A NEUROECONOMIC INVESTIGATION OF RISKY DECISION-MAKING AND LOSS IN THE RAT

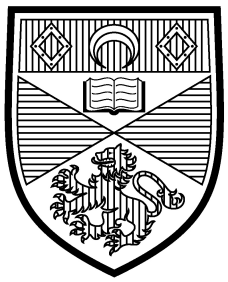
Annamarie Wheeler Huttunen

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

2016

A neuroeconomic investigation of risky decision-making and loss in the rat

Annamarie Wheeler Huttunen



University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of PhD
at the
University of St Andrews

March 2016

Declarations

I, Annamarie Wheeler Huttunen, hereby certify that this thesis, which is approximately 48,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2012 and as a candidate for the degree of PhD in September 2012; the higher study for which this is a record was carried out in the University of St Andrews between 2012 and 2016.

Date: Signature of candidate:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: Signature of supervisor:

Ethics

All procedures were carried out under the Project License numbers 60/4040 and 70/8077, conformed to the United Kingdom Animals (Scientific Procedures) Act (1986) and were approved by the Animal Welfare Ethics Committee of the University of St Andrews.

Date: Signature of candidate:

Date: Signature of supervisor:

Embargoes

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that

may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED COPY
No embargo on print copy

ELECTRONIC COPY
No embargo on electronic copy

Date: Signature of candidate:

Date: Signature of supervisor:

Table of Contents

Declarations ... ii

Dedications ... vii

Abstract ... 1

Chapter 1 ... 2

General Introduction ... 2

Abstract ... 3

Introduction ... 4

***Homo Economicus*: Rational Economics** ... 5

Expected Value ... 5

Expected Utility ... 9

General Axiom of Revealed Preferences (GARP) ... 11

Subjective Expected Utility ... 13

***Homo Sapiens*: Irrational Economics** ... 15

Prospect Theory ... 17

Efficient Markets: Rational Finance ... 22

Risk-Return Models ... 24

Efficient Markets ... 25

Inefficient Markets: 'Noise Traders' ... 28

The Disposition Effect ... 29

Economic Animals ... 32

Reward Learning and Expectation in the Brain ... 35

Value Representation in the Brain ... 40

Encoding Risk in the Brain ... 49

Discussion ... 52

Chapter 2 ... 55

A novel gambling task to capture resource loss in rats ... 55

Abstract ... 56

Introduction ... 57

Methods ... 62

Animals ... 62

Apparatus ... 62

Task Design ... 63

Discrimination task ... 66

Mixed contingencies ... 67

Behavioural measurements ... 67

Data Analysis ... 67

Results ... 69

Discrimination Task ... 69

Baseline Behaviour – Mixed Risk And Reward ... 70

Loss-related behaviour ... 73

Stay-shift behaviour ... 74

Effect of losing on poke duration ... 79

Discussion ... 82

Chapter 3 ... 89

Loss-stay behaviour in rats with more Substantia Nigra pars compacta neurons is mitigated by dopamine antagonism ... 89

Abstract ... 90

Introduction ... 91

Methods **92**
Pharmacological Manipulation 92
Histology 93
Data Analysis 94
Results **96**
Effects of Flupenthixol on behaviour 96
Effects of DA antagonism on poke duration 98
Effects of DA antagonism on choice 99
Discussion **106**
Stay-shift pattern of behaviour 106
Dopamine Antagonism 107

Chapter 4 ..... **112**

Disassociating effects of dopamine neurons, probability and reward in a rat gambling task **112**
Abstract **113**
Introduction **114**
Methods **116**
Animals 116
Apparatus 116
Training 117
Task Outline 117
Behavioural measurements 120
Histology 120
Data analysis 122
Results **123**
Contingency 123
Stay-Shift Strategy 129
Neuron counts and behaviour 130
Discussion **133**

Chapter 5 ..... **140**

Rats exhibit anchoring, loss aversion and the disposition effect in an experimental stock market task **140**
Abstract **141**
Introduction **142**
Methods **145**
Animals 145
Apparatus 146
Training 146
Testing 147
Task Outline 147
Block 1 148
Block 2 148
Pricing 150
Selling 150
Buying 151
Holding 152
Data Analysis **152**
Summary Measures 153
Proportion of Realized Gains & Losses 153
Cox Proportional-Hazard Modelling 154
Results **156**
Descriptive Statistics 156

Manipulation check: Trading payoffs are greater than dividend payoffs 158
The transaction cost of trading within the task 159
Rats exhibit reference-dependent behaviour 161
Rats are more risk-seeking after a loss than after a gain (loss-aversion) 163
Rats perform suboptimally when trading 165
Rats exhibit the disposition effect 172
Discussion 179
Reference Dependence (Anchoring) 180
Loss Aversion 181
Disposition Effect 181
Information salience & content 184
Future directions 186
General Discussion 189
Summary 190
Operationalizing Resource Loss 191
Loss-dependent behaviour 192
Brain and behaviour 196
Conclusions 201
Appendices 202
Appendix 1 202
Appendix 2 203
Appendix 3 206
Appendix 4 207

Dedications

This thesis is dedicated first and foremost to my husband, Julius. Thank you for supporting me when I needed strength, listening when I needed an ear, grounding me when I ventured too deep into the introspective abyss, and for loving me always.

I would like to thank my supervisor, Eric, not only for the incredible opportunity to embark on this scholarly journey, but also for the invaluable knowledge, advice and support that you have given me along the way. I would need to complete several more PhD's under your tutelage to claim even a fraction of your knowledge. Your sincere commitment to my growth as a young researcher has been tremendous.

I would like to thank my secondary supervisor, Verity, for the excellent career advice and guidance throughout my time in St Andrews. I would also like to thank David Tait for being ever willing (and ever patient) to teach me new techniques. I would especially like to thank the technicians and support staff of the St Mary's Animal Unit for their constant hard work and dedication to the animals involved in my studies.

I would also like to thank my mother, Jeanmarie, and my father, Robert. Mom, thank you for always believing in me and for giving me the drive to achieve. Dad, thank you for sharing your insatiable intellectual curiosity with me. You both have taught me to question the world, to challenge myself, to pursue my passions, and to love life. I could never have made it this far without your love and guidance.

I would also like to thank my family and friends. Without the diverse perspectives and deep friendship that each of you have given me, my journey would have lost a unique richness that I truly cherish. To Bo, Pete, Stephen, and Katy, I relish the times that we have shared growing up together, thank you for all of the laughter and the tears we have shared over the years. I would also like to thank Scott for the time-honoured traditions of our friendship and for the countless hours spent in thought-provoking conversation. I would like to thank Vlad for his unique sense of humour and unwavering logic despite his heritage, and Moritz for his relaxed ability to extract true joy from life, again despite his heritage. Finally, I would like to thank all of the friends and family that continue to have a profound influence on my life and whom I love very much despite the oceans between us.

Abstract

Humans exhibit a number of suboptimal behaviours in the wake of a loss. For example, gamblers often ‘chase’ their losses in an attempt to break even. Similarly, investors tend to hold on to losing stocks too long in the hope that the declining share price might make a recovery. However, the neural mechanisms that instantiate such behaviour are poorly understood. I begin the introductory chapter with a basic historical overview of fundamental economic concepts, interleaving intersecting ideas from psychology and neuroscience. This leads to a more in-depth exploration of the notion that loss-related behavioural biases might provide insight into the neural mechanisms that underlie risky choice. From this, I argue that rats represent a viable animal model of risky decision-making for neuroeconomic research. The original research presented in Chapters 2 – 5 pave the way toward advancing our current understanding of loss-related biases in behaviour with rat models of risky decision-making. By employing insight from psychology and economics, I developed two models of rat behaviour that can be used to study the neural substrates of loss valuation. I presented the experimental paradigms in Chapters 2 and 5, while demonstrating novel loss-related correlations between the midbrain dopamine system and observed loss behaviour in Chapters 3 and 4. The results presented in Chapter 5 demonstrate that rats are capable of producing behavioural patterns akin to loss aversion and the disposition effect. This work has also highlighted a number of areas for future research. In Chapter 6, I explore potential theoretical implications of the results discussed in previous chapters. In summary, this thesis uses experimental risky decision-making tasks in rats to advance our current knowledge of the ways in which concepts such as loss aversion critically influence our internal representation of value.

Chapter 1

General Introduction

Abstract

When deciding between two risky options, it is often the case that we change our behaviour in the face of a loss. If the resulting outcome leaves the individual worse off than an alternative, then it is considered suboptimal. Neoclassical economic theories typically label the resulting suboptimal choice as ‘irrational.’ This introductory chapter outlines key topics in understanding ‘rational’ and ‘irrational’ economic choices. Beginning with a basic historical overview of fundamental economic concepts, intersecting ideas from psychology and neuroscience are discussed. Kahneman and Tversky’s (1979) Prospect Theory stands as a turning point, marking the beginning of a truly interdisciplinary pursuit toward understanding the ways in which concepts such as loss aversion and risk aversion critically influence our internal representation of value.

Introduction

While neuroeconomics as a discipline remains only in its infancy, the intersection of the biological sciences with the economic sciences has a longer history than one might expect. In 1828, the Scottish Botanist Robert Brown first observed the natural oscillations of microscopic pollen grains that occurred after the granules had been submerged in water (R. Brown, 1828). The mathematics of this continuous random motion was later developed by Albert Einstein (1905) to become a pillar of physics, but not before French mathematician Louis Bachelier (1900) applied the principle to the stock market. Thus the notion that a 'random walk' could describe the movements of tiny submerged particles of biological matter just as well as the fluctuations of asset prices in the stock market was born.

From pollen grains to neurons, this thesis is written with the specific intention of informing the study of economics with experimental evidence from psychology and neuroscience, and *vice versa*. In an effort to engage the reader in a critical interdisciplinary enterprise, this chapter will begin by offering an introductory historical review of the relevant concepts, models and assumptions from economics and finance. While not exhaustive, this introduction is intended to provide the non-specialist with a general impression of the field and its formalisms, while simultaneously highlighting areas of economic theory that have been (or could be) improved by incorporating a behaviourist tradition. Beginning with conceptual and formal notions of valuation and risk, this section will guide the reader through progressively more comprehensive theories of economic decisions making. Throughout, theoretical prescriptions of rationality

will be juxtaposed with empirical observations of behaviour. This is done with the intention of imparting the reader with a clearer understanding of the fundamental struggle in neuroeconomics: finding the optimal point between the prescriptive power of formal theoretical models and the descriptive accuracy of behavioural models. The chapter will then shift focus to a collection of studies that implicate neurobiological mechanisms in the modulation of the economic parameters and behavioural biases reviewed in the preceding sections. Taken together, these studies constitute the backbone of the newly established neuroeconomics literature and reveal an obvious dearth in their ability to describe 'loss' and 'risk' using current behavioural and neuroscientific paradigms.

Homo Economicus: Rational Economics

The field of economics represents a long tradition of breaking decisions down into quantifiable variables with the goal of objectively representing decision outcomes and mathematically prescribing the best course of action to achieve that outcome. In both psychology and economics, there is an intuitive understanding that what is good in the immediate sense is not necessarily good in the long run. Thus, the first step in terms of modelling an optimal decision becomes defining a *value* function that estimates how good or bad an action will be in the future.

Expected Value

Pascal (1670) was the first to formalize the objective measurement of decision outcomes in terms of expected value (EV), which has formed the basis of normative economic models.

Table 1: Pascal's Wager

	God exists (50%)	God does not exist (50%)	Expected Value of Choice Set
Believe	$\infty \times 0.5$ (Infinite)	$-x \times 0.5$ (Finite)	$(\infty \times 0.5) + (-x \times 0.5) = \infty$
Do not believe	$-\infty \times 0.5$ (Infinite)	$x \times 0.5$ (Finite)	$(-\infty \times 0.5) + (x \times 0.5) = -\infty$

Table 1 displays 'Pascal's Wager,' which examines the expected value, or payoff, of believing in God versus not believing in God. Given any probability that God exists (displayed here as a 50/50 chance), the infinite gain of believing is greater than the infinite loss of not believing.

Pascal theorized that a decision maker would maximize long-term future payoffs by choosing the action that leads to the outcome with the highest expected value, which is based on its current value (positive or negative) and the likelihood of it occurring (probability). A decision-maker can maximize her payoff by simply multiplying these two variables for each alternative within a decision set, and then by choosing the option with the greatest resulting expected value, where:

$$\mathrm{EV} = \mathrm{Value} \times \mathrm{Probability} \qquad \textbf{(1)}$$

Therefore, a payoff of high magnitude that has an extremely low likelihood of occurring in the future (e.g. winning the lottery) would have a lower EV than a second payoff with a smaller magnitude but that will occur with relative certainty (e.g. gaining interest on savings). As simple as it may appear, this formulation has far-reaching implications for economic theory both past and present.
The probability distribution that results from calculating EV can be used to describe a given outcome or a set of outcomes. Mathematically, the distribution's first moment (mean) represents the 'expected value' of the set of outcomes, while its second moment (variance or its square root, standard

deviation) denotes the 'expected risk.' Variance (σ^2) is a measure of a distribution's dispersion and is equal to the mean of the squared deviations from the expected value, formally expressed as:

$$\boldsymbol{\sigma^2} = \frac{\sum_{i=1}^{n}(V_i - EV)^2}{N} \qquad \textbf{(2)}$$

Above, the expected value (*EV*) is subtracted from the observed value (*V*) at time *i*, and the total sum of the squared result from time *i* to *n* is divided by the total number of observations, *N*. Thus, variance is a measure of the relative certainty with which one can say that an outcome will be near the mean, with low variances reflecting a group of numerically similar outcomes. The variance of the outcome distribution is therefore conventionally regarded as an objective measure of a decision's riskiness. Importantly, this definition represents a divergence of the term 'risk' from its more colloquial meaning of the potential for a loss.[1] Instead, economic risk denotes how certain one is about whether the

---

[1] *A considerable muddling of concepts arose as psychologists, neuroscientists and economists began integrating theory and research. A particularly confusing disconnect can be seen between the definition of the term 'risk.' Psychologists' view of risk often hinges upon the magnitude of a potential loss (e.g. loss of job, loss of life). Risk within economics or finance has a much narrower, mathematical definition, and is often conceptually closer to the colloquial notion of 'uncertainty.' While the two definitions do partially overlap, this oversight may underlie several lines of diverging evidence, especially with regard to aberrant risk processing. For example, Attention Deficit Hyperactivity Disorder (ADHD) is often invoked as stereotypical impulsive behaviour, which is defined by poor risk assessment. However, the definition of risk assessment clearly depends on one's definition of risk—an impulsive individual is often understood as someone who acts with a sense of spontaneity and often with disregard to the magnitude of any potential losses that may result. Thus, an economist is more likely to attribute ADHD to abnormal discounting of delays or cost-benefit analyses (first moment) rather than incorrect risk evaluations (second moment), per se. Since pathological conditions are often advantageous in revealing the hidden structure and connectivity of many functions in the brain, it is imperative that the field collaborate in order to remedy such ambiguity.*

outcome will be close to the mean expected value in a given range of possibilities (the known unknowns). If there is a narrow range of possible outcomes in a decision set, then one can be relatively certain that the outcome will lie near the mean (expected value). Alternatively, a decision with outcomes spanning a wide range of expected values would be considered risky, because one is relatively uncertain about whether the actual outcome will be similar to the mean expected value. It is also of note that economists formally distinguish decisions under uncertainty, which indicates a known probability distribution but an unknown outcome, from decisions under 'ambiguity,' which implies that the probability distribution itself is at least partially unknown (the unknown unknowns).

While Pascal's formulation of expected value was able to explain a great deal of behaviour, it failed to describe the circumstance typified by the 'St Petersburg Paradox,' in which a probability distribution had an infinite expected value. The St Petersburg paradox was described as a game in which a coin was flipped, and the player must bet on how many flips were needed before it landed on heads. The game ends when heads is flipped, and the player wins £2 for every toss that occurs before then (i.e. £2 for each tails flip). According to Pascal's formulation, the expected value of this gamble is the sum of all the possible outcomes multiplied by their respective probabilities, or:

$$\boldsymbol{EV = \frac{1}{2} \times 2 + \frac{1}{4} \times 4 + \frac{1}{8} \times 8 + \cdots = 1 + 1 + 1 + \cdots = \infty} \qquad \textbf{(3)}$$

So if a player could play the game enough times, she could bet any finite amount of money and still beat the house. Put differently, a player should pay

an infinite amount of money to play the game. Since this answer obviously did not describe true behaviour, it was clear that expected value alone was not sufficient to describe realistic economic behaviour.

Expected Utility

Figure 1: Marginal Utility

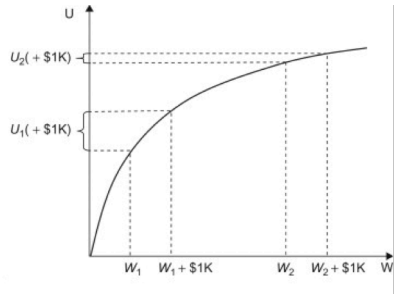


Figure with permissions from (Craig R. Fox & Poldrack, 2014) demonstrates the marginal utility of a $1K gamble experienced by an individual with wealth W_1 compared to an individual with a higher starting wealth, W_2. The marginal increase of the *same objective amount* is experienced differently between the two individuals. The person with W_1 perceives the $1K increase (and any respective decrease) as large relative to the person with W_2.

