(eq. 3.A3).}$$

To exemplify this, I took four contingencies as presented in **Table 3.A1**. Here I showed that the $I(S1, S2)$ is the same for any of the four contingencies. Irrespective of the presence of the absence of either $S1$ or $S2$. For example, for both contingencies with ΔP of 0.25, *i.e.*, the A event frequency is 15 with B and C 5 and D 15, I also obtained the same $I(S1, S2)$.

Table 3.A1. Contingency and Mutual Information between S1 and S2

Contingency Table Frequency	p(S2|S1)	p(S2|~S1)	ΔP	I(S1,S2)
A15; B5; C5; D5	0.75	0.5	0.25	0.04
A5; B15; C5; D5	0.25	0.5	-0.25	0.04
A5; B5; C15; D5	0.5	0.75	-0.25	0.04
A5; B5; C5; D15	0.5	0.25	0.25	0.04

Note: p(S2|S1) = probability of S2 given S1; p(S2|~S1) = probability of S2 given the absence of S1; ΔP = p(S2|S1) - p(S2|~S1); I(S1,S2) = mutual information between S1 and S2.

3.6.2. The Task

3.6.2.1. Screen: instructions

At the beginning of the experiment, the participant was instructed to provide their age and gender, after which the instructions will read:

In this experiment, you will be watching numerous series of rapidly presented shapes and drawings. After each series, a question screen will appear and you will be asked to rate the degree of relatedness between the shape and drawing on a scale from -10 to +10. Please keep your eyes on the cross in the center of the screen.

A STRONG POSITIVE RATING should be given when the shape and picture are always presented together and when one is absent the other is also absent.

A STRONG NEGATIVE RATING should be given when the shape is always presented without the picture and the picture is always presented without the shape.

3.6.2.2. Screen: rating

After each condition, participants were asked to rate the relatedness of the S1 and S2. The question read:

Please indicate the degree of relatedness between the [S1 picture inserted here] *and the* [S2 picture inserted here] *in the series that you just saw. Use the rating scale below to enter the degree of relatedness:*

-10 = strongly negatively related

0 = not related

+10 = strongly positively related

Subjects recorded their ratings by using the mouse and clicking (Experiment 2 and 3) on a Likert scale ranging from values -10 to +10, increasing by a unit of 1 (i.e., 21 options). After the question was answered by the participant, a screen appeared saying:

Now prepare for the next sequence of shapes and drawings. Press the {spacebar} when you are ready to begin.

3.6.2.3. Screen: debriefing

At the end of the experiment, a debriefing screen read:

Thank you for participating in this experiment. The success of our research depends on the cooperation of people like you.

This experiment is part of a larger study to understand how people change their relatedness judgements based on their perception of absent events. In this experiment, we used a fully within-subjects design, meaning that you were exposed to all the conditions in the experiment. Your ratings in each condition will be compared to your ratings in each of the other conditions. Across conditions, we varied the degree of relatedness between pairs of shapes and drawings, as well as the nature of events during which the shapes and drawings were both absent. We measured how your judgements of the degree of relatedness between the shapes and drawings varied accordingly. Of particular interest to us was the influence on your judgements of different variations in the nature of the events on which the shape and drawing were absent. Findings that you judged a higher degree of relatedness between pairs of shapes and drawings because of such manipulations may be useful in enhancing learning situations in tasks as varied as formal education settings and exposure therapy for anxiety disorders.

Please be quiet as you leave the laboratory by the same route through which you entered so that other participants are not disturbed. So as to not influence data from future subjects in this experiment, please do not share this information with any other student.

3.6.2.4. Visual references

For the three experiments, I used the same ePrime 2.0 program parameters. The only difference was that Experiment 2 did not include the distractor stimulus (Greek letter).

The size of the stimuli and frames were 130 x 130 pixels for both stimuli, and 240 x 190 pixels for the trial markers (TM) borders. Their respective positions were centered at different XY coordinates. For the TM borders: [42%, 40%] for the top left border; [58%, 40%] for the top right border; [42%, 60%] for the bottom left border; and [58%, 60%] for the bottom right border. The stimuli had the same coordinates due to their being centered at the same locations.

I used 25 different contextual stimuli (one per condition). These stimuli were the first 25 letters in the Greek alphabet: α, β, γ, δ, ε, ζ, η, θ, ι, κ, λ, μ, ν, ξ, ο, π, ρ, ς, σ, τ, υ, φ, χ, ψ, and ω. See a visual reference for 4 trials: A, B, C, and D, in **Figure 3.A1**.

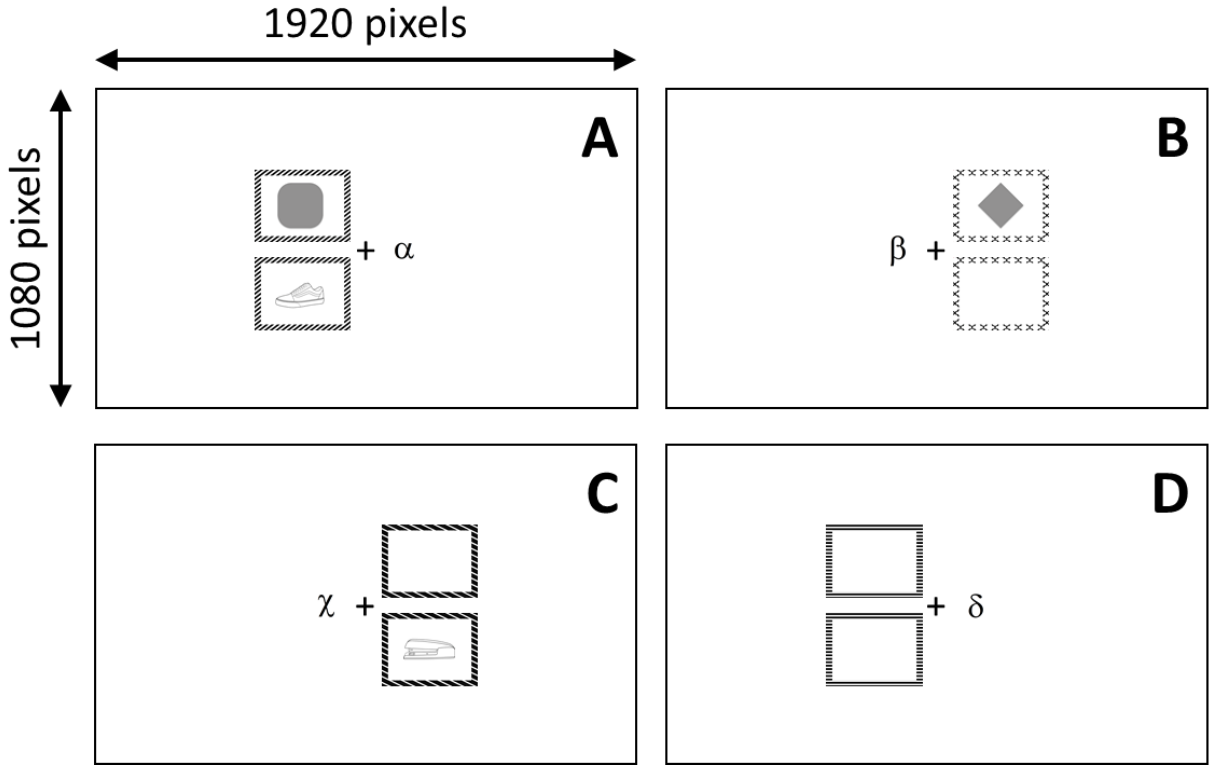


Figure 3.A1. Visual Reference for Experiments 2 and 3. A visual reference example is shown for 4 different conditions which used different: stimuli, trial marker, and contextual stimulus. Within each condition, trials alternated between left and right sequentially.

3.6.3. Sensitivity Analysis Experiment 3

Here I show the Full model fitted without the conditions with 0 D trials.

Table 3.A2. Experiment 3 Sensitivity Analysis: Full Model Fixed Effect Estimates

Variable	Estimate	*t* value	df	*p* value	Std. Est.	lower CI	higher CI
(Intercept)	1.11	2.27	751	<.05*	0.01	-0.11	0.13
Frequency (F)	0.09	4.30	751	<.001***	0.20	0.11	0.29
Contingency (C)	4.14	11.41	751	<.001***	0.62	0.52	0.73
Adj. Duration (AD)	-0.10	-0.45	751	ns	-0.02	-0.10	0.06
Schizotypy	-0.05	-0.14	751	ns	-0.01	-0.13	0.12
Stage Order	0.03	0.17	751	ns	0.00	-0.05	0.05
F x C	-0.03	-2.26	751	<.05*	-0.06	-0.12	-0.01
F x AD	0.00	0.07	751	ns	0.00	-0.07	0.08
C x AD	0.18	0.70	751	ns	0.03	-0.05	0.11
F x S	-0.03	-1.37	751	ns	-0.06	-0.16	0.03
C x S	0.13	0.35	751	ns	0.02	-0.09	0.12
AD x S	0.13	0.57	751	ns	0.02	-0.06	0.10
F x C x AD	0.00	-0.06	751	ns	0.00	-0.08	0.07
F x C x S	0.01	0.95	751	ns	0.03	-0.03	0.08
F x AD x S	0.01	0.63	751	ns	0.02	-0.05	0.10
C x AD x S	0.34	1.25	751	ns	0.05	-0.03	0.13
F x C x AD x S	0.01	0.47	751	ns	0.02	-0.06	0.10

Note: df = degree of freedom; Std. Est. = Standardized Estimate; lower CI and higher CI = 95% confidence interval; ns = non-significant.

Chapter Four

The Agency Paradox: Illusion of Control and Schizotypy

4.1. Introduction

4.1.1. Behavioural Control and Contingency Learning Theory

Humans, as agents with nervous systems (*cf.*, Burgos, 2018), can be described as controllers (from Control Theory; Powers, 1973a; Power, 1973b; Sosa & Alcalá, 2022). Within this view an agent is a system that controls its inputs with its outputs. For example, an air conditioning fan's goal is to control its thermostats (input) by cooling the room, *i.e.*, exerting cold air in it (output). In the case of living animals, to control the incoming sensations (inputs), agents use a discrepancy between a reference signal (internally produced) and a perceptual signal (externally produced). The agents' aim is to reduce the discrepancy signal (see also Seth et al., 2014). The discrepancies are used and transformed by an output function to produce output quantities, *e.g.*, response (*R*). Output quantities provide feedback as input quantities; therefore *R*s are also part of the system's input quantities. In addition, input quantities are affected by disturbances (*i.e.*, unexpected changes in the world) which are transformed by an input function to feed the agent with perceptual signals (Sosa & Alcalá, 2022). Finally, the feedback control loop is closed, and this dynamic adjusting process of behaviour continue over time.

There is an important difference between this view and classical models, such as associative or reinforcement learning models (Rescorla & Wagner, 1972; Sutton & Barto, 2018). In control system, the agents' *R* is a function of more than one way to use the discrepancy signal, also known as PE. First, a weighting of the current PE. Second, a mechanism to keep track of steady-state or

environmental stability, in other words higher learning to volatility environments and less learning to stable ones. This second idea has been implemented in different computational learning models (Behrens et al., 2007; Mathys et al., 2011; Pulcu & Browning, 2019).

Behavioural Control Theory (BCT) provides a framework to understand human behaviour and may be complemented with the CLT. For instance, ΔP between Rs and Os can be used as a rule to produce continuous judgments given a controller learnt set of values. Another approach to behavioural control −which is not mutually exclusive with BCT− suggests that R controls O, if a R reduces the entropy of O (Huys & Dayan, 2009), where entropy (H) is a measure of uncertainty (Shannon, 1948). If O and R were Bernoulli random variables (success = present), and $\mathrm{p}(O)$ is programmed by the experimenter and $\mathrm{p}(R)$ is internally generated by the agent, then, R controls O when the conditional entropy of O given R, *i.e.*, $H(O|R)$, is not maximal. In other words, $H(O|R)$ refers to the amount of information needed to predict O with R. To see a more detailed analysis, go to **4.5.1. Appendix**. In summary, $H(O|R) < 1$ is equivalent to $|\Delta P| > 0$. Therefore $H(O|R)$ and $|\Delta P|$ are negatively correlated. Finally, the implication is that when R reduce the $H(O|R)$, then R controls O and $|\Delta P| > 0$ is true. In sum, R controls an O when $|\Delta P| > 0$ and $H(O|R) < 1$. In next section I review literature relating to SoA and schizotypy.

4.1.2. Learning Control, Sense of Agency, and Schizotypy

From a CLT perspective, SoA is the subjective experience of outcomes' achievability (Moore, 2016). In other words, how well an agent learns to predict the occurrence or absence of an O given their R (Msetfi et al., 2005; Salgado-Pineda et al., 2021); thus, a reduction of $H(O|R)$ would be equivalent to increased SoA. Furthermore, unexpected Os tend to be more surprising −*i.e.*, large discrepancies between expected (reference signal) and observed (perceptual signal)−, then an agent's surprise reduction can be used as a measure of SoA.

In general population, delusional scores are negatively correlated with the error between expected and actual occurrence of a tone (Moore, Dickinson, et al., 2011). This suggested that delusional proneness is related to a reduction of contingency surprise (*i.e.*, SoA). In addition,

depressed people have been claimed to be more accurate at judging null contingencies (no-control); *i.e.*, depressive realism (Alloy & Abramson, 1979). Although this latter result has been showed to be related, not to accuracy, but a reduction with how people with depressive traits use contextual information (Byrom et al., 2015; Castiello, Senan, et al., 2020).

In addition to psychological traits, perceived control (PC) has been linked with the amount of behaviour performed by individuals. High probability of responding [p(R)] is related with an increase of PC in free-operant tasks (Blanco et al., 2012, 2013; Byrom et al., 2015; Matute, 1996). Also, using causal mediation analysis, probability of response increases PC (Byrom et al., 2015). Therefore, in free-operant tasks p(R) should be taken in consideration for analysis including moderation.

In this chapter I investigated how people perceived their own control and external source of control (another 'agent'; a computer agent), and its relationship with schizotypy. In a contingency learning task Baker et al (1993) found that estimating an external source of control interacts with participants' PC. A goal of this chapter is to replicate the interaction between perceived self-control and perceived external control. SoA is related with an adequate PC, and evidence suggests that high schizotypy scorers (and people with SZ) have a disrupted SoA and PC (Asai & Tanno, 2008; Balzan et al., 2013; Daprati et al., 1997; Moore, Dickinson, et al., 2011; Salgado-Pineda et al., 2021). In this chapter, I tested whether schizotypy would display enhanced perception of control and referential behaviours. Although, these behaviours can be expressed in different ways, for example, upweight of own control or downweight of external source of control.

The Illusion of Control is an overestimation of generating or preventing an O (Langer, 1975). This illusion has advantages, for instance it may be better to have a sense of control over uncontrollable situations (positive feelings and hope), rather than feeling that we cannot control anything, *e.g.*, learn helplessness (Blanco, 2017). At the same time, an elevated illusion of control may contribute to superstitious and odd beliefs formation (Blanco et al., 2015; Griffiths et al., 2019). Altogether, if an exacerbated illusion of control and odd beliefs are enhanced in schizotypy, this would match with what Claridge used to referred as *happy schizotypes*, high schizotypy scores with

no disfunctional symptoms (Claridge, 1997) also heightened in creativity (Claridge, 2018; Mohr & Claridge, 2015). For instance, optimistic people that feels controlling aspects of the world even when they do not, some sort of true control "ignorance is bliss". Therefore, I hypothesized that illusion of control is enhanced in schizotypy.

It is not clear if high schizotypy scores perceive an enhanced control in $\Delta P = 0$ conditions (*e.g.*, Blanco et al., 2013) or if they exhibit a bias across the whole ΔP dimension. In addition, one study found that positive symptoms of SZ are positively related with illusion of control (Moritz et al., 2014). Also, illusory correlation −assessing contingency problems in a non-instrumental task− is related with SZ and people with no diagnosis but with delusion proneness (Balzan et al., 2013). Assuming the continuous spectrum schizotypy and SZ as I have done throughout the dissertation (Claridge, 1994; Corlett & Fletcher, 2012), the results of the previous studies (Balzan et al., 2013; Moritz et al., 2014) provide evidence for a more specific first hypothesis, where I expected a positive correlation between positive schizotypy (*e.g.*, UE) and illusion of control. Nevertheless, given the R-O encoding deficiency found in SZ (Morris et al., 2018), and enhanced faulty causal relationship in schizotypy (Lee et al., 2005).

In summary, I hypothesized: *i)* that there will be an enhanced illusion of control in schizotypy (probably positive schizotypy); and *ii)* the possible action weighting disruption in schizotypy would moderate $p(R)$ and PC. The latter would indicate a possible reduced sensitivity to using proprioception in PC, but it may be also possible that referential behaviour is displayed by a reduced sensitivity to external sources of control. Finally, in the next section, I described the evidence and theoretical context for the development of the task used in this chapter.

4.1.3. Contingency Judgments with Two Agents

In the previous chapter, I used ΔP to measure the change in likelihood two stimuli, $S2$ and $S1$. But what happens when more than one predictive event (stimulus or response) is associated with the occurrence of an O, as usually displayed in the real world? Relevant studies for the current chapter

come from a selective association task, like the *Tank and Plane task* (Baker et al., 1993, 2000; Dickinson et al., 1984a; Vallée-Tourangeau et al., 1994).

Baker et al., (1993) asked participants to estimate PC and between two causes and one effect (O): a tank successfully crossing a mined field. The potential causes of the tank survival (O) were: *i)* camouflaging a tank, depending on participants responses (R), and *ii)* a plane, occurring independently of participants −*i.e.*, a visual stimulus (S)−. They observed a discounting effect, a more predictive factor reduces PC over a less predictive one (Baker et al., 1993). This effect was replicated by using different cover story and neutral stimuli (Vallée-Tourangeau et al., 1994). These causal competition effects can be used to developed mental models of causation which are based on associative learning computations (Baker et al., 1996, 2005; Dickinson, 2001; cf., Matute et al., 2019).

The following two experiments were similar in concept, wherein contingencies changed across conditions. In Experiment 4, participants were instructed to learn their control in the presence (*Stim.*) and in absence (~*Stim.*) of a stimulus (S). Therefore, the consequences of R were conditional upon the presence or absence of S, similar in concept to an occasion setter in instrumental learning literature (*e.g.*, Bonardi & Ward-Robinson, 2001). A cue signals the occasion to expect or not a change in the contingencies between responding and the O. Experiment 4 aimed to show that participants can learn simultaneously two control estimates, and that illusion of control (measured as a PC bias) will be related with positive schizotypy. In Experiment 5, participants were instructed to estimate their own control (*Hum.*; human), and another agent's control, here a simulated agent (the computer; *Comp.*) while playing the same free-operant task. Thus, I expected to observe the contrast effect previously reported between competing causes (Baker et al., 1993; Vallée-Tourangeau et al., 2005a). Finally, the moderation effect of schizotypy will be tested in both experiments. The differences in contingency calculations between experiments are stated in **Table 4.1**.

Table 4.1. Two Agents' Contingency Table and Formulas for Experiments 4 (E4) and 5 (E5)

	Agents:	**E4: Occasion Setter; E5: Computer (Comp.)**			
	Events:		**Stim. (S)**	**~Stim (~S)**	
Contingency Table	**Human (Hum.)**	**Response (R)**	p(*O*|*R*&*S*)	p(*O*|*R*&~*S*)	p(*O*|*R*)
		~Response (~R)	p(*O*|~*R*&*S*)	p(*O*|~*R*&~*S*)	p(*O*|~*R*)
			p(*O*|*S*)	p(*O*|~*S*)	
Contingency Formulas		**ΔPs**		**Average Probabilities**	
	E4 - Conditional	$\Delta P_{\text{Stim.}}$ = p(*O*|*R*&*S*) - p(*O*|~*R*&*S*)		p(*O*|*R*) = [p(*O*|*R*&*S*) + p(*O*|*R*&~*S*)]/2	
		$\Delta P_{\text{~Stim.}}$ = p(*O*|*R*&~*S*) - p(*O*|~*R*&~*S*)		p(*O*|~*R*) = [p(*O*|~*R*&*S*) + p(*O*|~*R*&~*S*)]/2	
	E5 - Marginals	$\Delta P_{\text{Hum.}}$ = p(*O*|*R*) - p(*O*|~*R*)		p(*O*|*S*) = [p(*O*|*R*&*S*) + p(*O*|~*R*&*S*)]/2	
		$\Delta P_{\text{Comp.}}$ = p(*O*|*S*) - p(*O*|~*S*)		p(*O*|~*S*) = [p(*O*|*R*&~*S*) + p(*O*|~*R*&~*S*)]/2	

Note: *O* = Outcome; *R* = Response; *S* = visual stimulus; ~ = absence of.

4.2. Experiment 4

In this experiment, participants learnt their control in the presence (*Stim.*) and absence (~*Stim.*) of the occasion setting stimulus (*S*; see **Figure 4.1C**) in a scenario designed to test causal control of an outcome, to turning on a light bulb as much as possible. I manipulated the R-O contingency, measured as conditional ΔP, for each *Stim.* and ~*Stim.* The correlation between both contingencies $\Delta P_{Stim.}$ and $\Delta P_{\sim Stim.}$ was null, to reduce the possibility that learning the control of one state (*e.g.*, *Stim.*) inform the control of the other state (~*Stim.*) in further conditions. This experiment investigates the illusion of control when participants learnt two independent contingencies conditional to a visual stimulus and its relationship with positive schizotypy (*cf.*, Baker et al., 1993). This experiment aimed to show that people learn two independent PC estimates, and that they do not depend on visual characteristic of *S*.

4.2.1. Methods

4.2.1.1. Participants

We ran this experiment in 4 samples, each one presented in **Table 4.A1**. In total, we collected all sociodemographic and psychometric data from N = 98. We used Gorilla Experiment Builder® (www.gorilla.sc) to create and host our experiment (Anwyl-Irvine et al., 2020). Data was collected between January 28th and March 14th 2021. Participants were recruited through Amazon Mechanical

Turk® (www.MTurk.com). In total 60 were females, 37 males, and 1 non-binary, between 24 and 65 years old; M = 37.48; SD = 9.96.

4.2.1.2. Questionnaires

At the end of the task, participants were asked to respond to two validated questionnaires, one for schizotypy personality and the second for depressive traits. For schizotypy we used the short O-LIFE (Mason & Claridge, 2006), where a larger score indicates higher schizotypy. For depression, I used the BDI-2 (Beck et al., 1961), where a higher score refers to lower mood and greater likelihood of depression. I included depression, to see the specificity of our schizotypy results as I know that contingency learning is highly affected by depressive traits (Byrom et al., 2015; Msetfi et al., 2013, 2017).

4.2.1.3. Design

Participants estimated their PC with a rating ranging between -10 to 10. Negative values meant that pressing a key (R) prevented the outcome (O; negative control); a 0 rating is for a null effect on O by R; and positive values represent that R increased the likelihood of generating the O (positive control). Participants rate their control for *Stim.* and ~*Stim*. Therefore, participants had to learn two contingencies simultaneously: $\Delta P_{Stim.}$ and $\Delta P_{\sim Stim.}$. Both of which were changing across eight conditions, randomized across participants. See **Figure 4.1A** to see an example of conditions for one random participant.

The experiment consisted of eight conditions, each condition with different configuration of the probabilities of the outcomes given the 4-events (*RS, R*~*S*, ~*RS*, and ~*R*~*S*) shown in **Table 4.1**. To see the exact probabilities of each of the four contingency table cells, see **Table 4.2**. The different conditions reflect experiences with positive, zero, and negative control for both *Stim.* and ~*Stim*. Each condition lasted for 40 seconds (or 40 trials), and at the end of each, people rated their PC for *Stim.* and ~*Stim.* Within a condition, in each 1-s trial, people can either press (R) or not press ($\sim R$). Participants instructions are presented in the **4.5.4. Appendix**.

Table 4.2. Conditions for Experiments 4 and 5

		Outcomes' Probabilities				Ratings' requirement				
Experiment	**Conditions' names**	**p(O|R&S)**	**p(O|R&~S)**	**p(O|~R&S)**	**p(O|~R&~S)**	**ΔP[Stim.]**	**ΔP[~Stim.]**	**ΔP[Hum.]**	**ΔP[Com.]**	**Samples**
4	A	1	0.5	0	0.5	1	0	*NA*	*NA*	1,2,3,4
	C	0	0.5	1	0.5	-1	0	*NA*	*NA*	1,2,3,4
	B	0.5	1	0.5	0	0	1	*NA*	*NA*	1,2,3,4
	D	0.5	0	0.5	1	0	-1	*NA*	*NA*	1,2,3,4
	A_and_B	0.75	0.75	0.25	0.25	0.5	0.5	*NA*	*NA*	1,2
	A_and_B	1	1	0	0	1	1	*NA*	*NA*	3,4
	C_and_D	0.25	0.25	0.75	0.75	-0.5	-0.5	*NA*	*NA*	1,2
	C_and_D	0	0	1	1	-1	-1	*NA*	*NA*	3,4
	A_and_D	0.75	0.25	0.25	0.75	0.5	-0.5	*NA*	*NA*	1,2
	A_and_D	1	0	0	1	1	-1	*NA*	*NA*	3,4
	B_and_C	0.25	0.75	0.75	0.25	-0.5	0.5	*NA*	*NA*	1,2
	B_and_C	0	1	1	0	-1	1	*NA*	*NA*	3,4
5	No Control	0.5	0.5	0.5	0.5	*NA*	*NA*	0	0	1,2,3,4
	PH	1	1	0	0	*NA*	*NA*	1	0	1,2,3,4
	NH	0	0	1	1	*NA*	*NA*	-1	0	1,2,3,4
	PS	1	0	1	0	*NA*	*NA*	0	1	1,2,3,4
	NS	0	1	0	1	*NA*	*NA*	0	-1	1,2,3,4
	Eq	1	0	0	1	*NA*	*NA*	0	0	1,2,3
	Di	0	1	1	0	*NA*	*NA*	0	0	1,2,3,4
	PH/PS	1	0	0	0	*NA*	*NA*	0.5	0.5	2,3
	PH/NS	0	1	0	0	*NA*	*NA*	0.5	-0.5	1,2,3
	NH/PS	0	0	1	0	*NA*	*NA*	-0.5	0.5	1,2,3
	NH/NS	0	0	0	1	*NA*	*NA*	-0.5	-0.5	2,3

Note: NA = Not Available; PH = Positive Human; NH = Negative Human; PS = Positive Simulation; NS = Negative Simulation; Eq = Equal; Di = Different. Numbers from "Samples" column comes from the samples in **Table 4.A1**.

2275

Finally, **Figure 4.1C** introduces each of the 4 event types occurring during every 1-second trial. *RS* events is when participants *R* in the presence of *S*; *R*~*S* events are participants respond (*R*) in the absence of the occasion setter (~*S*); ~*RS* events were defined as no participant response (~*R*) in the presence of *S*; and finally, ~*R*~*S* is when ~*R* in ~*S*, *i.e.*, the outcome occurring only in the context.

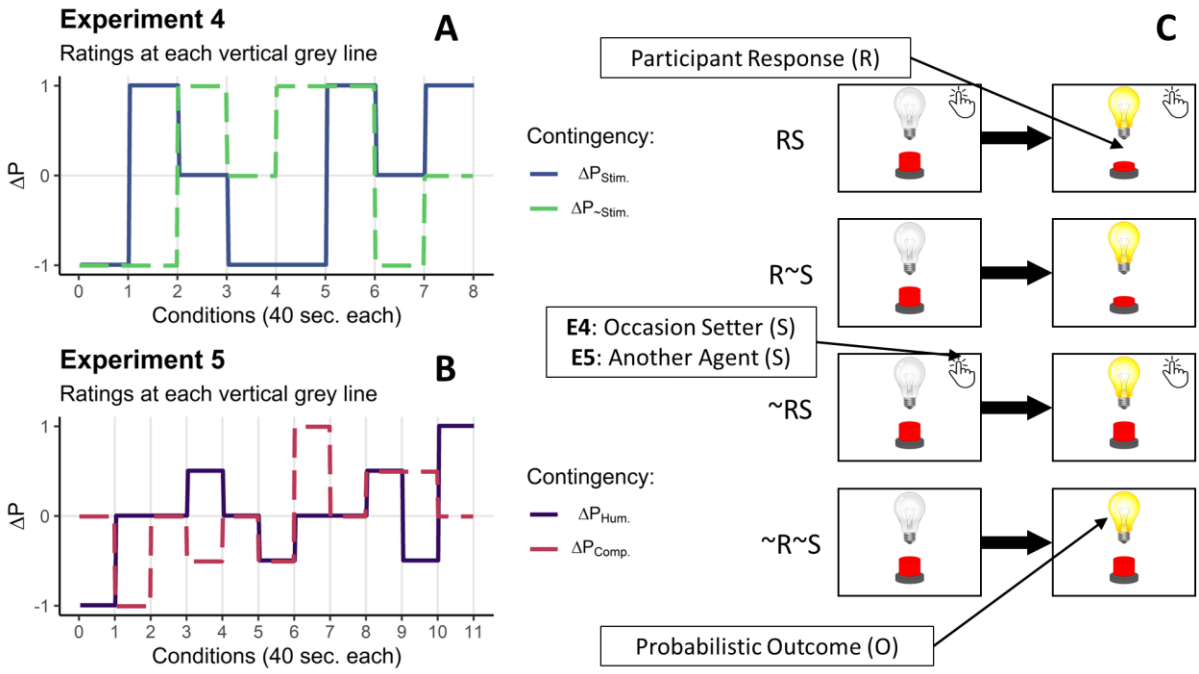


Figure 4.1. Paradigm Description for Experiments 4 and 5 (E4 and E5, respectively). Shows the experimental conditions and paradigm. **A** and **B**: shows the example of contingency condition sequences for one random participant, for both Experiment 4 and respectively Experiment 5. All participants received randomized conditions orders. **C**: shows the four possible events given the contingency table from **Table 4.1**.

Four different samples are compared which themselves constitute a 2 (O type) × 2 (deterministic, probabilistic) factorial design. Two samples received a light bulb turning on as the O and the other two a chicken hatching from an egg as the O; all the samples received the same four conditions, but for half of the samples we used four deterministic conditions and the second half four probabilistic conditions. I change the probabilistic/deterministic conditions across samples to collect a wider range of the ΔP dimension (see **Table 4.2**). Changing the quality of the O and the probabilistic

aspect of the O did not have any important effect, therefore we pooled the 4 samples into the main analysis.

4.2.1.4. Analysis

4.2.1.4.1. Dependent variables

The dependent variables were: *i)* PC (for both *S* present: *Stim.*; and *S* absent: ~*Stim.*); this was measure with ratings, which we rescaled by dividing them by 10, so PC was in the same dimension as the ΔP, *i.e.*, from -1 to 1; *ii)* PC error, which is the difference between each PC type (Stim. and ~Stim.) and their respective experienced ΔPs for each condition; *iii)* probability of responding during each condition, p(R); and finally, *iv)* behavioural conditional entropy (Treviño, Castiello, Arias-Carrion, et al., 2021; Treviño, Castiello, Torre-Valdovinos, et al., 2021), this is the classic conditional entropy from information theory (Shannon, 1948). The behavioural conditional entropy (BCE) measured how unpredictable next R is given the current trial. The more BCE, the more unpredictable is a behavioural pattern for a condition. I used this metric to explore the behavioural patterns, under the assumption that participants developed stereotypical (predictable) behaviours as long they interact with the task. Therefore, as we previously showed (Treviño, Castiello, Arias-Carrion, et al., 2021) this metric is a good predictor of choice biases (*i.e.*, predictable) in mice. I calculated BCE using classical information theory:

$$H(X|Y) = H(Y,X) - H(Y) \quad \text{(eq. 4.1)},$$

where X is R_{t+1} from bin 2 to bin 40, and Y is R_t from bin 1 to bin 39; and $H(Y,X)$ is the joint entropy which is calculated by:

$$H(X,Y) = -\sum_{x\in X}\sum_{y\in Y} p(x,y) * log_2[p(x,y)] \quad \text{(eq. 4.2)}.$$

Finally, $H(Y)$ is the Shannon entropy calculated as:

$$H(Y) = -\sum_{y\in Y} p(y) * log_2[p(y)] \quad \text{(eq. 4.3)}.$$

In sum, the BCE, $H(R_{t+1}|R_t)$, is a measure of how unpredictable a pattern of behaviour is. For instance, if the $p(R_{t+1}|R_t) = p(R_{t+1}|{\sim}R_t) = p({\sim}R_{t+1}|R_t) = p({\sim}R_{t+1}|{\sim}R_t)$, then BCE is maximum. This would make it impossible to predict the next $({\sim})R_{t+1}$ given the current $({\sim})R_t$. And would suggest that a participant is responding completely at random. Evidence suggests that mice develop stereotypical behaviours (Treviño et al., 2020). Therefore, if stereotypical behaviours emerged, we would expect a reduction in BCE. Finally, we normalized $H(R_{t+1}|R_t)$ to get values between 0 and 1 by dividing it with $ln(2)$ (Cuve et al., 2021; Shiferaw et al., 2019; Shiferaw et al., 2018).

4.2.1.4.2. Perceived Control Model

PC was measured by two ratings at the end of each condition *c*. PC was modelled using a hierarchical Bayesian model. So, each set of parameters were fit by participant as well as the group level parameter. The former was used for correlations with questionnaires, and the latter for the behavioural task inferences.