In 1738 a mathematician named Daniel Bernoulli published a solution to the St Petersburg Paradox, which entailed the transformation of expected value into expected utility based on a person's current wealth (Bernoulli, 1738). Bernoulli observed that expected value did not take into consideration whether or not the player was a pauper or a prince – and that one's state of wealth had implications for (a) the pleasure one gained from a marginal increase in wealth, and therefore (b) one's willingness to take a financial risk. In other words, a pauper would receive more utility from a £100 outcome, which might double his current wealth; whereas a prince would receive marginally less utility from the same £100 increase in wealth. This formulation has become known as 'utility,' (*u*) and is often expressed as a logarithmic[2] transformation of wealth:

[2] *Importantly, any nonlinear transformation that results in a concave utility function is allowed here (some work better than others in specific circumstances), but it is common to represent utility with the logarithmic function as a tribute to Bernoulli's original proposal, which used the logarithmic function.*

$$\boldsymbol{u(x) = \Theta \log(\omega + x)} \qquad \textbf{(4)}$$

This function takes the log of the value of the increase in wealth, ω, from a possible outcome, x, and multiplies it by some parameter Θ. The Θ parameter dictates the curvature of the line when x is mapped onto the x-axis. Here, the utility of a gamble would take the form of a concave value function seen in Figure 1, where the utility of a gamble increases at a decreasing rate. To make this clear, if the utility function were linear, a gamble that would increase a person's wealth from £10 to £20 would provide the same amount of pleasure as it would when increasing wealth from £100,000 to £100,010. In contrast, a logarithmic function implies that the hedonic experience gained from an increase in wealth decreases as wealth increases. As will be explained further in the section on Prospect Theory, if an individual's behaviour fits with a logarithmic utility function, the set of choices can be described as risk averse.

The resulting relative units of wealth, referred to as 'marginal wealth,' are expressed as 'utils'. Given this utility function, one can calculate the *expected utility* (*EU*) of a future action (*A*) by summing all the utilities of each possible outcome (*x*) multiplied by the respective probabilities of *x* occurring given *A* (P_A):

$$\boldsymbol{EU(A) = \sum P_A(x) U(x)} \qquad \textbf{(5)}$$

This allows for the direct comparison of actions based on their expected utility, and highlights economic models' reliance on a 'common currency' in order to compare individuals' subjective utility across all types of goods, services,

actions, and contexts. For example, common abstract units are necessary to evaluate a worker's choice between a cash bonus and extra vacation days. Although the cash bonus may hold more expected monetary *value*, the extra vacation day may hold more expected *utility* to the overworked employee. In order to achieve this common scale mathematically, economists transform cardinal *values* into ordinal *utilities*. In other words, the expected utility model changes objective values into rank-ordered preferences. For this transformation to hold true in expected utility theory, a number of axioms, or formal assumptions, have been developed and refined over the past three centuries since its original formation by Bernoulli (1738). With the conclusion that behaviour is subject to an individual's preferences that cannot be explained purely by objective measures, economists had (perhaps unknowingly) created a common interest with the cognitive sciences.

The separate fields forged independent pursuits of many of the same questions – only from different angles and with different terminology. For example, B.F. Skinner (1953) formulated his theory about operant conditioning (i.e. how individuals form preferences) at the same time that Houthakker's (1950) General Axiom of Revealed Preferences (GARP) formalized how individual's preferences form behaviour.

General Axiom of Revealed Preferences (GARP)

Of particular interest to behaviourists is Houthakker's (1950) General Axiom of Revealed Preferences (GARP). GARP defines a rational decision-maker as

one that is internally consistent across decisions.[3] To clarify GARP rationality, take for example the situation in which little Stevie must choose how to spend his £15 allowance. At the movie theatre, tickets to see Star Wars and Star Trek both cost £5. Over the course of the week, Stevie is observed buying 2 tickets for Star Wars and 1 ticket for Star Trek. The next week, Star Wars tickets are offered at the matinee price for £3 while Star Trek tickets increase to prime time price at £6. Little Stevie would be acting irrationally if he decided to spend his next £15 allowance on 2 Star Trek tickets and 1 Star Wars ticket, since it is inconsistent with his previously revealed preference for Star Wars over Star Trek. Satisfaction of this axiom is both essential and sufficient in order for individuals to be described with a single, continuous and monotonic utility function. This is critical in that it allows economists to infer that an individual's preferences—which are inherently unobservable—are *revealed* by observable choices. A person who satisfies GARP behaves *as if* she had multiplied utility by probability in her head and chosen the outcome with the highest expected value. In other words, unquantifiable internal subjective utilities become quantifiable (or at least able to be rank-ordered) when GARP is satisfied. It is of note that psychologists make similar inferences when assessing individuals' preferences. For example, 'real-life' risk preferences are inferred from scores on experimental tasks such as the Balloon Analogue Risk Task (BART) or from questionnaires like the Domain-Specific Risk-Taking (DOSPERT) Scale. The key difference being that experimental analyses of such psychological preferences remain largely unconstrained by the demands of mathematical

---

[3] *It is of note that satiety effects are explicitly assumed away in GARP. In other words, Houthakker recognized that an individual's preferences may change with increased consumption of a good, but chose to ignore that in the model.*

formalisms, and thus hold dissimilar validation requirements (i.e. replication of results vs. mathematical proof)[4]. Note that the level of risk aversion implied by any given utility formation described above is based on the *objective* probability of the event occurring.

Subjective Expected Utility

Savage (1954) took expected utility theory a step further by incorporating subjective probabilities and utilities into the model. Savage's model implies that each individual has a unique expectation of how likely an event is to actually occur, and this expectation influences the likelihood of choosing a particular outcome. This accounts for the fact that individuals rarely if ever have complete knowledge of the true likelihood of an event, which causally attributes to variance in preferences. The model is as follows, where Subjective Expected Value (SEV) simply takes the subjective transformation of probability multiplied by value. This is extended to Subjective Expected Utility (SEU):

$$\boldsymbol{SEV = S(probability) \times value} \qquad \textbf{(6)}$$

$$\boldsymbol{SEU(A) = \sum S(p_a) \times u(x_a)} \qquad \textbf{(7)}$$

In SEU, the utility of an action, A, is the aggregate of all possible outcome utilities ($u(x_a)$) multiplied by their respective subjective probabilities ($S(p_a)$). Note the transformation into standard notation from equation 6 to equation 7. In plain words, subjective expected utility is the combination of a person's preferences

[4] *Notable exceptions to this exist (e.g. Reinforcement learning, signal detection theory), but generally the argument holds true given that there is yet no universal, all-encompassing mathematical 'theorem of psychology.'*

(outcome utilities) with his or her personal beliefs (subjective probabilities). The fact that subjective probability is distinct from utility implies context-independent risk attitudes. In other words, subjective utility theory allows for trait risk preferences. This also represents an important link to reinforcement learning theories, which reason that personal beliefs/subjective probabilities are the culmination of past experiences.

Subjective Expected Utility Theory, like Expected Utility Theory, requires that a number of axioms be satisfied for a decision-maker to be considered rational (see Fishburn, 1986). Again, these axioms are simply a set of criteria that need be fulfilled in order for the mathematical calculation of utility to hold true. Some axioms are stronger than others, and some can be relaxed if suggested by empirical data. However, in general when an axiom is violated this indicates that no single monotonic utility function exists to describe the individual's behaviour. Thus, the definition of a 'rational' economic agent lies in the conformity of the agent's behaviour to a single utility function and not in the 'fit' of an agent's behaviour to any particular function or model. Satisfaction of the axioms simply confirms that a single utility function *exists* for the individual – it implies nothing about functional form. This point will be demonstrated further in the following section. Axiomatic models can therefore be viewed as tools to assess whether or not a function exists to describe an observation – a tool that has recently been extended to the behaviour of neurons (Caplin & Glimcher, 2014). However, when deviations from axioms are systematic and multiplicative, this implies that the model itself (i.e. Subjective Expected Utility Model) likely requires modification.

Homo Sapiens: Irrational Economics

Despite the appealing simplicity and power of GARP as a means of mathematically modelling utility preferences, an alternative body of empirical research demonstrates that humans and animals systematically deviate from rational decision-making. It quickly became apparent that people consistently violated the axioms of Savage's subjective expected utility theory. Two famous examples of this are known after their proposers as the 'Allais paradox' (Allais, 1953) and the 'Ellsberg paradox' (Ellsberg, 1961), which reliably elicit 'irrational' choices when decisions are made under risk.

There exist a number of well-documented psychological effects known to robustly bias economic behaviour, such as: when choices are framed as a loss compared to a gain, called *loss aversion* or more generally the *framing effect* (Johnson, Hershey, Meszaros, & Kunreuther, 1993; Krupenye, Rosati, & Hare, 2015; Levin, Schneider, & Gaeth, 1998), *temporal discounting*, when choice outcomes are delivered at different points in the future (Ainslie, 1975; Berns, Laibson, & Loewenstein, 2007; Kable, 2013), and the 'sure thing principle,' when information about previous outcomes are known vs. unknown (Savage, 1954; Shafir, 1994; Tversky & Shafir, 1992; Waite, 2001). From this extensive—yet not nearly exhaustive—list of examples, it is apparent that choices often cannot be modelled as context-independent. It is generally argued that the observed choices in these examples represent decision making errors – which may have arisen from a variety of factors such as cognitive limitations or inattention – and therefore such decisions should not be considered true revealed preferences.

Such decision-making errors are exemplified by reverse-reward experiments in psychology, where the delivery of a larger volume of reward requires the subject to choose the smaller reward option (e.g. choosing the 3 candy option results in 5 candies being delivered), and *vice versa*. Preschool children (Carlson, Davis, & Leach, 2005), chimpanzees (Boysen & Berntson, 1995), and cockerel chicks (Hershberger, 1986) all demonstrate immense difficulty in learning to overcome this conflict between Pavlovian approach and instrumental response. Thus, this represents a decision-making error rather than a revealed preference. Since the inference that preschool children generally prefer 3 candies to 5 candies is obviously incorrect here, it should be concluded that cognitive factors directly impose constraints upon ones ability to reveal a true preference through choice.

An analogous dissociation between internal preferences and revealed choices also arises in the context of addiction. Although an individual may no longer *want* to consume a substance, his or her actions directly conflict with this desire. Berridge (1996) makes the distinction between *wanting* and *liking*. Berridge demonstrates that incentive motivation, or wanting, arises from separable psychological and neural processes as incentive palatability, or liking. The dopaminergic midbrain is purported to modulate wanting, while the opioid system is implicated in hedonic pleasure of reward (Symmonds & Dolan, 2012). Further studies have corroborated this, providing evidence for pharmacologically and neurophysiologically dissociable behaviours arising from reward salience/motivation and hedonic pleasure in rodents (Wilson, Laidlaw, Butler, Hofmann, & Bowman, 2006) and primates (Rolls, Sienkiewicz, & Yaxley, 1989). Furthermore, Tindell and colleagues (2009) demonstrate that

manipulations of physiological state can alter the dominant value mechanism regulating choice. In this study, a previously non-preferred salt reinforcer triggered neurons in the ventral pallidum (previously associated with the preferred sucrose reward) to fire when rats were put into a salt-deprived state. The neuronal activity here, if equated to a measure of subjective utility, reflects a preference reversal where firing rates do not represent the preferred sucrose reward but rather the item that fulfils a homeostatic need. This suggests that the brain's representation of value is based on information from multiple systems that have the ability to conflict and create reversals depending on psychological and physiological state. In summary, these examples demonstrate that psychological and physiological limitations often prevent an individual from revealing his or her true preferences, leading to behaviour that would be characterised as 'irrational' under Expected Utility Theory. Given these constraints, the universal applicability of the Expected Utility model has been called into question, which highlights an opportunity for coordinated efforts toward modification with the psychological sciences.

Prospect Theory

The multiple-systems approach has also been argued for decisions resulting in asymmetrical effects of gains and losses. Both Expected Utility Theory and Subjective Expected Utility Theory assume that preferences among potential prospects reflect 'description invariance.' For example, imagine an individual is asked to choose between option A) a 50/50 gamble of either £1000 or £0 and option B) £500 for sure. The individual's stated preference is assumed not to change based on the manner in which the available options are described.

Given that both options have the same expected value ($EV_A = (0.5 \times £1000) + (0.5 \times £0) = £500 \; and \; EV_B = 1 \times £500 = £500$), a preference for option B implies that the subject is risk averse to the variability of outcomes in option A. Kahneman and Tversky (1979) tested the postulated description invariance of 70 subjects by comparing subjects' preference between A and B to preferences between C and D, where option C) is a 50/50 gamble of either £0 or *losing* £1000 and option D) represents losing £500 for sure. As before, options C and D have the same expected value and therefore any preference between the two will reveal either a relative preference or aversion toward risk. The only difference in the choice between A and B and the choice between C and D is that the former is framed as a gain while the latter is framed as a loss. Kahneman and Tversky (1979) found that while 84% of subjects indicated a preference for option B over option A (i.e. risk aversion), 69% of subjects preferred option C over option D (risk seeking).

Using a similar paradigm, the authors also found evidence in support of a second violation of description invariance whereby subjects exhibited risk-seeking behaviour for large, but highly unlikely gains (e.g. winning the lottery), yet risk-averse behaviour for large, but highly unlikely losses (e.g. purchasing home insurance). These dichotomies can be conceptualized as a 'four-fold pattern' of risk attitudes, outlined in Table 2 below:

Table 2: Prospect Theory risk attitudes

	Medium – High Probability of Occurring	**Low Probability of Occurring**
Loss	Risk Seeking	Risk Averse
Gain	Risk Averse	Risk Seeking

Table 2: modified from Kahneman & Tversky (1979), the four-fold pattern of risk attitudes for gambles

framed as losses or as gains with a medium-high or low likelihood of occurring.

In response to this, Kahneman and Tversky (1979) proposed Prospect Theory as a framework for utility maximization under risk that accommodated for cognitive biases and boasted a better fit to empirical data. Prospect Theory also received a considerable amount of attention for its ability to account for the Allais paradox. Tversky and Kahneman (1992) later developed an axiomatic version, Cumulative Prospect Theory, which is arguably the most successful approach to behavioural economics thus far (cf. Gigerenzer & Selten, 2001; H. A. Simon, 1959).

By drawing heavily from the field of psychology, Kahneman and Tversky (1979) constructed Prospect Theory with three distinctive attributes that allow it to better describe individual behaviour. First, Prospect Theory incorporates a variable *reference point* (Figure 2a), which typically refers to the 'status quo' rather than to wealth as does Bernoulli's utility formulation. To make clear the

Figure 2: Prospect Theory

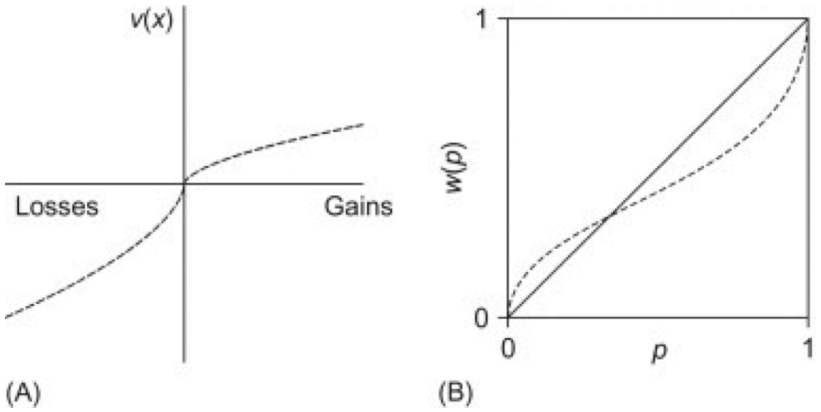


Figure with permissions from Fox and Poldrack (2009, p. 149), (A) represents a value function over losses and gains, with the 'kink' at the origin representing a steeper curve for losses than for gains, and (B) a weighting function for probabilities with 'S-shape,' implying that values at the extremes behave differently. The value function results in a utility curve that is concave for gains (implying risk-averse behaviour) and convex for losses (implying risk-seeking behaviour). Furthermore, the origin is not fixed at zero, but rather is variable, subjective, and often equal to the *status quo*. The weighting function captures the observation that individuals tend to overweight probabilities near zero (e.g. impossible) and underweight probabilities near one (e.g. certain).

notion of a reference point, consider the example where Robert unexpectedly earns a bonus of £1,000. Robert is quite satisfied with his bonus until he discovers that colleagues of his, Peter and Katy, both received bonuses of £3,000. As a result, Robert now finds himself quite unsatisfied with his bonus. Tversky and Kahneman (1981) might assert that, in comparing himself with Peter and Katy, Robert's reference point for gains and losses shifts from a baseline of zero (as the bonus was unexpected) to £3000 (the status quo). This implies that what was previously modelled as a gain of £1000 is now treated as a loss of £2000. Prospect Theory's second important distinction is that outcome probabilities are weighted by an 'S-shaped' probability function so that very unlikely outcomes have a stronger effect relative to very likely outcomes (Figure 2b). This fits with the 'Four-Fold pattern' of risk attitudes in **Table 2** above. Finally, loss aversion is modelled by computing utility for losses and gains with separable functions, as seen in equation 8 below:

$$u(x) = \begin{cases} -\lambda(-x)^{\beta}, & x < 0 \\ x^{\alpha}, & x \geq 0 \end{cases} \qquad (8)$$

The utility function detailed here dictates that gains are subjected to a (risk aversion) coefficient (α), while losses are influenced by a (risk-seeking) coefficient (β) *and* a loss-aversion coefficient, λ. In sum, these parameters result in a concave utility function over gains and a convex function over losses, with a kinked shape around the reference point indicating a steeper slope for losses than for gains (Figure 2a).

In equation 8, note the distinct parameters for loss-aversion and risk-aversion.

While the two terms can certainly interact[5], they should not be confused. Risk aversion distinguishes between option probabilities while holding value constant, whereby choice reflects a preference for a certain option over an uncertain option. By contrast, loss aversion describes the robust tendency of individuals to make decisions as if negatively valenced outcomes were twice the amount than they actual are with respect to comparable gains. With regard to equation 8 above, consider the individual who is evaluating the utility of the outcomes of a gamble that had a 50/50 chance of winning or losing £5. Assuming a risk-neutral attitude and a loss aversion coefficient of 2 (i.e. $\lambda = 2$, $\alpha = 1$ and $\beta = 1$), the individual will multiply the utility of a £5 gain $(u(x) = 5)$ and a £5 loss $(u(x) = -2\times5 = -10)$ by their respective probabilities of 50% in order to establish an expected utility for the gamble. Therefore, an individual is more sensitive to the prospect of a loss than to the prospect of a gain in Prospect Theory.

The λ coefficient of loss aversion is commonly cited at 2.25, which is the average provided by Tversky and Kahneman (1992). Gächter, Johnson, and Herrmann (2010) report that individuals' loss aversion estimates remain constant between decisions involving risk (e.g. gambles) and decisions that are riskless,(e.g. contrasting willingness to pay vs. willingness to accept for a good, see equation 10). However, numerous studies report substantial variation in λ across individuals (Haigh & List, 2005; Johnson, Gächter, & Herrmann, 2006) and decision attributes (Neumann & Böckenholt, 2014; Sayman & Öncüler,

[5] *Indeed, the effects of the two terms can interact in a way that one can even fully account for the other. For example, Novemsky and Kahneman (2005) demonstrate that loss aversion accounts for a large proportion of observed risk aversion in gambles involving small losses and gains.*

2005). Thus, despite its prevalence, it is unclear what loss aversion *is* exactly. Some researchers, such as Johnson and colleagues (2006) and Camerer (2005), suggest that loss aversion could represent: 1) a stable attribute of preference decisions, 2) something akin to a personality trait, 3) an affective response , or 4) the result of conflicting dual systems underlying losses and gains. Addressing the true nature of loss aversion presents an exciting opportunity for the collaboration of psychology, neuroscience, and economics. All three fields stand to gain by establishing the cognitive and neural mechanisms underpinning this source of variability. In order to fully consider this question, it would be useful to first gain an understanding of how loss aversion manifests in financial markets as well as the tools that have been utilized to measure it. Thus, we first shift our focus to the marketplace, after which point we will revisit the nature of loss aversion in the context of current neuroscientific research thereafter.

Efficient Markets: Rational Finance

Here, instead of goods and gambles, we consider assets (stocks and bonds) and trades. It is interesting to note that while economics often considers the prices of goods and services, the primary focus in financial markets lies in highly abstract representations of value and ownership. For example, buying 100 shares of Apple, Inc. stock in July of 2015 would have cost an investor $1,275.00. In return, that investor would be the proud owner of 0.0000000175^{th} of Apple Inc. Thus, although a stock fundamentally represents a share of ownership in a company, an investor may struggle to fully comprehend it as a tangible good. From a psychology standpoint, assets are far closer in nature to

a secondary reinforcer (e.g. currency) than a primary reinforcer (e.g. food). Moreover, given that fluctuations in stock prices generally approximate random walks (see Introduction), and that there is no contract associated with a stock (i.e. the company is under no contractual obligation to pay out dividends, etc.), the true value of a holding is neither fixed nor certain. An asset can become virtually valueless in an instant. Given this unique set of features, financial markets are particularly conducive to the study of cognitive mechanisms that bias loss- and risk-related behaviour.

Risk-Return Models

Finance is defined as a subfield of economics; while the two share similar normative ideals, finance has naturally developed its own traditions and field-specific methodologies. While expected utility is the primary computation of future subjective value in economics, the finance tradition generally relies on a similar but mathematically distinct construct: namely, the risk-return model proposed by Markowitz (1959). In its most basic form, the return, *r*, on an investment is calculated as the percentage of profit gained from a trade, by:

$$\boldsymbol{R = \frac{Price_{Sale} - Price_{Buy}}{Price_{Buy}} \times 100} \qquad \textbf{(9)}$$

The return is therefore a measure of the change in price between buying and selling with relation to the original purchase price. As with expected value, taking the product of an investment's returns and respective probabilities provides an estimation of expected return (ER). Importantly, the variability in returns on an investment over a given time period is regarded as a measure of the asset's *risk*. This is expressed as either the standard deviation (σ) or variance (σ^2), and represents the spread of the distribution of returns around the mean. It is therefore a general truism that greater risk confers the potential for greater return.

Markowitz (1959) states that one can construct an ideal portfolio based on preferred risks by eliciting the amount that an investor is willing to pay (WTP) for assets over varying degrees of risk and return (see Figure 3), using:

$$\boldsymbol{WTP = ER - b(\sigma^2)} \qquad \textbf{(10)}$$

Thus, when *b* represents an individual's risk preference, an investor's WTP for one unit of decreased risk will be reflected in the difference between expected return (ER) and subjectively weighted risk (σ or σ^2). Take, for example, the decision in which an investor chooses how to divide her portfolio between 'risk-free' government bonds that offer a certain return of 5%, and stocks that have a higher rate of return (between 5-15%) but also greater associated risk. This example is depicted in Figure 3, where the investor chooses point *X*, dividing her money between the two assets. A risk-averse investor would position *X* toward the left, and a risk-tolerant investor would position *X* toward the right. Therefore, *b*-Values offer a direct measure of an individual's risk preference from observed choice behaviour.

Figure 3: Risk-Return Models

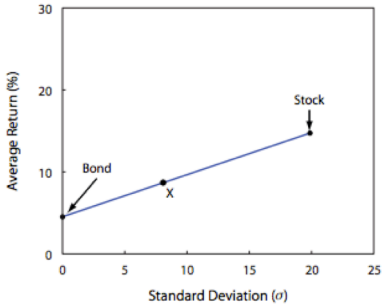


Figure 3 with permissions from Glimcher (2008). An investor chooses between how to divide her portfolio between 'risk-free' government bonds that offer a sure return of 5%, and stocks that have a higher rate of return but also a greater associated risk. The investor chooses point *X*, dividing her money between the two assets. A risk-averse investor would position X toward the left, and a risk-seeking investor would position X toward the right. *b*-Values offer a direct measure of an individual's risk preference from observed choice behaviour.

Efficient Markets

In parallel with the normative economic models implying the existence of a rational *Homo Economicus*, there also existed a similar sentiment regarding financial markets. Although it was not formally defined until the 1970's, much of the financial discourse centred around the notion that markets were *efficient* (Sewell, 2011; Shiller, 2003).

Fama (1970, p. 383) made the fundamental assertion that the price of a stock

acts as a signal to the market about the economic health of the company, and that in an efficient market, this signal "fully reflects" *all available* information. Importantly, this notion includes all types of information (e.g. financial, social, political, etc.) that might have a direct or indirect effect on the marketplace. As an example, consider the situation where the CEO of a company is publically accused of cheating on his wife. Although this information may have little to do with the financial abilities of the company itself (and indeed, the accusation need not necessarily even be true), an ensuing fall in stock prices would reflect a loss of shareholders' trust in the company's leadership. Thus prices are a reflection of the information that investors have, as well as their expectations about how other investors will react to that information.

By this logic, changes in price will reflect a combination of: new information, investors' own reactions to new information, and investors' expectations about other investors' reactions to the new information. While the composite reaction itself may be unpredictable, it can be assumed that the price change will fully depend on the new information and thus be independent of previous information (and previous price). Therefore, the idea of efficient markets became coupled with the concept introduced in the beginning of this chapter, namely that stocks prices follow the 'random walk' pattern of Brownian motion defined by Einstein in 1905. In other words, stock prices are completely random, denoting that future prices are completely independent of current prices. This represents a particularly grim conclusion for the individual whose job it is to predict future market trends. Indeed, given an efficient market in which prices are a reflection of all information, all investors privy to all prices by definition have access to all information. Therefore, 'bargain' stock prices do

not exist, and any attempt of one investor to outperform another investor is futile, and only adds to the overall randomness of prices. Fama (1965, p. 10) stated bluntly:

> If, as the empirical evidence seems to suggest, the random-walk theory is valid, then chartist theories are akin to astrology and of no real value to the investor.

Despite these rather self-defeating implications, the efficient markets hypothesis found a great deal of success in the field and remains a strong theoretical pillar. In support of the efficient market hypothesis, a number of researchers have highlighted the fact that even professional money managers do *not* beat the market (Rubenstein, 2001).

The efficient market hypothesis refutes the idea that irrational investors (often referred to as ‘noise traders’) can influence market prices in any meaningful way, since rational investors would quickly identify and take advantage of any deviation from fundamental value (Friedman, 1953). There also exists the possibility of beating the average by reacting to new or insider information more quickly than the time it takes for the market to adjust to its new average. It is generally accepted that while an investor may occasionally beat the market by reacting quickly to new information, it is not possible to sustain this on the average. Proponents of the efficient market theory and the Capital Asset Pricing Model (Markowitz, 1959) conclude that although it may not be possible for investors to beat the market on average, it is still possible to perform at average by maintaining a sufficiently diversified portfolio.

Inefficient Markets: 'Noise Traders'

Fama (1970) himself admits that a purely efficient market, where all information is freely known, is not an accurate description of any real world markets. In fact, he supplied a number of examples of markets exhibiting variable levels of efficiency in his (1970) paper, including the scenario in which information is monopolistically held and exploited. In general, behaviourists such as Barberis and Thaler (2003) argue that the effect that 'noise traders' have on market prices is *not* always cancelled out by quick-acting rational investors (otherwise referred to as 'arbitrage traders' in the finance literature). The authors argue that arbitrage, or the ability of rational investors to take advantage of mispricing from noise trading, is both risky and costly. In real markets, arbitrage does not fully cancel out the effects arising from psychological biases, as asserted by Friedman (1953).

There also exists a long-standing debate about whether or not the average stock's movement over time carries momentum, meaning it is more likely to continue on its current trajectory than to reverse directions, or whether its movement more closely approximates a random walk. While many argue that asset fluctuations are entirely random (notably, Malkiel, 1973), others provide evidence for serial autocorrelation (Bondt & Thaler, 1985; Campbell, Lo, & MacKinlay, 1997; Shiller, 2003), suggesting that psychological factors such as herding behaviour and heuristics lead to serial correlation in prices over time.

Within psychology, the term 'heuristic' refers to an intuitive reduction in choice sets due to the limited computational or attentional capacity of a decision-maker. Tversky and Kahneman (1974, p. 1124) assert that:

> [P]eople rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations. In general, these heuristics are quite useful, but sometimes they lead to severe and systematic errors.

In colloquial terms, a heuristic is a rule of thumb based on subjective beliefs. When a heuristic consistently leads to a departure from rationality, it creates what is called a 'cognitive bias.' Cognitive biases tend to be extremely robust – even prior knowledge of the bias often cannot effectively preclude its expression. This is akin to perceptual illusions such as Jastrow's (1899) duck-rabbit, whereby one's previous knowledge of cognitive psychology and perception does not change one's inability to see both the duck and the rabbit at the same time. When even very small departures from rationality arise as systematic reactions to market dynamics, the aggregate influence across thousands of investors can be powerful. Thus, many finance experts are eager to capture the effect that these systematic deviations have in predictive asset pricing models.

The Disposition Effect

Shefrin and Statman (1985) labelled one such bias the 'disposition effect,' which describes investors' inclination to sell profitable stocks too soon while holding on to losing stocks for too long. This effect is extremely robust and has been observed in both professionally managed and individual investment accounts (Shapira & Venezia, 2001), as well as across investor classes (P. Brown, Chappel, da Silva Rosa, & Walter, 2006), cultures (G. Chen, Kim, Nofsinger, & Rui, 2007) and genders (Feng & Seasholes, 2008). However, the

underlying causes of the disposition effect remain unclear.

In the past, it was argued that the disposition effect could be attributed to investors' irrational belief in 'mean-reversion,' i.e. that the direction of the stock will eventually flip and revert back to the mean (for review, see Mukherji, 2011; Poterba & Summers, 1988). If this explanation were correct, investors would be just as likely to realize gains and losses, as the investor would assume both an increase and a decrease in price would quickly return to average. However, the observation that investors tend to realize gains at a much higher rate (about 50%) than losses allowed Odean (1998) to refute this explanation.

Shefrin and Statman (1985) initially argued that this behaviour was instead a product of Prospect Theory utility, whereby an investor holding a stock that goes down in price after purchase would use the original price as a reference point. This places the investor in the steeper, convex loss domain of the Prospect Theory utility curve (i.e. the lower left-hand quadrant of Figure 2a). Compared to the concave gain domain, the now risk-seeking investor requires an even lower price before she is willing to sell. This process could potentially perpetuate itself until the stock no longer had value, or until the investor changes the reference point (e.g. by lowering expectations). The opposite would be true for individuals holding a winning stock, whereby the investor would find herself in the concave risk-averse gain domain (i.e. the upper right-hand quadrant of Figure 2a). Less willing to risk losing the current gain, the investor would be biased toward selling the winning stock too soon. Through this process, Prospect Theory provides an explanation for why individuals hold losses too long while selling winners too soon.

A number of alternative explanations for the disposition effect have also arisen. For example, Thaler (1998) highlighted the affective processes that impede action in the face of a loss, stating that selling a losing stock is more painful than holding on to a 'paper loss.' Alternatively, others have hypothesized that the act of selling at a gain bears inherent utility, while respective disutility is derived from selling at a loss (Barberis & Xiong, 2012; Barberis & Xiong, 2009; Frydman, Barberis, Camerer, Bossaerts, & Rangel, 2014). Hirshleifer (2001) postulates a third alternative in which investors avoid selling losing stocks because the act of doing so represents a self-signal that they have performed poorly. Hirshleifer's hypothesis aligns with the concept of self-justification in cognitive dissonance theory (Festinger, 1957). Given that poor performance may be seen as a reflection of low ability, the investor is motivated toward self-deception in order to maintain a notion of high self-regard. In all three of the alternative explanations introduced above, the investor disassociates the outcome of holding a gain/loss from the outcome of selling a gain/loss. In other words, even though an investor has already lost real money when she is holding a stock that has plummeted in value, this notion of financial loss is qualitatively different than that experienced upon actually selling the losing stock.

The 'Realization Hypothesis,' in which individuals conceptually separate a paper gain/loss from a realized gain/loss, has found traction in recent neurobiological research. Frydman and colleagues (2014) reported a large spike in fMRI BOLD activity in the ventral striatum (an area critically involved in reward processing) after participants decided to sell at a gain compared to when they decided to hold at a gain. In short, the researchers found that a

realized gain elicits quantifiably distinct reward-related neural activity compared to a paper gain. This study constitutes an important link between economics, psychology and neuroscience, because if clues to the origins of the disposition effect lie in subcortical structures such as the ventral striatum, then further testing in the laboratory may reveal important predictions about behaviour in the marketplace.

Economic Animals

In order to fully explore the neural mechanisms underlying behaviour in the laboratory, it is often ethically and economically preferable to use simpler organisms (e.g. rodents or birds). This is generally justified with regard to reward-related behaviour, because while cortical structures between mammalian species show potentially meaningful functional and cytoarchitechtonic differences, subcortical brain areas responsible for representing and learning from reward are remarkably evolutionarily preserved (this concept is developed in detail in the following section, and the reader is referred to Figure 5 for illustrative comparison of the reward system between species). Indeed, these underlying similarities may even help explain the successful migration of rational economic models of the early 20th Century into models of optimal reward-related decision making behaviour in the ecological literature (e.g. Optimal Foraging Theory) beginning in the 1960's (Cowie, 1977; Mäki, 2009; Stephens & Krebs, 1986). This also suggests that laboratory tasks with animals such as rats could potentially represent valid behavioural and neural models of economic decision-making.

In order to explore the viability of modelling economic behaviour in rats, researchers (e.g. Collier, Hirsch, & Hamlin, 1972; Kagel et al., 1975) in the 1970's began explicitly developing laboratory-based economic tasks. In one experiment, which Kagel and colleagues (1981) called a 'consumer demand' task (Figure 4), rats bar-pressed for food pellets up to a limit of 350 presses (Z'). The experimenter either decreased 'budget' by decreasing the amount of lever presses allowed to 250 while holding the payoff magnitude constant (Z), or increased the 'price' by requiring more presses per unit of reward (Z''). Doubling the price resulted in the same effect on demand for water as did decreasing the allowable budget. Thus the authors concluded that rat consumption patterns obey the basic principles of consumer demand theory. Similar studies report analogous findings with preferred and non-preferred substitutes, essential and non-essential commodities, and changes in wealth and labour-supply (Collier et al., 1972; Kagel & Battalio, 1980). These studies, which represent some of the first direct tests of economic theory using laboratory rats, provide robust evidence that maximizing behaviour under environmental constraints is not a capability unique to humans.

Figure 4: Rat Demand Curves

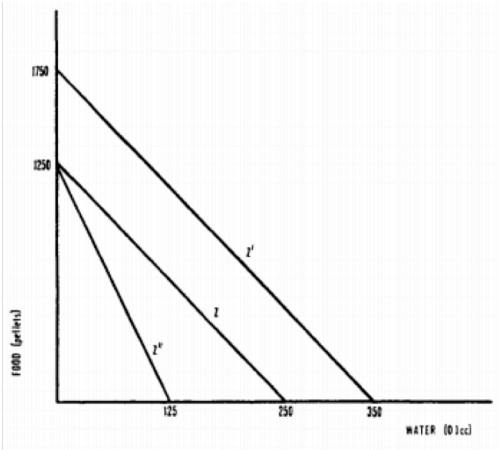


Figure 4 with permission from Kagel, Battalio, Rachlin, and Green (1981, p. 4): Demand curve for rats between two 'goods', where doubling the required bar presses for a given magnitude of water (x-axis) from Z to Z'' had the same effect on demand for water as halving the amount of fluid that paid out per bar press. This bar press per unit of reward measure corresponds with the definition of 'price' in consumer demand theory.

More recent research demonstrates that we may also share our 'irrational'

behavioural biases with our evolutionary ancestors. For example, Chen and colleagues (2006) reveal that non-human primates are also similar to humans when they behave irrationally. By creating a token economy with fruit, the authors find that while capuchin monkeys respond rationally to changes in wealth and price, they exhibit both reference dependence and loss aversion (refer to previous section on 'Prospect Theory') when risk is also incorporated. Work by the same group found that capuchins also exhibit framing effects – becoming risk seeking when gambles were presented as a loss and risk-averse if the gamble was presented as a gain (Lakshminaryanan, Chen, & Santos, 2008). These studies further support the notion that the mechanisms underlying choices that involve the potential for loss extend beyond the human species.

Researchers can take advantage of the fact that animals behave similarly to humans in many contexts. Like Kagel (1975) and his contemporaries, one can test theoretical assumptions of economic theories without the confounding 'human factors', such as prior assumptions about how an economy works or differing levels of numeracy amongst participants. In this way, animal paradigms offer clear advantages to neuroeconomic study. However, there are also disadvantages and operational obstacles to working with laboratory animals. For example, it is not possible to describe the outcome values and probabilities of a potential gamble to an animal – it must learn the contingencies of an action through trial and error. Furthermore, it is extremely difficult to operationalize resource loss in an animal task, for one can easily reward an animal with food or drink, but it is difficult to then retract the reward once it has been consumed. The development of tasks that overcome such challenges will be key in facilitating a comprehensive characterization of the

neural structures and functions that give rise to economic choices.