I assumed that ratings were drawn from a normal distribution with a mean μ and a SD or σ. Hence, PC in the presence of the stimulus (S) at the end of every condition *c* was:

$$Perceived\ Control = PC_c^S {\sim} N(\mu_c^S, \sigma^S) \qquad \text{(eq. 4.4)},$$

where σ is the fitted SD for the PC model. Also, PC for the absence of the stimulus (${\sim}S$) at every *c* was:

$$PC_c^{{\sim}S} {\sim} N(\mu_c^{{\sim}S}, \sigma^{{\sim}S}) \qquad \text{(eq. 4.5)}.$$

Each model μ was a weighted linear function of the current condition experienced control, ΔP_c; the previous experienced control, ΔP_{c-1}; the previous PC, PC_{c-1}. Experienced ΔP_c is different from the programmed ΔP given that participant's responses changed what participant's experience was at every c. Therefore, ΔP is a function of the outcomes that participants saw given their responses rather than what experimenters, programmed to happen. Then, PC μ (for both *Stim.* and ~*Stim.*) was:

$$\mu_c = b + m1 \times \Delta P_c + m2 \times \Delta P_{c-1} + m3 \times PC_{c-1}, \quad \text{(eq. 4.6)},$$

where b represents the overall PC bias, if $b > 0$ then participants tend to a positive perception of control over generating the O, otherwise $b < 0$ participants perceived a negative control preventing the O. $m1$ is the weight for the current experienced control or control sensitivity; $m2$ is the weight for the previous experienced control; and $m3$ weights the PC priors.

4.2.1.4.3. Moderation analysis

A moderator M is a variable that affects the relation between a cause X and an effect Y. Given previous evidence (Blanco et al., 2012; Byrom et al., 2015; Matute, 1996), I wasaware that there is a positive cause-effect relation between $\mathrm{p}(R)$ and PC. The second hypothesis was whether positive schizotypy would moderate the $\mathrm{p}(R)$ → PC relation. To test this, we fitted the next LMM:

$$PC \sim \beta_0 + \beta_1 * p(R) + \beta_2 * UE + \beta_3 * p(R) * UE, \quad \text{(eq. 4.7)}$$

where the moderation effect is defined as a significant β_3 (Zhang & Wang, 2017), *i.e.*, the interaction. The LMM had random intercept effects for participants nested for each sample. For this experiment each rating (for *Stim.* and ~*Stim.*) has an associated $\mathrm{p}(R)$, therefore two ratings per condition were considered for the analysis.

4.2.1.4.4. Summary and types of analysis

I conducted four stages of analyses. The first one, was a time series analysis, aimed to explore potential systematic bias (or formation of stereotypical behaviour) as a function of people gaining experience with the task. The second, aimed to model PC to estimate b, a measure of illusion of control. The third one correlated illusion of control (also as the grand mean of PC) with questionnaires, where I tested the illusion of control hypothesis and explored other correlations. Finally, the fourth was a moderation analysis of UE with $\mathrm{p}(R)$ and PC.

Time series. I tested whether the order of condition affected the dependent variables. For instance, if participants have a systematic bias toward perceiving more control as a function of conditions, PC will increase as a function of conditions order. If participants get better at estimating

experienced ΔP, a reduction of PC error would emerge as a function of conditions order. Also, if participants get bored across the experiment, I should expect a reduction in $\mathrm{p}(R)$ as a function of condition order. If participants developed specific behavioural patterns (see Treviño et al., 2018, 2020) I would expect to see a reduction in BCE as a function of conditions order.

Perceived control modelling. I modelled PC as previously described in section 4.2.1.4.2. In sum I estimated a bias toward perceiving generative or preventative O; the effect of the current experienced control (ΔP_c); the effect of previous experience control (ΔP_{c-1}); and the effect of previous PC (PC_{c-1}). I fitted this model using the Stan R package *rstan* (Stan Development Team, 2021) in R studio (R Core Team, 2021). I compared the hyperparameters when fitting the model to get the posterior densities. I did not expect any differences in the bias nor sensitivity between the two PC ratings (*Stim.* and ~*Stim.*), if S does not interfere with learning.

Individual differences analysis. This analysis was used to corroborate the first hypothesis whether illusion of control was related with positive schizotypy as well as to explore the relation with questionnaires and any other behavioural characteristics that were associated with schizotypy. If continuous variables were not normally distributed, I used Spearman rank-order correlation, otherwise it was Pearson correlation. To assess statistical significance, I used an $\alpha = 0.05$. To adjust for multiple comparisons, I used the Bonferroni correction, hence α / (8 dependent variables × 6 psychometric scales) = 0.001041667 ≈ 0.001.

Moderation analysis. There is a known relation between $\mathrm{p}(R)$ → PC (Blanco et al., 2012; Byrom et al., 2015; Matute, 1996). This method was used to test whether positive schizotypy reduced the effect of the relation between R and PC.

4.2.1.5. Software used and open access declaration

The R packages that we used were the following. LMMs were run with *lme4* (Bates et al., 2015)and *lmerTest* (Kuznetsova et al., 2017). Additional packages used for data processing were *reshape2* (Wickham, 2007) and *dplyr* (Wickham et al., 2022); models' standardized estimates and

export csv tables were obtained by *report* (Makowski et al., 2021), and data visualization used *ggplot2* (Wickham, 2016), *viridis* (Garnier et al., 2021), and *ggpubr* (Kassambara, 2020). The analysis codes and data are available under request.

4.2.2. Results

4.2.2.1. Time series

In **Figure 4.2** we present the PC ratings, PC error, p(R), and behavioural conditional entropy [$H(R_{t+1}|R_t)$] as a function of condition order. I ran LMMs for each subpanel and included participants as a random intercept and condition order as random slope, and for the case of PC ratings and PC error, I included rating type as a dummy variable. I did not find an effect of condition order for PC rating (**Figure 4.2A**; *t*s < 0.17, *p*s > 0.5), PC error (**Figure 4.2B**; *t*s < 0.7, *p*s > 0.5), nor p(R) (**Figure 4.2C**; $t < 0.9$, $p > 0.3$). However, I observed a reduction in $H(R_{t+1}|R_t)$ (**Figure 4.2D**; $t = -3.3$, $p < 0.005$) which suggests that participants developed predictable button press patterns as a function of training, confirming previous findings (Treviño et al., 2018, 2020).

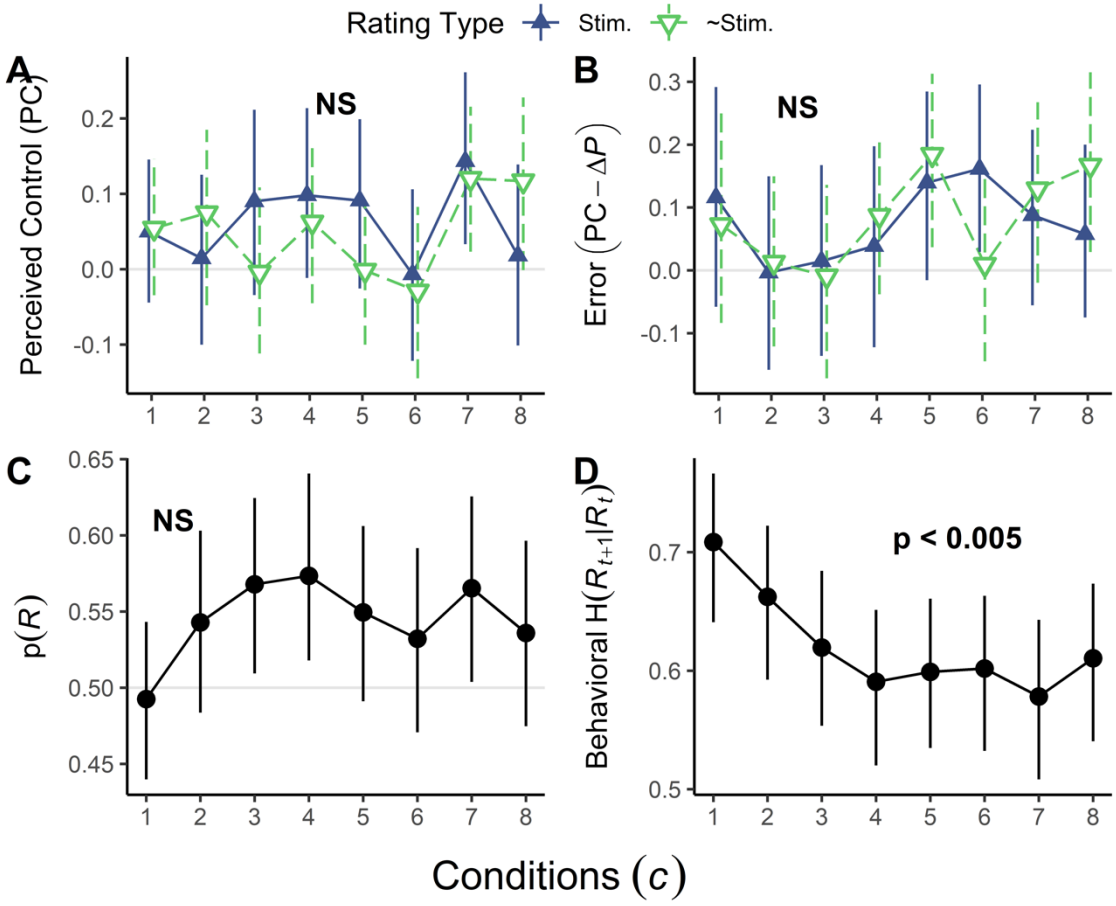


Figure 4.2. Time Series for Experiment 4. Shows the four dependent variables as a function of condition order. Here we did not consider *R-O* contingencies. **A**: PC. **B**: PC error, PC minus experienced Δ*P*. **C**: probability of response, p(*R*). **D**: behavioural conditional entropy, more = unpredictable response patterns. Error bars are 95% CI with bootstrap method as previously used. The presented p values and NS referred to the effect of condition order on the dependent variables based on a LMM per panel.

4.2.2.2. Delta P and Ratings

Figure 4.3A displays both PC types as a function of the two programmed Δ*P*s: for stimulus present (*Stim.*) and stimulus absent (*~Stim.*). Visually it seems that ratings correlate positively with their respective Δ*P*s, but this was not the case for ratings with opposite Δ*P*. **Figure 4.3B** shows the probability of response [p(*R*)] as a function of Δ*P*, this relation was linear low p(*R*) when preventing the *O* and high p(*R*) when producing *O*, a result that has been previously reported (Blanco et al., 2011; Byrom et al., 2015). Finally, **Figure 4.3C** introduced the hyperparameters (group level) of the PC model. For each PC type (*Stim.* and *~Stim.*) parameters' pairs, I also presented the mean to assess the main effect (independent of the occasion setter) and the difference between *Stim.* and *~Stim.*

Next, I conducted a Bayesian inference analysis (**Figure 4.3C**), where for each parameter estimation I presented the mean (M) and 95% credible intervals (CrI). Hence, if the 0 is not included in the 95% CrI I assume a credible difference from 0, which represent no bias, no effect, or no difference.

Based on all the data set, I found that participants have a bias (b) toward perceiving positive control on generating an outcome (*i.e.*, illusion of control), M = 0.055 (95% CrI; from 0.019 to 0.089), however no difference was found between *Stim.* and ~*Stim.* biases, M = -0.012 (95% CrI; -0.081 to 0.059). In addition, confirming the visual impression from **Figure 4.3A** I found a positive association between PC and sensitivity ($m1$) to experienced control ΔP_c, M = 0.259 (95% CrI; 0.202 to 0.313), but no difference in $m1$ between *Stim.* and ~*Stim.*, M = 0.028 (95% CrI; -0.082 to 0.131). Also, I did not find any effect on the current ratings from prior experience of control ($m2$) or ΔP_{c-1}, M = -0.016 (95% CrI; -0.047 to 0.016), nor a difference in $m2$ between *Stim.* and ~*Stim.*, M = 0.036 (95% CrI; -0.026 to 0.1). However, I found a mean positive association between ($m3$) the previous PC(*c-1*) and current ratings, M = 0.089 (95% CrI; 0.027 to 0.154), but no difference in $m3$ between *Stim.* and ~*Stim.*, M = 0.05 (95% CI; -0.077 to 0.177). Finally, given that the SD should be positive, it does not make sense to infer 95% CI from SD against 0. But the relevant comparison is the difference in rating types of SD, which they were not different, M = -0.121 (95% CrI; -0.286 to 0.031).

Overall, I found evidence that people learn the cuing value of the occasion setter in a free-operant task and that responding was related with the likelihood of generating and preventing the outcome, perhaps due to occasion setter inhibitory associated with the outcome (*e.g.*, Castiello et al., 2021). In addition, I found no difference in PC when an occasion setter is present. However, I found that participants used their current experience control (ΔP_c), but they did not use the prior experienced control (ΔP_{c-1}), and interestingly people's PC is influenced by the previous PC (PC_{c-1}). Experiment 4 suggests that people were able to discriminate and correctly PC in two independent states (signalled by a visual stimulus). To study how people, attribute control to external sources, in the next experiment, we used the S not as an occasion setter but as another agent competing for the O control.

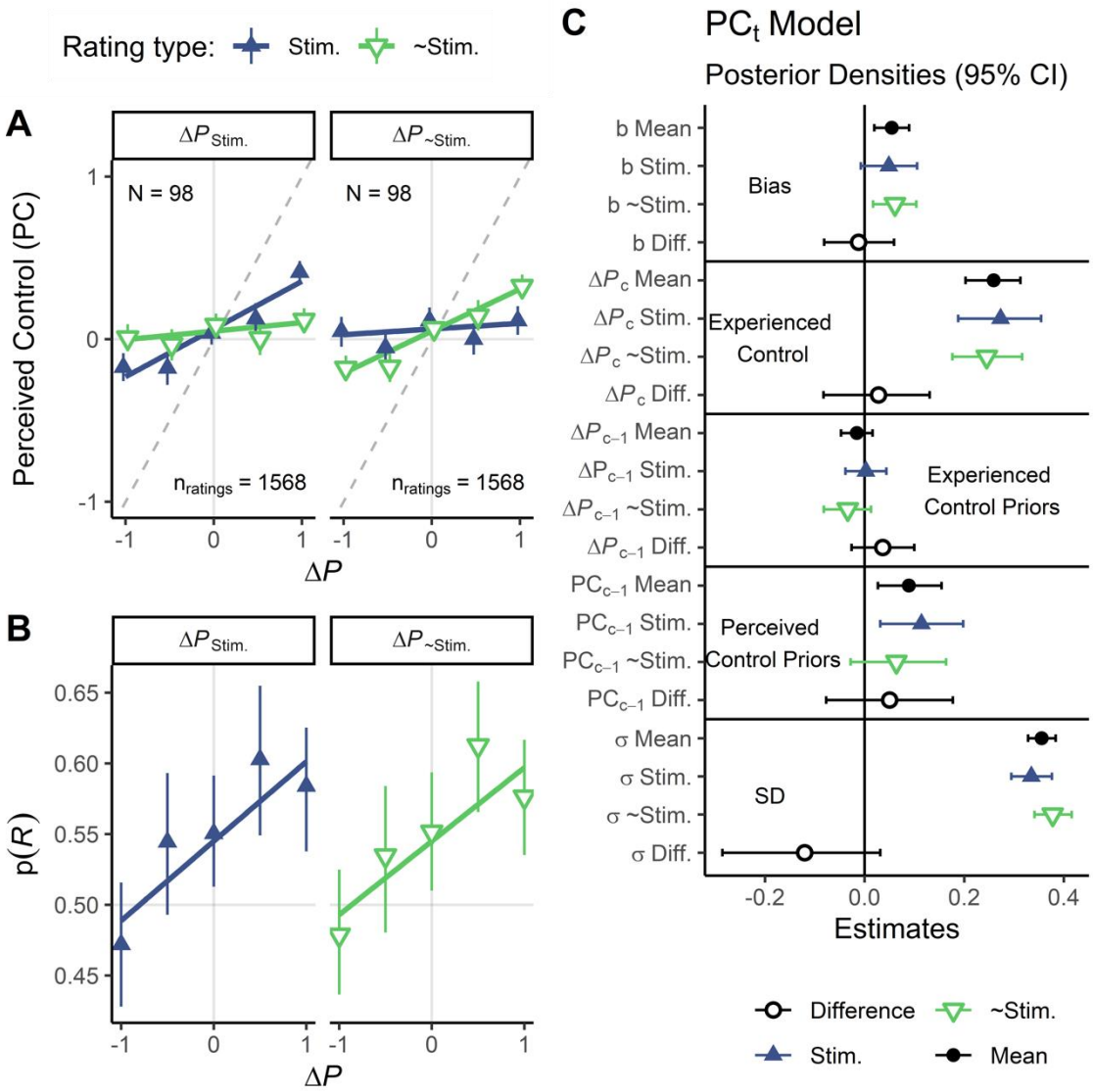


Figure 4.3. Main Results for Experiment 4. **A**: PC ratings as a function of conditions programmed Δ*P*s. **B**: probability of responding (*R*) as a function of conditions programmed Δ*P*s. Error bars are 95% CI with bootstrap method. **C**: Bayesian Hierarchical model fit hyperparameters with posterior densities, for: the bias, current experienced control (ΔP_c), the prior experienced control (ΔP_{c-1}), prior PC (PC_c) and SD (σ) of PC ratings. Error bars represents the 95% Credible Intervals from the posterior densities. The black dot represents the mean between the two states (*Stim.* and *~Stim.*) and empty dot is the difference the two states' parameters.

4.2.2.3. Correlations: Behaviour, Models, and Questionnaires

In this section, I correlated participants' behavioural variables and model parameters with the questionnaires. Due to lack of normal distribution in the variables, I used Spearman rank-order correlations. Also, I ran simple correlations with Bonferroni correction (6 questionnaires x 8

behavioural and model variables; given that Mean PC and Bias are practically the same). The results are presented in **Table 4.3**.

After the Bonferroni adjustment I observed two significant correlations in two subscales and one dependent variable and one model parameter. I observed that UE are positively correlated with the mean of PC ratings [$rho(96) = 0.34$; $p < 0.001$]; and IN positively correlates with and bias [$r(96) = 0.34$; $p < 0.001$]. Both mean PC and bias are practically the same. These results suggest that UE and IN are associated with an increase illusion of control.

4.2.2.4. Moderation analysis: Positive Schizotypy, Responses and Perceived Control

As expected, and replicating previous findings (Blanco et al., 2013; Byrom et al., 2015; Matute, 1996), I observed a significant and positive effect on $p(R)$ in PC, $\beta_1 = 0.52$, $p < 0.001$. Also, there was an effect in UE and PC, $\beta_2 = 0.05$, $p < 0.001$, this effect is the same as the illusion of control hypothesized for the previous results section. Finally, the moderation effect was negative $\beta_3 = -0.05$, $p < 0.05$ (see Experiment 4 in **Figure 4.4**). This may suggest that UE is related with a reduced sensitivity to use responses to inform their PC.

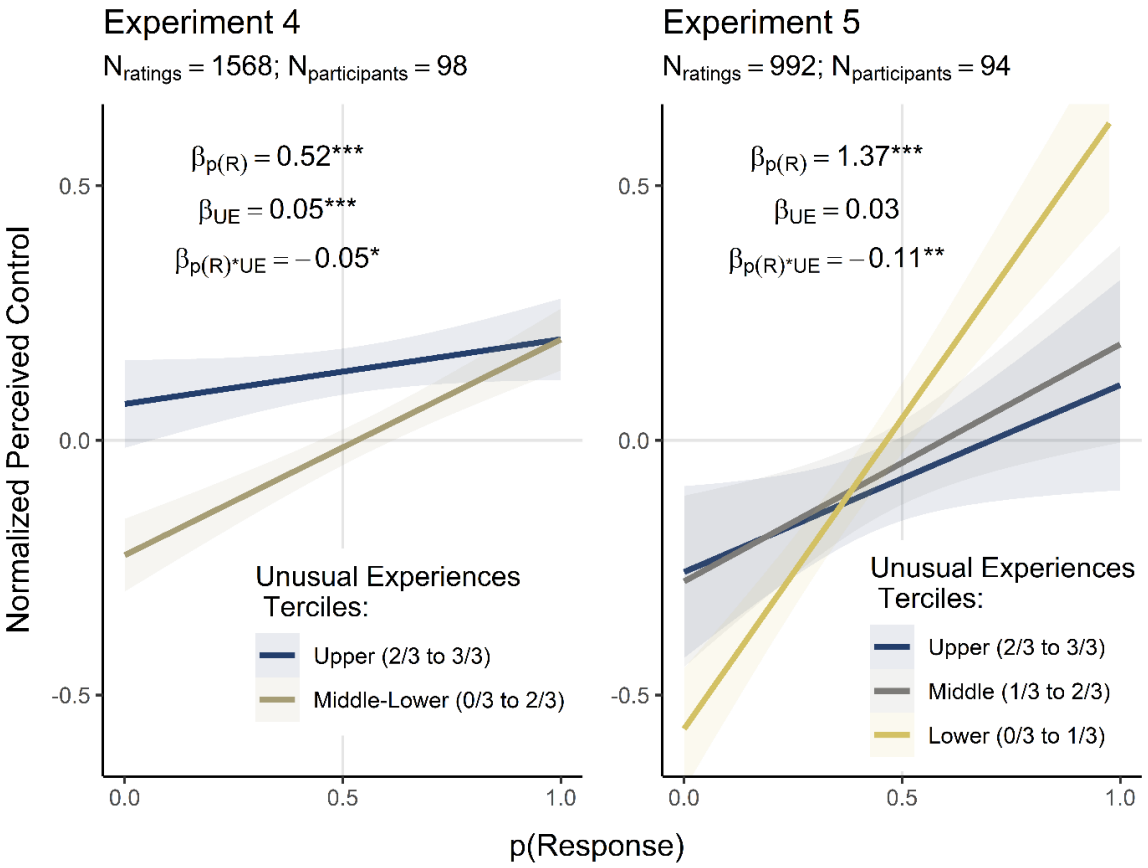


Figure 4.4. Moderation Analysis Results for Experiments 4 and 5. Effect of PC (y-axis) as a function of probability of responding (x-axis), for matter of visualization we showed data set split by UE terciles, In E4, middle and lower boundary were the same, therefore we combine both terciles. Left panel E4, and Right panel E5. A moderation analysis was conducted for each experiment. Probability of responding [p(Response) = p(*R*)], UE, and their interaction were regressed against the normalized PC estimate.

4.2.3. Discussion and Summary

Experiment 4 revealed that people learn simultaneously two different and non-correlated Δ*P*s. Also, I found that participants developed predictable behavioural patterns as a function of exposure with the task. But, the most important findings were the negative positive correlation between UE and illusion of control and the negative mediation effect of UE in p(*R*) and PC. In Experiment 5 we aimed to replicate the relation between UE and illusion of control when a competing source of control is in play. Finally, in Experiment 4 the control estimates were both for the participant's control, but conditional to the occasion setter, signalled by *S*. In Experiment 5, participants are asked to integrate information competing information from another agent (albeit artificial) to PC (see **Table 4.1**).

2490

Table 4.3. Experiment 4: correlations between behavioural task outputs and questionnaires

	Behavioural signatures	Psychometric self-reported scales					
		OLIFE	UE	CD	IA	IN	BDI2
Behavioural Variables	Mean PC	r = 0.24; p < 0.05	r = 0.34; p < 0.001*	r = -0.02; NS	r = 0.03; NS	r = 0.32; p < 0.01	r = 0.01; NS
	Mean Error	r = 0.2; p < 0.05	r = 0.25; p < 0.05	r = 0.02; NS	r = 0.04; NS	r = 0.26; p < 0.01	r = 0.03; NS
	Mean p(R)	r = -0.13; NS	r = -0.02; NS	r = -0.22; p < 0.05	r = -0.18; NS	r = 0.06; NS	r = -0.22; p < 0.05
	Mean BCE	r = 0.05; NS	r = 0; NS	r = 0.23; p < 0.05	r = -0.05; NS	r = -0.15; NS	r = 0.26; p < 0.01
	H(O|R)	r = 0.14; NS	r = 0.13; NS	r = 0.05; NS	r = 0.1; NS	r = 0.16; NS	r = -0.05; NS
Model Parameters	Bias	r = 0.23; p < 0.05	r = 0.31; p < 0.01	r = -0.05; NS	r = 0.05; NS	r = 0.34; p < 0.001*	r = -0.01; NS
	ΔP(c)	r = -0.11; NS	r = -0.19; NS	r = 0.07; NS	r = -0.07; NS	r = -0.18; NS	r = 0.04; NS
	ΔP(c-1)	r = -0.04; NS	r = -0.05; NS	r = 0.06; NS	r = -0.13; NS	r = 0.04; NS	r = -0.06; NS
	PC(c-1)	r = 0.03; NS	r = 0.05; NS	r = 0.04; NS	r = -0.13; NS	r = 0.12; NS	r = 0.03; NS

Notes: BCE = Behavioural Conditional Entropy; H(O|R) = entropy of outcome (O) given response (R); PC = perceptual control; Error = difference between PC and experienced ΔP; p(R) = probability of response; Bias = mean rating under the PC model, can also be interpreted as a bias toward perceiving positive or negative control; ΔP(c) =experienced control; ΔP (c-1) = experienced control priors; PC(c-1) = perceived control priors. O-LIFE = short schizotypy self-reported questionnaire, UE = Unusual Experiences (O-LIFE subscale); CD = Cognitive Disorganization (O-LIFE subscale); IA = Introvertive Anhedonia (O-LIFE subscale); IN = Impulsive non-conformity (O-LIFE subscale); and BDI2 = depression self-reported questionnaire; NS = non-significant; * means significant results with Bonferroni adjustment, $\alpha = 0.05/(8*6) = 0.001041667$.

2491

4.3. Experiment 5

Given the findings of Experiment 4, I aimed to replicate the positive correlation between schizotypy UE and PC (Balzan et al., 2013; Moritz et al., 2014) and the moderation effect of UE in p(R) and PC. Although, in Experiment 4 participants were instructed to learn their control and another agent control (**Figure 4.1C**), as well as try to turn on a bulb as much as they can. In Experiment 4 participants always rated their own control, but in two different states (*Stim.* and ~*Stim.*).

Experiment 5 involved a second putative agent (*e.g.*, Baker et al., 1993), where participants learnt two causal factors: their own responses (R) and an uncontrollable external stimulus (S). Previous research found that stronger contingent factors diminish the weaker ones. Baker et al. (1993), for example, found that when the external cue (S) was a stronger predictor of the O, participants' PC on their own responses was diminished or discounted. This was assessed as an interaction of PC type and contingency (ΔP) and was accounted for by a model of associative competition (Rescorla & Wagner, 1972). Thus, given Baker et al. (1993) results, we expected to find an interaction between self-PC and external PC. The aim of this experiment is to replicate the enhanced illusion of control in schizotypy and to test the source of referential behaviour (internal or external).

4.3.1. Methods

4.3.1.1. Participants

We tested 133 participants from four different samples (see **Table 4.A1)**, from which we have sociodemographic data from three samples. Therefore, complete data from 94 participants, 61 females and 33 males, between 18 and 50 years old (M: 24.18, SD: 7.91). The *bp* sample (n=68) were people who living Oxfordshire and were not paid for their participation. The *ch* sample (n=38) was a sample recruited online from China. The *lab* sample (n=17) where participants who assisted to the lab before the COVID-19 pandemic and who conducted this task as a second activity during a one-hour session. Finally, the *web* sample (n=10) were recruited by snowball method during one week for collecting the

rest of samples. For the web collection, we used the same program as Experiment 4. However, for the participants that were ran in the *lab*, the task was built in PsychoPy2® (Peirce et al., 2019).

4.3.1.2. Questionnaires

The same questionnaires were used as Experiment 4. However, I did not collect questionnaires for the fourth (*ch*) sample, nor BDI-2 for the first (*lab*) sample.

4.3.1.3. Design

The design that we used for this experiment was like Experiment 4. Participants received multiple conditions and were asked to estimate two non-correlated contingencies which changed every 40 seconds. At the end of each 40-second condition, participants reported, *i)* their PC (*Hum.*), *ii)* the computer (*Comp.*) control, and *iii)* the control from other factors (*Oth.*). See **Figure 4.1B** to see the example of one participant contingencies across the task.

The difference between Experiment 5 and Experiment 4, is that participants were told that another agent (*Comp.*) was playing the same task and when *Comp.* 'pressed the key', participants saw $S-$which in this experiment was the *Comp.'s* $R-$. Also, the contingencies (ΔP) in Experiment 5 were calculated as stated in **Table 4.1**, this is, the marginal conditional probabilities of O given R. Therefore, participants estimated the overall control for themselves and the computer agent. Participants instructions are presented in the **4.5.4. Appendix**.

4.3.1.4. Analysis

4.3.1.4.1. Dependent variables

I used the same four dependent variables to explore time series: *i)* PC, *ii)* PC error (PC minus experienced ΔP), *iii)* probability of responding, and *iv)* behavioural conditional entropy (BCE), which provides a measure how predictable is $(\sim)R_{t+1}$ with $(\sim)R_t$.

4.3.1.4.2. Perceived Control Model

I modelled PC in the same way as Experiment 4:

$$PC_c \sim N(b + m1 \times \Delta P_c + m2 \times \Delta P_{c-1} + m3 \times PC_{c-1}, \sigma), \qquad \text{(eq. 4.8)}.$$

The previous model was used for the self-control (*Hum.*) rating and for the external agent (*Comp.*) rating. In that model PC is a normal distribution centred in a linear combination of experienced control (ΔP_c), previous experienced control (ΔP_{c-1}), previous PC (PC_{c-1}), plus a bias b. Each factor was weighted by a parameter m (1 to 3). σ is the fitted SD of the PC model.

4.3.1.4.3. Moderation analysis

I conducted the same moderation analysis as Experiment 4. Although, for Experiment 5, the first PC rating assessed the experienced control by the participant (*Hum.*), and the second assessed the external cause (*Comp.*). Therefore, for the mediation analysis we only used the PC ratings provided for the human control.

4.3.1.4.4. Summary and types of analysis

As Experiment 4, I conducted four analyses. First, time series analysis (**Figure 4.5**) where I assessed the effect the four-dependent variables as a function of condition order. Second, I modelled PC with the factors mentioned in section 4.3.1.4.2. Third, individual difference section, I aimed to replicate the correlation between PC and UE, as well as explore PC fitted parameters and questionnaires. Finally, the fourth was a moderation analysis of UE with p(R) and PC.

4.3.1.5. Software used and open access declaration

I used the same software and packages declared in Experiment 4.

4.3.2. Results and discussion

4.3.2.1. Time series

Similar to **Figure 4.2**, in **Figure 4.5** I present PC, PC error, p(R) and BCE [$H(R_{t+1}|R_t)$] as a function of the condition order. I fitted LMMs for each subpanel and include participants as random intercept and condition order as random slope. In the case of PC and PC error I include rating type (*Hum.* vs *Comp.*; but no ratings for other factors) as a dummy variable. I did not find an effect on condition order in PC (**Figure 4.5A**; *t*s < 0.54, *p*s > 0.1), PC error (**Figure 4.5B**; *t*s < 0.11, *p*s > 0.1),

nor probability of responding (**Figure 4.5C**; $t < 1.2$, $p > 0.1$). However, I observed a reduction in BCE which suggests that across conditions participants sequences of responses become more predictable and can be interpreted as the development of a behavioural stereotypical (**Figure 4.5D**; $t = -3.85$, $p < 0.001$) replicating Experiment 4 results. In addition to the condition order effect on the dependent variables, the only significant effect was found in other source of control ratings (*Oth.*) in **Figure 4.5A**, where PC in *Oth.* Factors were overall larger than PC from *Hum.* and *Comp.*, $t = 4.4$, $p < 0.001$.

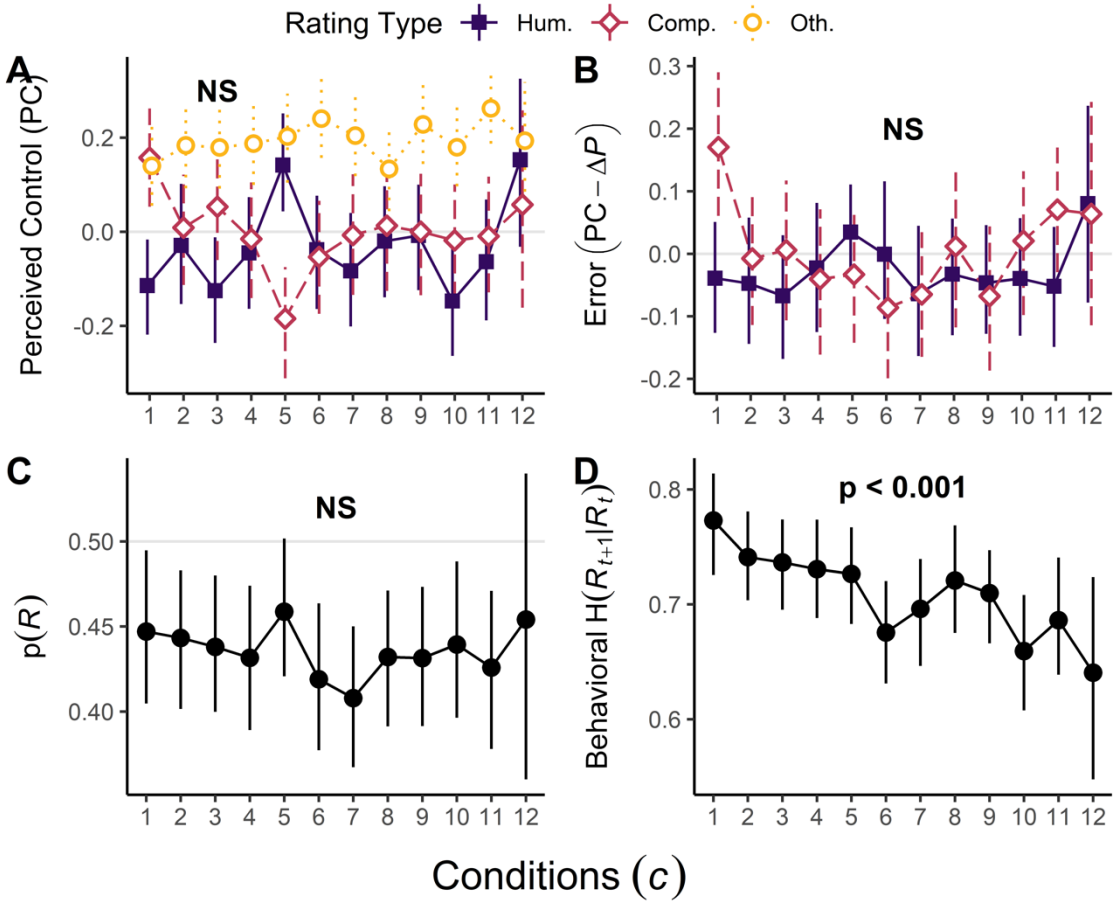


Figure 4.5. Time Series for Experiment 5. Shows the four dependent variables as a function of condition order. Here, I did not consider *R-O* contingencies. **A**: PC. **B**: PC error, PC minus experienced ΔP. **C**: probability of response, p(R). **D**: behavioural conditional entropy, more = unpredictable response patterns. Error bars are 95% CI with bootstrap method used previously. The presented p values and NS referred to the effect of condition order on the dependent variables based on a LMM per panel.