Reward Learning and Expectation in the Brain

Up to present, animal work has been instrumental in achieving our current understanding of how rewards are learned and represented in the brain. While a number of mechanisms underlie valuation and choice during decision-making, the dopamine system plays a central role in the neural processing of reward-related behaviour (Fibiger & Phillips, 1986). From an anatomical standpoint, the rat (and to a lesser extent the pigeon) dopaminergic system shows striking similarities to the human system (see Figure 5). Dopamine neurons primarily originate in the ventral tegmental area (VTA) and the substantia nigra pars compacta (SNc). Bilaterally, the rat has between 40,000 and 45,000 dopamine neurons in the VTA and SNc combined, while healthy young adult humans have 400,000 – 600,000 (Puig et al., 2014). The striatum represents the most densely innervated target area of midbrain dopamine projections (Björklund & Dunnett). Postsynaptic target cells express D_1- (low dopamine affinity) or D_2-like (high dopamine affinity) receptors that function as slow-acting

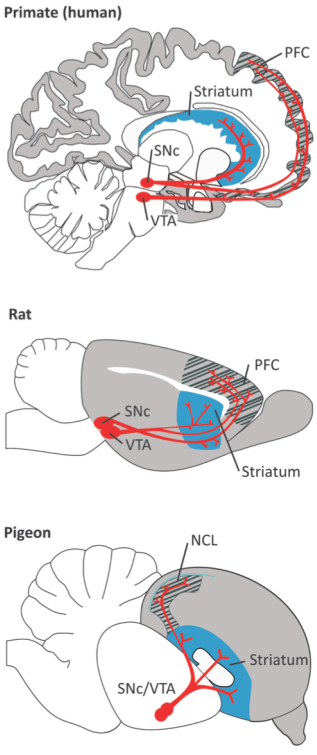


Figure 5 reproduced* from (Puig, Rose, Schmidt, & Freund, 2014): Comparative neuroanatomy of the dopamine system in the primate, rat and pigeon. Dopamine neurons originate in two main midbrain nuclei, the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc), and project to many areas (projections shown in red), especially the striatum and prefrontal cortex (PFC). Striatal areas are shaded in blue, cortical areas are shaded in gray, and the hatched areas represent PFC (or its structural equivalent in birds, the nidopallium caudolaterale; NCL). *Figure reproduced under Creative Commons Attribution License (CC BY).

G-protein coupled neuromodulators (Santana, Mengod, & Artigas, 2009). Dopamine neurons demonstrate two types of firing patterns: 1) phasic, quick bursts of action potentials and 2) tonic, slow ramping currents. It is theorized that the two modes of firing fulfil separate functions (i.e. transmit different types of information) with respect to reward prediction and motivation (Howe, Tierney, Sandberg, Phillips, & Graybiel, 2013; Tobler, Fiorillo, & Schultz, 2003), which may also be facilitated by the asymmetrical D_1- and D_2-like receptor affinities for dopamine.

Robust evidence supports the theory that the midbrain dopamine system elicits a phasic learning signal, or Reward Prediction Error (RPE), in response to expectations about reward (Schultz, Apicella, & Ljungberg, 1993; Schultz, Dayan, & Montague, 1997; Schultz & Dickinson, 2000). Unexpected rewards reliably evoke an increase in phasic dopamine activity (positive reward prediction error) in dopaminergic midbrain areas such as the VTA and nucleus accumbens (NAc), while the omission of an expected reward consistently inhibits activity in these areas, resulting in a negative reward prediction error. The error signal has been shown to equate to the expected (mean) value of a reward distribution divided by its standard deviation (Preuschoff & Bossaerts, 2007; Schultz, 2010; Tobler, Fiorillo, & Schultz, 2005). Thus, the prediction error signals a normalized value of how much an obtained reward differs from expectations. By modulating synaptic plasticity in midbrain dopamine neurons, it is theorized that this signal selectively reinforces rewarded behaviour while discouraging unrewarded behaviour (Suri & Schultz, 1999; Sutton & Barto, 1998). In this manner, dopaminergic modulation allows an organism to reliably update expectations about an outcome that lead to better predictions (and

thereby better information on which to base decisions) about similar encounters in the future.

Given that the majority of research presented above was in animals, Zaghloul and colleagues (2009) used single neuron recording during deep-brain stimulation in Parkinson's disease patients to establish whether reward prediction errors could be observed in humans as well. In this study, the authors employed a probabilistic gambling task where participants could choose from a relatively safe or risky deck of cards to earn hypothetical money. As has been observed in previous animal studies, Zaghloul and colleagues found that neurons in the substantia nigra responded to unexpected gains and losses with increased or decreased activity, respectively (see Figure 6). However, since Parkinson's Disease is characterized by extensive targeted cell loss in the substantia nigra, it could be argued that the results depicted here may not reflect those of a healthy individual (Aarts, 2012). Furthermore, the authors are not explicit about whether subjects perceived losses within the task as a monetary loss (e.g. Begin: $50 ➔ Outcome: -$5 ➔ End: $45) or as a non-reward (e.g. Begin: $50 ➔ Outcome: 'Loss' ➔ End: $50). The later would more closely approximate animal studies on reward prediction error. Either way, this study provides

Figure 6: Activity of Single Neurons in the Substantia Nigra of Parkinson's Disease Patients after a Financial Gain and Loss

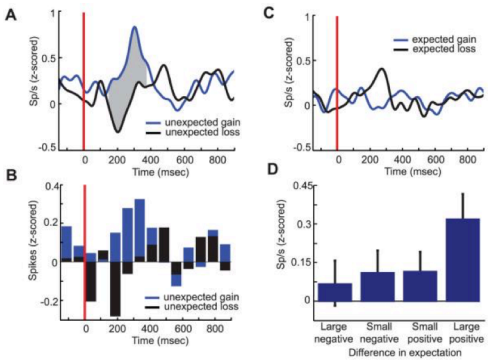


Figure 6 with permissions from Zaghloul, et al. (2009) depicts the single unit activity of substantia nigra neurons of Parkinson's Disease patients during a probabilistic choice task. (A-B) Increased activity after unexpected gains and suppressed activity after unexpected losses, (C-D) Small increase activity after an expected loss – no change after expected gain.

evidence that the human brain holds some representation of the expected value of an outcome, and that midbrain dopamine activity responds to any deviation from that expected value.

Put in the context of the stock market, reward prediction errors may offer some intriguing insight. Similar to midbrain activity in gambling tasks, one would predict that phasic dopamine signals would respond with a relative increase in firing rates after a large unexpected gain on a trade, and *vice versa* after a large unexpected loss. Dopamine neurons then encode the difference between expected and experienced reward, which provides a learning signal for updating expectations about subsequent events. The dopamine-modulated relative increase in expected outcome could potentially be responsible for the fluctuation of an individual's reference point. In the case of gain omission, as is the case when firms choose to withhold dividends for a certain period for example, dopamine activity would be suppressed, and a negative reward prediction error would provide a learning signal *not* to repeat the investment action that lead to the omission outcome. Indeed, Michaely, Thaler, and Womack (1995) found that firms significantly underperformed compared to the market one year after announcement of such a dividend omission. The opposite was also true for those announcing the initiation of dividend payments. Given that past market performance is a notoriously poor predictor of future prices, and that monitoring of individual stocks does not accurately represent portfolio-level performance, this reinforcement-learning signal may lead to maladaptive investment behaviour.

Since stock prices are in constant motion and few reliable signals of

increasing/decreasing future prices exist, observed prices would rarely – if ever – perfectly match predicted values. This implies that reinforcement learning prediction errors may have a sustained effect on investor behaviour, as price stochasticity requires constant updating by large increments. In reinforcement learning (Sutton & Barto, 1998), a primary reinforcement signal can itself become reinforcing (i.e. a secondary reinforcer). When applied to the market place, reinforcement learning theory suggests that the act of trading itself could become a reinforced behaviour. If this were the case, successful investors would be biased to trade more frequently and unsuccessful investors to trade less often. Evidence for this hypothesis presents itself in work by De, Gondhi, and Pochiraju (2010), who find that investors were indeed more active after experiencing recent success. In this way, reinforcement learning may act as a fundamental motor of investment decisions.

Over time, repeated maladaptive learning signals from the nervous system may become more established as investors learn action-outcome associations. In laboratory tasks, rats learn to associate actions (such as running down a runway to retrieve reward) with outcomes that lead to a rewarding state (e.g. consuming a sugar pellet), which is called an action-outcome association (Tolman, 1932). Similarly, Karlsson, Loewenstein, and Seppi (2009) demonstrate that investors log on to examine their trading accounts more often when markets are doing well than when markets are doing poorly. This suggests that investors also have a neural representation of an action-outcome association between accessing the online portfolio and the likelihood of experiencing a gain or loss. While action-outcome association learning is necessary and adaptive in many contexts, the resulting behaviour in the above

example would potentially bias investors toward holding on to losing stocks too long and selling winning stocks too quickly (i.e. exhibit the disposition effect).

Value Representation in the Brain

When given a choice between outcomes, including mixed gambles (i.e. those involving both wins and losses) and decisions that involve costs (e.g. effort, monetary, or foraging costs), outcome valuation allows an individual to ascribe a subjective value to a possible outcome based on its attributes (e.g. desirability, valence, salience and risk) as well as any previous experiences associated with that outcome. Thus, valuation represents the first stage of a simple three-stage neural decision process (see Figure 7) proposed by Platt and Plassmann (2014).

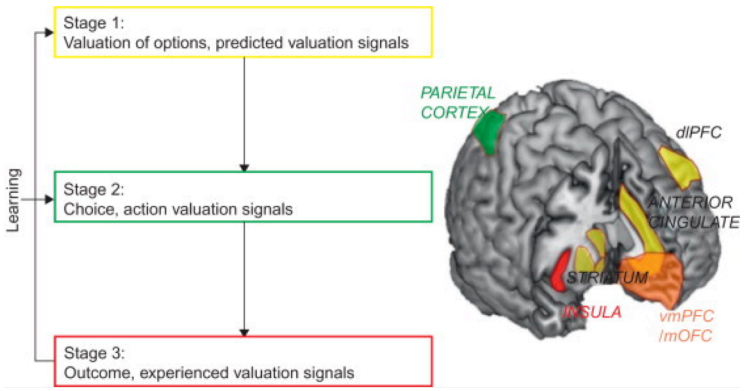


Figure 7 with permissions from Platt and Plassmann (2014, p. 239): In stage 1, observed and predicted attributes of options are consolidated into a subjective value signal. In primates, this occurs primarily in areas highlighted in yellow such as the dorsolateral prefrontal cortex (dlPFC), the anterior cingulate cortex, the orbitofrontal cortex and the striatum. In stage 2, the subjective value signal is transformed into action signals in the parietal cortex (green) for the motor systems to carry out the choice. Once the outcome is actually experienced, outcome values are signaled in stage 3 in areas highlighted in red, such as the insula striatum and orbitofrontal cortex. These stages are not required to operate in the order presented, and could potentially operate in parallel.

Once the value signals are computed, they are passed on to comparator regions and subsequently converted to action values so that motor areas can realize choice. In the final stage, the brain encodes the value of actually experiencing (e.g. of receiving or consuming) the chosen outcome. This final stage then feeds back to the first stage to update predicted values as reinforcement learning. It has been proposed that the vmPFC, OFC and striatum encode the subjective value of rewards, while the lateral interparietal area (LIP) translates subjective values into action values. The insula, striatum and OFC are implicated in signalling experienced reward.

When evaluating the value of one option within a set of options, it is important to note that one option's value will invariably be affected by the other alternative options' values. Tobler, Fiorillo, and Schultz (2005) observed electrophysiological recordings of single midbrain dopamine neuron activity in awake Macaque monkeys responding to liquid rewards. The researchers found that activation of dopamine neurons increased monotonically as a function of reward magnitude and probability (i.e. the components of expected value) with both predicted and unpredicted rewards. Thus, when given a choice between a set of alternatives, the outcomes associated with the other choices become reference points for the subjective value that one assigns to a chosen outcome. It is of note that dopaminergic activity did not adapt to the absolute value of rewards, but rather to the standard deviation of the most probable potential outcomes. This evidence suggests that the brain consolidates information by representing reward in a context-depend manner within the dopamine system, offering critical insight as to why one’s choices are often subconsciously influenced by the options surrounding it. For example, given a choice between

three competing brands, consumers will often choose the middle option – which can mean either the option with the middle price point or the option literally spaced in the middle of the visual field (Sigurdsson, Saevarsson, & Foxall, 2009; Simonson, Nowlis, & Lemon, 1993).

While research regarding the neural encoding of reward consistently points to targets of dopaminergic midbrain areas as primary modulatory substrates, much less is known about how the brain encodes resource losses. Taking insight from Prospect Theory's (Kahneman & Tversky, 1979) oft cited concept that losses loom twice as large as gains (see Figure 2), researchers have proposed two different hypotheses regarding the neural representation of losses. Loss aversion suggests that either: a) there is an affective response that biases evaluations involving losses in a dual system (Ashraf, Camerer, & Loewenstein, 2005; Kuhnen & Knutson, 2005), or b) gains and losses are handled asymmetrically within a single system (for review, see Kable & Glimcher, 2009). These two competing hypotheses of whether the valuation system is guided by a single system or dual-systems are fuelled by contradictory findings that have sparked a heated debate within the field.

On the one side, studies such as Tom *et al.* (2007) provide human neuroimaging evidence in support of the single-process theory that focus on valuation centres in the ventromedial prefrontal cortex (vmPFC) and striatum. Tom and colleagues (2007) gave participants a series of gambles that they chose to either accept or reject. Each gamble was associated with an equal (50/50) probability of winning or losing. The authors suggest that BOLD responses in the vmPFC and ventral striatum at the time of decision selection

exhibit ‘neural loss aversion.’ In other words, the same areas that encode reward not only show decreased activity to losses, but the decreases in loss-elicited activity were also greater than equivalent increases in gain-elicited activity. Furthermore, the study did not find any correlation between losses and activity in regions associated with negative emotions (e.g. amygdala or anterior insula).

Contrary to these findings, Gan and colleagues (2010) investigated whether costs were tracked by dopamine release in the rat nucleus accumbens with the use of fast-scan cyclic voltammetry. Rats were placed in a decision-making paradigm, whereby reward magnitude and effort-based cost were each manipulated. The results demonstrated that, while phasic dopamine in the nucleus accumbens of rats did track probability-weighted reward magnitudes as in Tobler *et al.* (2005), extracellular dopamine levels did not correlate with changing costs or a net cost-benefit utility function. Therefore, midbrain dopamine neurons appear to encode information about the benefits of an outcome but not specifically its costs or net utility.

Instead, a number of studies implicate the amygdala and anterior insula as the counterpart to reward in neural loss processing, particularly with respect to loss aversion. A recent study by McHugh and colleagues (2014) recorded both local field potentials and hemodynamic tissue oxygen signals, which are putatively equivalent to human BOLD signals (Lowry et al., 2010), in the basolateral amygdala of freely moving mice during a fear-conditioning (foot shock) task. The authors found an increased hemodynamic response to an unexpected foot shock and suppressed activity following an unexpected foot shock omission.

Furthermore, greater evoked hemodynamic and theta signals were predictive of better discrimination ability between aversive and nonaversive stimuli in subsequent sessions. This is in line with theories implicating theta oscillations in enhanced information transfer (Buzasaki, 2002) and attentional gain (Sejnowski & Paulsen, 2006). While these findings are intriguing in that they parallel dopaminergic reward prediction errors in the midbrain, it is unclear whether the (pre-synaptic) measurements in this study truly confirm that the aversive prediction errors generate from the amygdala.

De Martino, Camerer, and Adolphs (2010) found that patients with Urbach-Wiethe disease, a rare neurological disease that causes selective bilateral lesions of the amygdala, have difficulty processing fear. Similarly, rhesus monkeys with amygdala lesions have been shown to exhibit a lack of fear in approaching novel stimuli compared to non-lesioned monkeys (Mason, Capitanio, Machado, Mendoza, & Amaral, 2006). This led the authors to investigate whether the patients also exhibited loss aversion differently than healthy controls. Indeed, the study showed that amygdala-lesioned patients did not exhibit loss aversion under an experimental paradigm closely paralleling that of Tom *et al.* (2007), whereas healthy controls did. Interestingly, the authors also showed that despite this absence of loss aversion, patients with amygdala damage did still exhibit risk aversion (a preference for safer rather than riskier gambles) similarly to healthy controls. The researchers were able to conclude, therefore, that the amygdala is critical in the processing of losses in a manner that is independent of risk evaluation.

Loss aversion is also fully predicted by the somatic marker theory, which

implicates anterior insula in the integration and experience of emotional and somatosensory information (Bechara & Damasio, 2005; Craig, 2002; Damasio, 1994). Furthermore, the anterior insula has been associated with fear (Adolphs, Tranel, & Damasio, 1995), anticipatory anxiety (Chua, Krams, Toni, Passingham, & Dolan, 1999), as well as the anticipation of monetary losses (Kahn et al., 2002) and aversive physical (Buchel & Dolan, 2000) and visual stimuli (Simmons, Matthews, Stein, & Paulus, 2004). Given that the anterior insula mediates negative affect and the anticipation of negative outcomes, emotion may therefore contribute to a greater impact in the perception of losses than of gains, which supports the dual-system perspective. In further support of this hypothesis, Sokol-Hessner and colleagues (2009) found that losses elicited greater physiological arousal than gains in human participants, consistent with loss aversion. Moreover, participants were able to attenuate behavioural and physiological effects of loss aversion through intentional cognitive-regulation efforts. A second study not only correlated behavioural and physiological expressions of loss aversion with BOLD activation in the amygdala, but also demonstrated that successful cognitive-regulation strategies reduced activity in the amygdala (Sokol-Hessner, Camerer, & Phelps, 2013). Importantly, the changes in BOLD signals in the amygdala associated with cognitive regulation correlated with responses to losses, but not to gains, and also coincided with increased activity in prefrontal regions and the striatum.