4.3.2.2. Delta P and Ratings

Figure 4.6 presents the results from Experiment 5 results and are similar to **Figure 4.3** from Experiment 4. **Figure 4.6A** displays rating types as a function of the two ΔP: participant control

(*Hum.*) and computer control (*Comp.*). It seems that PC correlates positively with the respective ΔPs, and an interaction (or contrast) effect for the opposite PC, *e.g.*, when $PC_{Hum.}$ increased as function of $\Delta P_{Hum.}$, $PC_{Comp.}$ decreased, and vice versa. This is that an increase in PC_{Hum} was related with a decrease in PC_{Comp}, and vice versa. Finally, the PC for *Oth.* factors seem to be negatively correlated with both ΔPs. **Figure 4.6B** shows that the $p(R)$ increase as a function of $\Delta P_{Hum.}$; however, this was not the case for $\Delta P_{Comp.}$. Finally, **Figure 4.6C** shows the hyperparameters (group level) from the rating model: bias (intercept or *b*), experienced control (ΔP_c), prior experience control (ΔP_{c-1}), prior PC (PC_{c-1}), and rating SD (σ). To assess an overall PC effect, I presented the mean effect between human (*Hum.*) and computer (*Comp.*) PC models. In addition, to compare both PCs I obtained the parameters difference between *Hum.* minus *Comp*.

I found that participants did not show a bias toward perceiving positive action- contingencies, M = -0.01 (95% CrI; from -0.036 to 0.016), nor a specific bias for their own control and the computer control; and no difference between human and computer biases, M = -0.032 (95% CrI; -0.085 to 0.019). Therefore, no illusion of control was found as a main effect. In addition, confirming the visual impression from **Figure 4.6A**, I found a positive relation between PC and experienced control, ΔP_c, M = 0.693 (95% CrI; 0.651 to 0.736), and a larger sensitivity to experienced human control ($\Delta P_{Hum.}$) than to computer control ($\Delta P_{Comp.}$), M = 0.162 (95% CrI; 0.074 to 0.246). This suggests that participants were more sensitive estimating their control (*i.e.*, the effect of R over O) than another agent's control (*i.e.*, the effect of S over O).

Furthermore, I found a negative association between prior experienced control or ΔP_{c-1} and current PC, M = -0.061 (95% CrI; -0.111 to -0.011; **Figure 4.6C**, ΔP_{c-1} Mean), but no difference between prior experience control between *Hum.* and *Comp.*, M = 0.078 (95% CrI; -0.021 to 0.174; **Figure 4.6C**, ΔP_{c-1} Diff.). In concordance with Experiment 4, I found an overall positive association between prior PC(*c-1*) on the current PC, M = 0.062 (95% CrI; 0.017 to 0.11; **Figure 4.6C**, PC(*c-1*) Mean); however, no difference between human and computer prior PC influence, M = -0.088 (95% CrI; -0.179 to 0.002; **Figure 4.6C**, PC(*c-1*) Diff.). Finally, given that SD should be positive, it does

not make sense to infer 95% CI from SD against 0, but the relevant comparison is the difference SDs. I found that SD for human was smaller than SD for computer, M = -0.122 (95% CrI; -0.207 to -0.036; **Figure4. 6C**, σ Diff.).

Overall, in Experiment 5 I did not find an illusion of control, *i.e.*, bias towards generating the outcome (O). Not surprisingly, I found that PC was sensitive to the current experience control (ΔP_c). Notably, participants were more sensitive to their PC than the other agent control (**Figure 4.6C**, ΔP_{c-1} Diff.). In addition, I found a negative relation between previous experienced control (ΔP_{c-1}) and the current PC, especially for *Comp*. This means that $PC_{Comp.}$ was reduced when previous *Comp.* ΔP_{c-1} was high. Further, I found a positive effect in PC_{c-1} and PC, so when participants perceived positive control on the previous condition (*i.e.*, c-1) they more likely perceived positive control in the next condition. The same effect was found in Experiment 4. Finally, I found a lower SD for $PC_{Hum.}$ than $PC_{Comp.}$, which means that the people are less variable when PC their own control than another and external source of control.

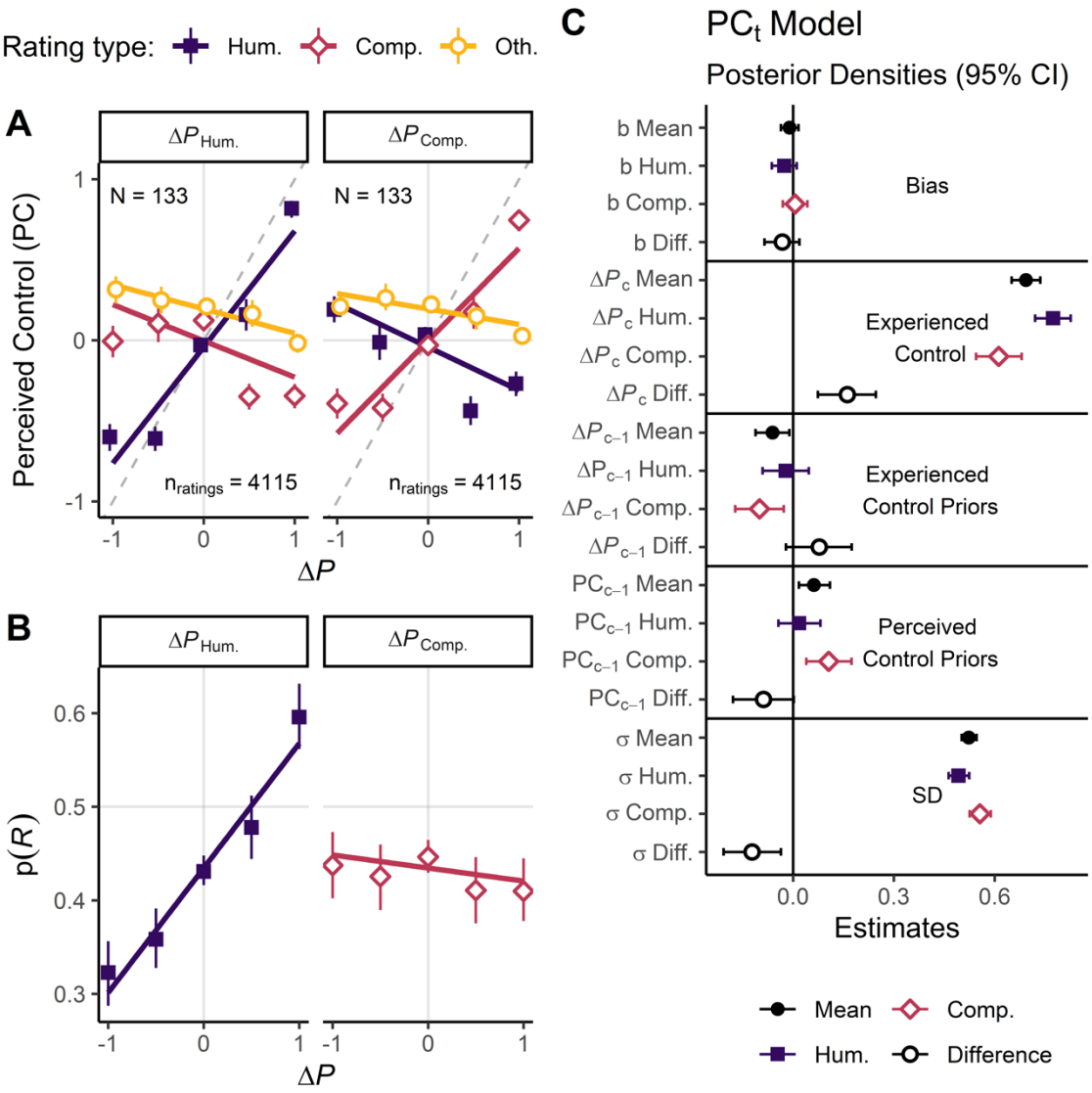


Figure 4.6. Main Results for Experiment 5. **A**: PC as a function of programmed Δ*P*s. **B**: probability of responses (*R*) as a function of programmed Δ*P*s. Error bars are 95% CI with bootstrap method. **C**: Bayesian Hierarchical model fit hyperparameters with posterior densities, for: the bias, current experienced control (ΔP_c), the prior experienced control (ΔP_{c-1}), prior PC (PC_c) and SD (or σ) of PC ratings. Error bars represents the 95% Credible Intervals from the posterior densities. The black dot represents the mean between the two estimates (*Hum.* and *Comp.*) and empty dot is the difference between *Hum.* and *Comp.* parameters.

4.3.2.3. Correlations: Behaviour, Models, and Questionnaires

In this section of Experiment 5, I conducted the same correlational analysis as I did with Experiment 4 (**Table 4.4**). Correlations between behaviour and model parameters with questionnaires. Given that I did not find an illusion of control in Experiment 5 it may be that the lack of replication may be due to the task differences. For instance, in Experiment 4 participants were not required to

estimate another source of control as in Experiment 5. It may be that introducing another causal factor (*e.g.*, *Comp.*) prevents people from perceiving an illusion of control.

Nonetheless, there is an alternative explanation, a study found that when mild depressed people focused on the context, their PC increased resulting in an illusion of control (Msetfi, Brosnan, et al., 2016). It is possible to think that the computer agent in Experiment 5 reduced participants' attention to the context, and therefore may have supressed the illusion of control. If this could be confirmed, then these findings, including Msetfi et al (2016), points out a way to remove the illusion of control by splitting attention to the context, or adding a competing causal factor (*e.g.*, another agent). A study testing these institutions shall unveil the uncertainty emerging from our results.

2645

Table 4.4. Experiment 5: correlations between behavioural task outputs and questionnaires

	Behavioural signatures	Psychometric self-reported scales					
		OLIFE	UE	CD	IA	IN	BDI2
Behavioural Variables	Mean PC	r = -0.03; NS	r = -0.01; NS	r = -0.06; NS	r = 0.03; NS	r = 0; NS	r = -0.06; NS
	Mean Error	r = 0.02; NS	r = 0.07; NS	r = -0.02; NS	r = -0.01; NS	r = 0; NS	r = 0.01; NS
	Mean p(R)	r = 0.03; NS	r = 0.13; NS	r = 0; NS	r = -0.01; NS	r = 0.03; NS	r = -0.01; NS
	Mean BCE	r = 0.01; NS	r = 0.04; NS	r = 0; NS	r = -0.04; NS	r = -0.04; NS	r = -0.12; NS
	H(O|R)	r = -0.05; NS	r = -0.14; NS	r = 0.01; NS	r = -0.01; NS	r = 0.03; NS	r = -0.12; NS
Model Parameters	Bias (Hum.)	r = -0.05; NS	r = -0.06; NS	r = -0.07; NS	r = -0.08; NS	r = -0.01; NS	r = -0.19; NS
	ΔP(c) (Hum.)	r = 0.12; NS	r = -0.06; NS	r = 0.14; NS	r = 0.1; NS	r = 0.2; NS	r = 0.02; NS
	ΔP(c-1) (Hum.)	r = -0.1; NS	r = -0.08; NS	r = -0.1; NS	r = -0.01; NS	r = -0.12; NS	r = -0.06; NS
	PC(c-1) (Hum.)	r = -0.04; NS	r = -0.04; NS	r = -0.14; NS	r = 0.03; NS	r = 0; NS	r = 0.01; NS
	Bias (Comp.)	r = 0.08; NS	r = 0.15; NS	r = 0.05; NS	r = 0.01; NS	r = 0.06; NS	r = 0.14; NS
	ΔP(c) (Comp.)	r = -0.15; NS	r = -0.21; $p < 0.05$	r = -0.12; NS	r = 0.07; NS	r = -0.15; NS	r = 0.01; NS
	ΔP(c-1) (Comp.)	r = -0.06; NS	r = 0.03; NS	r = -0.07; NS	r = -0.05; NS	r = -0.06; NS	r = 0.06; NS
	PC(c-1) (Comp.)	r = 0.14; NS	r = 0.11; NS	r = 0.13; NS	r = 0.06; NS	r = 0.03; NS	r = -0.1; NS

Notes: BCE = Behavioural Conditional Entropy; H(O|R) = entropy of outcome (O) given response (R); PC = perceptual control; Error = difference between PC and experienced ΔP; p(R) = probability of response; Bias = mean rating under the PC model, can also be interpreted as a bias toward perceiving positive or negative control; ΔP(c) =experienced control; ΔP (c-1) = experienced control priors; PC(c-1) = perceived control priors. O-LIFE = short schizotypy self-reported questionnaire, UE = Unusual Experiences (O-LIFE subscale); CD = Cognitive Disorganization (O-LIFE subscale); IA = Introvertive Anhedonia (O-LIFE subscale); IN = Impulsive non-conformity (O-LIFE subscale); and BDI2 = depression self-reported questionnaire; NS = non-significant; * means significant results with Bonferroni adjustment, $\alpha = 0.05/(8*6) = 0.001041667$.

2646

4.3.2.4. Moderation analysis: Schizotypy, Responses and Perceived Control

The moderation results are presented in **Figure 4.4**. The findings are like those reported in Experiment 4. I found a negative moderation effect. UE moderates negatively the effect between $p(R)$ and PC, β_3 = -0.11, p < 0.01. Nevertheless, the effect $p(R)$ on PC was positive, β_1 = 1.37, p < 0.001. However, no main effect on UE on PC, β_2 = 0.03, p = NS. In sum, the negative moderation may suggest that UE reduce the impact of responding in PC.

4.3.3. Discussion and Summary

Experiment 5 did not show an illusion of control nor the correlation with UE. However, it unveiled the larger sensitivity to perceived self-control (*Hum.*) than external source of control (*Comp.*; **Figure 4.6C**, ΔP_c Diff.), and the same moderation effect of UE in $p(R)$ and PC. This could mean that: *i)* the PC from high UE individuals is less influenced by responses; or *ii)* they were less motivated to produce the outcome. However, if the latter was confirmed, UE should be related with $p(O)$, but that was not case (no significant estimate in linear model between $p(O)$ and UE; p > 0.1). Notably, as in Experiment 4, I found that participants developed predictable behavioural patterns as a function of exposure with the task, *i.e.*, reduction of behavioural conditional entropy (BCE).

4.4. Discussion

In this chapter I presented two experiments using a free-operant paradigm in general population. I hypothesized that UE −a subscale of schizotypy 4-dimensional Claridge model− would be positively related with illusion of control. In Experiment 4, and confirming the literature (Balzan et al., 2013; Blanco et al., 2015; Griffiths et al., 2019; Matute et al., 2011), I found positive correlation between UE and illusion of control (higher perception of control). However, the data of Experiment 5 suggests that when introducing a competing external source of control (*e.g.*, a computer agent), the illusion of control may disappear as well as its relationship with UE or this was simply a failure to replicate the effect of Experiment 4. Notably, in both experiments I found that UE is related with a reduction in the effect of actions on perceiving control (UE moderates' probability of responding and PC).

The discussion section is divided in five sections. First, I present the results with respect to schizotypy and how UE relates with response-outcome (R-O) contingency estimation. Second, I described what I call the agency paradox in a competing control situation. Third, I proposed a two-agent computational model using Behavioural Control Theory and classic associative learning models. In the fourth section, I discussed limitations and some closing remarks, and fifth and final my conclusions.

4.4.1. Illusion of Control, Unusual Experiences, and Referentiality?

Experiments 4 and 5 suggest that a competing source of control may prevent illusion of control. Although this would be a between experiment result. Previous studies have found the illusion of control only in conditions when participants estimate their (self-)control with no competing cause (Balzan et al., 2013; Blanco et al., 2015; Griffiths et al., 2019; Matute et al., 2011), as in the current Experiment 4 although not in Experiment 5. The only difference between Experiment 4 and 5 was that in the latter the other agent (represented as a visual stimulus S) competed for the control over O. Thus, a competing source of control may be the reason why Experiment 5 did not show the correlation higher illusion of control in high UE scorers. An interesting phenomenon from Experiment 5, although not surprising, is that the PC sensitivity to the other agent is lower than the self-PC sensitivity.

The relation between p(R) and PC has been shown previously (Byrom et al., 2015; Matute, 1996). If p(R) maps directly onto PC then in addition to the UE moderation effect I would expect a negative correlation between UE and sensitivity to $\Delta P_{Hum.}$, but it this correlation is not significant [r(92) = -0.08, p = NS]. Hence, the relation between p(R) and PC may not be as strong as evidence suggests or an alternative explanation may be needed. Curiously, even if Experiment 5 did not show illusion of control and the overall PC bias was not correlated with UE, I found a negative correlation between UE and sensitivity to external control, *i.e.*, $\Delta P_{Comp.}$ (**Table 4.4**). This could suggest a referential behavioural trait, which is common in schizotypy (Lenzenweger et al., 1997). Altogether, this could suggest that referential thoughts are formed by a reduced weight to external sources of

control. This result is in concordance with results suggesting that people scoring high in schizotypy rate similar sensations produced by themselves and by external causes (Lemaitre et al., 2016; Whitford et al., 2017b). Thus, if that would be the case, participants with high UE scores would not increase $p(R)$ because they already perceive themselves to be the cause of the O, at least more so than the low scorers. So, there is no tendency to increase response rates.

To see the specificity of these results we conducted sensitivity analysis with depression and overall schizotypy (see **4.5.3. Appendix**). In summary, I did not find that either depression or overall schizotypy score moderated the effect $p(R)$ → PC. This suggests a specific UE effect.

4.4.2. Agency Paradox as a Contrast Effect

In this section, I describe what the agency paradox may be and discuss the relevance of making the illusion of control disappear by adding a competing factor. Thus, this section discusses only results from Experiment 5.

Baker et al (1993) found that an external but more predictive cue reduced PC on a self-generated cue, *i.e.*, interaction between PC and external event (Vallée-Tourangeau et al., 1998). Here, I confirmed and generalize those findings, where the interaction was found in the whole of the contingency dimension (ΔP). This interaction can be seen as a contrast effect. For example, when the other agent has control generating the O, the PC over the self becomes negative (as if preventing the O). Interestingly, this also happens the opposite way, when the other agent had no control and participants had positive control, they perceived the other agent as preventing the O. Given the previous findings, I consider that the agency paradox is a contrast effect between self and external PCs, and viceversa. One of the potential causes of this contrast effect is a differential weighting to the event types. For instance, to observe this effect, the weight ($\alpha_{\sim R \sim S}$, see next section) of the events where no agent responded, *i.e.*, observing the outcome only in the presence of the context, must be 0. I formalized this idea bellow.

4.4.3. Behavioural Control Theory for two agents: simulations

In this section based on Behavioural Control Theory (Sosa & Alcalá, 2022), I theorized and described two agents (as feedback learners) interact in the task implemented in Experiment 5. For the next simulation we assumed two exact similar agents (Agent 1 and Agent 2) who graphically presented in **Figure 4.7**. In this simulation, agents learn from each, not as the 'simulated' computer from Experiment 5; *Comp.* was a Bernoulli experiment with probability of 0.5.

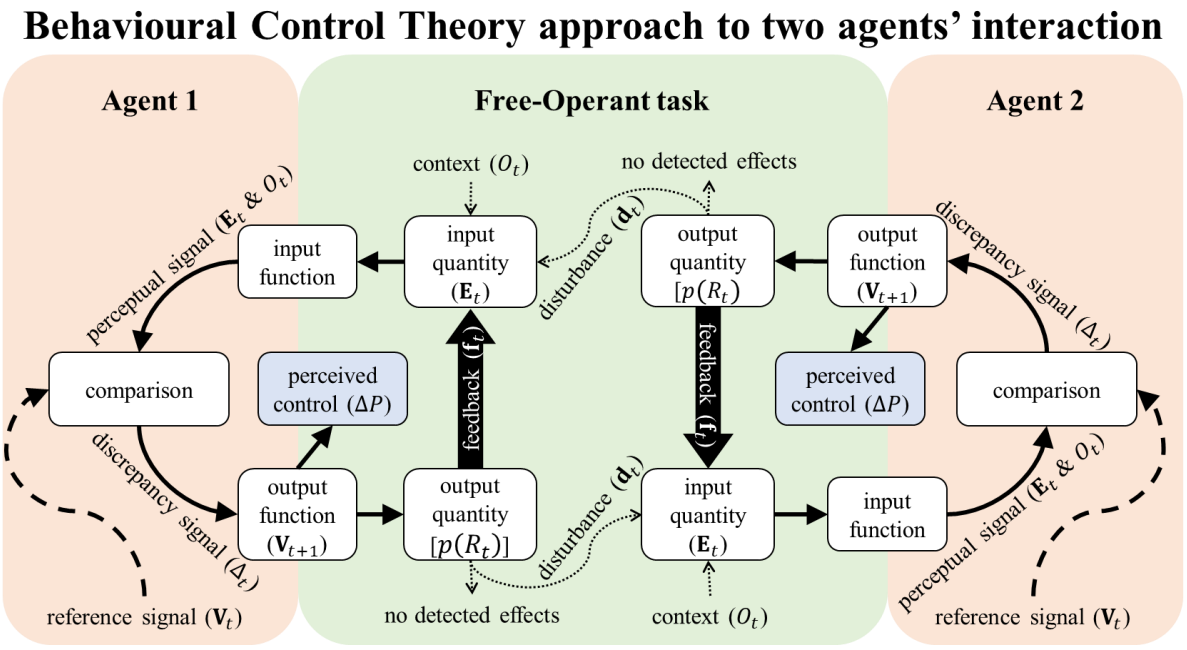


Figure 4.7. Behavioural Control Theory for Two Agents. Assuming an agent as a feedback controller (Power, 1973; Powers, 1973; Sosa & Alcalá, 2022). This figure presents all the elements to consider for a two agents' interaction in a specific shared environment (*e.g.*, two agents free-operant task). The light red rectangles represent 'inside' the (nervous/silicon) system of each agent. In light green the behavioural task or shared context between the agents. In blue the contingency judgments or PC by each agent, which does not affect the other agent, *i.e.*, they are not disturbances.

The current model of two interactive agents assumes that an agent's behaviour is a feedback control loop (Sosa & Alcalá, 2022). The more important idea is that the *disturbances* to the *input quantities* for each agent are the *output quantities* of the other agent, and vice versa (**Figure 4.7**). For a matter of simplicity, in this simulation I did not model *context* disturbances nor *no detected effects*.

Therefore, the *input quantities* from one agent are *disturbances* ($\mathbf{d}_t$) from the other agent (*e.g.*, Agent 2) and the *feedback* produced from the self-agent or agent of reference (*e.g.*, Agent 1; $\mathbf{f}_t$), where:

$$\mathbf{d}_t = \begin{vmatrix} S \\ \sim S \end{vmatrix};\ \mathbf{d}_t = \{0,1\}, \text{ and } \mathbf{f}_t = \begin{vmatrix} R \\ \sim R \end{vmatrix};\ \mathbf{f}_t = \{0,1\} \quad \text{(eq. 4.9)},$$

which *input quantity* (or observed state), represented by the event type matrix at t ($\mathbf{E}_t$; see **Table 4.1**):

$$\mathbf{E}_t = \mathbf{f}_t \bullet \mathbf{d}_t{}^{\mathbf{T}} = \begin{vmatrix} RS & R\sim S \\ \sim RS & \sim R\sim S \end{vmatrix};\ \mathbf{E}_t = \{0,1\} \quad \text{(eq. 4.10)}.$$

The *input function* could transform $\mathbf{E}_t$; however, to simplify things we assumed that $f(\mathbf{E}_t) = \mathbf{E}_t$. Then the *comparison* produces the *discrepancy signal* (Δ; Stout & Miller, 2007) which is a function of the elements conforming the *perceptual signal* ($\mathbf{E}_t$ and O_t), and the *reference signal* ($\mathbf{V}_t$), hence:

$$\Delta_t = f(O_t, \mathbf{V}_t, \mathbf{E}_t) = [\mathbf{E}_t O_t] - [\mathbf{E}_t \odot \mathbf{V}_t] \quad \text{(eq. 4.11)},$$

where $\odot$ is the Hadamard product (element wise multiplication); $O_t = \{0,1\}$ is the outcome; and Δ_t is a matrix with the *discrepancies* at t. In addition, $\mathbf{V}_t$ is the value matrix (*i.e.*, O probability tracker) for each of the events:

$$\mathbf{V}_t = \begin{vmatrix} v_{RS} & v_{R\sim S} \\ v_{\sim RS} & v_{\sim R\sim S} \end{vmatrix} \quad \text{(eq. 4.12)},$$

and follows the delta update learning rule (Bush & Mosteller, 1951; Rescorla & Wagner, 1972; Sutton & Barto, 2018):

$$\mathbf{V}_{t+1} = \mathbf{V}_t + \mathbf{A} \odot \Delta_t \quad \text{(eq. 4.13)}.$$

Here, $\mathbf{A}$ is a matrix with LRs or weights (α) for each event type (Kao & Wasserman, 1993; Wasserman et al., 1993):

$$\mathbf{A} = \begin{vmatrix} \alpha_{RS} & \alpha_{R\sim S} \\ \alpha_{\sim RS} & \alpha_{\sim R\sim S} \end{vmatrix} \quad \text{(eq. 4.14)}.$$

Weights are free parameters, therefore αs can be estimated by participant based on PC *output quantity* (blue rectangle in **Figure 4.7**). Finally, the outputs of this model are two: *i)* the probability of response [p(R)], and *ii)* the PC ratings. For the former, I used SoftMax function with an inverse temperature parameter (β) and an input based on the difference between $v_{RS} - v_{\sim RS}$ or $v_{R\sim S} - v_{\sim R\sim S}$ depending on the other agent behaviour (*i.e.*, S on the screen). Then the probability of R at every *t* is:

$$p(R_t) = \frac{1}{1+exp\langle -\beta*(v_R - v_{\sim R})\rangle}, \quad \text{(eq. 4.15).}$$

For the latter, I used the same approach to model PC ratings as in the experiments. This is, PC were drawn from a normal distribution with mean centred at a weighted experienced control (ΔP_c) which is a function of **V** at the end of a condition, *i.e.*, $\Delta P_c = f(\mathbf{V}_c)$ (see **Table 4.1**); wherefore:

$$PC \sim N(m1 * f(\mathbf{V}_c), \sigma) \quad \text{(eq. 4.16).}$$

I used the above-described model to simulate Experiment 5 with two agents. The weight for experienced control, $m1$, were set empirically based on the slopes from **Figure 4.6A**. Therefore, the self-PC weight was larger than the external-PC. Both agents had the same **A** and β where the values were 0.2 and 2, respectively. The results are presented in **Figure 4.8**. Agent 1 and 2 showed inverse results, as expected. 'They' both learnt the different levels of control and the p(R) was sensitive to their own control as shown in Experiment 5. For both agents, α were 0.2, except for $\alpha_{\sim R\sim S}$ for Agent 1 which was 0.

In sum, Behavioural Control Theory (BCT) provides a platform of how participants integrate external and internal feedback and CLT provides the function to judge control. Therefore, BCT and CLT are not mutually exclusive, and the combination of them provides a framework to model multiagent scenarios with agent responses and judgments of control. The agency paradox name used to describe the contrast effect found in competing PC scenarios implies a difference in the context weighting. However, it is important to remark that the two interacting agents used here do not incorporate the complete learning mechanisms of feedback control (Sosa & Alcalá, 2022), future investigation should improve this multiagent framework scenario.

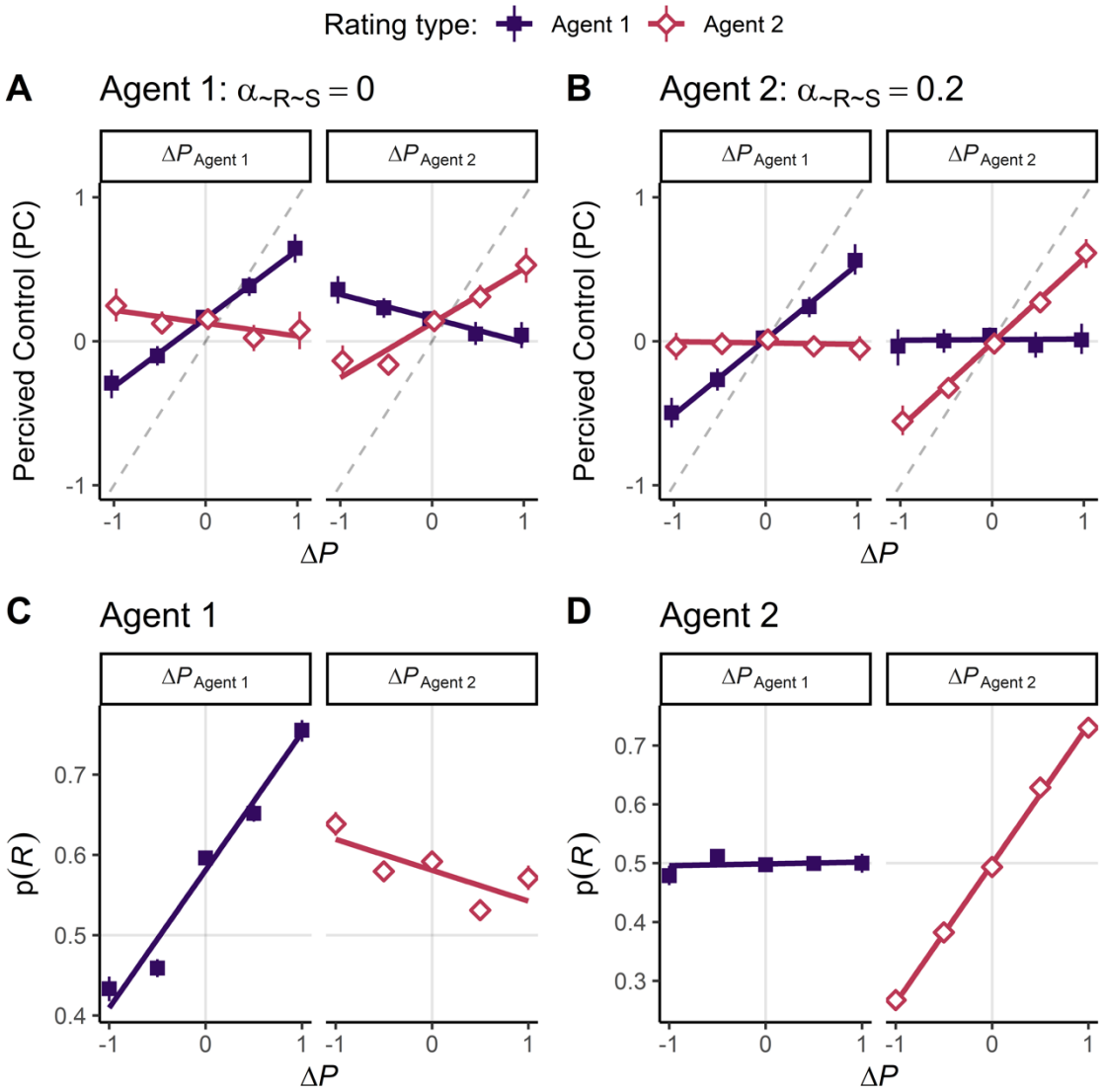


Figure 4.8. Simulation of Two Agents in a Free Operant Task. **A** and **B** show the PC and **C** and **D** the probability or responding [p(R)] of two artificial agents. Agent 1, left panels, and Agent 2, right panels. These results are congruent with results from Experiment 5.

4.4.4. Closing remarks and Limitations

An interesting finding observed in both experiments is the reduction of behavioural conditional entropy (BCE) as a function of experiencing more conditions. BCE measures how unpredictable the next response is given the current response, and the reduction of it suggests that participants' response patterns are more predictable. This result could motivate a whole dissertation. However, that result was not the main hypothesis of this chapter. Therefore, we only considered it as interesting and possible to motivate future research.

Behavioural Control Theory (BCT) is an old idea of what is the nature of behaviour (Powers, 1973; Sosa & Alcalá, 2022). This an approach that links associative learning and behaviourism with Control Theory and provides some heuristics to think about human behaviour. Nonetheless, in this chapter I also used a definition of behavioural control that it is not completely described BCT but is defined by Huys & Dayan (2009). They define behavioural control as actions reducing outcome entropy (see **4.5.1. Appendix**).

The critical limitation is that I interpreted that the correlation between illusion of control and UE disappeared when introducing another agent. This is because that conclusion is based on comparing two independent experiments (4 and 5). Ideally, a future experiment should point this limitation and assess both experiments (Experiment 4: two self-control ratings vs Experiment 5: one self-control and one external-control) in one. Thus, a within-subject design would be ideal for future research on illusion of control and schizotypy.

4.4.5. Conclusions

I found that schizotypy UE were positively correlated with an illusion of control only when participants estimated their own control (Experiment 4). Although, when another agent or competing source of control was introduced the illusion of control disappeared as well as its correlation with UE. Experiment 5 found a negative correlation between UE and sensitivity to the external agent. In addition, in both experiments I found a negative moderation effect of UE in the positive relation between probability of response and PC. Finally, the agency paradox is a contrast effect between oppositive experience and PC, and to emerge agents downweight the event type with outcome occurring in the background.