Upon experiencing an outcome, a rat study conducted by Steiner and Redish (2012) argues that the OFC encodes information about counterfactual options when a loss is incurred, thereby representing something akin to ‘regret.’ The authors designed an economics-based task to induce regret in rats, while

recording from single neurons in the OFC and ventral striatum. Regret was defined as the revaluation of a previous choice in terms of the current choice. A four-armed “Restaurant Row” task was implemented, wherein each arm contained an equal amount of different flavoured pellets (e.g. cherry, chocolate). As a rat approached the entrance to one arm, the pitch of a tone indicated the length of wait time required before the reward could be consumed. The rat then had the choice of waiting the given amount of time or moving on to the next arm. The wait time varied randomly between 1-45 seconds. Due to time constraints, any skipped arm that was comparably better than the subsequent arm was considered a missed opportunity, and cause for regret. Disappointing sequences, or sequences where a non-preferred outcome resulted from chance rather than from the rat’s decision, were analysed as controls to regret. For example, a disappointment-inducing sequence would result when a non-preferred outcome was (correctly) skipped, but followed by a similarly non-preferred outcome. By contrast, a regret-inducing sequence would occur when the rat (incorrectly) skipped a preferred outcome for a comparably less-preferred outcome. Intriguingly, in regret-inducing circumstances (as opposed to disappointing circumstances as controls), representations of the previous zone/arm were most strongly signalled in the OFC and ventral striatum. The representation of regret in these areas more closely related to the missed action rather than the missed outcome. This was likened to human regret, in which people tend to ruminate over actions taken or not taken rather than the missed outcome itself (Gilovich & Medvec, 1995). Together, these studies represent a robust argument in favour of neural separation of losses and gains.

Thus it is possible that rewards and losses are initially encoded in distinct brain areas and subsequently integrated into one comparison signal via a third region. Evidence from fMRI and lesion studies would suggest that a decision's benefits are signalled primarily by the ventral striatum (Basten, Biele, Heekeren, & Fiebach, 2009; B. Knutson & Cooper, 2005; Roesch, Singh, Brown, Mullins, & Schoenbaum, 2009), while the costs associated with a particular outcome are relayed by the amygdala and insula (De Martino et al., 2010; Yacubian & al, 2006). The combined signal would then act as a net value, akin to the economic concept of expected value. Thereafter, researchers have hypothesized that these two signals are combined in the vmPFC or OFC into a reward- or action-value signal, depending on the type of decision[6] (Gläscher, Hampton, & O'Doherty, 2009; Kolling, Behrens, Mars, & Rushworth, 2012; Lebreton, Jorge, Michel, Thirion, & Pessiglione, 2009). Further fMRI evidence shows that the combined valuation signal then accumulates in the parietal cortex until a decision threshold is met (Basten et al., 2009).

Clark and Dagher (2014) recently incorporated the separable dopamine signals into a Prospect Theory utility model of risky decision-making. In this model (see Figure 8a), the utility of potential gains and losses are computed separately (gains in the vmPFC and striatum and losses in the amygdala and insula) and then integrated into a decision value in the striatum. The degree of loss aversion, characterized by the steeper slope of the value function over losses, is determined by the balance between opposing tonic and phasic action-

[6] *There is some debate regarding value assignment in the prefrontal cortex, especially with regard to different types of decisions (e.g. economically constructed binary goods-based decisions vs. explore-exploit decisions with greater ecological plausibility). See Rushworth and colleagues (2012) for a comprehensive discussion.*

selection signals.

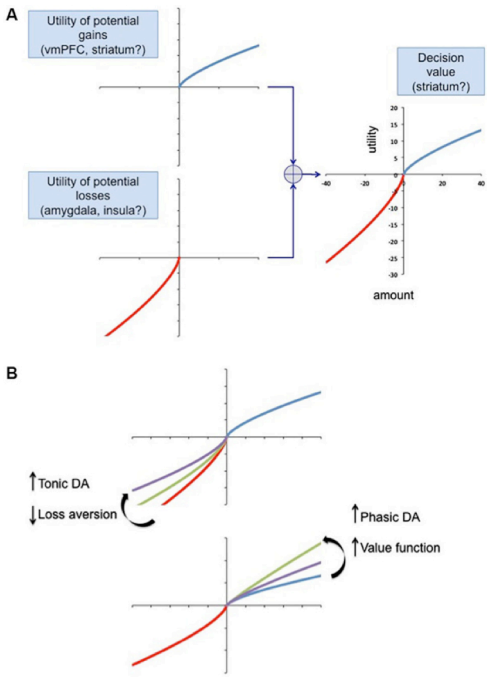


Figure 8 reproduced* from (Clark & Dagher, 2014): (A) The brain is hypothesized to compute utility from potential gains and losses in different substrates. In the model, utility from gains is computed in the vmPFC/striatum and utility from losses is computed in the amygdala/ insula. These separate utility signals are then integrated into a decision value in the striatum. Note that the likelihood (i.e. probability) and risk (i.e. variance) are either 1) also represented separately for gains and losses or 2) are computed elsewhere upstream. The combined decision value in the striatum reflects loss aversion, whereby the losses in the steeper red domain result in greater disutility than the utility of equivalent gains in the flatter blue domain. (B) The model hypothesizes that tonic and phasic dopamine activity fulfil distinct modulatory functions, whereby tonic dopamine controls the steepness of the value curve in the loss domain and phasic dopamine controls the steepness of the value curve in the gain domain.

*Figure reproduced under Creative Commons Attribution License (CC BY).

This model offers an intriguing application of the dual dopamine signals in loss aversion (i.e. the steeper curve for losses than for gains). However, its authors make the assumption that risk is equivalent to the proportion of potential losses, and not to the uncertainty, or variance, of the potential outcomes. Thus, the model does not explicitly incorporate dopamine as a risk signal in the economic sense, *per se*. Unfortunately, the authors also refrain from speculating about how dopamine might play a role in the two other key aspects of Prospect Theory, namely: 1) the 's-curved' weighting function (see Figure 2b) and 2) the variable reference-point. Despite its apparent incompleteness, the model offers a good starting point from which to base further investigations.

Note that in the Clark and Dagher (2014) model above, it remains unclear how a given outcome's risk is represented in the system. For example, the model

does not detail how the riskiness of an outcome reduces/increases phasic or tonic dopamine activity. The following section will explore the implications of different model classes on the representation of risk in the brain.

Encoding Risk in the Brain

Two competing models of risky decision-making are utility-based models (see Expected Utility & Prospect Theory), and risk-return models (refer to Risk & Return Models). Utility-based theories assume that decision makers weight the value of various options and then sum the weighted value of all available outcomes in order to decide which option is best. In contrast, risk-return models focus first on the average return of an option and its associated risk, and thereafter undertake a comparison of all options based on risk-corrected mean returns. Often times, the two models' behavioural predictions are very nearly the same, which makes it difficult to determine which one more closely resembles the true neural processes underlying risky decision making—or even if either of the two models are biologically plausible.

Although relatively little is known about the neural encoding of risk, a number of studies measuring hemodynamic responses of risky decisions point to the anterior insula, anterior cingulate cortex, nucleus accumbens, and inferior frontal gyrus as key areas in mediating risk (e.g. Mohr, Biele, Krugel, Li, & Heekeren, 2010; Paulus, Rogalsky, Simmons, Feinstein, & Stein, 2003; Preuschoff, Bossaerts, & Quartz, 2006; Tobler, O'Doherty, Dolan, & Schultz, 2007). The amygdala has also been implicated in risk tracking, but primarily in decision-making under ambiguity when probabilities are unknown (M. Hsu, Bhatt, Adolphs, Tranel, & Camerer, 2005; Rutishauser, Mamelak, & Schuman,

2006). Significant activations typically disappear when probability distributions are learned over time or explicitly stated (Bossaerts, 2010), but nonetheless represent important excitatory inputs for risky decision-making.

Tonic activation of the dopamine neurons within these areas may play an important role in the physiological modulation of risk in expected reward. Fiorillo and colleagues (2003) found that for binary reward outcomes (e.g. reward or no reward), both reward value and reward risk (as variance or entropy) were encoded by monkey midbrain dopamine neurons. Whereas the value of a conditioned stimulus correlated as expected with phasic burst firing, the authors found that sustained tonic responses encoded risk between stimulus onset and reward delivery. This tonic activity was highest when a reward was maximally uncertain. Kuhnen and Knutson (2005) used functional imaging to investigate the neural mechanisms underlying financial risk taking in humans. The authors asked participants to choose between two risky stock options with the potential for both large gains or large losses, or a safe bond option corresponding to a certain but small gain. The study showed that nucleus accumbens activity increased prior to risky choices and risk-seeking mistakes, serving as a predictor of future risky decisions. Further studies have corroborated these findings, suggesting that increased activation of the nucleus accumbens may increase one's tendency to choose options associated with higher risk and greater reward (B. Knutson, Wimmer, Kuhnen, & Winkielman, 2008; Matthews, Simmons, Lane, & Paulus, 2004; Rao, Korczykowski, Pluta, Hoang, & Detre, 2008). In support of this association, selective inactivation of the rat nucleus accumbens has been shown to elicit the opposite reaction. Stopper and Floresco (2011) found that inactivation of the nucleus accumbens shell using

localized microinjections of a dopamine antagonist reduced the tendency of rats to choose large/risky options versus small/safe options. This evidence suggests that dopamine release in the nucleus accumbens is not limited to a singular role of signalling reward prediction error via phasic bursts. Instead, tonic dopaminergic activity in the nucleus accumbens offers a putative signal about the variance, or uncertainty, of receiving an expected reward. The shared presence of these two separable signals in a common neurotransmitter and brain area is in line with the Markowitz (1959) mean-variance approach – although not specifically at the exclusion of utility models. In conclusion, although much work is being done in this area, it remains unclear which class of model (if any) more nearly approximates the brain's integration of risk and reward into a decision value.

Discussion

In the preceding sections, it has been established that decision-making under uncertainty requires an individual to dynamically update and *contextualize* information about the potential magnitude, valence, likelihood, and desirability of an outcome. It is hypothesized that in the brain, this information is integrated into a coherent standardized representation of subjective action value that an individual then maps over all potential outcomes in a preference-weighted manner. However, the subjective treatment of objective variables such as magnitude and probability often make it unlikely that an individual will achieve a utility-maximizing behavioural response, as defined by classic rational models (Expected Utility). For example, a decision maker in a risk-seeking frame (Levin et al., 1998; Tversky & Kahneman, 1981) may preferentially weight the subjective value of an uncertain option to bias its selection over more certain options with equal or greater expected value. It remains unclear how valuation mechanisms within the brain instantiate such behaviour.

While interdisciplinary collaborations in neuroeconomics have made great strides in understanding the neural representations of expectations about rewards and punishments, little is understood about the mechanisms that encode financial losses as well as the mechanisms that allow the prospect of a loss to bias representations of value, risk and ultimately behaviour. With this in mind, the following chapters represent a collection of original research intended to facilitate the study of resource loss on the brain and behaviour.

We sought to achieve a number of specific goals with the research presented in the following chapters. Most generally, we intended to establish the viability of

rat decision-making tasks to neuroeconomics research. The majority of neuroeconomic research is carried out in primates (humans and to some extent in monkeys). This may be due to an implicit assumption that higher order cognitive faculties are required for the economic decisions of interest, or possibly to the difficulties that arise in operationalizing economic decisions. For example, it is not difficult to elicit a mental representation of gains and losses from human participants. One must only signal that a loss (e.g. -$5.00) has occurred within the task. However, eliciting an abstract representation of resource loss is much more difficult in animal work. In general, it is not possible to retract a reward once it has been consumed. Previous researchers (e.g. N. W. Simon, Gilbert, Mayse, Bizon, & Setlow, 2009; Zeeb, Robbins, & Winstanley, 2009) have resolved this issue by substituting punishers (e.g. footshocks) and opportunities costs (e.g. timeouts). Given our lack of knowledge regarding the neural substrates that encode loss, it is unclear whether such substitutes are supported by the same mechanisms as is resource loss in the brain. Therefore, it could be argued that neither pain nor frustrative non-reward represent valid substitutes.

The research presented in the upcoming chapters aims to address this specifically by developing and validating two novel implementations of resource loss. The first operationalization of loss is based on expectations of potential gain, while the second is based on notions of a reference point (see ‘Prospect Theory’) between perceived gains and losses. Once it is established that rats form a representation of loss in each of these tasks, we also aim to determine whether or not rats exhibit similar loss-related behavioural biases (e.g. loss aversion) to humans and primates (Barberis & Xiong, 2009; M. K. Chen et al.,

2006). While rational economic behaviour has been observed in rat tasks (Kagel et al., 1975), suboptimal behavioural biases such as loss aversion and the disposition have never before been explicitly elicited from rats. The preceding review suggests that such behavioural biases may occur due to conflicting neural processes in the dopaminergic midbrain, and that there are marked similarities between the rat and primate midbrains. Therefore, a final aim of this research is to implement the novel rat paradigms to further elucidate neural mechanisms that instantiate suboptimal behaviour in decisions that involve losses.

Chapter 2

A novel gambling task to capture resource loss in rats

Abstract

Previous research has established a role for the dopaminergic midbrain in reward-related behaviour (e.g. reward prediction errors), but evidence implicating the dopamine system in the encoding of losses is conflicting. This critical lack of understanding may be attributable to difficulties in operationalizing resource loss in laboratory tasks – especially those with animal subjects. Thus, we developed a rat gambling task that utilizes a novel operationalization of resource loss in order to facilitate better translations between behavioural and neural research in animals and humans. In the task, thirsty rats (N=29) were trained to sustain a nosepoke for up to a maximum of 2 seconds in order to receive liquid reward. At each 100-millisecond interval of the nosepoke, the total volume of potential reward increased while the cumulative probability of winning that potential reward decreased. Thus, animals decided between longer poke durations for larger uncertain rewards or shorter poke durations for smaller certain rewards over a 0-2 second continuum. Rats also chose between three different contingencies of reward magnitude and probability on free-choice trials. The experimental results indicated that rats predictably altered behaviour to changes in either reward probability accrual rates or reward magnitude accrual rates, which suggests that rats were sensitive to contingency manipulations in the task. Furthermore, rats spent less time poking in error and moving to collect reward on trials immediately preceded by a loss compared to those preceded by a win. This supports the notion that rats adjusted behaviour to compensate for a loss. Rats also alter contingency choice and subsequent stay-shift behaviour after a loss, exhibiting a loss-stay/win-shift pattern of behaviour.

Introduction

With the knowledge that individuals often act as if losses are twice as impactful as equivalent gains (Kahneman & Tversky, 1979), a topic of considerable debate amongst neuroeconomics researchers is whether the brain encodes gains and losses via a single, bivalent system or via multiple competing systems. Evidence for both the single (e.g. Tom et al., 2007) and competing (e.g. Kuhnen & Knutson, 2005) system theories can be found. Methodological differences between human and animal research on loss aversion may contribute to the lack of resolution over this topic. On the one hand, human research benefits from its ability to exact abstract representations of monetary loss from participants, while animal research relies on operationalized notions of loss such as pain (e.g. a foot-shock or an air-puff) or opportunity cost (e.g. a time-out). On the other hand, animal research allows researchers to investigate the neural mechanisms driving behaviour at the level of single neurons, while commonly employed measures of neural activity in humans generally must sacrifice either temporal (e.g. functional Magnetic Resonance Imaging, or fMRI) or spatial (e.g. Electroencephalograph, or EEG) resolution. Given these trade-offs, animal research has been limited in its ability to translate to studies of human loss aversion up to this point, and *vice versa*. Thus, the development of an animal model of risky decision-making that incorporates resource loss represents a critical undertaking in allowing researchers to better understand the neural mechanisms subserving loss aversion.

Previous rodent models of risky decision-making demonstrate limited face and construct validity to human paradigms with regard to their operationalization of

losses. Whereas outcome gains are easily integrated into an animal model by resource gain (e.g. by varying the number of food pellets given to hungry rats), simulating outcome losses has proven to be a critical obstacle given that animals tend to consume outcome gains immediately. For example, rats in the previous risky decision making tasks could gain food pellets, but they could not arrive at an overall loss by the end of testing session, as would be the case if humans were to lose money in a gambling paradigm.

One way in which previous studies have attempted to model loss in rats is by substituting punishers for rewards,[7] such as food pellets saturated with quinine to make them less palatable than sugary rewards (e.g. van den Bos, 2006). Alternatively, some task designs introduce opportunity costs in the form of 'time-outs,' during which time rats cannot work for reward (e.g. Zeeb et al., 2009). In using either punishment or frustrative nonreward, the previous rat models do not incorporate true resource loss, which may confound translations of results into humans. By means of a simplistic example, consider an experiment in rats in which footshocks are employed as a substitute for resource loss. In the event that neural measures taken during the task suggest a significant interaction exists between one system encoding gains and a separate system encoding losses, the experimenters would not be able to rule out the possibility that one system measures value and the other measures pain. While one might argue that the emotional response elicited in the contexts of resource loss and punishment both demonstrate substantial overlap (Prelec

[7] In the reinforcement learning literature, a clear distinction is made between the term 'reward', which is associated with hedonic experience, and the term 'reinforcer', which refers to learning invoked by reward. Both terms are applicable in this task, but we use the term reward in order to generalize to the wider interdisciplinary audience.

& Loewenstein, 1998), one cannot conclude from this that the same mechanisms encoding the (negative) value of a punishment also encode the (negative) value of a financial loss.

However, this is not to say that emotion should or even could be decoupled from risky decision-making tasks. Positive and negative states of affect and arousal play an important role in guiding decision making in many contexts, including financial decisions (Kuhnen & Knutson, 2011). Indeed, tasks that elicit emotionally engaging responses to choice sets and outcomes are often more predictive of naturalistic everyday risky decision making than those that do not (for review, see Schonberg, Fox, & Poldrack, 2011). Thus, while it is important to incorporate affect and arousal in the task, it is still necessary to dissociate them from value.