4.5. Appendix 4

4.5.1. Information Theory and Contingency Learning Theory

When can we say that a response R controls an outcome O? Theories of behavioural control (Huys & Dayan, 2009) point out that one way to interpret that an O is being controlled by a R is in the case where R reduce the uncertainty or entropy of the O. Although, this conceptualization of control is not completely linked (yet) with the Behavioural Control Theory presented in the first section of the chapter. By entropy (H) we refer to the classical Shannon formulations (Shannon, 1948). So, the entropy of an outcome is calculated as follows:

$$H(O) = -\sum_{o \in O} p(o) * log_2[p(o)] \quad \text{(eq. 4.A1),}$$

where higher H implies more bits of information (can be considered as sampling stages in discreate temporal moments) required to describe O. Although, in a bivariate case, when another variable is in play, *e.g.*, R, it is possible to conditionalize H of O given R. As following:

$$H(O|R) = H(R, O) - H(R) \quad \text{(eq. 4.A2),}$$

where $H(O, R)$ is the bivariate join entropy, and $H(R)$ is the entropy of the R (calculated as eq. 4.A1). The joint entropy is:

$$H(R, O) = -\sum_{r \in R} \sum_{o \in O} p(r, o) * log_2[p(r, o)] \quad \text{(eq. 4.A3).}$$

In CLT, when I say that the presence or the absence of R does not provide information about O, as in the case of $\Delta P = 0$, then this is because the $H(O|R)$ is maximal. To explore the relation between Information Theory and ΔP. I conducted a parametric exploration of the 2×2 contingency table (A, B, C, and D), by running all the possibilities in a 40 trials free operant block where participants can either R or $\sim R$ and can either see an O or not $\sim O$. To do this, I ran permutations with replacement of vectors from 1 to 40 trials in 4 (given the four cells, see previous chapters). This is $P(n = 40, k = 4)$, however this number includes conditions where the sum of the 4 cells is not 40. Therefore, I removed the conditions where the sum of the 4 cells were not 40 and the conditions with the exact cell values.

Finally, the total number of contingency conditions with 40 trials was 7694. Each value with a unique frequency of the 2×2 contingency table and sum of 40.

Figure 4.A1A shows the $\mathrm{p}(O)$ as a function of the $\mathrm{p}(R)$ for all the permutated contingency tables; this shows the limits of the parametric exploration. **Figure 4.A1B** visualize $H(O|R)$ as a function of $\mathrm{p}(R)$, it seems that lower H is found around $\mathrm{p}(R) \approx 0.5$. **Figure 4.A1C** present the effect of O in $I(O,R)$, is evident the maximum case of mutual information is when $\mathrm{p}(O) \approx 0.5$. Finally, **Figure 4.A1D** shows $H(O|R)$ as a function of ΔP, as anticipated in the main text, when $\Delta P \approx 0$, there is maximal entropy and the farther away from 0 the lower values of entropy can be seen.

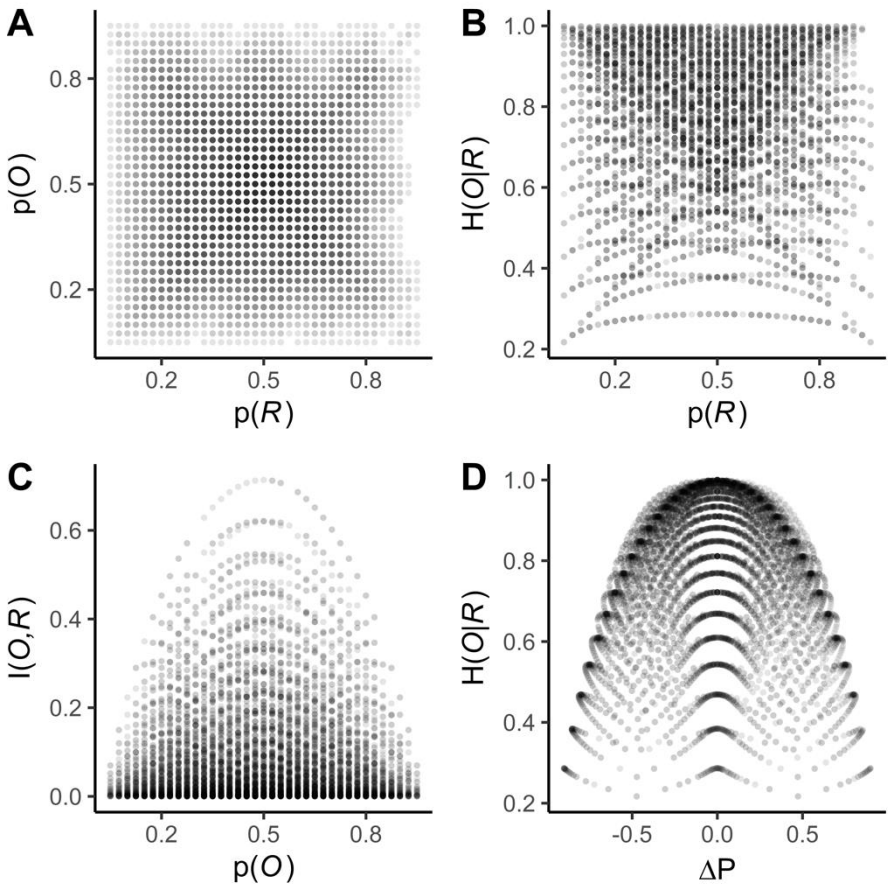


Figure 4.A1. Information Theory and Contingency Learning Theory. **A**: probability of outcome [$\mathrm{p}(O)$] as a function of probability of response [$\mathrm{p}(R)$]. **B**: entropy of outcomes given responses [$H(O|R)$] as a function of $\mathrm{p}(R)$. **C**: mutual information between outcomes and responses [$I(O,R)$] as a function of $\mathrm{p}(O)$. **D**: $H(O|R)$ as a function of ΔP.

4.5.2. Samples Characteristic

Table 4.A1. Sociodemographic and psychological characteristics from Experiments 4 and 5

			Experiment 4 (free-operant cue)				
			Bulb (1)	**Egg (2)**	**Egg det. (3)**	**Bulb det. (4)**	**Total**
	sample sizes:		n = 24	n =25	n = 24	n = 25	N = 98
Age	M (SD); [range]		40.25 (10.09); [24 - 58]	38.96 (11.66); [24 - 65]	37.54 (10.69); [28 - 65]	33.28 (5.37); [26 - 45]	37.48 (9.95); [24 - 65]
Sex	Freq. (%)	F	16 (66.67)	10 (40)	6 (25)	5 (20)	37 (37.76)
		M	8 (33.33)	14 (56)	18 (75)	20 (80)	60 (61.22)
		NB	NA	1 (4)	NA	NA	1 (1.02)
CD	M (SD); [range]		3.29 (2.71); [0 - 9]	2.72 (2.64); [0 - 8]	2.71 (2.91); [0 - 10]	1.72 (2.57); [0 - 8]	2.6 (2.73); [0 - 10]
IA	M (SD); [range]		3.62 (2.2); [0 - 9]	3.48 (2.12); [0 - 8]	3.17 (2.14); [0 - 9]	3.72 (2.13); [0 - 8]	3.5 (2.13); [0 - 9]
IN	M (SD); [range]		1.75 (1.65); [0 - 6]	2.32 (1.89); [0 - 7]	1.79 (1.77); [0 - 6]	3.08 (2.02); [0 - 8]	2.24 (1.89); [0 - 8]
UE	M (SD); [range]		0.96 (1.6); [0 - 6]	1.96 (2.94); [0 - 10]	1.71 (2.73); [0 - 9]	2.96 (2.85); [0 - 9]	1.91 (2.65); [0 - 10]
O-LIFE	M (SD); [range]		9.62 (5.05); [0 - 20]	10.48 (6.79); [0 - 26]	9.38 (6.34); [1 - 21]	11.48 (7.47); [2 - 27]	10.26 (6.44); [0 - 27]
BDI2	M (SD); [range]		11.83 (13.06); [0 - 47]	13.92 (15.83); [0 - 54]	9.92 (11.03); [0 - 35]	8.12 (11.07); [0 - 38]	10.95 (12.9); [0 - 54]
			Experiment 5 (agency paradox)				
			Pilot lab. (1)	**Pilot online (2)**	**Block Practical (3)**	**China (4)**	**Total**
	sample sizes:		n = 17	n = 10	n = 68	n = 38	n = 95; N = 133
Age	M (SD); [range]		27.06 (7.17); [20 - 49]	24.67 (4.77); [18 - 30]	23.4 (8.32); [18 - 50]	*NA*	24.18 (7.92); [18 - 50]
Sex	Freq. (%)	F	11 (64.71)	4 (40)	46 (67.65)	*NA*	61 (64.21)
		M	6 (35.29)	5 (50)	22 (32.35)	*NA*	33 (34.74)
CD	M (SD); [range]		5.44 (3.33); [0 - 10]	4.8 (2.78); [0 - 9]	5.63 (2.86); [0 - 11]	*NA*	5.51 (2.91); [0 - 11]
IA	M (SD); [range]		2.62 (2.09); [0 - 7]	1.3 (1.57); [0 - 5]	2.46 (2.17); [0 - 8]	*NA*	2.36 (2.12); [0 - 8]
IN	M (SD); [range]		3 (2.1); [0 - 8]	4.3 (2.75); [0 - 8]	3.72 (2.44); [0 - 10]	*NA*	3.66 (2.42); [0 - 10]
UE	M (SD); [range]		3.31 (2.8); [0 - 9]	4.2 (2.2); [0 - 8]	3.19 (2.72); [0 - 11]	*NA*	3.32 (2.68); [0 - 11]
O-LIFE	M (SD); [range]		14.38 (6.63); [3 - 25]	14.6 (6.8); [0 - 24]	15 (7.9); [2 - 35]	*NA*	14.85 (7.52); [0 - 35]
BDI2	M (SD); [range]		NA	10.3 (9.71); [0 - 27]	14.85 (11.65); [0 - 55]	*NA*	14.27 (11.47); [0 - 55]

Note: NB = Non-Binary; M = Mean; SD = Standard Deviation; CD = Cognitive Disorganization; IA = Introvertive Anhedonia; IN = Impulsive Nonconformity; UE = Unusual Experiences; O-LIFE = Oxford-Liverpool Inventory of Life and Experiences (schizotypy traits); BDI2 = Beck Depressive Inventory 2 (depressive traits).

4.5.3. Sensitivity Analysis of Moderation analyses effects

We conducted two sensitivity analysis for each experiment. On the first hand, the first analysis for Experiment 4 incorporated depression scores, which resulted in a null moderation, $\beta_3 <$ 0.1, p = NS. The second analysis included overall schizotypy score as the moderator, and the result was null as well, $\beta_3 <$ -0.1, p = NS. On the other hand, for Experiment 5 depression did not moderate, $\beta_3 >$ 0.1, p = NS, nor schizotypy, $\beta_3 <$ 0.1, p = NS. After this sensitivity analyses, we confirm that our main results are specific to UE.

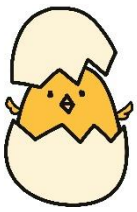
4.5.4. Participants instructions

Experiment 4 (E4)

E4. Instructions at the beginning of the task

This game will take approximately 11 minutes to complete.

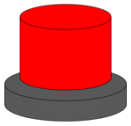
We want to test your ability to cause a baby chick to hatch from its egg by using a button. We have eight different buttons to test, one at a time during a 40 second 'testing period'. Here is what a chick in an egg looks like:



OR (depending on the sample)



Your job is to hatch as many baby chicks as you can. By clicking the 'B' key on your keyboard you can activate the red button:



The red button may or may not help chicks come out of their egg. Sometimes, a chick may come out even if you do not activate the red button. Your task is to learn whether it is best to activate or not activate the red button in order to release the chicks.

In addition, sometimes you will see a hand symbol. The hand may influence whether the chick hatches when you activate or do not activate the red button. Here is what the hand looks like:



At the end of each 40 seconds 'testing period', we will ask you to rate how much control you had over hatching chicks, both, when the hand was there and when it was not there.

Remember, try to release the chicks out of their egg as much as you can by either activating or not activating the red button.

Note that the control you have over hatching in each 'testing period' may differ. Similarly, the hand symbol may or may not change how much control you have over hatching chicks.

E4. Ratings after each condition

Rate the degree of control you had over helping the chicks to come out of the eggs when the hand symbol was present and when it was absent.

*To rate your control when the hand is present use the horizontal slider bars that say **Present hand **, and to rate the control when the hand is absent use the slider that says **Absent hand**.*

**+10 indicates that every time you pressed the key, a chick was released, and that every time you did not press the key, the chick was not released.*

**0 indicates that the chick emerged regardless of when you pressed or not pressed the button.*

**-10 indicates that every time you pressed the button, the chick did not come out, and that every time you did not press the key, the chick was released.*

You should choose a value between 0 and 10 and 0 and -10 to indicate that the relationship is not perfect (e.g., the chick comes out more when you press than do not press would indicate a value between 0 and 10, whereas the chick stays in the egg more when you do not press than when you press would indicate a value between 0 and -10)

Use the mouse and click on the horizontal lines.

Experiment 5 (E5)

E5. Instructions at the beginning of the task

The total duration of this game is approximately 12 minutes.

We are interested in how you learn to control different smart light bulbs by pressing a key on the keyboard.

While you are playing, another player (the computer) will also be trying to do control the lights.

There are 11 light bulbs and you will test each for 40 seconds.

Your goal is to try to keep each bulb lit much as possible.

To control the light you may press the 'b' key on the keyboard. The key may or may not make the light more likely to turn on (on the screen). Moreover, pressing the key might make it less likely to turn on. This can happen because sometimes the bulb will come on even if you do not press the key.

Furthermore while you are doing this the second player (computer) will be testing each bulb at the same time as you do. Every time the computer presses the key, a hand will appear on the screen.

The 11 different bulbs are independent of one another. That is to say what you have learned about your and the computer's control of one light will not help you learn about the next one. The "b" key will be used to control each light

*While **you must try to ensure the light is on as much as possible,** you should also try learn how much and what type of control you and the computer have over each light, because we will ask you this after you have tried to control each light.*

E5. Ratings after each condition

Rate the degree of control that you and the other player had over the bulb on, as well as other factors.

*To rate your control use the horizontal slider bars that say **my control**, and to rate the other player's control use the slider that say **computer control**, and to rate how much do you think that the outcome was produced by factors other than you and the other participant? use the **other factors** slider.*

** (+10) is when the light always came on when the key was pressed, and never when it was not pressed. You (or the computer player) made the light happen.*

** (0) is when it came on equally when the key was pressed and when it was not pressed.*

** (-10) is when the light never came on when the key was pressed, and always when it was not pressed. You (or the computer player) stopped the light from happening.*

Use the mouse and click on horizontal lines.

Chapter Five

Contingency Weighting in Schizotypy: General Discussion

In this last chapter, I summarise my results and interpretation at multiple levels of analysis: perceptual, associative, and instrumental (control and agency) learning. For my dissertation CLT provided an *umbrella* framework that unifies different time scales and relational qualities (seeing cues versus acting) into one theory that describes the statistical relations between events.

In Chapter Two, I focused on perceptual learning within the millisecond time frame. I used a perceptual contingency classification learning task, and I did not find evidence of hyperassociation −tendency to perceive positive correlation between stimuli−. Although, I found that schizotypy was related with larger Decision Time (DT) for correct trials. High schizotypy scorers seemed to gather more visual evidence, before they were confident to make a correct decision. Interestingly, this is counterintuitive with the JTC literature, in which schizotypy is related with less evidence gathering to make judgments (Dudley et al., 2016; Moritz et al., 2012; Ross et al., 2015a). However, I discussed that my results can be interpreted from a reduced confidence in sensory evidence perspective, with similar results in people with SZ (Moustafa et al., 2015). Given that schizotypy is associated with odd sensory experiences (*e.g.*, UE), people learn not to trust their sensory stream, therefore, to obtain a good performance in the task they needed to gather more sensory evidence, though I have no independent measure of trust in this task. Nonetheless, there was no difference in performance accuracy as a function of schizotypy, but only in the DT. Finally, in an exploratory analysis I found a moderation effect on information type and event weights, this suggests that high UE scores attribute less relevance to evidence that reduces perceiving positive contingencies.

In Chapter Three, I also used the rapid-trial streaming procedure. But I explored the second to minutes timeframe, and contingency judgments instead of perceptual contingency classification. Contingency judgments implied the use the prior experience and recall the events to produce a continuous rating describing the experienced contingency. The aim was to study the relevance of mutual absent cues, *i.e.*, D trials, no cue and no outcome. My hypothesis was that schizotypy would display a reduced sensitivity to the mutual absence of cues, *i.e.*, interaction between schizotypy and D trials [co-absent events] frequency. In two experiments I modified the duration and frequency of co-absent events and found that people were more sensitive to frequency than duration of the absent information. Moreover, schizotypy UE were less sensitive to the influence of co-absent events in their contingency judgments. This is potentially relevant for understanding how beliefs are formed, for example superstitions or odd beliefs may be produced given a reduced weights of information that is not present. Absent information in contingency learning can be considered as disconfirmatory evidence (Wason, 1968), which has been associated with SZ (Juárez-Ramos et al., 2014) and schizotypy (Buchy et al., 2007; Woodward et al., 2007).

In Chapter Four, I investigated how instrumental actions were related with learning one's agency versus another agent and its relationship with schizotypy, specifically UE. Higher UE scores exhibited an enhanced illusion of control and moderates how they weight responses to inform judgments. On one hand, in Experiment 4, I hypothesised that illusion of control was positively correlated with UE, this would be supported by empirical evidence associating SZ and illusion of control (Balzan et al., 2013). On the other hand, in Experiment 5 I found that when another 'agent' (computer simulated) was introduced, the illusion of control disappeared. Although I found evidence of self-referential behaviour in high UE scorers (Cicero & Kerns, 2011; Lenzenweger et al., 1997). This was shown with a negatively correlation between UE and the sensitivity to perceive the other agent control. Nonetheless, the moderation effect with UE on probability of responding and PC was replicated. Finally, I proposed a model based on Behavioural Control Theory (Powers, 1973; Sosa & Alcalá, 2022) which described the interaction between two agents learning in the same shared

contingencies, and with this model I proposed that the agency paradox may emerge when participants downweight the event when no control sources are present.

Finally, in the last chapter I explore the meaning of weighting in computational models, compare qualitatively different models' families, and conclude that computational models implement the concept of weighting in a similar way. Then, I go on to summarize learning evidence in schizotypy and SZ and interpret my results considering weighting different features of my experiments. Finally, I discuss limitations and future directions.

5.1. What do Computational Models mean by 'Weighting'?

Throughout this thesis I have used the term weighting as synonym of learning rate (LR) or as a multiplier of some experience. My focus here is *i)* to analyse the concept of weighting in general computational models, and *ii)* to understand schizotypy from this approach.

Associative theories involve weighting different sensory inputs (Mackintosh, 1975; Pearce, 1987; Stout & Miller, 2007). Furthermore, weights depend on environmental criteria presented to the agents, which suggests that agents adapt (or change) their weights to optimally match the environmental criteria. Therefore, what is relevant or important now may not be in the future, *i.e.*, weights change (variable weights) or distinct weights for different times (multiple fixed weights). For example, two people experiencing a complex sensory scene (*e.g.*, a setting sun on the horizon) will both differentially attend to the salient features (*e.g.*, the colour of the sky) instead of the less salient (*e.g.*, proprioceptive pressure on the soles of their feet). However, our understanding of mental processing suggests that not only do people differentially attend to these features, but that there are striking mental health effects in how people encode, represent, and process current experiences and combine this with prior experiences (see Byrom & Murphy, 2018). As a result (to reference the previous example), some may focus on their heart rate or other internal proprioceptive stimuli instead of the setting sun even though the stimulus input for two different people is similar or identical.

In one study illustrative of this idea with depression, Gur et al., (1992) found that people with depressive symptoms were more likely to classify faces with ambiguous expressions as negative. The

result is explained by either prior biasing or enhanced sensitivity to stimuli consistent with a negative processing bias (Gur et al., 1992). In computational terms this difference can be modelled in terms of weighting positive and negative valanced sensory evidence. Pulcu & Browning (2019) suggested that people learnt outcomes with positive and negative valence independently (with two different linear equations 5.1). They also found that an affective bias favours better learning with positive outcomes in volatile conditions (Pulcu & Browning, 2017). An affective bias can be quasi-formalized, where the change (Δ) of positive (P) and negative (N) valences are:

$$\begin{matrix} \Delta valence_P \sim weight_P * learning\ signal_P \\ \Delta valence_N \sim weight_N * learning\ signal_N, \end{matrix} \qquad \text{(eq. 5.1)}$$

Describing the problem in this way involves a comparison between the quantitative variables extracted from participants behaviour data –$weight_P$ and $weight_N$–. In general terms, the positive bias is therefore $weight_P > weight_N$ and the negative bias could be $weight_P < weight_N$. The learning signal is also often referred as (PE or δ; see **Section 5.3.**). Therefore, psychological differences can be captured by a quantitative difference in processing and weighting of the learning input. Weighting in a form of a scalar multiplying a discrepancy learning signal or PE.

5.1.1 Differential Weighting of Sensory Evidence

Selecting relevant information from the world is a critical cognitive tool for any living agent with any kind of condition. For instance, so-called 'aberrant salience attribution' is associated with SZ (Kapur, 2003) and the reduced use of contextual cues with depressive traits (Castiello, Senan, et al., 2020; Msetfi et al., 2005). In computational models, the salience, or the relevance of a cue or outcome, is usually modelled as a multiplicative scalar parameter (Esber & Haselgrove, 2011). This parameter also is known as LR (or precision) and changes the value or association of some world state (cue or set of cues). In many computational models, LRs have been related with individual differences associated with psychopathologies.

LR is a widely used as a mechanism to weight especially the discrepancy between an observed sensory state and an expected sensory state. For example, anxious individuals have lower

LR in volatile environments (Browning et al., 2015). Ketamine is related with a reduction in LR confidence (Vinckier et al., 2016). However, some evidence suggests no differences in LR between SZ and general population (Murray, Corlett, et al., 2008). Nonetheless, in a different set of theories (Bayesian), LR is known as precision a measure of sensory uncertainty (see below). SZ dysfunction in smooth pursuing eye movement is believed to occur due to a reduced precision weighting of PE (δ) in a second level of a Bayesian filtering (Adams et al., 2013). Using different precision weighting configuration is a powerful tool to describe psychological conditions, even hysteria (Edwards et al., 2012).

The aim of the next two sections is to describe the general structure of cognitive computational models (decision and learning rules). Then compare the learning rules between two different types of models (Bayesian and reinforcement/associative), this comparison aims to explore is the meaning of weighting in learning rules and address claims that type of model rather than the weighting principle within models is an important conceptual difference. In addition, I do not pretend to explore all the computational models of learning, but in **Table 5.A1** I present an exhaustive list. Overall, I found that computational models are more alike than distinct in terms of weighting and that there is little computational difference between a class of Bayesian and reinforcement/associative models. Moreover, I discuss the intersection between these two types of models and use them to understand my contingency schizotypy results in a descriptive way. Furthermore, in **Section 5.4.** I present a compilation related to this specific literature and conclude how weighting occurs in people with SZ and schizotypy. Finally, in **Section 5.5.** I combine the interpretations of my results with the findings of weighting.

5.2. Overview of Computational Models

There are many shared features in two distinct families of computational models: Bayesian Learning (BL) and Reinforcement/Associative Learning (RAL). To explain complex systems such as brains, cognition, or behaviour, models require simplifications. How much simplification? The Bonini Paradox states that model usability and its realism are inversely related (Bonini, 1962; Burgos &

García-Leal, 2015). Therefore, on one hand, a model can be too complex, in an exhaustive attempt to describe reality (*e.g.*, neural networks for conditioning; Burgos, 1997; Delamater, 2012; Donahoe et al., 1993; Schmajuk et al., 1996), creating difficulty in using these models to interpret reality (*cf.* Burgos et al., 2008; Castiello et al., 2021). On the other hand, a simple model offers more usable interpretations, though they may be less likely to accurately reflect reality (*e.g.*, simple conditioning models; Bush & Mosteller, 1951; Rescorla. & Wagner, 1972). Overall, cognitive computational models of learning and decision-making (BL and RAL) may be described by two rules or submodels: *i)* learning (or perceptual model), and *ii)* expression of behaviour (decision, activation, or response model; Burgos, 2007; Daunizeau et al., 2010; Delamater, 2012; Ghirlanda et al., 2020; Honey et al., 2020; Mathys et al., 2014; Stout & Miller, 2007). To see more details and models go to **Table 5.A1**.

The previous computational distinction between learning and behaviour is rooted in early psychological studies (Tolman, 1948). Tolman (1948) suggested that learning could be latent, *i.e.*, unexpressed with musculoskeletical activity, therefore what is learnt and what is expressed by musculoskeletical activities are different (see Baker & Mercier, 1989). Studies provided evidence for this distinction (Miller & Springer, 1972a, 1972b) and how learning is expressed, *i.e.*, the functions which map what is learnt with behaviour (*e.g.*, the comparator hypothesis; Miller & Matzel, 1988) has been proposed. Even though, in this chapter I mainly focused on learning rules (weight changes), rather than decisions rules. In the following two subsections I briefly describe both rules.

5.2.1. Learning rule (or perceptual model)

The learning rule determines the changes of a connection, association, or state value as a result of experience. There are multiple examples of learning rules − although some models have more than one (Ghirlanda et al., 2020), others are based on an error-correction mechanism. For instance, the delta rule (Rumelhart, McClelland, et al., 1986), RW (Rescorla & Wagner, 1972), temporal difference (Sutton & Barto, 2018), sometimes-competing retrieval (Stout & Miller, 2007), among others (Delamater, 2012; Donahoe et al., 1993; Esber & Haselgrove, 2011; Ghirlanda et al., 2020; le Pelley, 2004; Luzardo et al., 2017; Mackintosh, 1975; Pearce & Hall, 1980).

5.2.2. Decision Rule (or response model)

In the case of the decision or activation rule, on one hand, in artificial neural networks[4] the activation rule is the function which maps the net input (receiving information) for a specific neurocomputational unit and the output of that unit (Burgos, 2007; Castiello, Burgos, et al., 2020; Delamater, 2012; Rumelhart, McClelland, et al., 1986). On the other hand, in RAL (models which maximize expected rewards) or BL (models using Bayes Theorem to update beliefs), the decision rule maps state values, associative strengths, or beliefs, with actions or responses. For example, a well-known rule is the softmax function, which transforms value representations into the probability of responding (*e.g.*, Collins & Frank, 2012; Pessiglione et al., 2006). In the following section, I describe the learning rules for both models' families (BL and RAL).

5.3. Two Families of Computational Models

5.3.1. Bayesian Learning (BL)

In recent years, BL models (Behrens et al., 2007; Daunizeau et al., 2010; Gershman, 2015; Kruschke, 2008; Mathys et al., 2011; Mirza et al., 2019; Rao & Ballard, 1999) have been used to attempt to understand multiple psychological and neurobiological phenomena (including psychopathologies, see above). Next, I describe in general terms how these models conceptually use 'weight' to change the relevance of world features. This framework provides a way to estimate changes in beliefs (*i.e.*, learning) for specific world-states. Beliefs are defined by probability density distributions and expressions of sensory evidence, updated by experiences and contrasted with sensory expectation (priors). For instance, an updated belief (posterior probability density distribution) is the product of prediction of world-states (prior beliefs) and observed sensory data (likelihood). This is:

[4] Feedforward Artificial Neural Network (FANN) are computational models comprised of two levels: the network architecture and the neurocomputational rules (activation and learning rules). The architecture of a FANN is determined by the number of neurocomputational units which are connected forward in layers (input, hidden, and output) and their connectivity. The first neurocomputational rule, activation, determines the units are activated given the sets of inputs at each layer; the second rule, learning, determines how the connections between units are calculated at each time step (e.g., Burgos & Donahoe, 2016).

$$Posterior\ Belief \propto Prior\ Belief * Likelihood, \quad \text{(eq. 5.2)}$$

updating beliefs is proportional to ($\propto$) the product of prior beliefs and likelihood of sensory data. If the priors and likelihoods are normally distributed (N), then the posteriors should be as well. Hence:

$$N(\mu_{post}, \sigma_{post}) \propto N(\mu_{prior}, \sigma_{prior}) * N(\mu_{lik}, \sigma_{lik}), \quad \text{(eq. 5.3)}$$

Each N has two sufficient statistics (parameters): the mean (μ) and the SD (σ). From this framework, the inverse variance is known as precision ($\pi = 1/\sigma^2$), and π are the weights of priors and likelihoods, *i.e.*, how precise is μ. If π of one of the two distributions (π_{prior} or π_{lik}) is low, then that distribution effect to μ_{post} will be reduced (see below). The posterior belief μ_{post} is a prediction of the next state of the world. This family of models can be stacked in a hierarchical manner, allowing higher levels to estimate lower-level parameters (Mathys et al., 2011). In the case of one level and one time step, the posterior parameters estimation of eq. 5.3 is (Mathys et al., 2014; Sterzer et al., 2018):

$$\pi_{post} = \pi_{prior} + \pi_{lik}, \quad \text{(eq. 5.4)}$$

$$\mu_{post} = \mu_{prior}(\pi_{prior}/\pi_{post}) + \mu_{lik}(\pi_{lik}/\pi_{post}), \quad \text{(eq. 5.5)}$$

$$\mu_{post} = \mu_{prior} + (\pi_{lik}/\pi_{post}) * (\mu_{lik} - \mu_{prior}). \quad \text{(eq. 5.6)}$$

Here, the likelihoods μ_{lik} and π_{lik} are directly observed by our sensory systems; μ_{prior} and π_{prior} are representative of our expectations of the world and for a particular task are considered 'free' parameters, *i.e.*, $t = 0$. Finally, the posteriors will be the priors at $t + 1$. In this case, sensory evidence (μ_{lik}) is compared with predicted sensory inputs (μ_{prior}), also known as, PE. This PE is weighted by a ratio of πs; therefore, the weight is:

$$\frac{\pi_{lik}}{\pi_{post}} = \frac{\pi_{lik}}{\pi_{prior}+\pi_{lik}}, \quad \text{(eq. 5.7)}$$

$\pi \in \mathbb{R}^+$; therefore when $\lim_{\pi_{lik}\to 0} \frac{\pi_{lik}}{\pi_{prior}+\pi_{lik}} = 0$, implies low certainty of sensory data, *i.e.*, downweight PE. However, when $\lim_{\pi_{lik}\to \infty} \frac{\pi_{lik}}{\pi_{prior}+\pi_{lik}} = 1$, the weight implies a high certainty in sensory input and

the PE is more relevant, *i.e.*, upweighted. An inverse effect occurs with π_{prior}, *i.e.*, $\lim_{\pi_{prior}\to 0} \frac{\pi_{lik}}{\pi_{prior}+\pi_{lik}} = 1$. As a result, zero π_{prior} produce an upweight PE, indicating greater relevance to previously observed sensory information. Moreover, $\lim_{\pi_{prior}\to \infty} \frac{\pi_{lik}}{\pi_{prior}+\pi_{lik}} = 0$, implies that the PE is less relevant for updating beliefs and is multiplied by a very low weight.

An underlying idea is that when the environment is volatile or uncertain, the PE is downweighted (Behrens et al., 2007). Associative theories have employed a similar idea. For instance, reducing the weight for uncertain states or cues was postulated in an early theory of selective attention (Mackintosh, 1975). Mackintosh stated that predictive cues (or states) evoke an increase in salience, like BL when there is high certainty in the sensory world (high π_{lik}). Therefore, predictive cues or states increase attention or saliency. However, this contrasts with other models like the Pearce-Hall model (Pearce & Hall, 1980), where uncertainty drives up weighting of evidence. And some models combine these two contradictory ideas into hybrid models (Esber & Haselgrove, 2011; Le Pelley, 2004) to capture how different forms of evidence for different behaviours may reflect these two processes.

In summary, BL suggests that π of observed sensations and expected sensations (priors) multiplies (or weight) the PE, which then updates beliefs. The advantage of this approach is that agents represent uncertainty in the world and uncertainty in their own beliefs. The general assumption is that learning for uncertain situation is reduced, although this assumption may struggle to explain experimental findings in partial reinforcement (*e.g.*, Haselgrove et al., 2010).

5.3.2. Reinforcement and Associative Learning (RAL)

Associative learning (AL; Shanks, 1995) and its related field reinforcement learning (RL; Sutton & Barto, 1992) models do not use probability density distributions to represent uncertainty of the learnt association (Gershman, 2015). I refer to both associative and reinforcement learning as one family, RAL. The main pillars of RAL are: *i)* events are associated, those associations have a strength which changes over time (Esber & Haselgrove, 2011; Le Pelley, 2004; Mackintosh, 1975; Pearce &

Hall, 1980; Rescorla & Wagner, 1972); and *ii)* there are qualitatively different types of associations: cue-outcome (Pavlovian), cue-response (Pavlovian), and response-outcome (instrumental; see **Figure 5.1**; similar to the Skinnerian "three-term contingency"; Skinner, 1965).

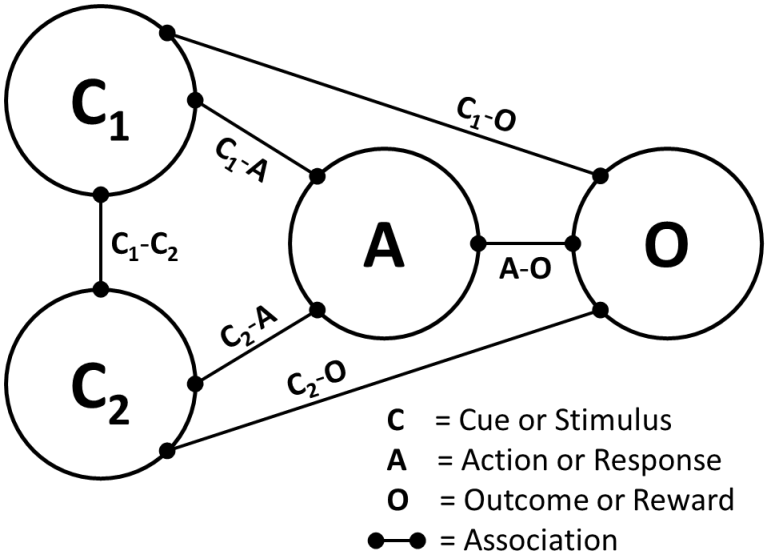


Figure 5.1. Skinner's Three-Term Contingency. This diagram represents the different associations than can be learnt by animas. Cue-Cue (C_1-C_2), Cue-Action (C_1-A and C_2-A), Cue-Outcome (C_1-O and C_2-O), and Action-Outcome (A-O).

5.3.2.1. Fixed weights

Influential RAL models include the RW model and the real time extension Temporal Difference model (Ludvig et al., 2012). These models allow for both individual learning of cues but also cue competition, *i.e.*, what happens when two cues are paired simultaneously with the same outcome. However, in some RL literature, the cue competition component is not invoked (Behrens et al., 2007, 2009; Browning et al., 2015; Mathys et al., 2011; Pulcu & Browning, 2017, 2019; cf., Murray, Corlett, et al., 2008; Vinckier et al., 2016), but invoke a similar but learning model (Bush & Mosteller, 1951), which is more alike to the update rule for tracking a nonstationary problem or Q-learning (Sutton & Barto, 2018).

The Q-learning model states that the value of a state in the future ($v_{s,t+1}$) is the sum of the current value ($v_{s,t}$) and the PE multiplied by a LR (α). The PE is the difference between an observed reward or outcome (o_t) and $v_{s,t}$, *i.e.*, $\delta_{s,t} = o_t - v_{s,t}$, and the update value equation at $t + 1$ is:

$$v_{s,t+1} = v_{s,t} + \alpha\delta_{s,t}, \quad \text{(eq. 5.8)}$$

where $\alpha \in \mathbb{R}^{+}: \alpha < 1$. An implication of Q-learning is that the value calculation for the states is independent. Curiously, on the surface (eq. 5.8) resembles (eq. 5.6); they both update a prior value (μ_{prior} and $v_{s,t}$) with a weighted (π_{lik}/π_{post} and α) PE (*i.e.*, outcome or sensory evidence minus priors or expectations). However, for the BL account, LR is not fixed, akin to RAL rules that permit a dynamic LR (Esber & Haselgrove, 2011; Le Pelley, 2004; Pearce & Hall, 1980). Before describing the Pearce-Hall (PH) learning rule (Le Pelley, 2004; Pearce & Hall, 1980), I describe a comprehensive form of the RW learning rule, crucially including the cue competition component. As noted above, this component is not incorporated in previously described BL accounts (*e.g.*, Mathys et al., 2011; Sterzer et al., 2018), though cue competition is possible in BL (Gershman, 2015; Kruschke, 2008; Ludvig et al., 2012).

The key feature of the RW model is that it estimates the associative strength (V) from multiple cues paired with an outcome. The exact equations in the original paper (Rescorla & Wagner, 1972) are:

$$\begin{aligned} \Delta V_A &= \alpha_A \beta_1 (\lambda_1 - V_{AX}) \\ \Delta V_X &= \alpha_X \beta_1 (\lambda_1 - V_{AX})' \end{aligned} \quad \text{(eq. 5.9)}$$

Where Δ means change, A and X refers to distinct cues, αs are LRs, βs outcome salience, λ is the reward (and learning asymptote), and finally $V_{AX} = V_A + V_X$. From an algebraic point of view, this model is difficult to implement computationally. In addition, cue competition in RW is analogous the *netinput* calculation in other algorithms such as backpropagation (Castiello et al., 2021; Delamater, 2012; Rumelhart, Hinton, et al., 1986).

Next, I rewrite the RW learning rule in a different way (Gershman, 2015; Kruschke, 2008). The sum of associative strength (sigma; ς) −also referred to as the overall expectation of the outcome−, is the dot product between **v**, a vector of the cue's associative strength (or values), and **c**, a Boolean vector (with {0,1}) representing the (presence/absence) of the cues:

$$\varsigma_t = \mathbf{c}_t \bullet \mathbf{v}_t, \quad \text{(eq. 5.10)}$$

therefore, the PE is: $\delta_t = o_t - \varsigma_t$. Then, the update of associative strength is:

$$\mathbf{v}_{t+1} = \mathbf{v}_t + \alpha\delta_t\mathbf{c}_t, \quad \text{(eq. 511)}$$

Here, I formally express the temporal update, *i.e.*, $t + 1$, and use linear algebra to simplify the description of eq. 5.9. Also, α represents the LR parameter(s). Usually, RW assumes that α is the product of $\alpha\beta_0\beta_1$, which is the product between LR (α), non-reward (β_0), and reward sensitivity (β_1). An implication is that RW describes the possibility of using different weights for reward and non-reward, but does not describe formally how this would work (I describe one possibility in Chapter Four, section **4.4.3.**). In the following section I present RAL models with variable LRs, which has been demonstrated to be superior to fixed LRs in mice (Treviño, 2016). Although, comparing LR between species and even within a species but between different tasks is not ideal, given that LR from one task does not predict LR in different task in one same individual (Eckstein et al., 2022).

5.3.2.2. Variable weights

Cue competition is an important component of computational models because it provides the mechanism to explain phenomena such as KB, which requires an interaction between cues (*e.g.*, cue competition) and might be considered a better model for cognition in environments which are filled with more than one cue. In the case of RW the dot product in ς produce the interaction between cues. However, other mechanisms such as variable LR also can be used to explain interaction between cues.

To describe a variable LR rule, I used the hybrid PH learning rule (Le Pelley, 2004; Li et al., 2011). This rule is comprised of two equations, *i)* associative strength changes and *ii)* LR changes. The dynamic LR or α is calculated for each time step t and for each state (or cue) s as:

$$\alpha_{s,t} = \eta\left|\delta_{s,t-1}\right| + (1-\eta)\alpha_{s,t-1}, \quad \text{(eq. 5.12)}$$

where $\eta \in \mathbb{R}^+ : \eta < 1$, and $\delta_{s,t} = o_t - v_{s,t}$. If $\eta = 0$, then LR is fixed. The $v_{s,t}$ for each state follows the same error-correction mechanism:

$$v_{s,t+1} = v_{s,t} + \kappa\alpha_{s,t}\delta_{s,t}, \quad \text{(eq. 5.13)}$$

equal to η, $\kappa \in \mathbb{R}^{+}: \kappa < 1$. It is important to recognise, as noted by Li et al (2011), that $\delta_{s,t}$ is not counted twice, given that to calculate $\alpha_{s,t}$ the previous $\delta_{s,t-1}$ is used. To my knowledge this model has not been described with vectors, which may facilitates its implementation.

I presented three learning rules with reference to RAL. Two with fixed LR: Q-learning (Sutton & Barto, 2018) and RW (Rescorla & Wagner, 1972), and one with variable LR, the PH adaptation (Li et al., 2011). Q-learning cannot model multiple cue competition learning phenomena (*e.g.*, Baker et al., 2000; Miller & Escobar, 2002), but RW and PH can. I do not present other associative models in detail (Esber & Haselgrove, 2011; Ghirlanda et al., 2020; Le Pelley, 2004), as our focus was on the analysis of weighting mechanisms, here noted as fixed and variable LRs.

In summary, in the RAL family of models, the weighting method is produced by multiplying the PE (error correction mechanism) by a scalar. The scalar is referred to as LR, and this can either be fixed or variable. Cue competition mechanisms can emerge by variable LR (changing cue salience or predictiveness; Pearce & Hall, 1980) or because of the cue-competition due to the overall outcome expectation (ς; Rescorla. & Wagner, 1972). Describing models with linear algebra may clarify their implementation.

5.3.3. The intersection of Bayesian and Associative theories

A main finding from the previous section is that BL overlaps with a classical associative learning model with variable LRs: the Theory of Attention (Mackintosh, 1975). Mackintosh (1975) suggested that predictive cues (that provide information to predict outcomes) tend to be upweighted by the agent, *i.e.*, predictive cues are more salient. This agrees with Bayesian theories where the LR is increased when the sensory evidence is precise (or predictive). Therefore, Bayes style learning can be considered as an extension of associative selective attention (Mackintosh, 1975). But this type of theory involving upweighting predictive cues does not account for experimental evidence suggesting that upweighting of unpredictive cues seems to occur in some situations. For example, partially

reinforced cues (uncertain), can support faster learning than continuously reinforced cues (certain) suggesting upweighting of non-predictive cues (Haselgrove et al., 2010). Those results support an opposite idea –that uncertain cues have a baseline of increased salience (Pearce & Hall, 1980). Thus, hybrid models incorporate these two mechanisms (Esber & Haselgrove, 2011; Le Pelley, 2004). It is unclear how BL (*e.g.*, Behrens et al., 2007; Mathys et al., 2011) can account for these effects.

Nevertheless, the critical distinction between these two families of models is that LRs in BL are composed by two precisions. Thus, the same LR can be influenced by both prior knowledge and sensory experience. Which is a useful heuristic tool for thinking about differential weighting between internal and external events. In summary, both families of models use mathematically the concept of weight in a similar way, multiplying PEs. But BL has the advantage of changing the LR by two subfactors: prior precision (internal), and likelihood precision (external).

5.4. Summary of Learning in SZ and Schizotypy

I analysed and compared the weighting mechanisms between BL and RAL. Weighting is often referred to as the degree in which a sensory signal is increased or decreased relative to a particular context. I now describe how weighting relates to my findings. First, I describe general findings and exemplified weighting mechanisms in clinical SZ and subclinical schizotypy based on a quite exhaustive review (**Table 5.A2.**; although I did not use a systematic review method, such as the Cochrane handbook; Higgins & Green, 2011). Second, I discuss my results considering the concept of weighting sensory evidence.

An account of positive symptoms of SZ, such as hallucinations, has been explained by high precision priors, *i.e.*, a strong expectation of world states (Corlett et al., 2019). In BL, the high precision of priors relative to sensory data refers to downweighting the relevance of sensations, meaning beliefs are not easily updated by sensory data. Although this discussion is by no means complete, evidence and computational models also support the idea that psychosis is a condition of weak priors (Adams et al., 2013). However, in a hierarchical structure such as predictive coding (Rao & Ballard, 1999), it is possible that both mechanisms operate simultaneously on different levels of the

hierarchy and affect different classes of symptoms, *e.g.*, delusions versus hallucinations (Adams et al., 2013; Corlett et al., 2019; Sterzer et al., 2018). Accordingly, weights are the changing precision which multiply PEs (see **Table 5.A1.**).

5.4.1. Clinical Population

Non-paranoid people with SZ display a reduced or absence of KB (Bender et al., 2001; Jones et al., 1997; Oades et al., 1996, 2000; Serra et al., 2001), that is, they learn from the 'blocked' cue (*cf.*, Oades & Eggers, 1992). Similarly, reduced KB is found in acute SZ (Jones et al., 1992b). A similar task that relies on the relative relevance of cues aimed at testing learning found a positive correlation with positive symptoms in people with SZ (Morris et al., 2013). Therefore, with more positive symptoms, there is evidence for enhanced learning of irrelevant cues. This is consistent with the aberrant salience theory of SZ (Kapur, 2003), which states that 'random' (high baseline firing) dopaminergic signalling may produce increased salience to irrelevant aspects of the world, *e.g.*, non-predictive cues. This is confirmed by reports supporting a general attenuation of PE in the right prefrontal cortex reduce the PE weighting (Corlett et al., 2007).

In addition, from an instrumental learning perspective, SZ showed illusory correlation and illusion of control (Balzan et al., 2013), also correlating with positive symptoms (Moritz et al., 2014). This could suggest an extreme weighing of action-outcome or a reduced weight to counterfactuals (*e.g.*, no action → no outcome). In fact, in one study SZ did not show deficits in learning by reward (Morris et al., 2015, 2018), but rather a deficit in experience for the correlation action-outcome and a lack of sensitivity to outcome degradation (Morris et al., 2018). In sum, evidence with SZ in instrumental learning tasks supports the idea that in SZ there is a lack of understanding of the contingency action-outcome, although potentially caused by working memory deficits (Collins et al., 2017). Nonetheless, instrumental learning in SZ raises a critical point – what does it mean to weight a correlation between action-outcome? Before I can respond to this question, I revise evidence in subclinical schizotypy in the following subsection.

5.4.2. Subclinical Population

The KB effect has been studied in relatives of people with SZ which may indicate a possible genetic overlap at a subclinical level. Research has found a reduced KB effect in relatives of people with SZ (Jones et al., 1997; Serra et al., 2001). A similar effect as in SZ is found in people with no diagnosis by administration of amphetamines (Jones et al., 1997; *cf.*, Gray et al., 1997). In the general population, there is little relation between KB and psychoticism (Jones et al., 1990). But a marginal effect has been found between a reduced KB and magical ideation and unusual perceptual experiences, *i.e.*, positive schizotypy (Jones et al., 1992a). The evidence from the subclinical population seems less convincing than from the clinical population. However, when measuring schizotypy instead of psychoticism, the impact on learning is demonstrated.

KB and other learning tasks designs such as LI and LIRR, have been used to study attention to irrelevant cues and positive schizotypy is related to increased learning of irrelevant and non-predictive cues (Le Pelley et al., 2010; Schmidt-Hansen et al., 2009), as well as reduced LI, *i.e.*, increased learning to a pre-exposed non-predictive cue in comparison to a novel cue (Evans et al., 2007; Granger et al., 2012, 2016; Schmidt-Hansen et al., 2009). Also, people with proneness to delusions show diminished PE weighting (Moore, Dickinson, et al., 2011). A final piece of evidence using instrumental learning techniques may provide another interpretation of weighting. UE relates to a reduced awareness in detecting reinforcement schedules based on responses (reinforcement ratio; RR) and on time (reinforcement intervals; RI; Randell et al., 2009, 2012). This points to a lack of contingency action-outcome sensitivity.

Furthermore, negative schizotypy, measured by IA, is related with a reduced KB effect (Haselgrove & Evans, 2010; Moran et al., 2012; *cf.*, Humpston et al., 2017). Using an irrelevance learning task, IA exhibited higher learning in relation to irrelevant cues due to lower attention to relevant stimuli (Haselgrove et al., 2016). Overall, in the general population, weighting irrelevant cues in a compound, such as KB, may be related to negative schizotypy such as IA. Yet weighting a non-predictive cue, as in LI, may be related with positive schizotypy such as UE.

5.4.3. Summary: Clinical and Subclinical Schizotypy and Learning

In summary, the schizotypy data in clinical and subclinical people with psychotic experiences support an upweighting of irrelevant sensory evidence or a downweighting of relevant evidence. I use irrelevant in a predictive way. In the clinical population this seems to be related more often with acute and non-paranoid manifestations. But in subclinical, there is no consensus with respect the type of symptoms-like (positive or negative). In addition, it seems that people with SZ and high schizotypy scores have intact reward-learning acquisition, although they may be insensitive to the understanding of the goal directed or causal action-outcome relationship.

5.5. Schizotypy and Weighting

5.5.1. Interpretation of my Findings

In my experimental chapters, I found evidence indicating that schizotypy, particularly high UE scorers, gather more sensory evidence to make correct decisions (Chapter Two); they may be less sensitive to sensory counterfactuals (*i.e.*, absent events; Chapter Three); display an enhanced illusion of control when there is no competition with an external source of control; and they are less sensitive to estimate control of another agent (Chapter Four).

For example, imagine the case of a person J scoring high in UE, *i.e.*, that may experience seeing shapes and forms in the dark that are not there −reinterpretation of the questionnaire item: "*When in the dark do you often see shapes and forms even though there is nothing there?*" −. During J's life almost every time that J shared their experience of such sensations with someone else, it may be the case that others tend to invalidate J's sensations. If this process is iterative in J's life, at some point, J would learn that their sensory information may not be in concordance with reality, given that in previous situations other people does not feel or see the same things. Hence, J reduces the weight in her sensations. Therefore, J may need to gather more sensory evidence to make correct assessment (Chapter Two).

The aberrant state of salience model of psychosis suggests an exacerbated dopaminergic activity that signals 'everything' as important (Kapur, 2003); in a patient's words: "*Everything seems to grip my attention although I am not particularly interested in anything* […]". (L. Sass et al., 2017). If this applies to people with psychotic experiences and without diagnosis of SZ, *e.g.*., high UE scorers like person J, then given that J tends to attend irrelevant information and experience a state of aberrant saliency, J may compensate the exacerbated irrelevant signalling (attention to low salient features) by increasing a saliency detection threshold. This would imply an impaired detection of low salient information, as the mutual absence of stimuli (or D trials). Thus, this compensation mechanism, *i.e.*, an increase in the detection threshold for low salient features, explains why J would exhibit a reduction in sensitivity for the absent information (Chapter Three).

Moreover, agents upweight their actions in comparison to another agent's actions. For example, one's actions imply proprioception and visual feedback (among other processes), and another agent actions imply visual feedback (among other processes). In Experiment 5 I present evidence of this with a larger sensitivity to self-control than another agent's control. Thus, I can say that the other agent is a low salient feature. Interestingly, I found a negative correlation between UE and sensitivity to the other agent in Chapter Four. This is compatible with an increased detection threshold for low salient features, where people scoring high in UE are less sensitive to the other agent because external factors are low salient. Thus, continuing with J as an example, if J has an increased threshold to miss detection of low salient features, then J would not integrate information from the other agent because the other agent is a low salient feature.

However, there are two points which does not match completely with my findings. First, I recognise that gathering more sensory evidence in schizotypy is completely the opposite to the well stablish JTC bias in SZ (Dudley et al., 2016). Second, evidence supports the idea that people with SZ and people with high scores on schizotypy learn more from irrelevant information (**Table 5.A.2.**), and irrelevant information can be considered as low salient information. Thus, in the next section I develop these two seemingly counter views.

5.5.2. Jumping-to-Conclusion and Low Saliency Types

Even though, my results seemed counterintuitive in the light of JTC (Dudley et al., 2016). I believe that a new interpretation of JTC can be achieved. One of the most used dependent variables in JTC is the extreme responders, this is when a participant after seeing one piece of evidence is 'confident' enough and make the choice and 'know' the correct beads' jar. For instance, Moutoussis et al. (2011) showed clinical data and a Bayesian model suggesting that the JTC bias may not be related to the cost of gathering information but to the noise at the level of action selection (Moutoussis et al., 2011). It may be the case that extreme responders are a product of an 'impulsive' action (or decisional noise) and that the decision is not based on the development of their belief or some higher cognitive phenomenon. The relationship between impulsivity and JTC has been shown previously (Korkmaz & Can, 2020; *cf.*, Strube et al., 2022). If that is the case, then extreme responders are not focusing on the amount of evidence, but they are responding quickly to finish the task. Therefore, the JTC task, in spite of its wide acceptance and uses as a measure of beliefs, and specifically beliefs that relate to delusions, may be particularly problematic due to the decision noise that may not permit the assessment of sensory data gathering. Finally, there are two differences between my perceptual contingency task and the JTC task. First, in my task participants' decision occurs in the millisecond to second timeframe, in contrast of the second to minutes timeframe in the original task. Second, in the JTC task, to get more evidence participant must require it, implying an action cost, which is not the case in my task, where the sensory evidence is provided automatically.

For the case of the heighten attention to irrelevant information in schizotypy (**Table 5.A2.**). If irrelevant information is low salient, how is it possible that high schizotypy scorers upweight low salient irrelevant information and downweight low salient absent information? Thus, one possible answer is that there are qualitative different saliences. For instance, *i)* low salient events as counterfactuals or absent events (low saliency = absent information), and *ii)* low salient events in a statistical or predictive sense (low saliency = no predictive), for instance, events with a $\Delta P = 0$. Thus, my findings support that positive schizotypy have a reduced sensitivity to the former low saliency concept, *i.e.*, absence of information; and the literature review (section **5.4.**) support an attentional

increase to to the latter low saliency concept, *i.e.*, in a non-predictive sense. Another possibility may be that the lack of sensitivity for absent information may increase the overall sensory sensitivity, as a contrast effect. Perhaps, an experiment assessing both types of saliencies would clarify this point. For example, using the same paradigm employed in Chapter Three but adding conditions where D trials are not 'empty' but are filled with predictive or non-predictive cues.

5.5.3. Consistency with the Two-Stage Cognitive Theory

The two-stage cognitive theory of Mortiz et al (2017) states that psychosis is characterized by a lower decision threshold. First, during belief formation (Stage 1), a decisional threshold is reduced, *i.e.*, liberal acceptance. Second, during belief maintenance (Stage 2) *i)* there is an augmented confidence in delusional beliefs (confirmation biases); and *ii)* alternative hypotheses are considered less for creating judgments (BADE; Moritz et al., 2017). I believe my findings are consistent with both stages of this theory.

Stage 1, liberal acceptance may be due to decision noise (Moutoussis et al., 2011). Thus, this stage is at the decision rule level and not perceiving and learning, as suggested by my data and Moustafa et al. (2015). Yet, there are important differences between Moustafa's task and Experiment 1, they used a reward and punishment learning task whereas I used a contingency perceptual task where participants' action stopped the information stream. Whereas they found a larger sensory evidence threshold and lower discriminability to punishment trials. This could have two interpretations: first, different tasks and similar results may suggest a robust latent construct, thus a possible generalization of the results and the continuous nature of schizotypy; second, the interpretations given to the findings may be wrong, in which the increased sensory threshold in Moustafa's results does not translate to my slower DT findings. However, using a simple reaction time task another group found that people diagnosed with SZ exhibited slower DT (Strube et al., 2022), which provides reassurance to support the first interpretation.

Stage 2, higher confidence in delusional beliefs may be due to a high precision priors which downweight the sensory experience. This implies that higher confidence in delusional beliefs disrupts

updating beliefs and maintains delusional beliefs. This may relate with my findings indicating a weighting reduction in absent information (Chapter Three), and a reduced impact of responses in beliefs formation (Chapter Four). Given that absent information could be seen as a logical counterfactual (Johnson-Laird & Wason, 1970), reducing the weight to absent information may promote delusion maintenance. Supported by the finding that people with high schizotypy scorers do not search evidence against their beliefs (*e.g.*, BADE; Juárez-Ramos et al., 2014).

Finally, this theory has been postulated for clinical psychosis any generalization to schizotypy may be limited. Therefore, if both stages are observed also in the general population with high schizotypy scores, then evidence would support a continuum between SZ and schizotypy. Otherwise, if they are not observed in high schizotypy scorers, then the differences in stages may be a hint for cognitive protective factors for developing clinical psychosis or SZ.

5.5.4. Summary: What Schizotypy Weights?

It seems that at least there are two types of saliences, one is about sensory intensity (presence and absence) and the second is about predictiveness (statistical regularities). It is possible that these two can be weighted in models independently. However, these different weighting mechanisms are not differentially described in computational models. With the interpretation of my results, I suggest that people scoring high in schizotypy gather more sensory evidence to make correct decisions. This is backed up by clinical evidence with SZ, where in a decision-making task SZ exhibits a higher sensory evidence threshold (Moustafa et al., 2015) and slower decision times (Strube et al., 2022). This could be interpreted as a reduce confidence in sensory evidence which may lead to an increase in threshold for detecting low intensity sensory information (*e.g.*, absent of target cues). I found a downweight of target absent information, thus only context. This may be due to the same heighten detection threshold mechanism, where a high sensory detection threshold implies a reduction in processing contextual information, *i.e.*, information with the absence of target stimuli. Furthermore, I suggest that high UE scorers, downweight their responses when informing their judgments, which can

be seen as a referential trait, if assuming they have a baseline heighten sense of generating the outcomes instead of external sources of control.

5.6. Limitations

One of the main limitations of the experiments presented in my dissertation is that given their cross-sectional nature, causal assessments are not possible. With this type of designs causal claims are difficult to assess given that behavioural tasks and questionnaires were conducted in one day for each participant. However, the within subject design that I employed has the advantage that changes in behaviour can be compare at the individual level, so they are useful to evaluate behavioural constraints (conditions) but not the causal relation between behaviour and questionnaires.

Another point for consideration is the fact that the population used for these studies may be quite restricted. Mostly participants involved in my studies were students and MTurk workers. Even though the goal of my dissertation was to study general population varying in schizotypy scores, my research population is not representative of the general population. MTurk workers tend to show emotional irregularities and undergraduate students in psychology tend to be particularly good to solve psychology behavioural tasks.

From a computational modelling perspective, conclusions with respect the reduced weight for absent events and action-consequences were only partially explored. Apart from Chapter Two where I fit a perceptual decision-making model (DDM), I used computational models as a heuristic tool, but I did not fit nor made comparisons between learning models. For example, Chapter Three I decided not to model the actual participants experience given that no responses were recorded at the trial level, and a prior model testing was not successful. Also, in Chapter Four I decided to model ratings and not responses within conditions because the free-operant task does not exhibit the standard RL behaviour. This is participants followed a completely distinct behavioural trajectories to perceive control and the probability of outcome did not dramatically increase as a function of experience (as one would expect in RL tasks). Thus, my use of learning models was at the theoretical level, *i.e.*, simulating, but not as a method to fit individual participants parameters and test hypothesis.

5.7. Future directions

In recent years, computer-based cognitive remediation (CBCR) intervention has been used to support treatment of people with mental health issues or to improve cognitive performance in general population. The evidence of the latter is not robust. Although, the evidence for the former suggests that people with SZ may benefit from CBCR. However, sometimes researchers' outcomes focus on aspects that are not always relevant for patients nor clinicians. For example, I would not be surprised that training memory in a memory task would improve memory performance in that task, but how does that impact overall wellbeing, clinical improvement, or the probability of starting a job or to start a romantic relationship (relevant outcomes for patients)?

To address the previous problem, we published a Cochrane review protocol that will only include Randomized Controlled Trials (RCT) addressing the efficacy of CBCR for people with SZ (Castiello-de Obeso et al., 2020). The goal was to summarize the evidence and conduct a meta-analysis of all the RCT testing CBCR for SZ. If research finds a non-significant effect between CBCR and clinically relevant outcomes, we will have an opportunity to improve the behavioural tasks used in CBCR with any other behavioural tasks. For instance, the nature of the CBCR effect as an intervention is about interacting with a computer and is independent of the type of task. Thus, any behavioural task may be used to remediate patients. The behavioural tasks that I present in this dissertation could be used as a control condition for further validation of CBCR or even as intervention against not using behavioural tasks in a hospital setting.

5.8. Conclusions

Clinical and subclinical schizotypy exhibit an enhanced learning to irrelevant cues and uncertain sensory data, mostly evaluated with associative learning paradigms. My data suggests that positive schizotypy, such as unusual experiences (UE), exhibited a reduction of confidence in sensory evidence, perhaps by a whole life full of moments where other people disconfirm their percepts. This was observed by larger decision times to perform correctly (Chapter Two). This can also be partially related with why high UE scorers and people with schizophrenia (SZ) are adamant with beliefs, *i.e.*,

perseverance. In addition, I identify two types of saliences, the first is about the predictive value and the second is about the sensory intensity. For instance, cues with low predictive value are heightened in SZ and schizotypy (associative learning results and SZ), but events with low sensory intensity may be ignored in people scoring high in UE. Thus, I suggest that low salient sensory information, such as the absent of target cues is downweighted by high UE scorers (Chapter Three).

Moreover, I found evidence of an exacerbated illusion of control in high UE scorers, but only when no other competing source of control is also being tested. I also found people scoring high in UE, were less sensitive to external events (Chapter Three). This relates with the previous paragraph, where external events in baseline are perceived with lower intensity (low salient). Thus, less importance was attributed to low salient events in high UE scorers. I also found that UE moderates the relation between judging self-control and the use of the actions. Perhaps as a referential mechanism where no responses are required to test their control over an outcome given an already exacerbated self-importance. Finally, and relevant for computational models, I showed that what different models refers as weighting is a concept, *i.e.*, multiplying a prediction-error by a numeric value. Nevertheless, I describe benefits of Bayesian learning (BL) and account for the parallelism between BL and the Mackintosh's theory of attention, where both models postulate higher weights for certain or predictive information.

5.9. Appendix 5

3525 **Table 5.A1.** Computational Models (Cognitive, Bayesian, Associative and Reinforcement)
3526

ID	Model	Type	# Free Parameters	Learning Rate (LR)	Weighting Method	Cue Competition	Real Time	Learning Rule	Response Rule	Notes
1	Bush-Mosteller (Bush & Mosteller, 1951)	Associative	2	Fixed	LR × PE	No	No	Yes	No	Usually referred as RW model.
2	Rescorla-Wagner (Rescorla & Wagner, 1972)	Associative	4	Fixed	LR × PE	Yes	No	Yes	No	Sum of Associative strength is principal component of PE.
3	Theory of Attention (Mackintosh, 1975)	Associative	NA	Variable	LR × PE	Yes	No	No	No	LR for predictive cues increase.
4	Pearce-Hall Pavlovian theory (Pearce & Hall, 1980)	Associative	3	Variable	LR × PE	Yes	No	Yes	No	LR for uncertain cues increase.
5	Q-Learning (Sutton & Barto, 1992, 2018)	Reinforcement	1	Fixed	LR × PE	No	No	Yes	Yes / No	Usually referred as RW model.
6	Temporal Difference (Sutton & Barto, 1992, 2018)	Reinforcement	1	Fixed	LR × PE	Yes	Yes	Yes	Yes	Q-learning extension for 'real' time + cue competition.
7	Discrepancy Signal Network (Donahoe et al., 1993)	Neural Network	7[1]	Fixed | Variable[3]	Multiple	Yes	Yes	Yes	Yes[4]	Backpropagation-like and temporal-difference-like for conditioning

ID	Model	Type	# Free Parameters	Learning Rate (LR)	Weighting Method	Cue Competition	Real Time	Learning Rule	Response Rule	Notes
8	SLG model (Schmajuk et al., 1996)	Neural Network	>10[1]	Fixed | Variable[3]	Multiple	Yes	Yes	Yes	Yes[4]	Developed to explain LI.
9	Hybrid Attentional and Salience associability (Le Pelley, 2004)	Associative	6	Fixed / Variable	LR × PE	Yes	No	Yes	Yes	Combine both PH and Mackintosh assumptions.
10	Ideal Bayesian Learner (Behrens et al., 2007)	Bayesian	1	Variable	Volatility (LR) × PE	No	No	Yes	Yes	Generative model, estimate reward probability and environment volatility
11	Sometimes-Competing Retrieval (SOCR) (Stout & Miller, 2007)	Associative	4	Fixed	LR × PE	Yes	No	Yes	Yes	Based on the comparator hypothesis for response rule.
12	Kalman Filter (Kruschke, 2008)	Bayesian	1[2]	Variable	LR ×PE	Yes	No	Yes	Yes	Bayesian version of RW and entropy for response.
13	Hybrid model (Esber & Haselgrove, 2011)	Associative	3	Variable	LR ×PE	Yes	No	Yes	No	LR is the sum of acquired and physical salience.
14	Distributed Pavlovian Representation (Delamater, 2012)	Neural Network	2[1]	Fixed | Variable[3]	Multiple	Yes	No	Yes	Yes[4]	Backpropagation is the learning algorithm
15	Hierarchical Gaussian Filter	Bayesian	4*	Variable	LR × PE	No	Yes	Yes	Yes	This paper present only one level HGF model. But can

ID	Model	Type	# Free Parameters	Learning Rate (LR)	Weighting Method	Cue Competition	Real Time	Learning Rule	Response Rule	Notes
	(Mathys et al., 2014)									be extended a more than one level
16	Kalman Temporal Difference (Gershman, 2015)	Bayesian | Reinforcement	3	Variable	LR × PE	Yes	Yes	Yes	No	Description of 4 model types (estimate, distributions, and real time)
17	HeiDI (Honey et al., 2020)	Associative	3	Fixed	LR × Total (CS or US) expected	Yes	No	Yes	Yes	Estimates associations CS-US and US-CS.
18	RW drift-diffusion (Luzardo et al., 2017)	Perceptual timing | Associative	6	Variable	Rate × Diffusion Process	Yes	Yes	Yes	Yes	Combine RW with DDM, simulate time invariant property.
19	Theory of Actions and Habits (Perez & Dickinson, 2020)	Associative | Reinforcement	5	Fixed (habit) | Variable (goal-dir.)	LR × PE | Scalar × Correlation	No	Yes	Yes	Yes	Used for free operant procedures. Agents detects correlations between responses and rewards.
20	A-Learning (Ghirlanda et al., 2020)	Associative | Reinforcement	3	Fixed	LR × PE	Yes	Yes	Yes	Yes	Ideal to address sequential learning, usually neglected by all conditioning models
21	Proportional-Integral-Derivate controller (Sosa & Alcalá, 2022)	Control System	3	Fixed	LR × PE	No	Yes	Yes	No	One parameter for each method: proportion, integrate, derivate.

Notes: PE = Prediction Error. Associative, Neural Network, Bayesian, and Reinforcement types (?).

*Taken from the example in the Mathys, et al (2011). But the parameters on this model depends on how many levels the experimenter incorporate.

ID	Model	Type	# Free Parameters	Learning Rate (LR)	Weighting Method	Cue Competition	Real Time	Learning Rule	Response Rule	Notes

[1]In Neural Networks is difficult to assess how ‘many’ free parameters are equivalent to the network architecture (hidden layers and connectivity).

[2]One parameter plus initial conditions’ uncertainty (covariance matrix).

[3]The ‘weights’ interpretations are not the same here. Weights are connections between neurocomputational units. But there are free parameters which multiplied PE.

[4]The response rule here reflects an activation function for neurocomputational units.

3529 **Table 5.A2.** Schizotypy and Characteristics of Learning Studies

Year of Publication	Reference	# Exp.	Sample	Paradigm	Outcome(s)	Diagnoses or Scales	Findings
1990	(Jones et al., 1990) – British Journal of Psychology	1	N(NT) = 58 (14 F); age, M 32.5, 18-55	KB and IL	Success criterion: 5 correct trials in Stage 3.	-Questionnaires: Eysenck Personality Questionnaire (EPQ); Launay-Slade Hallucination Scale (LS); Schizotypy Scale (STA); Rust Inventory of Schizotypal Cognition (RISC); Mill Hill Vocabulary Scale -Tasks: Incidental Learning	*i)* KB effect was found; *ii)* KB was unrelated to intelligence; *iii)* elevated incidental learning also showed less KB (same underlying process: selective attention); *iv)* no association with psychoticism; *v)* no relation incidental learning and extroversion; *vi)* high neuroticism lower incidental recall.
		2	[a]N(NT) = 30 (13 F); age M 31.6, 24-55	KB and IL Additional task control (white shapes)	Success criterion: 5 correct trials in Stage 3.		*i)* KB effect was found; *ii)* KB was unrelated to intelligence; *iii)* KB was unrelated psychoticism; *iv)* relationships between psychoticism were not consistent with those in E1; *v)* KB was stronger in introverts and not in extroverts.
1991	(Jones et al., 1991) – British Journal of Psychiatry	1	N(NT) = 34; age M 31.6, SD 7.3 N(SZ) = 30 (19 chronic; age M 39.8, SD 14; 11 acute, age M 34.4, SD 11.1)	Flanker Task (Choice Reaction Time Task)	Correctness (%) and Reaction Time (ms), for (1) the valid and (2) the invalid trials.	SZ based on the Research Diagnosis Criteria Mill Hill Synonym Test (intelligence); BPRS score (psychiatric symptoms)	Chronic SZ slower than acute SZ and NT. Chronic and NT RT invalid > RT valid. No effect in acute SZ. High BPRS median split groups mirrored acute SZ.