The Balloon Analogue Risk Task (BART), developed by Lejuez and colleagues (2002), represents one human risky decision-making task that is both particularly emotionally engaging and also reliably predictive of naturalistic risk taking such as stealing, smoking and substance abuse (Bornovalova, Daughters, Hernandez, Richards, & Lejuez, 2005; Lejuez et al., 2007; Lejuez et al., 2003). In this task, participants are asked to pump up a series of virtual balloons with the goal of cashing out before a given balloon 'pops'. The participant accrues reward with each successful pump of the balloon, but will lose any accrued reward in the event that the balloon explodes. Different coloured balloons represent differing probabilities of popping as the balloon is pumped up, which must be learned over the course of the session. The suspense of increasing pumps and the surprise of an exploding balloon

naturally elicit affective engagement from participants. Jentsch and colleagues (2010) adapted this idea in a rodent BART, but with an additional dimension of uncertainty. Rats were presented with two levers, one that added an increment of reward and one that cashed out for reward delivery. The researchers varied the risk of losing a trial with the 'add' lever and the 'cash out' lever. In other words, the cash out lever was not risk-free, as is the case with the human task. Therefore, in the decision moment, subjects did not always have the dilemma of taking reward in hand versus potential increased reward. Furthermore, each additional bar press resulted in only one unit increase in reward, which meant that rats had little incentive to continue bar-pressing beyond the first press. Jentsch and colleagues found that inactivation of the rat mPFC increased the variability of responding in the task, leading to increased suboptimal behaviour, while OFC inactivation decreased response rates altogether. However, the authors did not specifically contrast behaviour after a gain with behaviour after a loss in either the baseline or inactivation conditions. Therefore, it is unclear how losses affected subsequent behaviour within the task.

In the rat decision-making model carried out in this paper, strategic choices were based on manipulations of reward probability and magnitude. Similar to the BART, losses were operationalized as the omission of any accrued reward up to the point of an unsuccessful gamble. To implement this, we trained thirsty rats to nosepoke in a standard operant testing chamber to earn sweet liquid reward. The longer a rat poked, the greater the potential volume of reward, but also the greater the probability the reward would be cancelled and it would receive nothing. In essence, at each moment during the nosepoke the rat is faced with a dilemma, for it could hold the poke longer to earn more reward, but

in so doing it also risks losing the reward already earned in the trial. This is the equivalent of the decision to 'let a bet ride' in human gambling contexts and may elicit a closer decision-making scenario to that of the human BART than other rodent versions of the task. The specific goal of this design was to elicit an internal representation of the volume of reward that was lost (i.e. loss encoding). It is possible that such a representation would also generate a negative emotional state (Kuhnen & Knutson, 2011) as well as a fictive reward prediction error[8], i.e. a reward prediction error for an unchosen option (Boorman, Behrens, & Rushworth, 2011).

Within the task, rats were faced with two types of decisions involving uncertainty: 1) the decision among nosepoke holes associated with varying rates of reward magnitude and probability, and 2) the decision at each 100 msec interval of a poke to either unpoke and collect accrued reward or to continue poking for the chance at accruing more. Once rats' performance had stabilized, we made the following hypotheses regarding nosepoke durations and hole choices on free-choice trials in the task:

1. Each response will be contingent on manipulations of expected reward via probability of loss/reward-accrual rates, where risk-aversion will limit both poke durations and choice of the high-risk/reward contingency.
2. Following a loss, rats will change behaviour in a way that is consistent with loss aversion in repeated gambles (Heilbronner & Hayden, 2013; Kahneman & Tversky, 1979): increasing risk-seeking behaviour by poking longer and choosing the high-risk/reward contingency more often.

[8] *It should be noted that there is a psychological distinction between learning from fictive prediction errors that represent options that were not chosen (regret) vs. options that were chosen but not obtained (disappointment), and that these two likely have distinct neural substrates (for review, see Platt & Hayden, 2011).*

Methods

Animals

Subjects were 29 male outbred Lister Hooded rats (Harlan U.K.) that were housed in groups of four in a climate-controlled colony room on a reverse 12-hour light: 12-hour dark cycle (6PM lights off). Baseline data are pooled from two cohorts of 16 and 13 rats with similar training protocols. The discrimination task and pharmacological manipulation with *cis*-Flupenthixol described below were conducted with the cohort of 16 rats only. After three weeks of habituation to experimenter handling, rats were placed on restricted water access for the duration of behavioural training and testing with *ad libitum* food access in the home cage. Rats were tested 5 days a week. Water access was restricted to 1 hour on weekdays following testing, but was available *ad libitum* on weekends from Friday at 4PM until Sunday afternoon (typically between 2-4PM). Rats' weights were monitored so that no animal dropped below 85% of its maximum body weight and showed growth throughout the experiment. All procedures were carried out under the Project License number 60/4040, conformed to the United Kingdom Animals (Scientific Procedures) Act (1986) and were approved by the Animal Welfare Ethics Committee of the University of St Andrews.

Apparatus

Testing was carried out in four 34mm × 29mm × 25mm Perspex inner chambers with metal bar flooring that were located within 60cm × 74cm × 55cm sound-attenuating outer shell boxes (Med Associates Inc., St Albans, VT) with closed caption video cameras and ventilation fans. The right metal wall of the inner testing chamber contained five square nosepoke holes, each

accommodating a recessed green LED light as well as an infrared sensor to record nosepokes. A recessed custom-built liquid reward magazine delivering 0.3% w/v sodium saccharin solution at a rate of 0.05 ml/sec was located on the left metal wall of the inner testing chamber. The reward spigot was fitted with a lickometer (Med Associates Inc., St Albans, VT or Weignen 1989) to record licking behaviour, as well as a white LED (approximately 2072 mcd luminosity) and a piezoelectric buzzer (2900Hz, 85dB) to signal reward availability. Two electronically controlled syringe pumps (model PHM – 100, Med Associates Inc., St Albans, VT) dispensed liquid from 50 ml glass syringes with stainless steel plungers (Rocket Medical plc, Herts, U.K.) and an 18-gauge needle connected to the reward spigot by Teflon tubing. This setup allowed for precision in the timing and flow rate of reward delivery.

Behavioural testing was interfaced by the MED-PC™ data experimental control system (Med Associates Inc., St Albans, VT) with an IBM® computer running Windows™ 98 at a temporal resolution of 2 msec. Summary measures were also available in an online display on the computer screen along side real-time video feeds. Behavioural events were also time-stamped (2 msec resolution) and recorded for offline data analysis and session reconstruction.

Task Design

After three weeks in which rats were habituated to human handling, the animals were placed on water restriction and submitted to 30 min training sessions in the testing chambers with no fixed trial limit. First, thirsty rats were trained over 2 days to associate a tone-light cue with delivery of 0.15 ml sweet liquid reward (sodium saccharin 0.3% w/v). Rats were subsequently trained over two

sessions to nosepoke in the three middle nosepoke holes of the five-hole array in order to receive the tone/light cue followed by the reward. Rats were then trained to sustain gradually longer nosepokes for up to 2 sec over 18 sessions.

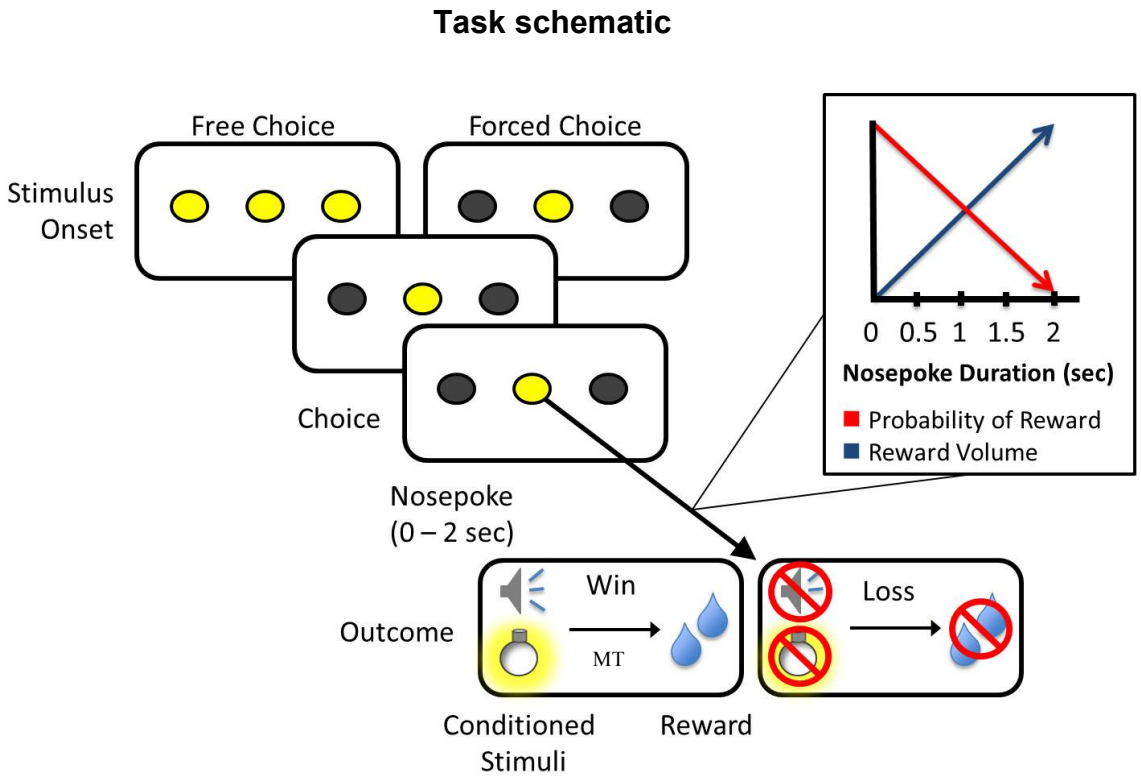


Figure 9: Task Schematic. Rats were presented with blocks of trials in which forced- and free-choice trials were pseudorandomly interleaved in a proportion of 3:1. Lit LEDs indicated the available hole(s) in which a rat could poke on any given trial, and the locations of the holes associated with each contingency were counterbalanced across testing chambers. All LEDs in nosepoke holes were extinguished after the rat began poking into one of the lit nosepoke holes, but no changes in stimuli occurred if an animal poked into an unlit hole (error). If the animal either withdrew its snout from the nosepoke hole before a loss, or it successfully reached the full 2 sec limit without losing, a tone-light cue emanating from the reward magazine would indicate the availability of reward for collection. Movement time (MT) to reward was measured from the onset of the conditioned stimuli to licking onset at the reward spigot. The amount of reward earned was a function of nosepoke length (refer to plot at top right of figure), and the animal could commence licking at the reward spigot until the full amount was delivered. If a loss occurred during the course of a nosepoke, both cue and reward were omitted. A 15 sec timer was activated from the onset of a nosepoke response, and a trial ended either when the timer elapsed or with the end of a lick bout (an inter-lick interval > 300 msec) after reward delivery, in which case a new trial began immediately. The animals were free to complete as many trials as possible over the course of the 30-minute session.

Figure 9 depicts the risky decision making task, in which rats could earn reward for any duration of nosepoke in a lit hole that lasted between 0.1 and 2 sec. The amount of reward delivered after a successful nosepoke depended on the

rate of reward accrual corresponding with the given nosepoke hole. Each nosepoke hole was associated with a rate of risk and reward accrual so that increasing nosepoke durations (max 2 sec) resulted in greater volumes of potential reward, but also a greater likelihood that that reward would be cancelled. Therefore, at each moment during a poke, a rat must consider whether to continue poking for more reward (akin to 'letting the bet ride'), or to unpoke and keep any reward that it had accrued up to that point (akin to 'cashing out'). The rats chose between a high risk/high reward hole, a medium risk/medium reward hole, and a low risk/low reward hole, depending on which holes were lit and therefore available on any given trial. Poking into unlit holes was counted as an error. Rats had two ways of adjusting their behaviour in order to maximize reward within the task: 1) they could vary the duration of a nosepoke in order to maximize reward volume and minimize the probability of loss on a given trial, and 2) on free-choice trials, they could choose the nosepoke hole with optimal contingencies of reward probability/volume based on previous experience. As in training, testing sessions lasted 30 minutes with no fixed trial limit.

The task was constructed so that even the steepest reward discounters would experience a discernable trade-off between the rate of reward accrual and the decreasing likelihood of receiving accrued reward. For example, a rat poking only until the first 100ms tick in the hole associated with low rate of reward accrual and low probability of losing had a 99.5% chance of success, whereas the probability of success dropped to 89% at the first 100ms tick in the hole associated with a high rate of reward accrual and high probability of losing. The expected value of poking in each contingency is displayed in Figure 10.

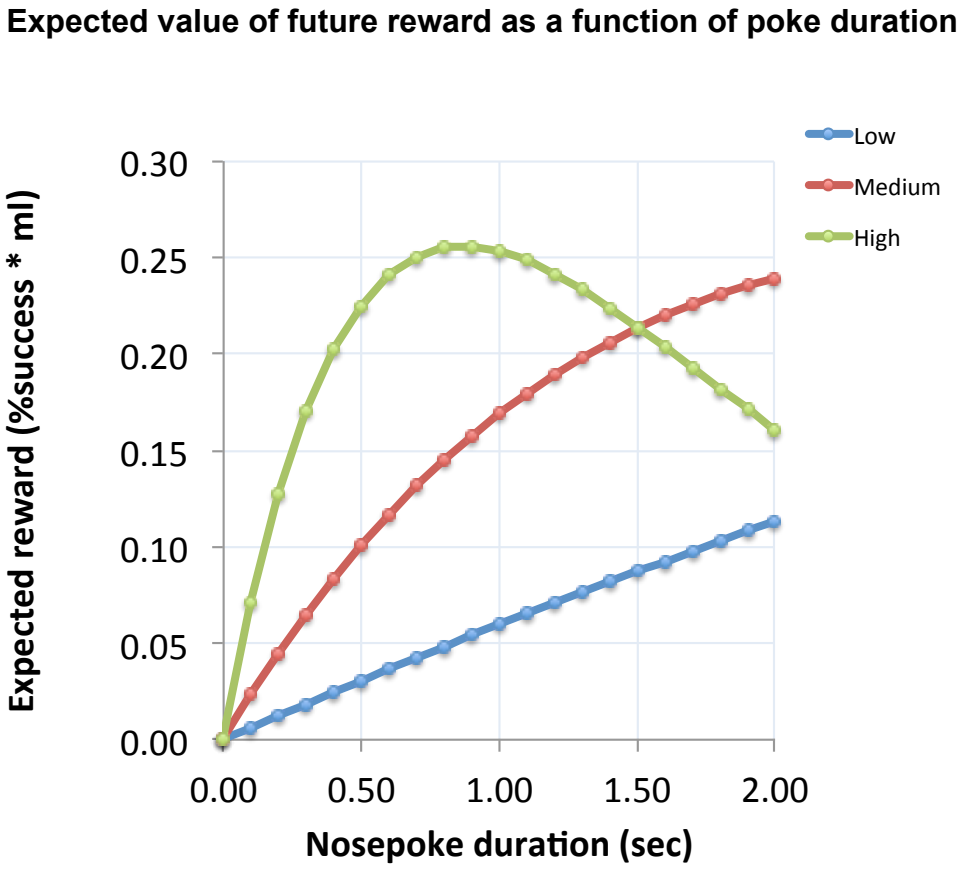


Figure 10: The expected value of future reward as a function of poke duration in stage III (mixed probability and magnitude of reward). In the low contingency, a rat would expect a low rate of reward accrual but also a low probability of losing over the duration of the nosepoke. Thus, one might expect a rat to poke for the full 2 sec duration in the low contingency in order to maximize expected reward. In the high contingency, the expected likelihood of losing outweighs any potential increase in reward accrual just before a 1 second poke duration. Thus, one might expect a rat to unpoke on average at about 900 msec. The medium contingency falls between the low and high contingencies.

Discrimination task

In order to confirm that rats were able to discriminate variation in rates of reward accrual and probability of winning between the contingencies, we manipulated the contingencies so that either only reward probability (stage I), only reward volume (stage II), or both reward probability and reward volume (stage III) varied across the holes in a subset of 15 rats. Stages I and II each took place over 15 sessions, and Stage III lasted 26 sessions.

Mixed contingencies

After the discrimination task was carried out, all testing was performed with contingencies of mixed probability of reward and volume accrual rates. For ease of comprehension, these contingencies will be referred to as:

- **Low Contingency**: Low reward accrual rate and low probability of losing
- **Medium Contingency**: Medium reward accrual rate and medium probability of losing
- **High Contingency**: High reward accrual rate and high probability of losing

Behavioural measurements

Baseline data were from the last five days of stable performance. Given that rats were not limited in the number of trials that they were able to complete in a session, we calculated percentage choice (rather than absolute number of choices) using the number of trials chosen in a given contingency over the 5-day period divided by the total number of free-choice trials over that period. The percentage of stay/shift trials was calculated as the total number of decisions to stay/switch after a previous trial across the 5 days, divided by the total number of free-choice trials across the 5 days. The following variables were also measured and analysed per rat per trial separately across conditions: error rate per forced-choice trial, time spent (sec) in incorrect nosepoke holes, lick rate (Hz), and movement time to reward (sec).

Data Analysis

Session reconstruction with time-stamped data was performed using a program written by EMB in AWK programming language. Subsequent data analysis was carried out using Microsoft® Excel for Mac 2011 as well as R version 3.2.2 and

SPSS® version 21 for Mac. Discrimination task behaviour was analysed using the average percentage each contingency was chosen (on free-choice trials) across rats and days during the last five testing sessions in each stage. Repeated-measures ANOVAs were performed with contingency (three levels: high, medium, low) as the within-subject variable. The medium contingency was often utilized as a point of contrast between the low and high contingencies, and therefore while it is omitted from many graphical depictions of the data, it is always included in the underlying analysis. Baseline behaviour was analysed using repeated measures ANOVAs with contingency (three levels: high, medium, low) and session (5 levels: five testing sessions) as within-subject variables. Baseline performance was considered to be stable once there was no significant main effect of session over the previous five days of testing (Figure 12a). Greenhouse-Geisser adjusted degrees of freedom and Sidak-corrected p-values were applied where appropriate. All means are reported with standard errors and any significant main effects are reported with associated planned contrasts.