Year of Publication	Reference	# Exp.	Sample	Paradigm	Outcome(s)	Diagnoses or Scales	Findings
1992	(Oades & Eggers, 1992) – Schizophrenia Research	1	N(SZ) = 18; N(compulsive-neurotic) = 8; and N(NT) = 18	Oades's task Conditioning Blocking (mouse in a house-plan)	NA	NA	No difference between groups. But urine sample revealed association between high limbic cortical DA and reduced KB.
1992	(SJones et al., 1992a) – Personality and Individual Differences	1	[a]N(NT) = 59 (see above)	KB & IL (see notes)	Success criterion: 5 correct trials in Stage 3.	Same as Jones et al (1990). STA: 'magical ideation', 'unusual perceptual experiences', and 'paranoid ideation and suspiciousness'.	Marginal effect on KB and 'magical ideation' and unusual perceptual experiences. Incidental learning recall associated with 'unusual perceptual experiences.'
1992	(Jones et al., 1992b) – Biological Psychiatry	1	[a]N(NT) = 40; age M 32.5, SD 10.6 N(SZ) = 30 (16 chronic; age M 37.8, SD 15.8; 14 acute, age M 36.8, SD 11.2)	KB (sensorial cross modality: CS1, visual, and CS2, auditory)	Trials for learning criterion: Mean ranks and Median.	SZ based on the Research Diagnosis Criteria Mill Hill Synonym Test (intelligence); BPRS score	No KB in acute SZ. Chronic SZ bad learning, hence difficult to interpret. KB may be a useful measure for cognitive dysfunction in positive symptoms of SZ.

Year of Publication	Reference	# Exp.	Sample	Paradigm	Outcome(s)	Diagnoses or Scales	Findings
		2	N(NT) = 30; age M 31.1, SD 6.8 N(SZ) = 29 (15 chronic; age M 38.1, SD 12.4; 14 acute, age M 31.5, SD 9.1)	KB (same sensorial modality CS1 and CS2 visual)		(psychiatric symptoms)	Same modality enhanced overall learning. No KB in acute SZ. Chronic SZ showed trend to KB. KB may be a useful measure for cognitive dysfunction in positive symptoms of SZ.
1996	(Oades et al., 1996) – Journal of Psychiatric Research	1	N(par.-hal.) = 14 (F 6); M 19.3, SD 3.6. N(non-par) = 12 (6 F); M 17.3, SD 3.6. N(OCD) = 13 (4 F); M 16.3, SD 2.9. N(NT) = 29 (14 F); M 18, SD 3.6.	Oades's task Conditioning Blocking (mouse in a house-plan)	Goal latency (sec) difference between pre expose colours and second phase compound colour.	Intelligence (Ravens' test); Minnesota Multiphasic Personality Inventory (MMPI); Hamburg Neuroticism Extroversion Scale (HANES); Urine collection.	Non-paranoid SZ showed enhanced learning to novel cue. Paranoid SZ showed KB attenuation. No association between neurotic personality and KB. Association between high DA and high KB. Psychotic patients spend more time exploring.
1997	(S. H. Jones et al., 1997) – Behavioural Brain Research	1	Review (experiments N = 89 (27 NT, 19 NT with SZ relatives; 21 SZ; 22 SZ relatives with SPD)	Between-subject KB	Means on ranked number of trials taken to learn that CS2 (white flanking squares) predicts UCS (yellow square) in Stage 3	SPD with DSM-IIIR	KB effect found NT. No KB in SPD. SZ and non-schizotypal relatives learnt faster the "blocked" cue. Schizotypal relatives reduced KB.

Year of Publication	Reference	# Exp.	Sample	Paradigm	Outcome(s)	Diagnoses or Scales	Findings
	[this is review based on Serra's Ph.D. thesis experiments]	2	N(NT) = 84 (no drug, placebo, 5 mg, or 10 mg of amphetamine).	Between-subject KB		NA	No drug and placebo showed KB. But, 5 and 10 mg of amphetamine remove the KB effect. Only 'clinical', not supported by stats.
		3	No sample characteristics available	Within-subject KB	Predictive ratings	RDC and BPRS	Attenuation of KB in patients with SZ.
		4	N(NT) = 48 (placebo, 5- and 10 mg)	Within-subject KB	Predictive ratings	NA	Amphetamine reduce KB and produce superconditioning to the 'irrelevant' cue.
1997	(Gray et al., 1997) – Journal of Psychopharmacology	1	N(NT) = 93 (F 49) N(0 mg) = 32 N(5mg) = 29 N(10mg) = 32	KB (same sensorial modality CS1 and CS2 visual)	Trials for learning criterion: Mean ranks and Median.	Mill Hill Synonym Test (intelligence); Standard Progressive Matrices.	KB effect was found, but it was not disrupted by amphetamines in NT (see discussion regarding SZ differences in KB and LI).
2000	(Oades et al., 2000) – Behavioural Pharmacology	1	N(SZ) = 108; M ~35, SD 10.5 N(NT) = 62; M 32.5, SD 10.9	Oades's task Conditioning Blocking (mouse in a house-plan)	Goal latency (sec) difference.	PANSS; SAPS; AIMS	Non-paranoid SZ showed impairment in KB task. D2 occupancy related to KB task impairment in non-paranoid with KB task impairment.
2001	(Serra et al., 2001) – Schizophrenia Research	1	N(NT) = 27 (F 14)	KB and LI (Jones, et al., 1992)	Learning criterion	NA	LI and KB were found.
		2	N(SZ) = 21 N(first-degree SPD) = 22	KB and LI (Jones, et al., 1992)	Learning criterion	BRPS; DSM-III-R for SPD	SZ and first order relatives with SPD and non-SPD showed impaired in both KB and LI.

Year of Publication	Reference	# Exp.	Sample	Paradigm	Outcome(s)	Diagnoses or Scales	Findings
			N(first-degree non-SPD) = 19				
2001	(Bender et al., 2001) – Schizophrenia Research	1	N(SZ learn) = 62 (F 17); M 30.4, SD 9.6 N(SZ not learn) = 39 (F 18); M 37.1, SD 12 N(NT) = 62 (F 29); M 32.5, SD 11	Oades's task Conditioning Blocking (mouse in a house-plan)	Goal latency (sec) difference.	PANSS; SAPS; AIMS	Paranoid SZ showed KB, but non-paranoid SZ did not exhibit KB. Self-referential ideation negatively correlates with KB score.
2003	(Moran et al., 2003) – Journal of Psychiatric Research	1	N(NT) = 30 (F 15); M 28.5, SD 9.84 N(SZ) = 30 (F 5); M 35.8, SD 10.65	Oades's task Conditioning Blocking (mouse in a house-plan)	Goal latency (sec) difference.	Patients diagnosed by DSM IV. Long (160 items) O-LIFE	Impairment in KB found in non-paranoid SZ and intact in paranoid SZ and controls. Negative correlation between UE, CD, and IN and KB (in both patients and controls).
2007	(Corlett et al., 2007) – Brain	1	N(NT) = 12 (F 4) N(SZ) = 12 (F 5)	Retrospective Revaluation (Backward Blocking)	Predictive response and duration of press for confidence.	Patients diagnosed by DSM IV and BRPS.	Unovershadowing occurred in both groups, but no group difference. Psychotic patients have a general attenuation to process PE in the rPFC. The attenuation was correlated with 'unusual though content'.
2007	(Evans et al., 2007) – Biological Psychology	1	N(NT) = 80 (F 46); M 20.39, SD 2.34	LI (predict target cue)	Predictive response; RT (sec)	Long (160 items) O-LIFE	No association between LI and schizotypy (RT nor % correct). Only a trend ($p = 0.08$) between LI and UE for both RT and % correct.

Year of Publication	Reference	# Exp.	Sample	Paradigm	Outcome(s)	Diagnoses or Scales	Findings
2009	(Schmidt-Hansen et al., 2009) – Cognitive Neuropsychiatry	1a & 1b	N(NT) = 69 (F 60); M 19.38, SD 2.02	LI	% Correct predictive response	Long (160 items) O-LIFE	Experiment 1a no differences in LI in UE. Experiment 1b showed reduce LI in high levels of UE.
		2	N(NT) = 79 (F 68); M 19.46, SD 1.99	LIRR			Learned irrelevance was abolished in high UE schizotypy scores (measured by correct predictive responses)
2009	(Randell et al., 2009) – Personality and Individual Differences	1	N(NT) = 73 (F 27); M 25.16, SD 4.28	Random-ratio (RR) and random-interval (RI) schedule.	Response rates in 4 min trials per reinforcement schedule.	Short (43 items) O-LIFE; PDI; BDI; STAI-T; PANAS	Higher response-rate to RR than RI. High UE and PDS were insensitive to the different reinforcement schedules.
		2	N(NT) = 72 (F 39); M 21.66, SD 4.83	Systematic replication. But higher RR.			Higher response-rate to RR than RI. High UE were insensitive to the different reinforcement schedules.
2010	(le Pelley et al., 2010) – Psychiatry Research	1	N(NT) = 115 (F 98)	LIRR	Difference scores (test minus stage 2)	Long (160 items) O-LIFE	UE exhibit a reduced LIRR. LIRR score (based on difference between cues) controlling by % of correct was negatively correlated with UE.
2010	(Haselgrove & Evans, 2010) – The Quarterly	1	N(NT) = 86 (F 53)	KB	Contingency Rating	Long (160 items) O-LIFE	IA showed a KB reduction. Schizotypy modelled with non-competition mechanism (Bush and Mosteller instead of RW model).

Year of Publication	Reference	# Exp.	Sample	Paradigm	Outcome(s)	Diagnoses or Scales	Findings
	Journal of Experimental Psychology	2	N(NT) = 66 (F 48)	Rescorla (2000)	Contingency Rating	Long (160 items) O-LIFE	High IA resulted in a greater loss of associative strength to A than to B in comparison of Low IA. Suggesting a non-competition mechanism for High. No UE findings.
2011	(Moore, Dickinson, et al., 2011) – Consciousness and Cognition	1	N(NT) = 16 (F 11); M 25	Outcome blocking procedure: Phase 1 (pre); Phase 2 (compound); Phase 3 (binding test)	Judgment errors (judged - actual time). Outcome surprise magnitude (tone judgement error – action judgement error)	PAS; MIS; PDI	Estimated time of responses shifted into produced tones (SoA). Magnitude of outcome surprise was negatively correlated with PAS, MIS and PDI. When people have less SoA this reduced the surprise, i.e., less PE, hence less believe updating.
2012	(Randell et al., 2012) – Learning and Individual Differences	1	N(NT) = 50 (F 25); M 21, SD 1.9	Reinforcement Schedules: Random-ratio (RR) and random-interval (RI) schedule.	Contingency awareness (CA); Response rates; Verbal description of the underlying rule	Short (43 items) O-LIFE; BDI; STAI-T	Replication of Randell et al. (2009) High UE did not respond differently between RR and RI, and they were not aware about the underlying contingency rules (CA nor verbally).
2012	(Corlett & Fletcher, 2012) – Neuropsychologia	1	N(NT) = 18 (8 F)	KB (Forward Blocking)	Predictive strength = R x (length of button push) + fMRI	Chapman scales (social anhedonia, physical anhedonia, magical ideation, perceptual aberration); PDI	KB was found. Magical ideation negatively correlated with striatum PE activation. "Distressing beliefs are associated with aberrant prefrontal prediction error." (Corlett, 2012)

Year of Publication	Reference	# Exp.	Sample	Paradigm	Outcome(s)	Diagnoses or Scales	Findings
2012	(Moran et al., 2012) – Plos One	1	N(NT) = 18; only 9 learn KM (F 4); M 22.8, SD 5.1	Oades's task Conditioning Blocking (mouse in a house-plan)	KB score: goal latency (sec) difference.	Short (43 items) O-LIFE	Only IA was negatively correlated with KB (n=18). fMRI: right medial frontal gyrus was negatively correlated with KB.
2012	(Granger et al., 2012) – Behavioural Brain Research	1	N(NT) = 33; M 21.2, SD 5.6	Overshadow task; LI. Exp. 1: individual testing; Exp. 2: group testing	% Correct; RT (sec)	Long (104 items) O-LIFE	Negative association between UE and LI, as well as overshadowing. Not the other subscales). Replication of Evans et al (2007).
		2	N(NT) = 34; M 21.8, SD 2.1				
2013	(Morris et al., 2013) – Schizophrenia Bulletin	1	N(NT) = 14 (F 6); M 36, SD 2.8 N(SZ) = 14 (F 5); M 36, SD 2.0	LIRR (Le Pelley et al, 2010)	Learning score or confidence accuracy (confidence rating multiply by -1 when incorrect)	SZ or schizoaffective (DSM-IV); PANSS; WAIS; WTAR	SZ did not show learned irrelevance effect, i.e., they attend more nonpredictive cues.
		2	N(NT) = 15 (F 7); M 32, SD 2.1 N(SZ) = 20 (F 8); M 33, SD 2.1	Easier version of LIRR (less cues in stage 1)			Positive correlation between positive symptoms and nonpredictive learning score. No correlation for predictive learning score.
2013	(Balzan et al., 2013) – Journal of Nervous and Mental Disease	1	N(SZ) = 25 (F 10); M 40, SD 10 N(NT delusion risk) = 25 (F 16); M 44, SD 16 N(NT) = 25 (F	Illusory Correlation Fertilizer Task (Kao & Wasserman, 1993); Illusion of Control Task (Alloy & Abramson,	PANSS; MINI; PDI-21; WASI; NART	*i)* SZ and delusion proneness higher rating to noncontingent problems (illusory correlation). *ii)* PDI-21 is correlated with	

Year of Publication	Reference	# Exp.	Sample	Paradigm	Outcome(s)	Diagnoses or Scales	Findings
			11); M 42, SD 15	1979)0 to +10); Contingency ratings (0 to 100)		both positive and negative noncontingent problems (more PDI, more extreme ratings). *iii)* SZ weight less cell D (no fertilizer and no bloom). *iv)* SZ higher button press. *v)* SZ higher PC (also correlated with PDI)	
2014	(Moritz et al., 2014) (Moritz et al., 2014) – Journal of Experimental Psychopathology	1	N(SZ) = 32 (F 13); M 34.41, SD 11.34 N(NT) = 31 (F 17); M 36.03, SD 15.83	Contingency Learning (adapted from Alloy & Abramson, 1979)	Button press and control rating	SZ with DSM-IV; PANSS; MINI Neuropsychiatric Interview	No difference in SZ and NT. In positive contingencies, the positive PANSS positively correlated with illusion of control.
2015	(Morris et al., 2015) – Biological Psychiatry	1	N(SZ) = 18 (F 9); M 45.3, SD 11.4 N(NT) = 18 (F 9); M 39.9, SD 12.9	Pavlovian-Instrumental Transfer Test; and Devaluation Test	Responses to specific and general transfer effect; Devaluated ratings	WASI; WTAR; DASS-21; BIS/BAS; SAPS; SANS	SZ and healthy similar contingency judgments. Goal directed action deficit in SZ, as well as inability to integrate devaluated actions. This may be associated with corticostriatal control. No deficit in reward learning motivation in SZ,

Year of Publication	Reference	# Exp.	Sample	Paradigm	Outcome(s)	Diagnoses or Scales	Findings
							but behavioural aberrant incentive motivation in SZ.
2016	(Haselgrove et al., 2016) – British Journal of Psychology	1	N(NT) = 86 (F 56)	LIRR (Le Pelley et al, 2010)	Difference in test- ratings; relevant and irrelevant seeing proportion.	Long (104 items) O-LIFE	IA learned better irrelevant cues and saw less (Eye Tracking) relevant cues.
		2	N(NT) = 60 (F 20)	LIRR and a target probe test (in RT)	% Correct; RT (sec)		Attentional bias towards relevant stimuli that was present in low-IA participants was absent in high-IA participants.
2016	(Granger et al., 2016) – Personality and Individual Differences	1	N(NT) = 57 (F 22)	LI (two versions; same effect)	Predictive response and RT	Long O-LIFE	UE and IN associated with high LI.
		2	N(NT) = 60 (F 50)	LI, but saying aloud each cue	Predictive response and RT	Long O-LIFE	UE predictor of higher RT to the pre-exposed cue in both Experiments: *slower learning*.
2017	(Humpston et al., 2017) – Cognitive Neuropsychiatry	1	N(NT) = 102 (F 81); M 21.96, SD 3.14.	KB; force-matching task; Reversal learning task;	KB score, Force-matching overcompensation; post-reversal perseveration and error switching	Long O-LIFE; PDI-21; CAPS	None of the tasks were associated with any schizotypy measure.
2017	(Collins et al., 2017) – Translational psychiatry	1	N(SZ) = 46 (F 18); M 37.8, SD 9.	Reinforcement learning with cognitive load task	RT; Correctness (%); Q-Learning values	NA	Slower learning in SZ than NT. In SZ Reinforcement Learning may be affected no by reward mechanisms but by working memory capacity.
		2	N(NT) = 32 (F 11); M 37.1, SD 10.2.				

Year of Publication	Reference	# Exp.	Sample	Paradigm	Outcome(s)	Diagnoses or Scales	Findings
		3	Medicated SZ				
2018	(Morris et al., 2018) – Translational psychiatry	1	N(NT) = 25 (F 11); M 41, SD 15. N(SZ) = 25 (F 11); M 45, SD 8.	Free-operant: reward learning; contingency degradation; outcome devaluation	Causal judgments; Causal actions; Goal-directed choices; Reinforced. choices.	WTAR; DASS-21; SAPS; SANS	*i)* Reward learning intact in SZ; *ii)* SZ had a lower correlation between causal judgments and action-outcome contingency than NT; *iii)* SZ no sensitive to contingency degradation unless feedback was provided.

Note: Kamin Blocking = KB; Latent Inhibition = LI; Learned Irrelevance = LIRR; Incidental Learning = IL; Reaction Time = RT; Schizophrenia = SZ; Neurotypicals = NT; Schizoid Personality Disorder = SPD; Dopamine = DA; Noradrenaline = NA; Brief Psychiatric Rating Scale = BPRS; Positive and Negative Symptoms Scale = PANSS; Scale for Assessment of Positive Symptoms = SAPS; Abnormal Involuntary Movement Scale = AIMS. Oxford-Liverpool Inventory of Feelings and Experiences = O-LIFE; Peter's Delusional Inventory = PDI; Beck's Depression Inventory = BDI; Spielberger Trait Anxiety Inventory = STAI-T; Positive and Negative Affect Schedule = PANAS; Perceptual Aberration Scale = PAS; Magical Ideation Scale = MIS; Sense of Agency = SoA; Wechsler Adult Intelligence Scale = WAIS; Wechsler Test for Adult Reading = WTAR; Unusual Experiences = UE; Introvertive Anhedonia = IA; Cognitive Disorganization = CD; Impulsive Nonconformity = IN; Depression, Anxiety and Stress Scale = DASS-21; Scale for the Assessment of Negative Symptoms = SANS. Wechsler Abbreviated Scale of Intelligence = WASI.

[a]These experiments are the same from Jones et al (1990) Experiment 1.

6. References

Adams, R. A., Huys, Q. J. M., & Roiser, J. P. (2016). Computational Psychiatry: towards a mathematically informed understanding of mental illness. *J Neurol Neurosurg Psychiatry*, *87*, 53–63. https://doi.org/10.1136/jnnp-2015-310737

Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., & Friston, K. J. (2013). The computational anatomy of psychosis. *Frontiers in Psychiatry*, *4*, 47. https://doi.org/10.3389/FPSYT.2013.00047/BIBTEX

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer New York.

Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, *15*(3), 147–149. https://doi.org/10.3758/BF03334492

Allan, L. G., Hannah, S. D., Crump, M. J. C., & Siegel, S. (2008). The psychophysics of contingency assessment. *Journal of Experimental Psychology: General*, *137*(2), 226–243. https://doi.org/10.1037/0096-3445.137.2.226

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, *108*(4), 441–485. https://doi.org/10.1037/0096-3445.108.4.441

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.

Andreasen, N. C., & Carpenter, W. T. (1993). Diagnosis and Classification of Schizophrenia. *Schizophrenia Bulletin*, *19*(2), 199–214. https://doi.org/10.1093/schbul/19.2.199

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/S13428-019-01237-X

Asai, T., & Tanno, Y. (2008). Highly schizotypal students have a weaker sense of self-agency. *Psychiatry and Clinical Neurosciences*, *62*(1), 115–119. https://doi.org/10.1111/J.1440-1819.2007.01768.X

Avramopoulos, D., Stefanis, N. C., Hantoumi, I., Smyrnis, N., Evdokimidis, I., & Stefanis, C. N. (2002). Higher scores of self reported schizotypy in healthy young males carrying the COMT high activity allele. *Molecular Psychiatry*, *7*(7), 706–711. https://doi.org/10.1038/SJ.MP.4001070

Baetu, I., & Baker, A. G. (2009). Human judgments of positive and negative causal chains. *Journal of Experimental Psychology. Animal Behavior Processes*, *35*(2), 153–168. https://doi.org/10.1037/a0013764

Baetu, I., Baker, A. G., Darredeau, C., & Murphy, R. A. (2005). A comparative approach to cue competition with one and two strong predictors. *Learning & Behavior 2005 33:2*, *33*(2), 160–171. https://doi.org/10.3758/BF03196060

Baker, A. G. (1976). Learned irrelevance and learned helplessness: Rats learn that stimuli, reinforcers, and responses are uncorrelated. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*(2), 130–141. https://doi.org/10.1037/0097-7403.2.2.130

References

Baker, A. G., Berbrier, M. W., & Vallee-Tourangeau, F. (1989). Judgements of a 2 × 2 Contingency Table: Sequential Processing and the Learning Curve. *The Quarterly Journal of Experimental Psychology Section B*, *41*(1), 65–97. https://doi.org/10.1080/14640748908401184

Baker, A. G., & Mackintosh, N. J. (1979). Preexposure to the CS alone, US alone, or CS and US uncorrelated: Latent inhibition, blocking by context or learned irrelevance? *Learning and Motivation*, *10*(3), 278–294. https://doi.org/10.1016/0023-9690(79)90034-1

Baker, A. G., & Mercier, P. (1989). Attention, retrospective processing and cognitive representations. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 85–116). Lawrence Erlbaum Associates, Inc.

Baker, A. G., Mercier, P., Vallée-Tourangeau, F., Frank, R., & Pan, M. (1993). Selective Associations and Causality Judgments: Presence of a Strong Causal Factor May Reduce Judgments of a Weaker One. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(2), 414–432. https://doi.org/10.1037/0278-7393.19.2.414

Baker, A. G., Murphy, R. A., Mehta, R., & Baetu, I. (2005). Mental models of causation: A comparative view. In A. J. Wills (Ed.), *New directions in human associative learning* (pp. 11–40). Lawrence Erlbaum Associates Publishers.

Baker, A. G., Murphy, R. A., & Vallée-Tourangeau, F. (1996). Associative and normative models of causal induction: Reacting to versus understanding cause. In D. R. Shanks, K. Holyoak, & D. L. Medin (Eds.), *Causal Learning* (pp. 1–45). Academic Press. https://psycnet.apa.org/record/2003-00368-001

Baker, A. G., Vallée-Tourangeau, F., & Murphy, R. A. (2000). Asymptotic judgment of cause in a relative validity paradigm. *Memory & Cognition*, *28*(3), 466–479. https://doi.org/10.3758/BF03198561

Balsam, P., Drew, M., & Gallistel, C. (2010). Time and Associative Learning. *Comparative Cognition & Behavior Reviews*, *5*, 1–22. https://doi.org/10.3819/CCBR.2010.50001

Balzan, R. P., Delfabbro, P. H., Galletly, C. A., & Woodward, T. S. (2013). Illusory correlations and control across the psychosis continuum: The contribution of hypersalient evidence-hypothesis matches. *Journal of Nervous and Mental Disease*, *201*(4), 319–327. https://doi.org/10.1097/NMD.0B013E318288E229

Bansal, S., Bae, G. Y., Robinson, B. M., Hahn, B., Waltz, J., Erickson, M., Leptourgos, P., Corlett, P., Luck, S. J., & Gold, J. M. (2022). Association Between Failures in Perceptual Updating and the Severity of Psychosis in Schizophrenia. *JAMA Psychiatry*, *79*(2), 169–177. https://doi.org/10.1001/JAMAPSYCHIATRY.2021.3482

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/J.JML.2012.11.001

Barrantes-Vidal, N., Gross, G. M., Sheinbaum, T., Mitjavila, M., Ballespí, S., & Kwapil, T. R. (2013). Positive and negative schizotypy are associated with prodromal and schizophrenia-spectrum symptoms. *Schizophrenia Research*, *145*(1–3), 50–55. https://doi.org/10.1016/J.SCHRES.2013.01.007

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

References

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An Inventory for Measuring Depression. *Archives of General Psychiatry*, *4*(6), 561. https://doi.org/10.1001/archpsyc.1961.01710120031004

Beedie, S. A., St.Clair, D. M., & Benson, P. J. (2011). Atypical scanpaths in schizophrenia: Evidence of a trait- or state-dependent phenomenon? In *Journal of Psychiatry and Neuroscience* (Vol. 36, Issue 3, pp. 150–164). Canadian Medical Association. https://doi.org/10.1503/jpn.090169

Beesley, T., Nguyen, K. P., Pearson, D., & le Pelley, M. E. (2015). Uncertainty and predictiveness determine attention to cues during human associative learning. *Quarterly Journal of Experimental Psychology*, *68*(11), 2175–2199. https://doi.org/10.1080/17470218.2015.1009919

Beesley, T., Pearson, D., & le Pelley, M. (2015). Implicit learning of gaze-contingent events. *Psychonomic Bulletin & Review*, *22*(3), 800–807. https://doi.org/10.3758/s13423-014-0720-4

Behrens, T. E. J., Hunt, L. T., & Rushworth, M. F. S. (2009). The computation of social behavior. *Science*, *324*(5931), 1160–1164. https://doi.org/10.1126/SCIENCE.1169694/ASSET/A83EA1D1-0D18-42CD-A64E-425665EEDD55/ASSETS/GRAPHIC/324_1160_F4.JPEG

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *NATURE NEUROSCIENCE*, *10*(9). https://doi.org/10.1038/nn1954

Bender, S., Müller, B., Oades, R. D., & Sartory, G. (2001). Conditioned blocking and schizophrenia: a replication and study of the role of symptoms, age, onset-age of psychosis and illness-duration. *Schizophrenia Research*, *49*(1–2), 157–170. https://doi.org/10.1016/S0920-9964(00)00040-2

Blain, S. D., Longenecker, J. M., Grazioplene, R. G., Klimes-Dougan, B., & DeYoung, C. G. (2020). Apophenia as the disposition to false positives: A unifying framework for openness and psychoticism. *Journal of Abnormal Psychology*, *129*(3), 279–292. https://doi.org/10.1037/ABN0000504

Blanco, F. (2017). Positive and negative implications of the causal illusion. *Consciousness and Cognition*, *50*, 56–68. https://doi.org/10.1016/J.CONCOG.2016.08.012

Blanco, F., Barberia, I., & Matute, H. (2015). Individuals Who Believe in the Paranormal Expose Themselves to Biased Information and Develop More Causal Illusions than Nonbelievers in the Laboratory. *PLOS ONE*, *10*(7), e0131378. https://doi.org/10.1371/JOURNAL.PONE.0131378

Blanco, F., Matute, H., & Vadillo, M. A. (2011). Making the uncontrollable seem controllable: The role of action in the illusion of control. *Quarterly Journal of Experimental Psychology*, *64*(7), 1290–1304. https://doi.org/10.1080/17470218.2011.552727

Blanco, F., Matute, H., & Vadillo, M. A. (2012). Mediating Role of Activity Level in the Depressive Realism Effect. *PLOS ONE*, *7*(9), e46203. https://doi.org/10.1371/JOURNAL.PONE.0046203

Blanco, F., Matute, H., & Vadillo, M. A. (2013). Interactive effects of the probability of the cue and the probability of the outcome on the overestimation of null contingency. *Learning & Behavior*, *41*(4), 333–340. https://doi.org/10.3758/S13420-013-0108-8

Blumer, D. (2002). The Illness of Vincent van Gogh. *American Journal of Psychiatry*, *159*(4), 519–526. https://doi.org/10.1176/appi.ajp.159.4.519

References

Bonardi, C., & Ward-Robinson, J. (2001). Occasion Setters: Specificity to the US and the CS–US Association. *Learning and Motivation*, *32*(3), 349–366. https://doi.org/10.1006/LMOT.2001.1089

Bonini, C. P. (1962). *Simulation of information and decision systems in the firm*. Markham Publishing Company.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, *114*(1), 80–99. https://doi.org/10.1037/0033-2909.114.1.80

Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, *18*(4), 590–596. https://doi.org/10.1038/nn.3961

Buchy, L., Woodward, T. S., & Liotti, M. (2007). A cognitive bias against disconfirmatory evidence (BADE) is associated with schizotypy. *Schizophrenia Research*, *90*(1–3), 334–337. https://doi.org/10.1016/J.SCHRES.2006.11.012

Burgos, J. E. (1997). Evolving Artificial Neural Networks in Pavlovian Environments. *Advances in Psychology*, *121*(C), 58–79. https://doi.org/10.1016/S0166-4115(97)80090-8

Burgos, J. E. (2007). Autoshaping and Automaintenance: A Neural-Network Approach. *Journal of the Experimental Analysis of Behavior*, *88*(1), 115–130. https://doi.org/10.1901/JEAB.2007.75-04

Burgos, J. E. (2018). Is a Nervous System Necessary for Learning? *Perspectives on Behavior Science*, *41*(2), 343–368. https://doi.org/10.1007/S40614-018-00179-7

Burgos, J. E., & Donahoe, J. W. (2016). Unified principle of reinforcement in a neural-network model: Reply to N. T. Calvin and J. J. McDowell. *Behavioural Processes*, *126*, 46–54. https://doi.org/10.1016/J.BEPROC.2016.03.003

Burgos, J. E., Flores, C., García, Ó., Díaz, C., & Cruz, Y. (2008). A simultaneous procedure facilitates acquisition under an optimal interstimulus interval in artificial neural networks and rats. *Behavioural Processes*, *78*(2), 302–309. https://doi.org/10.1016/J.BEPROC.2008.02.018

Burgos, J. E., & García-Leal, Ó. (2015). Autoshaped choice in artificial neural networks: Implications for behavioral economics and neuroeconomics. *Behavioural Processes*, *114*(1), 63–71. https://doi.org/10.1016/J.BEPROC.2015.01.010

Bürgy, M. (2008). The Concept of Psychosis: Historical and Phenomenological Aspects. *Schizophrenia Bulletin*, *34*(6), 1200–1210. https://doi.org/10.1093/SCHBUL/SBM136

Bush, R. R., & Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, *58*(5), 313–323. https://doi.org/10.1037/h0054388

Byrom, N. C., Msetfi, R. M., & Murphy, R. A. (2015). Two pathways to causal control: Use and availability of information in the environment in people with and without signs of depression. *Acta Psychologica*, *157*, 1–12. https://doi.org/10.1016/J.ACTPSY.2015.02.004

Byrom, N. C., & Murphy, R. A. (2018). Individual differences are more than a gene × environment interaction: The role of learning. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*(1), 36–55. https://doi.org/10.1037/xan0000157

Byrom, N. C., & Murphy, R. A. (2019). Cue competition influences biconditional discrimination. *Quarterly Journal of Experimental Psychology (2006)*, *72*(2), 182–192. https://doi.org/10.1080/17470218.2017.1363256

Castiello, S., Burgos, J. E., Buriticá, J., dos Santos, C. V., & Alcalá, J. E. (2020). Interaction between magnitude and probability of reinforcement on autoshaped choice. *Revista Mexicana de Analisis de La Conducta*, *46*(1). https://doi.org/10.5514/rmac.v46.i1.76949

Castiello, S., Miller, R. R., Witnauer, J. E., Alcaide, D. M., Fung, E., Pitliya, R. J., Morrissey, D. K. C., & Murphy, R. A. (2022). Benefiting from trial spacing without the cost of prolonged training: Frequency, not duration, of trials with absent stimuli enhances perceived contingency. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/XGE0001166

Castiello, S., Senan, S., Msetfi, R. M., & Murphy, R. A. (2020). Traits for depression related to agentic and external control. *Learning and Motivation*, *72*. https://doi.org/10.1016/j.lmot.2020.101684

Castiello, S., Zhang, W., & Delamater, A. R. (2021). The retrosplenial cortex as a possible "sensory integration" area: A neural network modeling approach of the differential outcomes effect in negative patterning. *Neurobiology of Learning and Memory*, *185*, 107527. https://doi.org/10.1016/J.NLM.2021.107527

Castiello-de Obeso, S., Aguayo Mendoza, M. A., Ortiz-Orendain, J., Itzaman, I., Landa-Ramírez, E., Carmona, J., & Murphy, R. A. (2020). Computer-based Cognitive Remediation Therapy plus standard care versus standard care for people with schizophrenia or related disorders. *Cochrane Database of Systematic Reviews*, *2020*(5). https://doi.org/10.1002/14651858.CD013625

Chan, C. K. J., & Harris, J. A. (2019). The partial reinforcement extinction effect: The proportion of trials reinforced during conditioning predicts the number of trials to extinction. *Journal of Experimental Psychology. Animal Learning and Cognition*, *45*(1), 43–58. https://doi.org/10.1037/XAN0000190

Cheng, P. W. (1997). From Covariation to Causation: A Causal Power Theory. *Psychological Review*, *104*(2), 367–405. https://doi.org/10.1037/0033-295X.104.2.367

Cicero, D. C., & Kerns, J. G. (2011). Unpleasant and pleasant referential thinking: Relations with self-processing, paranoia, and other schizotypal traits. *Journal of Research in Personality*, *45*(2), 208–218. https://doi.org/10.1016/J.JRP.2011.02.002

Ciullo, V., Piras, F., Vecchio, D., Banaj, N., Coull, J. T., & Spalletta, G. (2018). Predictive timing disturbance is a precise marker of schizophrenia. *Schizophrenia Research: Cognition*, *12*, 42. https://doi.org/10.1016/J.SCOG.2018.04.001

Claridge, G. (1994). Single Indicator of Risk for Schizophrenia: Probable Fact or Likely Myth? *Schizophrenia Bulletin*, *20*(1), 151–168. https://doi.org/10.1093/SCHBUL/20.1.151

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

Claridge, G. (2018). *Psychopathology and Personality Dimensions: The Selected Works of Gordon Claridge* (1st ed.). Routledge.

Claridge, G., McCreery, C., Mason, O., Bentall, R., Boyle, G., Slade, P., & Popplewell, D. (1996). The factor structure of 'schizotypal' traits: A large replication study. *British Journal of Clinical Psychology*, *35*(1), 103–115. https://doi.org/10.1111/J.2044-8260.1996.TB01166.X

Collins, A. G. E., Albrecht, M. A., Waltz, J. A., Gold, J. M., & Frank, M. J. (2017). Interactions Among Working Memory, Reinforcement Learning, and Effort in Value-Based Choice: A New Paradigm and Selective Deficits in Schizophrenia. *Biological Psychiatry*, *82*(6), 431–439. https://doi.org/10.1016/J.BIOPSYCH.2017.05.017

Collins, A. G. E., Brown, J. K., Gold, J. M., Waltz, J. A., & Frank, M. J. (2014). Working Memory Contributions to Reinforcement Learning Impairments in Schizophrenia. *The Journal of Neuroscience*, *34*(41), 13747. https://doi.org/10.1523/JNEUROSCI.0989-14.2014

Collins, A. G. E., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience*, *35*(7), 1024–1035. https://doi.org/10.1111/j.1460-9568.2011.07980.x

Corlett, P. R., & Fletcher, P. C. (2012). The neurobiology of schizotypy: Fronto-striatal prediction error signal correlates with delusion-like beliefs in healthy people. *Neuropsychologia*, *50*(14), 3612–3620. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2012.09.045

Corlett, P. R., Horga, G., Fletcher, P. C., Alderson-Day, B., Schmack, K., & Powers, A. R. (2019). Hallucinations and Strong Priors. *Trends in Cognitive Sciences*, *23*(2), 114–127. https://doi.org/10.1016/J.TICS.2018.12.001

Corlett, P. R., Murray, G. K., Honey, G. D., Aitken, M. R. F., Shanks, D. R., Robbins, T. W., Bullmore, E. T., Dickinson, A., & Fletcher, P. C. (2007). Disrupted prediction-error signal in psychosis: Evidence for an associative account of delusions. *Brain*, *130*(9), 2387–2400. https://doi.org/10.1093/brain/awm173

Crider, A., Solomon, P. R., & McMahon, M. A. (1982). Disruption of selective attention in the rat following chronic d-amphetamine administration: Relationship to schizophrenic attention disorder. *Biological Psychiatry*.

Crump, M. J. C., Hannah, S. D., Allan, L. G., & Hord, L. K. (2007). Contingency judgements on the fly. *Quarterly Journal of Experimental Psychology*, *60*(6), 753–761. https://doi.org/10.1080/17470210701257685

Cummins, D. D., Lubart, T., Alksnis, O., & Rist, R. (1991). Conditional reasoning and causation. *Memory & Cognition 1991 19:3*, *19*(3), 274–282. https://doi.org/10.3758/BF03211151

Cuve, H. C., Castiello, S., Shiferaw, B., Ichijo, E., Catmur, C., & Bird, G. (2021). Alexithymia explains atypical spatiotemporal dynamics of eye gaze in autism. *Cognition*, *212*. https://doi.org/10.1016/J.COGNITION.2021.104710

Daedelow, L. S., Banaschewski, T., Berning, M., Bokde, A. L. W., Brühl, R., Burke Quinlan, E., Curran, H. V., Desrivières, S., Flor, H., Grigis, A., Garavan, H., Hardon, A., Kaminski, J., Martinot, J. L., Paillère Martinot, M. L., Artiges, E., Murray, H., Nees, F., Oei, N. Y. L., … Heinz, A. (2021). Are psychotic-like experiences related to a discontinuation of cannabis consumption in young adults? *Schizophrenia Research*, *228*, 271–279. https://doi.org/10.1016/J.SCHRES.2021.01.002

Daprati, E., Franck, N., Georgieff, N., Proust, J., Pacherie, E., Dalery, J., & Jeannerod, M. (1997). Looking for the agent: an investigation into consciousness of action and self-consciousness in schizophrenic patients. *Cognition*, *65*(1), 71–86. https://doi.org/10.1016/S0010-0277(97)00039-5

Daunizeau, J., den Ouden, H. E. M., Pessiglione, M., Kiebel, S. J., Stephan, K. E., & Friston, K. J. (2010). Observing the observer (I): Meta-bayesian models of learning and decision-making. *PLoS ONE*, *5*(12). https://doi.org/10.1371/JOURNAL.PONE.0015554

Davis, K. L., Kahn, R. S., Ko, G., & Davidson, M. (2006). Dopamine in schizophrenia: a review and reconceptualization. *Https://Doi.Org/10.1176/Ajp.148.11.1474*, *148*(11), 1474–1486. https://doi.org/10.1176/AJP.148.11.1474

Delamater, A. R. (2012). On the nature of CS and US representations in Pavlovian learning. *Learning & Behavior*, *40*(1), 1–23. https://doi.org/10.3758/s13420-011-0036-4

Dickinson, A. (2001). Causal learning: Association versus computation. *Current Directions in Psychological Science*, *10*(4), 127–132. https://doi.org/10.1111/1467-8721.00132

Dickinson, A., Shanks, D., & Evenden, J. (1984a). Judgement of Act-Outcome Contingency: The Role of Selective Attribution. *Quarterly Journal of Experimental Psychology*, *36*(1), 29–50. https://doi.org/10.1080/14640748408401502

Dickinson, A., Shanks, D., & Evenden, J. (1984b). Judgement of Act-Outcome Contingency: The Role of Selective Attribution: *Quarterly Journal of Experimental Psychology*, *36*(1), 29–50. https://doi.org/10.1080/14640748408401502

Donahoe, J. W., Burgos, J. E., & Palmer, D. C. (1993). A Selectionist Approach To Reinforcement. *Journal of the Experimental Analysis of Behavior*, *60*(1), 17–40. https://doi.org/10.1901/JEAB.1993.60-17

D'Souza, D. C., Abi-Saab, W. M., Madonick, S., Forselius-Bielen, K., Doersch, A., Braley, G., Gueorguieva, R., Cooper, T. B., & Krystal, J. H. (2005). Delta-9-tetrahydrocannabinol effects in schizophrenia: Implications for cognition, psychosis, and addiction. *Biological Psychiatry*, *57*(6), 594–608. https://doi.org/10.1016/J.BIOPSYCH.2004.12.006

Dubovsky, S. L., Ghosh, B. M., Serotte, J. C., & Cranwell, V. (2021). Psychotic Depression: Diagnosis, Differential Diagnosis, and Treatment. *Psychotherapy and Psychosomatics*, *90*(3), 160–177. https://doi.org/10.1159/000511348

Dudley, R., Taylor, P., Wickham, S., & Hutton, P. (2016). Psychosis, Delusions and the "Jumping to Conclusions" Reasoning Bias: A Systematic Review and Meta-analysis. *Schizophrenia Bulletin*, *42*(3), 652–665. https://doi.org/10.1093/schbul/sbv150

Eckstein, M. K., Master, S. L., Xia, L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. (2022). The interpretation of computational model parameters depends on the context. *ELife*, *11*. https://doi.org/10.7554/ELIFE.75474

Edwards, M. J., Adams, R. A., Brown, H., Pareés, I., & Friston, K. J. (2012). A Bayesian account of "hysteria." *Brain*, *135*(Pt 11), 3495–3512. https://doi.org/10.1093/BRAIN/AWS129

Engle, R. W. (2016). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science*, *11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Esber, G. R., & Haselgrove, M. (2011). Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1718), 2553–2561. https://doi.org/10.1098/RSPB.2011.0836

Escobar, M., Arcediano, F., & Miller, R. R. (2002). Latent inhibition and contextual associations. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*(2), 123–136. https://doi.org/10.1037/0097-7403.28.2.123

Esterberg, M. L., & Compton, M. T. (2009). The psychosis continuum and categorical versus dimensional diagnostic approaches. *Current Psychiatry Reports*, *11*(3), 179–184. https://doi.org/10.1007/S11920-009-0028-7

Evans, L. H., Gray, N. S., & Snowden, R. J. (2007). A new continuous within-participants latent inhibition task: Examining associations with schizotypy dimensions, smoking status and gender. *Biological Psychology*, *74*(3), 365–373. https://doi.org/10.1016/J.BIOPSYCHO.2006.09.007

Eysenck, H. J. (1947). *Dimensions of Personality: A Record of Research*. Paul, Trench, Trubner & Co.

Eysenck, H. J. (1952). *The Scientific Study of Personality*. Routledge & Kegan Paul.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of Reinforcement*. Appleton-Century-Crofts.

Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., Akers, C. A., Clinton, S. M., Phillips, P. E. M., & Akil, H. (2011). A selective role for dopamine in stimulus-reward learning. *Nature*, *469*(7328), 53–59. https://doi.org/10.1038/nature09588

Friston, K., Brown, H. R., Siemerkus, J., & Stephan, K. E. (2016). The dysconnection hypothesis (2016). *Schizophrenia Research*, *176*(2–3), 83–94. https://doi.org/10.1016/J.SCHRES.2016.07.014

Fusar-Poli, P., & Politi, P. (2008). Paul Eugen Bleuler and the Birth of Schizophrenia (1908). *American Journal of Psychiatry*, *165*(11), 1407. https://doi.org/10.1176/appi.ajp.2008.08050714

Fyfe, S., Williams, C., Mason, O. J., & Pickup, G. J. (2008). Apophenia, theory of mind and schizotypy: Perceiving meaning and intentionality in randomness. *Cortex*, *44*(10), 1316–1325. https://doi.org/10.1016/J.CORTEX.2007.07.009

Gal, G., Mendlovic, S., Bloch, Y., Beitler, G., Levkovitz, Y., Young, A. M. J., Feldon, J., & Ratzoni, G. (2005). Learned irrelevance is disrupted in first-episode but not chronic schizophrenia patients. *Behavioural Brain Research*, *159*(2), 267–275. https://doi.org/10.1016/J.BBR.2004.11.017

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences*, *4*(1), 14–21. https://doi.org/10.1016/S1364-6613(99)01417-5

Gallistel, C. R., & Papachristos, E. B. (2020). Number and time in acquisition, extinction and recovery. *Journal of the Experimental Analysis of Behavior*, *113*(1), 15–36. https://doi.org/10.1002/JEAB.571

Garnier, Simon, Ross, Noam, Rudis, Robert, Camargo, Pedro, A., Sciaini, Marco, Scherer, & Cédric. (2021). *viridis - Colorblind-Friendly Color Maps for R*. https://doi.org/10.5281/zenodo.4679424

Gershman, S. J. (2015). A Unifying Probabilistic View of Associative Learning. *PLOS Computational Biology*, *11*(11), e1004567. https://doi.org/10.1371/JOURNAL.PCBI.1004567

Ghirlanda, S., Lind, J., & Enquist, M. (2020). A-learning: A new formulation of associative learning theory. *Psychonomic Bulletin and Review*, *27*(6), 1166–1194. https://doi.org/10.3758/S13423-020-01749-0/FIGURES/15

Gold, J. M. (2004). Cognitive deficits as treatment targets in schizophrenia. *Schizophrenia Research*, *72*(1), 21–28. https://doi.org/10.1016/J.SCHRES.2004.09.008

Goldman, A. I. (1967). A Causal Theory of Knowing. *The Journal of Philosophy*, *64*(12), 357. https://doi.org/10.2307/2024268

Goldman, A. I. (1979). What is Justified Belief? *Justification and Knowledge*, 1–23. https://doi.org/10.1007/978-94-009-9493-5_1

Gooding, D. C., Tallent, K. A., & Matts, C. W. (2005). Clinical status of at-risk individuals 5 years later: Further validation of the psychometric high-risk strategy. *Journal of Abnormal Psychology*, *114*(1), 170–175. https://doi.org/10.1037/0021-843X.114.1.170

Granger, K. T., Moran, P. M., Buckley, M. G., & Haselgrove, M. (2016). Enhanced latent inhibition in high schizotypy individuals. *Personality and Individual Differences*, *91*, 31–39. https://doi.org/10.1016/J.PAID.2015.11.040

Granger, K. T., Prados, J., & Young, A. M. J. (2012). Disruption of overshadowing and latent inhibition in high schizotypy individuals. *Behavioural Brain Research*, *233*(1), 201–208. https://doi.org/10.1016/J.BBR.2012.05.003

Grant, P., Green, M. J., & Mason, O. J. (2018). Models of Schizotypy: The Importance of Conceptual Clarity. *Schizophrenia Bulletin*, *44*(suppl_2), S556–S563. https://doi.org/10.1093/schbul/sby012

Gray, N. S., Pickering, A. D., Gray, J. A., Jones, S. H., Abrahams, S., & Hemsley, D. R. (1997). Kamin blocking is not disrupted by amphetamine in human subjects. *Journal of Psychopharmacology*, *11*(4), 301–311. https://doi.org/10.1177/026988119701100404

Griffiths, O., Langdon, R., le Pelley, M. E., & Coltheart, M. (2014). Delusions and prediction error: re-examining the behavioural evidence for disrupted error signalling in delusion formation. *Cognitive Neuropsychiatry*, *19*(5), 439–467. https://doi.org/10.1080/13546805.2014.897601

Griffiths, O., Shehabi, N., Murphy, R. A., & le Pelley, M. E. (2018). Superstition predicts perception of illusory control. *British Journal of Psychology*. https://doi.org/10.1111/bjop.12344

Griffiths, O., Shehabi, N., Murphy, R. A., & le Pelley, M. E. (2019). Superstition predicts perception of illusory control. *British Journal of Psychology*, *110*(3), 499–518. https://doi.org/10.1111/BJOP.12344

Gur, R. C., Erwin, R. J., Gur, R. E., Zwil, A. S., Heimberg, C., & Kraemer, H. C. (1992). Facial emotion discrimination: II. Behavioral findings in depression. *Psychiatry Research*, *42*(3), 241–251. https://doi.org/10.1016/0165-1781(92)90116-K

Hall, W., & Degenhardt, L. (2008). Cannabis use and the risk of developing a psychotic disorder. *World Psychiatry*, *7*(2), 68. https://doi.org/10.1002/J.2051-5545.2008.TB00158.X

Harper, L., Spencer, E., Davidson, C., & Hutchinson, C. v. (2020). Selectively reduced contrast sensitivity in high schizotypy. *Experimental Brain Research*, *238*(1), 51. https://doi.org/10.1007/S00221-019-05695-9

Harrell Jr, F. E. (2021). *Hmisc: Harrell Miscellaneous*. https://CRAN.R-project.org/package=Hmisc

Harris, J. A., & Andrew, B. J. (2017). Time, Trials, and Extinction. *Journal of Experimental Psychology: Animal Learning and Cognition*, *43*(1), 15–29. https://doi.org/10.1037/XAN0000125

Harris, J. A., & Bouton, M. E. (2020). Pavlovian conditioning under partial reinforcement: The effects of nonreinforced trials versus cumulative conditioned stimulus duration. *Journal of Experimental Psychology: Animal Learning and Cognition*, *46*(3), 256–272. https://doi.org/10.1037/XAN0000242

Harris, J. A., Patterson, A. E., & Gharaei, S. (2015). Pavlovian conditioning and cumulative reinforcement rate. *Journal of Experimental Psychology: Animal Learning and Cognition*, *41*(2), 137–151. https://doi.org/10.1037/XAN0000054

Haselgrove, M., Esber, G. R., Pearce, J. M., & Jones, P. M. (2010). Two Kinds of Attention in Pavlovian Conditioning: Evidence for a Hybrid Model of Learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*(4), 456–470. https://doi.org/10.1037/A0018528

Haselgrove, M., & Evans, L. H. (2010). Variations in selective and nonselective prediction error with the negative dimension of schizotypy. *Quarterly Journal of Experimental Psychology*, *63*(6), 1127–1149. https://doi.org/10.1080/17470210903229979

Haselgrove, M., le Pelley, M. E., Singh, N. K., Teow, H. Q., Morris, R. W., Green, M. J., Griffiths, O., & Killcross, S. (2016). Disrupted attentional learning in high schizotypy: Evidence of aberrant salience. *British Journal of Psychology*, *107*(4), 601–624. https://doi.org/10.1111/bjop.12175

Henriksen, M. G., & Parnas, J. (2014). Self-disorders and schizophrenia: a phenomenological reappraisal of poor insight and noncompliance. *Schizophrenia Bulletin*, *40*(3), 542–547. https://doi.org/10.1093/SCHBUL/SBT087

Hergovich, A., Schott, R., & Arendasy, M. (2008a). On the relationship between paranormal belief and schizotypy among adolescents. *Personality and Individual Differences*, *45*(2), 119–125. https://doi.org/10.1016/j.paid.2008.03.005

Hergovich, A., Schott, R., & Arendasy, M. (2008b). On the relationship between paranormal belief and schizotypy among adolescents. *Personality and Individual Differences*, *45*(2), 119–125. https://doi.org/10.1016/J.PAID.2008.03.005

Higgins, J. P. T., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions: Vol. Version 5.1.0* (J. P. T. Higgins & S. Green, Eds.). The Cochrane Collaboration. Available from www.handbook.cochrane.org.

Hipólito, I., Gonçalves, J., & Pereira, J. (2018). *Schizophrenia and Common Sense*. Springer Nature.

Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-analyses. *Cognitive Therapy and Research*, *36*(5), 427–440. https://doi.org/10.1007/s10608-012-9476-1

Honey, R. C., Dwyer, D. M., & Iliescu, A. F. (2020). HeiDI: A model for Pavlovian learning and performance with reciprocal associations. *Psychological Review*. https://doi.org/10.1037/REV0000196

Hua, J. P. Y., Karcher, N. R., & Kerns, J. G. (2020). Examining associations between two different jumping to conclusions scores with positive schizotypy and recent distress. *Cognitive Neuropsychiatry*, *25*(1), 45–56. https://doi.org/10.1080/13546805.2019.1682984

Humpston, C. S., Evans, L. H., Teufel, C., Ihssen, N., & Linden, D. E. J. (2017). Evidence of absence: no relationship between behaviourally measured prediction error response and schizotypy. *Cognitive Neuropsychiatry*, *22*(5), 373–390. https://doi.org/10.1080/13546805.2017.1348289

Huys, Q. J. M., & Dayan, P. (2009). A Bayesian formulation of behavioral control. *Cognition*, *113*(3), 314–328. https://doi.org/10.1016/J.COGNITION.2009.01.008

Jocham, G., Brodersen, K. H. H., Constantinescu, A. O. O., Kahn, M. C. C., Ianni, A. M., Walton, M. E. E., Rushworth, M. F. F. S., & Behrens, T. E. E. J. (2016). Reward-Guided Learning with and without Causal Attribution. *Neuron*, *90*(1), 177. https://doi.org/10.1016/J.NEURON.2016.02.018

Johnson-Laird, P. N., & Wason, P. C. (1970). Insight into a logical relation. *Quarterly Journal of Experimental Psychology*, *22*(1), 49–61. https://doi.org/10.1080/14640747008401901

Jones, C., Hacker, D., Meaden, A., Cormac, I., Irving, C. B., Xia, J., Zhao, S., Shi, C., & Chen, J. (2018). Cognitive behavioural therapy plus standard care versus standard care plus other psychosocial treatments for people with schizophrenia. *Cochrane Database of Systematic Reviews*, *2018*(11). https://doi.org/10.1002/14651858.CD008712.PUB3/MEDIA/CDSR/CD008712/IMAGE_N/NCD008712-CMP-004-13.PNG

Jones, C., Hacker, D., Xia, J., Meaden, A., Irving, C. B., Zhao, S., Chen, J., & Shi, C. (2018). Cognitive behavioural therapy plus standard care versus standard care for people with schizophrenia. *Cochrane Database of Systematic Reviews*, *2018*(12). https://doi.org/10.1002/14651858.CD007964.PUB2/MEDIA/CDSR/CD007964/IMAGE_N/NCD007964-CMP-004-01.PNG

Jones, S. H., Gray, J. A., & Hemsley, D. R. (1990). The Kamin blocking effect, incidental learning and psychoticism. *British Journal of Psychology*, *81*(1), 95–109. https://doi.org/10.1111/J.2044-8295.1990.TB02348.X

Jones, S. H., Gray, J. A., & Hemsley, D. R. (1992a). The Kamin blocking effect, incidental learning and schizotypy (a reanalysis). *Personality and Individual Differences*, *13*(1), 57–60. https://doi.org/10.1016/0191-8869(92)90218-E

Jones, S. H., Gray, J. A., & Hemsley, D. R. (1992b). Loss of the Kamin blocking effect in acute but not chronic schizophrenics. *Biological Psychiatry*, *32*(9), 739–755. https://doi.org/10.1016/0006-3223(92)90078-E

Jones, S. H., Hemsley, D., Ball, S., & Serra, A. (1997). Disruption of the Kamin blocking effect in schizophrenia and in normal subjects following amphetamine. *Behavioural Brain Research*, *88*(1), 103–114. https://doi.org/10.1016/S0166-4328(97)02312-7

Jones, S. H., Hemsley, D. R., & Gray, J. A. (1991). Contextual effects on choice reaction time and accuracy in acute and chronic schizophrenics. Impairment in selective attention or in the influence of prior learning? *British Journal of Psychiatry*, *159*(SEPT.), 415–421. https://doi.org/10.1192/BJP.159.3.415

Juárez-Ramos, V., Rubio, J. L., Delpero, C., Mioni, G., Stablum, F., & Gómez-Milán, E. (2014). Jumping to Conclusions bias, BADE and Feedback Sensitivity in schizophrenia and schizotypy. *Consciousness and Cognition*, *26*, 133–144. https://doi.org/10.1016/J.CONCOG.2014.03.006

Kahn, R. S., & Keefe, R. S. E. (2013). Schizophrenia Is a Cognitive Illness. *JAMA Psychiatry*, *70*(10), 1107. https://doi.org/10.1001/jamapsychiatry.2013.155

Kamin, L. J. (1967). Attention-like processes in classical conditioning. In *Symp. on Aversive Motivation*.

Kao, S. F., & Wasserman, E. A. (1993). Assessment of an Information Integration Account of Contingency Judgment With Examination of Subjective Cell Importance and Method of Information Presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(6), 1363–1386. https://doi.org/10.1037/0278-7393.19.6.1363

Kapur, S. (2003). Psychosis as a State of Aberrant Salience: A Framework Linking Biology, Phenomenology, and Pharmacology in Schizophrenia. *American Journal of Psychiatry*, *160*(1), 13–23. https://doi.org/10.1176/appi.ajp.160.1.13

Kapur, S., & Remington, G. (2001). Dopamine D2 receptors and their role in atypical antipsychotic action: Still necessary and may even be sufficient. *Biological Psychiatry*, *50*(11), 873–883. https://doi.org/10.1016/S0006-3223(01)01251-3

Kassambara, A. (2020). *ggpubr: "ggplot2" Based Publication Ready Plots*. https://CRAN.R-project.org/package=ggpubr

Keefe, R. S. E., Eesley, C. E., & Poe, M. P. (2005). Defining a cognitive function decrement in schizophrenia. *Biological Psychiatry*, *57*(6), 688–691. https://doi.org/10.1016/J.BIOPSYCH.2005.01.003

Klingberg, S., Wölwer, W., Engel, C., Wittorf, A., Herrlich, J., Meisner, C., Buchkremer, G., & Wiedemann, G. (2011). Negative symptoms of schizophrenia as primary target of cognitive behavioral therapy: Results of the randomized clinical TONES study. In *Schizophrenia Bulletin* (Vol. 37, Issue SUPPL. 2). https://doi.org/10.1093/schbul/sbr073

Korkmaz, Ş. A., & Can, S. S. (2020). The jumping to conclusions bias associated with symptoms in schizophrenia: which factors influence this bias? *Journal of Cognitive Psychology*, *32*(4), 449–459. https://doi.org/10.1080/20445911.2020.1764570

Kornreich, C., Delle-Vigne, D., Brevers, D., Tecco, J., Campanella, S., Noël, X., Verbanck, P., & Ermer, E. (2017). Conditional reasoning in schizophrenic patients. *Evolutionary Psychology*, *15*(3), 1–8. https://doi.org/10.1177/1474704917721713/ASSET/IMAGES/LARGE/10.1177_1474704917721713-FIG2.JPEG

Kotov, R., Waszczuk, M. A., Krueger, R. F., Forbes, M. K., Watson, D., Clark, L. A., Achenbach, T. M., Althoff, R. R., Ivanova, M. Y., Michael Bagby, R., Brown, T. A., Carpenter, W. T., Caspi, A., Moffitt, T. E., Eaton, N. R., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., … Zimmerman, M. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, *126*(4), 454–477. https://doi.org/10.1037/ABN0000258

Kraepelin, E. (1919). *Dementia praecox and paraphrenia*. Chicago Medical Book Co.

Kruschke, J. K. (2008). Bayesian approaches to associative learning: From passive to active learning. *Learning & Behavior*, *36*(3), 210–226. https://doi.org/10.3758/LB.36.3.210

Kruschke, J. K., & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin and Review*, *7*(4), 636–645. https://doi.org/10.3758/BF03213001

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Kwapil, T. R., & Barrantes-Vidal, N. (2015). Schizotypy: Looking Back and Moving Forward. *Schizophrenia Bulletin*, *41*(2), 366–373. https://doi.org/10.1093/schbul/sbu186

Laing, R. D. (1960). *The Divided Self: An existential study in sanity and madness.* Penguin.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, *32*(2), 311–328. https://doi.org/10.1037/0022-3514.32.2.311

Lasserre, V. (1991). Determination of optimal designs using linear models in crossover trials. *Statistics in Medicine*, *10*(6), 909–924. https://doi.org/10.1002/sim.4780100612

le Pelley, M. E. (2004). The role of associative history in models of associative learning: A selective review and a hybrid model. *Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology*, *57*(3), 193–243. https://doi.org/10.1080/02724990344000141

le Pelley, M. E., & McLaren, I. P. L. (2004). Associative History Affects the Associative Change Undergone by Both Presented and Absent Cues in Human Causal Learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*(1), 67–73. https://doi.org/10.1037/0097-7403.30.1.67

le Pelley, M. E., Schmidt-Hansen, M., Harris, N. J., Lunter, C. M., & Morris, C. S. (2010). Disentangling the attentional deficit in schizophrenia: Pointers from schizotypy. *Psychiatry Research*, *176*(2–3), 143–149. https://doi.org/10.1016/J.PSYCHRES.2009.03.027

le Pelley, M. E., Vadillo, M., & Luque, D. (2013). Learned predictiveness influences rapid attentional capture: Evidence from the dot probe task. *Journal of Experimental Psychology: Learning Memory and Cognition*, *39*(6), 1888–1900. https://doi.org/10.1037/A0033700

Le, T. P., Fedechko, T. L., Cohen, A. S., Allred, S., Pham, C., Lewis, S., & Barkus, E. (2019). Stress and cognitive biases in schizotypy: A two-site study of bias against disconfirmatory evidence and jumping to conclusions. *European Psychiatry*, *62*, 20–27. https://doi.org/10.1016/J.EURPSY.2019.08.006

Lee, H. J., Cougle, J. R., & Telch, M. J. (2005). Thought–action fusion and its relationship to schizotypy and OCD symptoms. *Behaviour Research and Therapy*, *43*(1), 29–41. https://doi.org/10.1016/J.BRAT.2003.11.002

Lemaitre, A. L., Luyat, M., & Lafargue, G. (2016). Individuals with pronounced schizotypal traits are particularly successful in tickling themselves. *Consciousness and Cognition*, *41*, 64–71. https://doi.org/10.1016/J.CONCOG.2016.02.005

Lenzenweger, M. F. (2010). *Schizotypy and schizophrenia: The view from experimental psychopathology*. Guilford Press.

Lenzenweger, M. F. (2018a). Schizotypy, schizotypic psychopathology and schizophrenia. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, *17*(1), 25–26. https://doi.org/10.1002/wps.20479

Lenzenweger, M. F. (2018b). Schizotypy, Schizotypic Psychopathology, and Schizophrenia: Hearing Echoes, Leveraging Prior Advances, and Probing New Angles. *Schizophrenia Bulletin*, *44*(suppl_2), S564–S569. https://doi.org/10.1093/SCHBUL/SBY083

Lenzenweger, M. F., Bennett, M. E., & Lilenfeld, L. R. (1997). The referential thinking scale as a measure of schizotypy: Scale development and initial construct validation. *Psychological Assessment*, *9*(4), 452–463. https://doi.org/10.1037/1040-3590.9.4.452

Lenzenweger, M. F., & O'Driscoll, G. A. (2006). Smooth pursuit eye movement and schizotypy in the community. *Journal of Abnormal Psychology*, *115*(4), 779–786. https://doi.org/10.1037/0021-843X.115.4.779

Levy, D. L., Sereno, A. B., Gooding, D. C., & O'Driscoll, G. A. (2010). Eye tracking dysfunction in schizophrenia: Characterization and pathophysiology. *Current Topics in Behavioral Neurosciences*, *4*, 311–347. https://doi.org/10.1007/7854_2010_60

Li, J., Schiller, D., Schoenbaum, G., Phelps, E. A., & Daw, N. D. (2011). Differential roles of human striatum and amygdala in associative learning. *Nature Neuroscience*, *14*(10), 1250–1252. https://doi.org/10.1038/nn.2904

Limongi, R., Bohaterewicz, B., Nowicka, M., Plewka, A., & Friston, K. J. (2018). Knowing when to stop: Aberrant precision and evidence accumulation in schizophrenia. *Schizophrenia Research*, *197*, 386–391. https://doi.org/10.1016/J.SCHRES.2017.12.018

Linscott, R. J., & van Os, J. (2010). Systematic Reviews of Categorical Versus Continuum Models in Psychosis: Evidence for Discontinuous Subpopulations Underlying a Psychometric Continuum. Implications for DSM-V, DSM-VI, and DSM-VII. *Annual Review of Clinical Psychology*, *6*, 391–419. https://doi.org/10.1146/ANNUREV.CLINPSY.032408.153506

Lo, S., & Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171. https://doi.org/10.3389/FPSYG.2015.01171/BIBTEX

Lubow, R. E. (1973). Latent inhibition. *Psychological Bulletin*, *79*(6), 398–407. https://doi.org/10.1037/H0034425

Lubow, R. E., & Moore, A. U. (1959). Latent inhibition: the effect of nonreinforced pre-exposure to the conditional stimulus. *Journal of Comparative and Physiological Psychology*, *52*(4), 415–419. https://doi.org/10.1037/H0046700

Ludvig, E. A., Sutton, R. S., & Kehoe, E. J. (2012). Evaluating the TD model of classical conditioning. *Learning and Behavior*, *40*(3), 305–319. https://doi.org/10.3758/S13420-012-0082-6/FIGURES/7

Luzardo, A., Alonso, E., & Mondragón, E. (2017). A Rescorla-Wagner drift-diffusion model of conditioning and timing. *PLOS Computational Biology*, *13*(11), e1005796. https://doi.org/10.1371/JOURNAL.PCBI.1005796

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*(4), 276–298. https://doi.org/10.1037/h0076778

Maes, E., Boddez, Y., Alfei, J. M., Krypotos, A. M., D'Hooge, R., de Houwer, J., & Beckers, T. (2016). The elusive nature of the blocking effect: 15 failures to replicate. *Journal of Experimental Psychology: General*, *145*(9), e49–e71. https://doi.org/10.1037/XGE0000200

Maia, S., Lefèvre, F., & Jozefowiez, J. (2018). Psychophysics of associative learning: Quantitative properties of subjective contingency. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*(1), 67–81. https://doi.org/10.1037/xan0000153

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Makowski, D., Ben-Shachar, M. S., Patil, I., & Lüdecke, D. (2021). Automated Results Reporting as a Practical Tool to Improve Reproducibility and Methodological Best Practices Adoption. *CRAN*. https://github.com/easystats/report

Mason, O., & Claridge, G. (2006). The Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE): Further description and extended norms. *Schizophrenia Research*, *82*(2–3), 203–211. https://doi.org/10.1016/J.SCHRES.2005.12.845

Mason, O., Claridge, G., & Jackson, M. (1995). New scales for the assessment of schizotypy. *Personality and Individual Differences*, *18*(1), 7–13. https://doi.org/10.1016/0191-8869(94)00132-C

Mason, O. J. (2014). The Duality of Schizotypy: Is it Both Dimensional and Categorical? *Frontiers in Psychiatry*, *5*(SEP). https://doi.org/10.3389/FPSYT.2014.00134

Mason, O., Linney, Y., & Claridge, G. (2005). Short scales for measuring schizotypy. *Schizophrenia Research*, *78*(2–3), 293–296. https://doi.org/10.1016/J.SCHRES.2005.06.020

Mathys, C. D., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A Bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience*, *0*(MAY), 9. https://doi.org/10.3389/FNHUM.2011.00039/BIBTEX

Mathys, C. D., Lomakina, E. I., Daunizeau, J., Iglesias, S., Brodersen, K. H., Friston, K. J., & Stephan, K. E. (2014). Uncertainty in perception and the Hierarchical Gaussian filter. *Frontiers in Human Neuroscience*, *8*, 825. https://doi.org/10.3389/FNHUM.2014.00825/BIBTEX