In order to avoid potential ceiling-effects using proportion data, arcsine transformations were used on all variables expressed as a percentage (Zeeb et al., 2009), although data are shown as raw values. Missing data were replaced with series means. Given that average poke durations could be biased by truncated loss trials, descriptive statistics and ANOVA's including average poke duration are calculated based on successful trials only. All analyses measuring responses to a previous win or loss are defined as those trials immediately preceded by a win or loss, omitting the first trial of a session.

We also conducted survival analyses of poke duration and contingency choice by fitting a Cox proportional hazard model to the data. This analysis is carried out on a trial-by-trial basis, rather than averaging across subjects and sessions. The Cox proportional hazard model is a semi-parametric model that makes no assumption about the shape of the baseline curve (e.g. linear), and takes the following form:

$$h\big(t, x(t)\big) = h_0(t)\exp\big(\beta_1 x_1 + \cdots + \beta_p x_p\big)$$

where the hazard rate of x occurring, h(t,x(t)), on trial t is conditional on p predictors. The β coefficients are estimated from the data. The associated survival function represents the cumulative proportion of the sample that has not experienced an event x by time t. Alternatively, this can be understood as the probability that an event will not occur until time t.

While the model's primary assumption is that the hazard associated with any given covariate is proportional across time, it can be extended to incorporate time-varying covariates and stratified to accommodate within-subject designs. By stratifying across subjects, the models fit here include individual baseline hazards for each animal, which accounts for the variance in survival rates contributed by individual subjects.

Results

Discrimination Task

Rats demonstrably altered choice behaviour on free-choice trials in stages I - III (Figure 11). Rats' choices minimized losses in stage I by choosing the low risk

nosepoke hole (Contingency: $F_{(2,8)}$=576.80, p<.001, η_p^2=.99), maximized reward in stage II by choosing the high reward-accrual nosepoke hole ($F_{(2,8)}$=79.50, p<.001, η_p^2=.95), and appeared to prefer a trade-off of high risk for high reward in stage III ($F_{(2,8)}$=262.00, p<.001, η_p^2=.99). It is of note that stage III most nearly resembles stage II (reward maximizing) rather than stage I (risk minimizing). These results confirmed that the rats were sensitive to manipulations of reward accrual and risk in the task.

Contingency choice during manipulations of probability, reward, and mixed probability & reward accrual rates

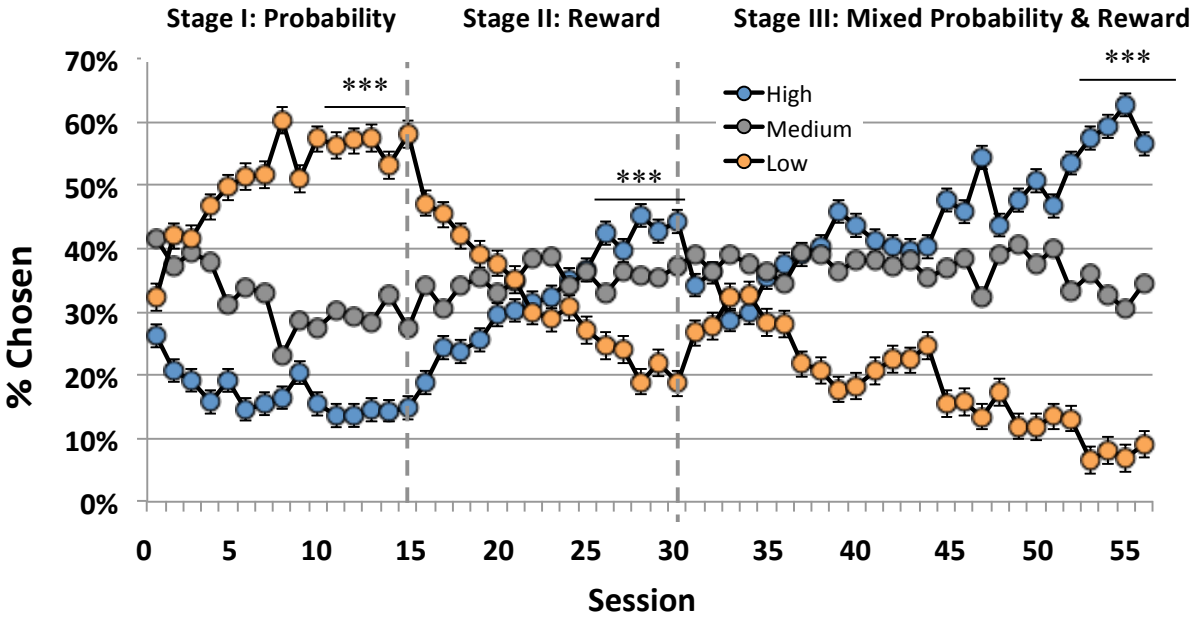


Figure 11: Rats (N=16) predictably and reliably preferred the low-probability of losing (orange circles) contingency in stage I and the high-reward contingency (blue circles) in stage II. Although there was some initial uncertainty in responses to the mixed risk and reward in stage III, rats quickly began to prefer the contingency with the trade-off of higher risk for higher reward (blue circles). Error bars represent SE. *p<.05, **p<.01, ***p<.001.

Baseline Behaviour – Mixed Risk And Reward

On free-choice trials, rats (N=29) demonstrated consistent choices across the three contingencies (Figure 12a), with a clear preference for the high contingency (Contingency: $F_{(2,56)}$=38.28, p<.001, η_p^2=.58). We found no

significant main effect of Session on the proportion of contingency choices across the last the last 5 days of testing ($F_{(4,112)}$=1.76, p=NS), which is suggestive of stable baseline behaviour. Average poke duration (Figure 12b) varied detectably among the three contingencies ($F_{(2,56)}$=3.70, p<.05, η_p^2=.12), whereby animals tended to poke slightly longer in the low (568 ± 77 msec) and medium contingency (574 ± 67 msec) compared to the high contingency (501 ± 64 msec). However, post-hoc tests revealed no significant difference between the high and low contingencies (p=.25). Average poke durations did not significantly differ across sessions ($F_{(4,112)}$=2.17, p=NS), indicating that the task had elicited consistent nosepoke behaviour across days.

Similarly, the rate at which rats licked the reward spigot was fastest in the high contingency (7.94 ± 0.15 Hz) and slowest in the low contingency, although this effect is weak (7.50 ± 0.24 Hz; Contingency: $F_{(2,56)}$=5.14, p<.05, η_p^2=.16, Session: $F_{(4,112)}$=2.28, p=NS). It is of note that all volumes of reward were delivered at the same rate on every trial (i.e. larger rewards equated to longer delivery times, see Methods), and therefore it was not strictly necessarily for animals to adjust lick rates according to reward volume, *per se*.

Baseline differences in behaviour between the high and low contingencies

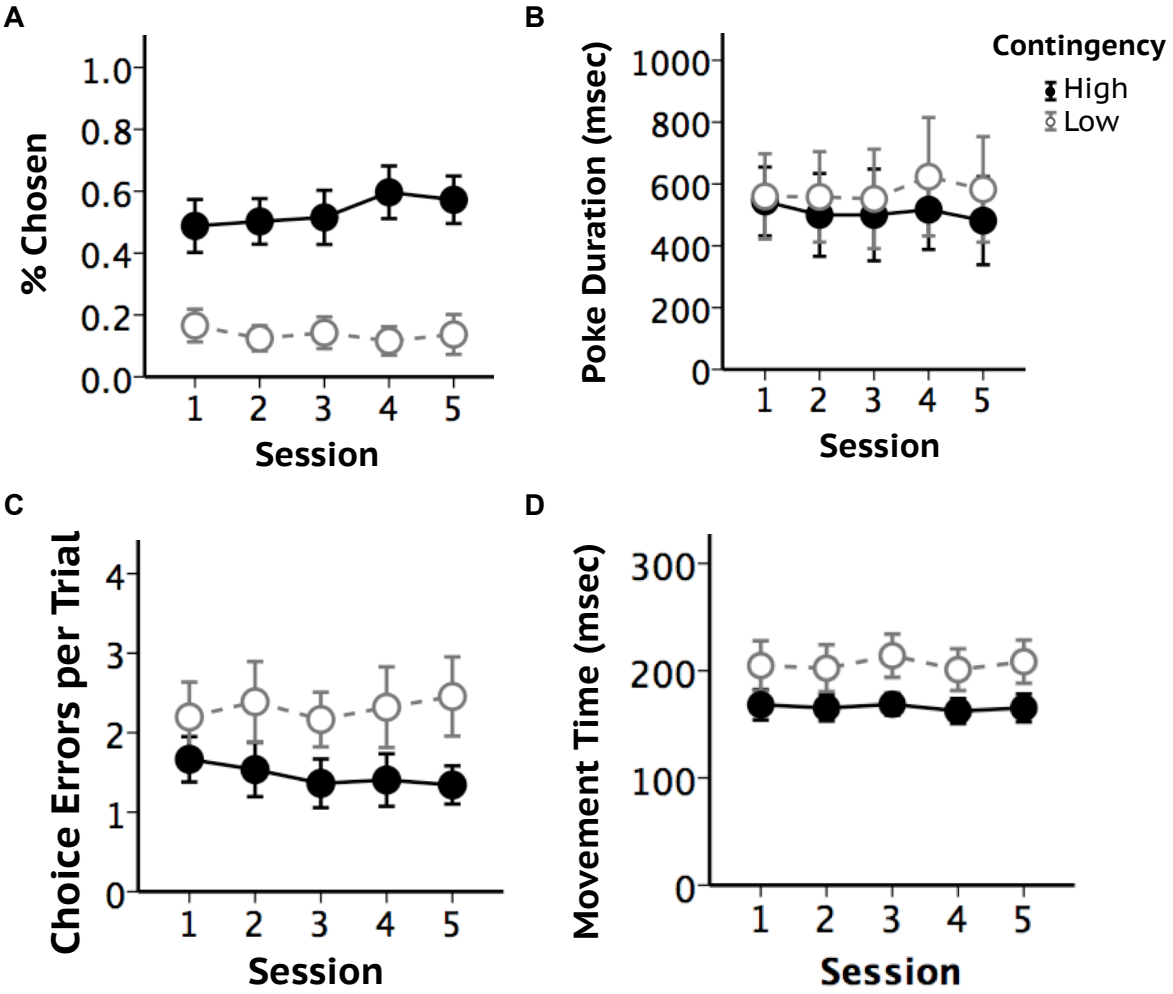


Figure 12: (**A**) Rats consistently preferred the high contingency on free-choice trials, despite the relative certainty of receiving reward in the low contingency. (**B**) Poke durations tended to be shorter in the high contingency, although this difference was not significant. (**C**) Rats made fewer errors on forced-choice trials and (**D**) moved more quickly to collect reward in the high contingency. Error Bars are 95% CI's.

On forced-choice trials, animals poked in error in an unlit hole (Figure 12c) about 1.5 times more often (2.31 ± 0.21 pokes/trial) during trials in the low contingency compared to the medium (1.68 ± 0.17 pokes/trial) and high (1.46 ± 0.13 pokes/trial) contingencies (Contingency, $F_{(2,56)} = 27.79$, $p < .001$, $\eta_p^2 = .50$), as though the rats were experiencing frustrative nonreward. This behaviour was stable throughout the last five baseline testing sessions (Session, $F_{(4,112)} = 1.95$, p = NS). As depicted in Figure 12d, movement time to collect reward also decreased monotonically as the level of reward accrual/probability of losing

increased, with rats moving significantly faster to collect reward in the high contingency and slower in the low contingency ($F_{(2,56)}$= 23.92, p<.001, η_p^2=.46).

Loss-related behaviour

Our primary topic of analysis focused on behavioural responses to losing in the task. We began by establishing whether or not rats' behaviour was significantly altered by a loss vs. a win in the task on the aggregate level, and we proceeded by conducting trial-by-trial analyses of any potential behavioural response strategies with respect to contingency choice (e.g. win-stay/lose-shift).

To begin, we compared behaviour on trials that were immediately preceded by a win trial vs. loss trial. To identify whether rats differentiated between wins and losses in the task, paired-sample t-tests were carried out by previous outcome on the following three behavioural measures: Lick rate (Hz), MT (sec), and time spent poking in error (sec). A visual comparison of these analyses is provided in Figure 13.

We found that two of the three measures reflected significant differences in behaviour on trials immediately preceded by a loss compared to those preceded by a win. Rats moved significantly more quickly to collect reward after a loss than a win (MT: $t(28) = 3.14$, $SEM = 0.05$, $p < .01$). Animals also spent more time poking in incorrect (i.e. unlit) holes on trials with a previous loss outcome vs. those with a previous win outcome (Error Time: $t(28) = 4.32$, $SEM = 0.06$, $p < .001$). However, we found no significant difference in lick rate between the two trial types (Lick Rate: $t(28) = 1.10$, $SEM = .05$, $p = .28$).

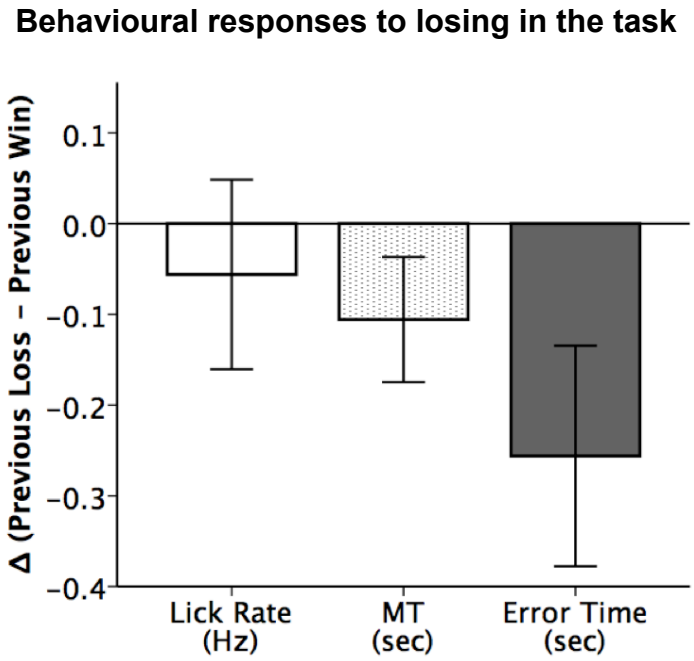


Figure 13: Behavioural responses to losing in the task. The difference between trials that are immediately preceded by a win and those preceded by a loss are displayed for three behavioural measures: Lick rate (Hz), Movement Time (MT, sec), and error time (sec). Negative numbers signify a reduction in the measure on a loss trial vs. win trial. Rats significantly reduced both MT (p<.01) and the time spent poking in error (p<.001) after a loss. There was no significant difference in lick rate after a loss vs. a win (p=.28). Error bars represent 95% CI's.

Stay-shift behaviour

We next analysed whether or not rats' choices followed a strategy that was contingent upon the previous outcome, such as a win-stay/lose-shift pattern. Here, we began by conducting a simple paired-sample t-test on the proportion of stay trials after a win vs. loss. As depicted in Figure 14a, the results indicated that rats' decision to return to the previous contingency was significantly affected by the previous outcome, $t(28) = 3.02$, $p < .01$. After a win, rats perform right around chance, choosing to stay on 32.4% (*SEM* = 1.4%). After a loss however, rats are 10% more likely to stay compared to a win. In other words, rats develop a win-shift/lose-stay strategy.

It was also possible that rats shifted more away from some contingencies than others. Therefore, we next analysed whether or not rats stayed/shifted more as

a function of the previous contingency. Figure 14b illustrates the results of a repeated-measures ANOVA of the percentage of stay trials with Previous Contingency (3 levels: low, medium, and high) as a within-subject factor. A rat's decision to return to the previous hole on a free-choice trial was strongly affected by the contingency associated with that hole ($F_{(2,56)}$ = 34.79, $\eta_p^2 = .55$, Greenhouse-Geisser adjusted $p < .001$). If the previous hole was the high contingency, rats stayed on 51% of trials. Compared to the high contingency, rats stayed on average 16.7% (*SEM* = 5.1%) fewer trials in the medium contingency ($p < .01$) and 37.5% (*SEM* = 5.2%) fewer trials in the low contingency ($p < .001$).

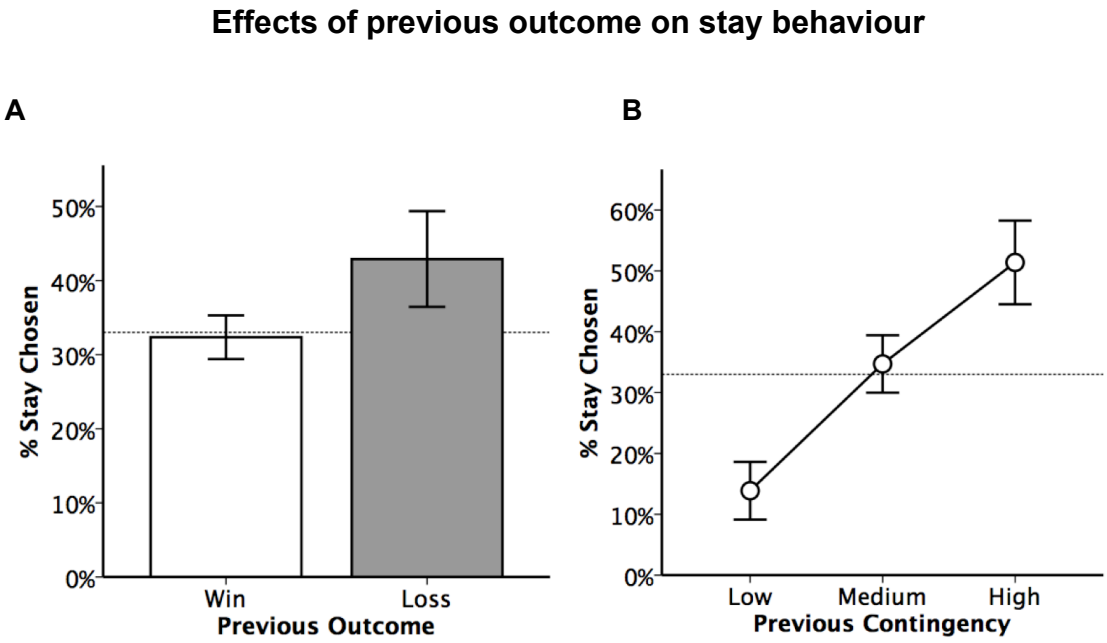


Figure 14: (**A**) In contrast to a typical win-stay/lose-shift strategy, we find that rats exhibit win-shift/lose-stay behaviour. Here, rats return to the same contingency about 10% more often after a previous loss than a previous gain (p<.01). (**B**) Whereas rats rarely returned to the low contingency after a loss (13.9% ± 2.3% of trials), this proportion increased linearly as the rate of risk/reward-accrual associated with each contingency increased. The linear trend was significant: $F_{(1,28)}$ = 10.62, $\eta_p^2 = .28$, p < .01. The dotted lines denote chance at 33.3. Error bars represent 95% CI's.