Matute, H. (1996). Illusion of Control: Detecting Response-Outcome Independence in Analytic but Not in Naturalistic Conditions. *Psychological Science*, *7*(5), 289–293. https://doi.org/10.1111/J.1467-9280.1996.TB00376.X

Matute, H., Blanco, F., & Díaz-Lago, M. (2019). Learning Mechanisms Underlying Accurate and Biased Contingency Judgments. *Journal of Experimental Psychology: Animal Learning and Cognition*. https://doi.org/10.1037/xan0000222

Matute, H., Yarritu, I., & Vadillo, M. A. (2011). Illusions of causality at the heart of pseudoscience. *British Journal of Psychology*, *102*(3), 392–405. https://doi.org/10.1348/000712610X532210

McLean, B. F., Mattiske, J. K., & Balzan, R. P. (2016). Association of the Jumping to Conclusions and Evidence Integration Biases With Delusions in Psychosis: A Detailed Meta-analysis. *Schizophrenia Bulletin*, *43*(2), sbw056–sbw056. https://doi.org/10.1093/schbul/sbw056

Meehl, P. E. (1962). Schizotaxia, schizotypy, schizophrenia. *American Psychologist*, *17*(12), 827–838. https://doi.org/10.1037/H0041029

Meehl, P. E. (1990). Toward an integrated theory of schizotaxia, schizotypy, and schizophrenia. *Journal of Personality Disorders*, *4*(1), 1–99. https://doi.org/10.1521/PEDI.1990.4.1.1

Mercier, P., & Parr, W. (1996). Inter-trial interval, stimulus duration and number of trials in contingency judgments. *British Journal of Psychology*, *87*(4), 549–566. https://doi.org/10.1111/J.2044-8295.1996.TB02608.X

Miller, R. R., & Barnet, R. C. (1993). The Role of Time in Elementary Associations. *Current Directions in Psychological Science*, *2*(4), 106–111. https://doi.org/10.1111/1467-8721.EP10772577

Miller, R. R., & Escobar, M. (2002). Associative interference between cues and between outcomes presented together and presented apart: an integration. *Behavioural Processes*, *57*(2–3), 163–185. https://doi.org/10.1016/S0376-6357(02)00012-8

Miller, R. R., & Matzel, L. D. (1988). The Comparator Hypothesis: A Response Rule for The Expression of Associations. *Psychology of Learning and Motivation - Advances in Research and Theory*, *22*(C), 51–92. https://doi.org/10.1016/S0079-7421(08)60038-9

Miller, R. R., & Springer, A. D. (1972a). Induced recovery of memory in rats following electroconvulsive shock. *Physiology & Behavior*, *8*(4), 645–651. https://doi.org/10.1016/0031-9384(72)90089-3

Miller, R. R., & Springer, A. D. (1972b). Recovery from amnesia following transcorneal electroconvulsive shock. *Psychonomic Science*, *28*(1), 7–9. https://doi.org/10.3758/BF03328645

Mirza, M. B., Adams, R. A., Friston, K., & Parr, T. (2019). Introducing a Bayesian model of selective attention based on active inference. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-50138-8

Mohr, C., & Claridge, G. (2015). Schizotypy--Do Not Worry, It Is Not All Worrisome. *Schizophrenia Bulletin*, *41*(suppl 2), S436–S443. https://doi.org/10.1093/schbul/sbu185

Molet, M., Jozefowiez, J., & Miller, R. R. (2010). Integration of spatial relationships and temporal relationships in humans. *Learning & Behavior*, *38*(1), 27–34. https://doi.org/10.3758/LB.38.1.27

Moore, J. W. (2016). What Is the Sense of Agency and Why Does it Matter? *Frontiers in Psychology*, *7*(AUG). https://doi.org/10.3389/FPSYG.2016.01272

Moore, J. W., Dickinson, A., & Fletcher, P. C. (2011). Sense of agency, associative learning, and schizotypy. *Consciousness and Cognition*, *20*(3–2), 792. https://doi.org/10.1016/J.CONCOG.2011.01.002

Moore, J. W., Turner, D. C., Corlett, P. R., Arana, F. S., Morgan, H. L., Absalom, A. R., Adapa, R., de Wit, S., Everitt, J. C., Gardner, J. M., Pigott, J. S., Haggard, P., & Fletcher, P. C. (2011). Ketamine administration in healthy volunteers reproduces aberrant agency experiences associated with schizophrenia. *Cognitive Neuropsychiatry*, *16*(4), 364–381. https://doi.org/10.1080/13546805.2010.546074

Moran, P. M., Al-Uzri, M. M., Watson, J., & Reveley, M. A. (2003). Reduced Kamin blocking in non paranoid schizophrenia: Associations with schizotypy. *Journal of Psychiatric Research*, *37*(2), 155–163. https://doi.org/10.1016/S0022-3956(02)00099-7

Moran, P. M., Rouse, J. L., Cross, B., Corcoran, R., & Schürmann, M. (2012). Kamin Blocking Is Associated with Reduced Medial-Frontal Gyrus Activation: Implications for Prediction Error Abnormality in Schizophrenia. *PLOS ONE*, *7*(8), e43905. https://doi.org/10.1371/JOURNAL.PONE.0043905

Moritz, S., Pfuhl, G., Lüdtke, T., Menon, M., Balzan, R. P., & Andreou, C. (2017). A two-stage cognitive theory of the positive symptoms of psychosis. Highlighting the role of lowered decision thresholds. *Journal of Behavior Therapy and Experimental Psychiatry*, *56*, 12–20. https://doi.org/10.1016/J.JBTEP.2016.07.004

Moritz, S., Thompson, S. C., & Andreou, C. (2014). Illusory Control in Schizophrenia. *Journal of Experimental Psychopathology*, *5*(2), 113–122. https://doi.org/10.5127/JEP.036113

Moritz, S., van Quaquebeke, N., & Lincoln, T. M. (2012). Jumping to Conclusions Is Associated with Paranoia but Not General Suspiciousness: A Comparison of Two Versions of the Probabilistic Reasoning Paradigm. *Schizophrenia Research and Treatment*, *2012*, 1–9. https://doi.org/10.1155/2012/384039

Moritz, S., & Woodward, T. S. (2006). A generalized bias against disconfirmatory evidence in schizophrenia. *Psychiatry Research*, *142*(2–3), 157–165. https://doi.org/10.1016/J.PSYCHRES.2005.08.016

Morris, R., Griffiths, O., le Pelley, M. E., & Weickert, T. W. (2013). Attention to Irrelevant Cues Is Related to Positive Symptoms in Schizophrenia. *Schizophrenia Bulletin*, *39*(3), 575–582. https://doi.org/10.1093/schbul/sbr192

Morris, R. W., Cyrzon, C., Green, M. J., le Pelley, M. E., & Balleine, B. W. (2018). Impairments in action–outcome learning in schizophrenia. *Translational Psychiatry*, *8*(1), 54. https://doi.org/10.1038/s41398-018-0103-0

Morris, R. W., Quail, S., Griffiths, K. R., Green, M. J., & Balleine, B. W. (2015a). Corticostriatal control of goal-directed action is impaired in schizophrenia. *Biological Psychiatry*, *77*(2), 187–195. https://doi.org/10.1016/j.biopsych.2014.06.005

Morris, R. W., Quail, S., Griffiths, K. R., Green, M. J., & Balleine, B. W. (2015b). Corticostriatal Control of Goal-Directed Action Is Impaired in Schizophrenia. *Biological Psychiatry*, *77*(2), 187–195. https://doi.org/10.1016/J.BIOPSYCH.2014.06.005

Moskowitz, A., & Heim, G. (2011). Eugen Bleuler's Dementia Praecox or the Group of Schizophrenias (1911): A Centenary Appreciation and Reconsideration. *Schizophrenia Bulletin*, *37*(3), 471. https://doi.org/10.1093/SCHBUL/SBR016

Moustafa, A. A., Kéri, S., Somlai, Z., Balsdon, T., Frydecka, D., Misiak, B., & White, C. (2015). Drift diffusion model of reward and punishment learning in schizophrenia: Modeling and experimental data. *Behavioural Brain Research*, *291*, 147–154. https://doi.org/10.1016/J.BBR.2015.05.024

Moutoussis, M., Bentall, R. P., El-Deredy, W., & Dayan, P. (2011). Bayesian modelling of Jumping-to-Conclusions bias in delusional patients. *Cognitive Neuropsychiatry*, *16*(5), 422–447. https://doi.org/10.1080/13546805.2010.548678

Msetfi, R. M., Brosnan, L., & Cavus, H. A. (2016). Enhanced attention to context: An intervention that increases perceived control in mild depression. *The Quarterly Journal of Experimental Psychology*, *69*(6), 1073–1081. https://doi.org/10.1080/17470218.2016.1138134

Msetfi, R. M., Byrom, N., & Murphy, R. A. (2017). To neglect or integrate contingency information from outside the task frame, that is the question! Effects of depressed mood. *Acta Psychologica*, *178*, 1–11. https://doi.org/10.1016/J.ACTPSY.2017.05.003

Msetfi, R. M., Kornbrot, D. E., Matute, H., & Murphy, R. A. (2015). The relationship between mood state and perceived control in contingency learning: effects of individualist and collectivist values. *Frontiers in Psychology*, *6*, 1430. https://doi.org/10.3389/FPSYG.2015.01430/BIBTEX

Msetfi, R. M., Kumar, P., Harmer, C. J., & Murphy, R. A. (2016). SSRI enhances sensitivity to background outcomes and modulates response rates: A randomized double blind study of

instrumental action and depression. *Neurobiology of Learning and Memory*, *131*, 76–82. https://doi.org/10.1016/J.NLM.2016.03.004

Msetfi, R. M., Murphy, R. A., & Kornbrot, D. E. (2012). Dysphoric mood states are related to sensitivity to temporal changes in contingency. *Frontiers in Psychology*, *3*(SEP), 368. https://doi.org/10.3389/FPSYG.2012.00368/BIBTEX

Msetfi, R. M., Murphy, R. A., & Simpson, J. (2007). Depressive realism and the effect of intertrial interval on judgements of zero, positive, and negative contingencies. *The Quarterly Journal of Experimental Psychology*, *60*(3), 461–481. https://doi.org/10.1080/17470210601002595

Msetfi, R. M., Simpson, J., Murphy, R. A., & Kornbrot, D. E. (2005). Depressive realism and outcome density bias in contingency judgments: The effect of the context and intertriai interval. In *Journal of Experimental Psychology: General* (Vol. 134, Issue 1, pp. 10–22). https://doi.org/10.1037/0096-3445.134.1.10

Msetfi, R. M., Wade, C., & Murphy, R. A. (2013). Context and Time in Causal Learning: Contingency and Mood Dependent Effects. *PLoS ONE*, *8*(5), e64063–e64063. https://doi.org/10.1371/journal.pone.0064063

Murphy, R. A., Baker, A. G., & Fouquet, N. (2001). Relative validity effects with either one or two more valid cues in Pavlovian and instrumental conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*(1), 59–67. https://doi.org/10.1037/0097-7403.27.1.59

Murphy, R. A., Byrom, N., & Msetfi, R. M. (2017). The problem with explaining symptoms: The origin of biases in causal processing. *European Journal for Person Centered Healthcare*, *5*(3), 344. https://doi.org/10.5750/EJPCH.V5I3.1318

Murphy, R. A., Mondragón, E., & Murphy, V. A. (2008). Rule learning by rats. *Science*, *319*(5871), 1849–1851. https://doi.org/10.1126/SCIENCE.1151564/SUPPL_FILE/MURPHY.SOM.PDF

Murphy, R. A., Mondragón, E., Murphy, V. A., & Fouquet, N. (2004). Serial order of conditional stimuli as a discriminative cue for Pavlovian conditioning. *Behavioural Processes*, *67*(2), 303–311. https://doi.org/10.1016/J.BEPROC.2004.05.003

Murphy, R. A., Schmeer, S., Vallée-Tourangeau, F., Mondragón, E., & Hilton, D. (2011). Making the illusory correlation effect appear and then disappear: The effects of increased learning. *Quarterly Journal of Experimental Psychology*, *64*(1), 24–40. https://doi.org/10.1080/17470218.2010.493615

Murphy, R. A., Witnauer, J. E., Castiello, S., Tsvetkov, A., Li, A., Alcaide, D. M., & Miller, R. R. (2021). More frequent, shorter trials enhance acquisition in a training session: There is a free lunch! *Journal of Experimental Psychology: General*. https://doi.org/10.1037/XGE0000910

Murray, G. K., Cheng, F., Clark, L., Barnett, J. H., Blackwell, A. D., Fletcher, P. C., Robbins, T. W., Bullmore, E. T., & Jones, P. B. (2008). Reinforcement and Reversal Learning in First-Episode Psychosis. *Schizophrenia Bulletin*, *34*(5), 848. https://doi.org/10.1093/SCHBUL/SBN078

Murray, G. K., Corlett, P. R., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., Jones, P. B., Bullmore, E. T., Robbins, T. W., & Fletcher, P. C. (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry*, *13*(3), 239. https://doi.org/10.1038/SJ.MP.4002058

Novick, D., Haro, J. M., Suarez, D., Vieta, E., & Naber, D. (2009). Recovery in the outpatient setting: 36-month results from the Schizophrenia Outpatients Health Outcomes (SOHO) study. *Schizophrenia Research*, *108*(1–3), 223–230. https://doi.org/10.1016/J.SCHRES.2008.11.007

Oades, R. D., & Eggers, C. (1992). Paranoid schizophrenics may not use irrelevant signals: Conditioned blocking and urinary measures of dopamine activity. *Schizophrenia Research*, *6*(2), 162–163. https://doi.org/10.1016/0920-9964(92)90256-5

Oades, R. D., Rao, M. L., Bender, S., Sartory, G., & Müller, B. W. (2000). Neuropsychological and conditioned blocking performance in patients with schizophrenia: assessment of the contribution of neuroleptic dose, serum levels and dopamine D2-receptor occupancy. *Behavioural Pharmacology*, *11*(3–4), 317–330. http://www.ncbi.nlm.nih.gov/pubmed/11103886

Oades, R. D., Zimmermann, B., & Eggers, C. (1996). Conditioned Blocking in patients with paranoid, non-paranoid psychosis or obsessive compulsive disorder: Associations with symptoms, personality and monoamine metabolism. *Journal of Psychiatric Research*, *30*(5), 369–390. https://doi.org/10.1016/0022-3956(96)00006-4

Oestreich, L. K. L., Mifsud, N. G., Ford, J. M., Roach, B. J., Mathalon, D. H., & Whitford, T. J. (2015a). Cortical Suppression to Delayed Self-Initiated Auditory Stimuli in Schizotypy. *Clinical EEG and Neuroscience*, *47*(1), 3–10. https://doi.org/10.1177/1550059415581708

Oestreich, L. K. L., Mifsud, N. G., Ford, J. M., Roach, B. J., Mathalon, D. H., & Whitford, T. J. (2015b). Subnormal Sensory Attenuation to Self-Generated Speech in Schizotypy: Electrophysiological Evidence for a 'Continuum of Psychosis.' *International Journal of Psychophysiology*, *97*(2), 131. https://doi.org/10.1016/J.IJPSYCHO.2015.05.014

Ortiz-Orendain, J., Castiello-de Obeso, S., Colunga-Lozano, L. E., Hu, Y., Maayan, N., & Adams, C. E. (2017). Antipsychotic combinations for schizophrenia. *Cochrane Database of Systematic Reviews*, *2017*(6). https://doi.org/10.1002/14651858.CD009005.pub2

Pareés, I., Brown, H., Nuruki, A., Adams, R. A., Davare, M., Bhatia, K. P., Friston, K., & Edwards, M. J. (2014). Loss of sensory attenuation in patients with functional (psychogenic) movement disorders. *Brain*, *137*(11), 2916–2921. https://doi.org/10.1093/BRAIN/AWU237

Patel, S. J., Khan, S., M, S., & Hamid, P. (2020). The Association Between Cannabis Use and Schizophrenia: Causative or Curative? A Systematic Review. *Cureus*, *12*(7). https://doi.org/10.7759/CUREUS.9309

Pearce, J. M. (1987). A Model for Stimulus Generalization in Pavlovian Conditioning. *Psychological Review*, *94*(1), 61–73. https://doi.org/10.1037/0033-295X.94.1.61

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*(6), 532–552. https://doi.org/10.1037/0033-295X.87.6.532

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/S13428-018-01193-Y/FIGURES/3

Perales, J. C., Catena, A., Gonzalez, J. A., & Shanks, D. R. (2005). Dissociation between judgments and outcome-expectancy measures in covariation learning: A signal detection theory approach. *Journal of Experimental Psychology: Learning Memory and Cognition*, *31*(5), 1105–1120. https://doi.org/10.1037/0278-7393.31.5.1105

Perez, O. D., & Dickinson, A. (2020). A theory of actions and habits: The interaction of rate correlation and contiguity systems in free-operant behavior. *Psychological Review*, *127*(6), 945–971. https://doi.org/10.1037/REV0000201

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–1045. https://doi.org/10.1038/nature05051

Power, W. T. (1973). *Behavior: The control of perception.* Aldine. https://psycnet.apa.org/record/1974-10192-000

Powers, W. T. (1973). Feedback: Beyond Behaviorism. *Science*, *179*(4071), 351–356. https://doi.org/10.1126/SCIENCE.179.4071.351

Pulcu, E., & Browning, M. (2017). Affective bias as a rational response to the statistics of rewards and punishments. *ELife*, *6*. https://doi.org/10.7554/eLife.27879

Pulcu, E., & Browning, M. (2019). The Misestimation of Uncertainty in Affective Disorders. *Trends in Cognitive Sciences*, *23*(10), 865–875. https://doi.org/10.1016/J.TICS.2019.07.007

R Core Team. (2021). *R A Language and Environment for Statistical Computing*. httpswww.R-project.org

Raballo, A., & Parnas, J. (2011). The Silent Side of the Spectrum: Schizotypy and the Schizotaxic Self. *Schizophrenia Bulletin*, *37*(5), 1017. https://doi.org/10.1093/SCHBUL/SBQ008

Randell, J., Ranjith-Kumar, A. C., Gupta, P., & Reed, P. (2009). Effect of schizotypy on responding maintained by free-operant schedules of reinforcement. *Personality and Individual Differences*, *47*(7), 783–788. https://doi.org/10.1016/J.PAID.2009.06.023

Randell, J., Searle, R., & Reed, P. (2012). Influence of schizotypy on responding and contingency awareness on free-operant schedules of reinforcement. *Learning and Individual Differences*, *22*(3), 425–428. https://doi.org/10.1016/J.LINDIF.2012.01.006

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience 1999 2:1*, *2*(1), 79–87. https://doi.org/10.1038/4580

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. B. and W. F. Prokasy (Ed.), *Classical Conditioning II* (pp. 64–99). Appleton-Century-Crofts.

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20–27. https://doi.org/10.1016/J.TICS.2010.09.003

Ross, R. M., McKay, R., Coltheart, M., & Langdon, R. (2015a). Jumping to conclusions about the beads task? A meta-analysis of delusional ideation and data-gathering. *Schizophrenia Bulletin*, *41*(5), 1183–1191. https://doi.org/10.1093/schbul/sbu187

Ross, R. M., McKay, R., Coltheart, M., & Langdon, R. (2015b). Jumping to Conclusions About the Beads Task? A Meta-analysis of Delusional Ideation and Data-Gathering. *Schizophrenia Bulletin*, *41*(5), 1183–1191. https://doi.org/10.1093/SCHBUL/SBU187

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1: Foundations* (pp. 318–362). MIT Press.

Rumelhart, D. E., McClelland, J. L., & the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. MIT Press.

Sahakyan, L., & Kwapil, T. R. (2018). Moving beyond summary scores: Decomposing free recall performance to understand episodic memory deficits in schizotypy. *Journal of Experimental Psychology. General*, *147*(12), 1919–1930. https://doi.org/10.1037/XGE0000401

Salgado-Pineda, P., Fuentes-Claramonte, P., Spanlang, B., Pomes, A., Landin-Romero, R., Portillo, F., Bosque, C., Franquelo, J. C., Teixido, C., Sarró, S., Salvador, R., & Pomarol-Clotet, E. (2021). Neural correlates of disturbance in the sense of agency in schizophrenia: An fMRI study using the 'enfacement' paradigm. *Schizophrenia Research*. https://doi.org/10.1016/J.SCHRES.2021.06.031

Sass, L. A., & Parnas, J. (2003). Schizophrenia, consciousness, and the self. *Schizophrenia Bulletin*, *29*(3), 427–444. https://doi.org/10.1093/OXFORDJOURNALS.SCHBUL.A007017

Sass, L., Pienkos, E., Skodlar, B., Stanghellini, G., Fuchs, T., Parnas, J., & Jones, N. (2017). EAWE: Examination of Anomalous World Experience. *Psychopathology*, *50*(1), 10–54. https://doi.org/10.1159/000454928

Savastano, H. I., & Miller, R. R. (1998). Time as content in Pavlovian conditioning. *Behavioural Processes*, *44*(2), 147–162. https://doi.org/10.1016/S0376-6357(98)00046-1

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038. https://doi.org/10.1016/J.JML.2019.104038

Schennach, R., Riedel, M., Obermeier, M., Spellmann, I., Musil, R., Jäger, M., Schmauss, M., Laux, G., Pfeiffer, H., Naber, D., Schmidt, L. G., Gaebel, W., Klosterkötter, J., Heuser, I., Maier, W., Lemke, M. R., Rü ther, E., Klingberg, S., Gastpar, M., … Maier, W. (2015). What are residual symptoms in schizophrenia spectrum disorder? Clinical description and 1-year persistence within a naturalistic trial. *Eur Arch Psychiatry Clin Neurosci*, *265*, 107–116. https://doi.org/10.1007/s00406-014-0528-2

Schlagenhauf, F., Huys, Q. J. M., Deserno, L., Rapp, M. A., Beck, A., Heinze, H. J., Dolan, R., & Heinz, A. (2014). Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *NeuroImage*, *89*, 171–180. https://doi.org/10.1016/J.NEUROIMAGE.2013.11.034

Schmajuk, N. A., Lam, Y.-W., & Gray, J. A. (1996). Latent inhibition: a neural network approach. *Journal of Experimental Psychology. Animal Behavior Processes*, *22*(3), 321–349. https://doi.org/10.1037//0097-7403.22.3.321

Schmidt-Hansen, M., Killcross, A. S., & Honey, R. C. (2009). Latent inhibition, learned irrelevance, and schizotypy: Assessing their relationship. *Cognitive Neuropsychiatry*, *14*(1), 11–29. https://doi.org/10.1080/13546800802664539

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461 – 464. https://doi.org/10.1214/aos/1176344136

Sellen, J. L., Oaksford, M., & Gray, N. S. (2005). Schizotypy and Conditional Reasoning. *Schizophrenia Bulletin*, *31*(1), 105–116. https://doi.org/10.1093/SCHBUL/SBI012

Serra, A. M., Jones, S. H., Toone, B., & Gray, J. A. (2001). Impaired associative learning in chronic schizophrenics and their first-degree relatives: A study of latent inhibition and the Kamin blocking effect. *Schizophrenia Research*. https://doi.org/10.1016/S0920-9964(00)00141-9

Seth, A. K., Wanja Wiese, C., Gutenberg-, J., Metzinger, T., & Windt, J. M. (2014). The Cybernetic Bayesian Brain. *Open MIND*, *35*, 1–24. https://doi.org/10.15502/9783958570108

Shanks, D. R. (1985). Forward and Backward Blocking in Human Contingency Judgement. *The Quarterly Journal of Experimental Psychology Section B*, *37*(1b), 1–21. https://doi.org/10.1080/14640748508402082

Shanks, D. R. (1995). The Psychology of Associative Learning. *The Psychology of Associative Learning*. https://doi.org/10.1017/CBO9780511623288

Shanks, D. R., & Dickinson, A. (1987). Associative Accounts of Causality Judgment. In G. H. Bower (Ed.), *Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 21, Issue C, pp. 229–261). Academic Press. https://doi.org/10.1016/S0079-7421(08)60030-4

Shannon, C. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, *27*(3), 379–423.

Shiferaw, B. A., Downey, L. A., Westlake, J., Stevens, B., Rajaratnam, S. M. W., Berlowitz, D. J., Swann, P., & Howard, M. E. (2018). Stationary gaze entropy predicts lane departure events in sleep-deprived drivers. *Scientific Reports*, *8*(1), 2220. https://doi.org/10.1038/s41598-018-20588-7

Shiferaw, B., Downey, L., & Crewther, D. (2019). A review of gaze entropy as a measure of visual scanning efficiency. In *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2018.12.007

Silverman, J. (1967). Shamans and Acute Schizophrenia. *American Anthropologist*, *69*(1), 21–31. https://doi.org/10.1525/aa.1967.69.1.02a00030

Skinner, B. F. (1965). *Science and Human Behavior*. Macmillan.

Sosa, R., & Alcalá, E. (2022). The nervous system as a solution for implementing closed negative feedback control loops. *Journal of the Experimental Analysis of Behavior*. https://doi.org/10.1002/JEAB.736

Stan Development Team. (2021). *RStan: the R interface to Stan*. https://mc-stan.org/

Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., & Corlett, P. R. (2018). The Predictive Coding Account of Psychosis. *Biological Psychiatry*, *84*(9), 634–643. https://doi.org/10.1016/J.BIOPSYCH.2018.05.015

Stout, S. C., & Miller, R. R. (2007). Sometimes-Competing Retrieval (SOCR): A Formalization of the Comparator Hypothesis. *Psychological Review*, *114*(3), 759–783. https://doi.org/10.1037/0033-295X.114.3.759

Strube, W., Cimpianu, C. L., Ulbrich, M., Öztürk, Ö. F., Schneider-Axmann, T., Falkai, P., Marshall, L., Bestmann, S., & Hasan, A. (2022). Unstable Belief Formation and Slowed Decision-making: Evidence That the Jumping-to-Conclusions Bias in Schizophrenia Is Not Linked to Impulsive

Decision-making. *Schizophrenia Bulletin*, *48*(2), 347. https://doi.org/10.1093/SCHBUL/SBAB108

Sutton, R. S., & Barto, A. (2018). *Reinforcement Learning: An Introduction* (A. Barto, Ed.; Second). The MIT Press.

Swerdlow, N. R., Braff, D. L., Hartston, H., Perry, W., & Geyer, M. A. (1996). Latent inhibition in schizophrenia. *Schizophrenia Research*, *20*(1–2), 91–103. https://doi.org/10.1016/0920-9964(95)00097-6

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208. https://doi.org/10.1037/H0061626

Treviño, M. (2016). Associative learning through acquired salience. *Frontiers in Behavioral Neuroscience*, *9*, 353. https://doi.org/10.3389/FNBEH.2015.00353/BIBTEX

Treviño, M., Castiello, S., Arias-Carrion, O., de La Torre-Valdovinos, B., & Coss Y Leon, R. M. (2021). Isomorphic decisional biases across perceptual tasks. *PLoS ONE*, *16*(1 January). https://doi.org/10.1371/journal.pone.0245890

Treviño, M., Castiello, S., Torre-Valdovinos, B. D. la, Carrasco, P. O., & León, R. M.-C. y. (2021). Multimetric behavioral profile and psychological traits. *PsyArXiv*. https://doi.org/10.31234/OSF.IO/PT8GB

Treviño, M., Fregoso, E., Sahagún, C., & Lezama, E. (2018). An Automated Water Task to Test Visual Discrimination Performance, Adaptive Strategies and Stereotyped Choices in Freely Moving Mice. *Frontiers in Behavioral Neuroscience*, *12*, 251. https://doi.org/10.3389/fnbeh.2018.00251

Treviño, M., Medina-Coss y León, R., & Haro, B. (2020). Adaptive Choice Biases in Mice and Humans. *Frontiers in Behavioral Neuroscience*, *14*, 99. https://doi.org/10.3389/FNBEH.2020.00099/BIBTEX

Tschacher, W., & Kupper, Z. (2006). Perception of Causality in Schizophrenia Spectrum Disorder. *Schizophrenia Bulletin*, *32*(suppl_1), S106–S112. https://doi.org/10.1093/SCHBUL/SBL018

Urcelay, G. P., Wheeler, D. S., & Miller, R. R. (2009). Spacing extinction trials alleviates renewal and spontaneous recovery. *Learning & Behavior*, *37*(1), 60–73. https://doi.org/10.3758/LB.37.1.60

Vaitl, D., Lipp, O., Bauer, U., Schüler, G., Stark, R., Zimmermann, M., & Kirsch, P. (2002). Latent inhibition and schizophrenia: Pavlovian conditioning of autonomic responses. *Schizophrenia Research*, *55*(1–2), 147–158. https://doi.org/10.1016/S0920-9964(01)00250-X

Vallée-Tourangeau, F., Baker, A. G., & Mercier, P. (1994). Discounting in Causality and Covariation Judgements. *Quarterly Journal of Experimental Psychology*, *47*(2), 151–171. https://doi.org/10.1080/14640749408401354

Vallée-Tourangeau, F., Murphy, R. A., & Baker, A. G. (2005a). Contiguity and the outcome density bias in action-outcome contingency judgements. *Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology*, *58*(2), 177–192. https://doi.org/10.1080/02724990444000104

Vallée-Tourangeau, F., Murphy, R. A., & Baker, A. G. (2005b). Contiguity and the outcome density bias in action-outcome contingency judgements. *The Quarterly Journal of Experimental*

Psychology. B, Comparative and Physiological Psychology, *58*(2), 177–192. https://doi.org/10.1080/02724990444000104

Vallée-Tourangeau, F., Murphy, R. A., Drew, S., & Baker, A. G. (1998). Judging the Importance of Constant and Variable Candidate Causes: A Test of the Power PC Theory. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *51*(1), 65–84. https://doi.org/10.1080/713755745

van Os, J., Linscott, R. J., Myin-Germeys, I., Delespaul, P., & Krabbendam, L. (2009). A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness–persistence–impairment model of psychotic disorder. *Psychological Medicine*, *39*(2), 179–195. https://doi.org/10.1017/S0033291708003814

Vinckier, F., Gaillard, R., Palminteri, S., Rigoux, L., Salvador, A., Fornito, A., Adapa, R., Krebs, M. O., Pessiglione, M., & Fletcher, P. C. (2016). Confidence and psychosis: a neuro-computational account of contingency learning disruption by NMDA blockade. *Molecular Psychiatry*, *21*(7), 946. https://doi.org/10.1038/MP.2015.73

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods 2007 39:4*, *39*(4), 767–775. https://doi.org/10.3758/BF03192967

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology*, *52*(1), 1–9. https://doi.org/10.1016/J.JMP.2007.09.005

Wagner, A. R., Logan, F. A., & Haberlandt, K. (1968). Stimulus Selection in Animal Discrimination Learning. *Journal of Experimental Psychology*, *76*(2 PART 1), 171–180. https://doi.org/10.1037/H0025414

Waltz, J. A., Demro, C., Schiffman, J., Thompson, E., Kline, E., Reeves, G., Xu, Z., & Gold, J. (2015). Reinforcement learning performance and risk for psychosis in youth. *The Journal of Nervous and Mental Disease*, *203*(12), 919. https://doi.org/10.1097/NMD.0000000000000420

Wason, P. C. (1968). Reasoning about a Rule. *Quarterly Journal of Experimental Psychology*, *20*(3), 273–281. https://doi.org/10.1080/14640746808400161

Wasserman, E. A., Elek, S. M., Chatlosh, D. L., & Baker, A. G. (1993). Rating causal relations: Role of probability in judgments of response^outcome contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(1), 174–188. https://doi.org/10.1037/0278-7393.19.1.174

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review*, *72*(2), 89–104. https://doi.org/10.1037/H0021797

White, P. A. (2004). Causal judgment from contingency information: A systematic test of the pCI rule. *Memory & Cognition*, *32*(3), 353–368. https://doi.org/10.3758/BF03195830

Whitford, T. J., Mitchell, A. M., & Mannion, D. J. (2017a). The ability to tickle oneself is associated with level of psychometric schizotypy in non-clinical individuals. *Consciousness and Cognition*, *52*, 93–103. https://doi.org/10.1016/J.CONCOG.2017.04.017

Whitford, T. J., Mitchell, A. M., & Mannion, D. J. (2017b). The ability to tickle oneself is associated with level of psychometric schizotypy in non-clinical individuals. *Consciousness and Cognition*, *52*, 93–103. https://doi.org/10.1016/J.CONCOG.2017.04.017

Wickham, H. (2007). Reshaping Data with the reshape Package. *Journal of Statistical Software*, *21*(12), 1–20. http://www.jstatsoft.org/v21/i12/

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H., François, R., Henry, L., & Müller, K. (2022). *dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr

Wickham H, Hester, J., Chang, W., & Bryan, J. (2021). devtools: Tools to Make Developing R Packages Easier. In *CRAN* (R package version 2.4.3,). R package version 2.4.3.

Witnauer, J. E., Castiello, S., Fung, E., Pitliya, R. J., Murphy, R. A., & Miller, R. R. (2022). Determinants of extinction in a streamed trial procedure. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/17470218221110827

Woodward, T. S., Buchy, L., Moritz, S., & Liotti, M. (2007). A Bias Against Disconfirmatory Evidence Is Associated With Delusion Proneness in a Nonclinical Sample. *Schizophrenia Bulletin*, *33*(4), 1023–1028. https://doi.org/10.1093/SCHBUL/SBM013

Young, A. M. J., Kumari, V., Mehrotra, R., Hemsley, D. R., Andrew, C., Sharma, T., Williams, S. C. R., & Gray, J. A. (2005). Disruption of learned irrelevance in acute schizophrenia in a novel continuous within-subject paradigm suitable for fMRI. *Behavioural Brain Research*, *156*(2), 277–288. https://doi.org/10.1016/J.BBR.2004.05.034

Zhang, Z., & Mai, Y. (2022). WebPower: Basic and Advanced Statistical Power Analysis. In *CRAN*. https://CRAN.R-project.org/package=WebPower

Zhang, Z., & Wang, L. (2017). *Advanced statistics using R*. ISDSA Press.