Given this unexpected behaviour, we sought to establish whether the lose-stay/win-shift strategy was attributable to rats' preference for the high

contingency. Given that rats had a higher probability of losing in the high contingency, a preference to return to the high contingency would naturally lead to lose-stay/win-shift behaviour. In order to establish whether lose-stay/win-shift behaviour was attributable to the previous contingency, we fit a Cox hazard model to the data based on the number of free-choice trials that occurred between 'stay' choices. For example, if a rat chose to switch on the first two free-choice trials and then to stay on the third, the analysis would model the likelihood of a rat's stay choice 'surviving' to the third consecutive free-choice trial, given any number of covariates (e.g. previous win/loss). If the rat chose to stay on a trial, the counting process would reset to zero. The model was stratified over subjects, and previous outcome (2 levels: win and loss) and previous contingency (3 levels: low, medium, and high) were entered as time-varying covariates in the model.[9] A PreviousOutcome*PreviousContingency interaction term was also added to the model. The reader is referred to Table 4 of Appendix 1 for further particulars of the model coefficients.

[9] *Due to their ability to change over the course of the counting process, these covariates would normally not satisfy the assumption of proportional hazards over time. A typical method of remedying this is to multiply each event by the log of time, represented in this model as the trial count.*

Survival curves for stay behaviour as a function of previous contingency

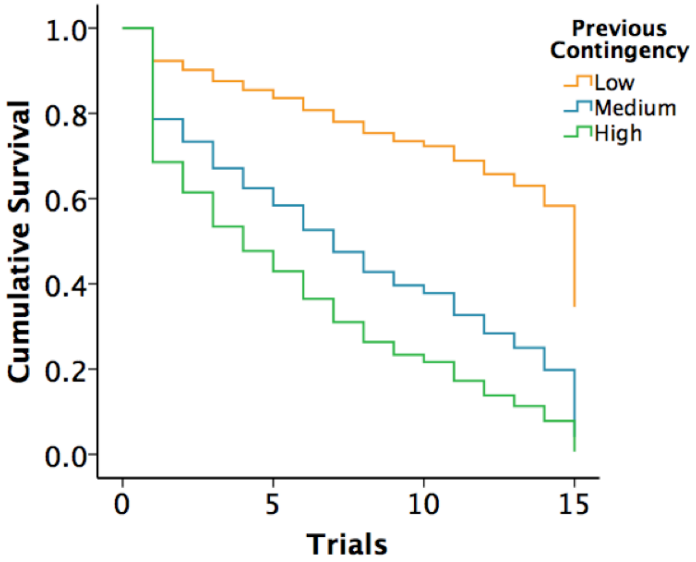


Figure 15: Survival curves for each contingency on free-choice trials are depicted above. Previous contingency was a significant predictor (p<.001) of stay 'survival', or the number of free-choice trials that elapsed before a rat was likely to return to the same contingency as the previous trial. The high contingency is associated with the steepest survival curve, which indicates that rats are more likely to choose to stay sooner in that contingency (p<.001) in contrast to the medium contingency. The relatively flat curve associated with the low contingency (p<.001) indicates that rats are only 40% likely to return to that contingency by the 15th free-choice trial.

As Illustrated in Figure 15, Previous Contingency was a significant predictor in the model (p < .001). In contrast to the medium contingency, rats were much more likely to return to the high contingency on a free-choice trial (p < .001), and much less likely to return to low contingency (p < .001). Therefore, the results of this analysis corroborate previous findings from the analysis of stay trial proportions (Figure 14b).

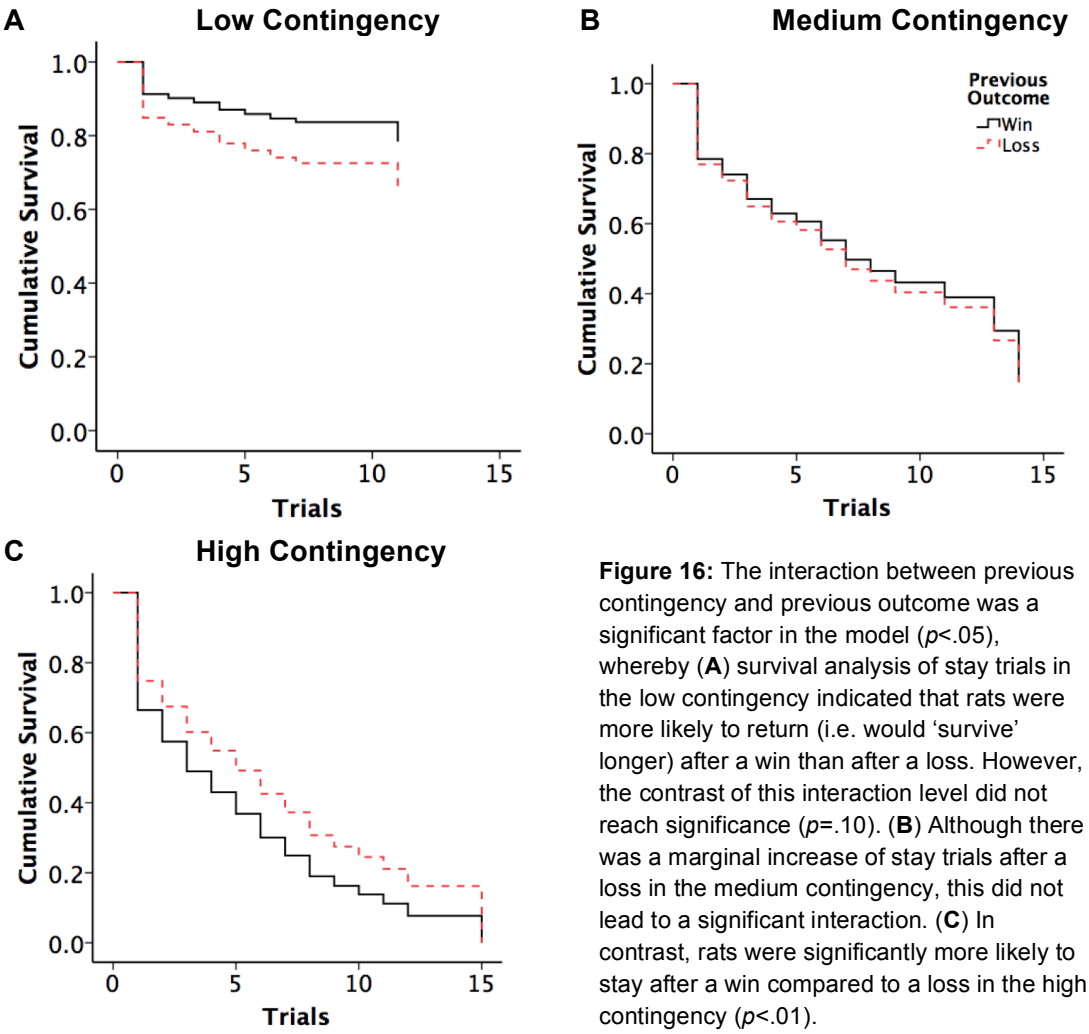


Figure 16: The interaction between previous contingency and previous outcome was a significant factor in the model (p<.05), whereby (**A**) survival analysis of stay trials in the low contingency indicated that rats were more likely to return (i.e. would 'survive' longer) after a win than after a loss. However, the contrast of this interaction level did not reach significance (p=.10). (**B**) Although there was a marginal increase of stay trials after a loss in the medium contingency, this did not lead to a significant interaction. (**C**) In contrast, rats were significantly more likely to stay after a win compared to a loss in the high contingency (p<.01).

While Previous Outcome did not emerge as a significant predictor in the model, there was a significant effect of the interaction term between Previous Outcome and Previous Contingency on stay behaviour (p < .05). Figures 16a-c depict the survival curves of staying on free-choice trials with a previous win vs. a previous loss in each different contingency. The analysis suggests that rats tended to increase stay behaviour after a loss in the low and medium contingencies, while decreasing staying more after a loss in the high contingency. However, only the contrast of the interaction at the high contingency was significant (p < .01), and therefore trends in the other two contingencies should be interpreted with caution. However, these results

suggest that loss-stay behaviour cannot be entirely explained by rats returning to the more preferred high contingency after a loss.

It should also be noted that the relatively low number of stay trials with a previous loss in the low contingency likely contributed to low power in modelling an interaction factor as depicted in Figure 16a. Of the 3525 free-choice trials used as data in the model, only 12 trials (0.34%) represent stay trials after a loss in the low contingency. This number increases to 52 trials (1.48%) in the medium contingency and to 172 trials (4.88%) in the high contingency.

Effect of losing on poke duration

Our final analysis of baseline loss-behaviour was carried out in order to determine whether a previous loss affected the duration a rat was willing to poke on a current trial. We performed a 2-way repeated-measures ANOVA on poke duration with Previous Outcome (2 levels: win and loss) and Contingency (3 levels: High, Medium, and Low) as within-subject factors. This analysis revealed that Contingency had a significant main effect on poke duration, $F_{(2,56)}$ = 5.48, η_p^2 = .16, Greenhouse-Geisser adjusted p = .01. Planned contrasts indicated that this significant effect was attributable to differences in poke duration between the high and medium contingencies (p < .01) rather than the low and medium (p = .12) or low and high (p = .96). The effect of Previous Outcome on poke duration fell short of significance ($F_{(1,28)}$ = 3.97, p = .06). Nor did we find a significant interaction between Previous Outcome and Contingency, $F_{(2,56)}$ = 1.14, p = .33.

We furthered this analysis by fitting a Cox proportional hazards model to the data. This analysis has the potential for greater power given that that: 1) it can account for censored poke durations (i.e. premature unpokes due to losses) and, 2) it is computed on a trial-by-trial basis. Therefore, we fit a model based on poke duration with Previous Outcome (2 levels: win and loss) and Contingency (3 levels: low, medium, and high) as time-varying covariates. For example, if the rat nosepoked for a duration of 1.2 sec, the analysis would model the likelihood that any given poke would 'survive' from 0 to 1.2 sec, given any number of covariates (e.g. contingency). The model was stratified over subjects. The model survival curves for each contingency are illustrated in Figure 17.

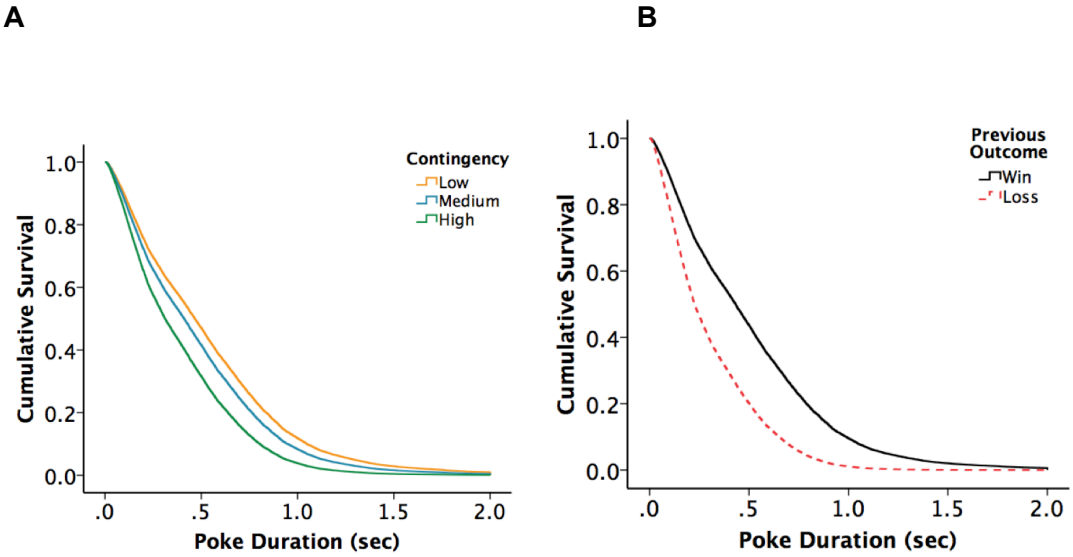


Figure 17: (**A**) Poke duration significantly varies as a function of contingency (p<.001). In contrast to the medium contingency, rats poke longer in the low contingency (p<.01) and shorter in the high contingency (p<.001). (**B**) Poke durations are significantly shorter on trials immediately preceded by a loss compared to those immediately preceded by a gain.

Contingency was found to be a highly significant factor in the model (p < .001), with longer expected poke durations in the low contingency (p < .01) and shorter expected poke durations in the high contingency (p < .001) in contrast

to the medium contingency. Note that this result (Figure 17a) differs to that of the ANOVA illustrated in Figure 12b. This is likely due to the increased power of the Cox proportional hazard model, which arises from its ability to incorporate censored data (i.e. both wins and losses). Previous Outcome also emerged as a significant factor in the model ($p < .01$), whereby poke durations were shorter on trials immediately preceded by a loss compared to a win. Although the contrast between previous outcome and the low contingency was significant ($p < .05$), the main effect of the interaction term was not significant ($p = .14$). A table detailing further particulars of the model coefficients can be found in Table 3 of Appendix 1.

Discussion

The research presented here has been carried out with the specific intention of increasing the translational validity of animal decision-making research to work on risk-taking in humans. We have developed a novel task (Figure 9) that operationalizes loss in a way that is on par with human tasks (e.g. the BART) that are highly predictive of naturalistic risky decision-making, such as the use of illegal substances (Bornovalova et al., 2005). We demonstrate that rats are sensitive to manipulations of reward volume and probability, and that losses significantly affect subsequent behaviour in the task. Interestingly, we find that rats develop a lose-stay/win-shift strategy in the task, and that this strategy cannot be fully explained by preference for returning to the high contingency (which is associated with a greater likelihood of losing).

We began by systematically manipulating either the magnitude or the probability of potential reward a rat could receive by nosepoking in one of three nosepoke holes. Although perhaps unsurprising given the rich history of operant work in rats (e.g. Chung & Herrnstein, 1967), we found that rats' choice behaviour between the nosepoke holes was exquisitely sensitive to our contingency manipulations of reward probability and magnitude (Figure 11). Rats predictably chose the low-risk contingency when we manipulated the probability of losing only (Stage I) and the high-reward contingency when we manipulated the magnitude of reward only (Stage II). This suggests that rats learned how to adjust their choices to maximise reward given these contingency variations. When both risk and reward were varied simultaneously,

rats preferred trading off a higher probability of loss for a higher magnitude of reward.

We predicted that rats' choice allocation between contingencies would be a reflection of individual preferences for the trade off between probability of loss and reward. For example, more risk averse rats would choose the low contingency more often than more risk seeking rats. Instead, we found that choice of the high contingency remained high for all rats throughout baseline testing (Figure 12a) despite high loss rates, and that choice of the low contingency remained very low despite very low loss rates. The observed choice allocation reflects expected reward rates in each contingency with respect to rats' rather short average poke durations (see Figures 12b and 17a). At poke durations of 500 – 600 msec, the expected reward was highest in the high contingency, followed by the medium and then the low contingencies. Instead of choosing the contingency with the highest rate of reward at that nosepoke, rats 'matched' their allocation of choices to the rate of reward in each contingency. Such deviations from optimal reinforcement maximization behaviour (i.e. not allocating 100% of choices to the option with the highest reward rate) have been observed previously in a number of tasks, species, and environments in an effect typically referred to as the 'matching law' (Chung & Herrnstein, 1967). This effect describes the robust tendency of an individual's choice allocations to reflect environmental contingencies of reinforcement rate rather than to reward maximization (Herrnstein, 1990). The model in Figure 10 suggests that, given a 500 msec poke, the matched allocation of contingency choices should be: 64.3% in the high hole, 28.6% in the medium hole, and 7.1% in the low hole. Rats' behaviour matches this very closely. Had rats

averaged longer nosepoke durations (e.g. the maximum 2 seconds), then one would predict the greatest proportion of choices in the medium contingency, followed by the high and then the low. Thus, rats appear to have matched choices with the reward rates at their preferred nosepoke duration rather than to allocate choices to maximise reward or according to their individual risk preferences.

We also predicted that rats would vary nosepoke lengths to maximise expected reward within each contingency, but also as a function of risk preference with shorter nosepokes reflecting risk-aversion and longer nosepokes reflecting risk-seeking preferences. We found that rats modestly varied nosepoke length as a function of rates of reward magnitude/probability (Figure 17a), although the discrepancy between contingencies was much smaller than expected. However, it is of note that this result was not clear from the analysis of session averages (Figure 12b), which excluded the censored poke durations of loss trials from analysis. 'Probe' trials with no chance of losing could be incorporated in order to facilitate such analyses in the future. While the survival analysis of nosepoke duration did reveal an optimal ordinal pattern of nosepoke durations given the average nosepoke length (compare to Figure 10), nosepoke durations fell unexpectedly short of optimal. For example, rats should have poked for the full 2 sec maximum in order to maximize reward in the low contingency instead of 575 msec. This rather large discrepancy between optimal nosepoke length and observed length suggests that there was a missing factor from the Expected Reward calculation modelled in Figure 10. It is of note that rats responded in a similarly suboptimal manner in the rat BART task presented by Jentsch and colleagues (2010), whereby rats completed

much fewer bar presses than optimal even in no-risk conditions. Cognitive effort costs and discounting of future reward represent two plausible explanations for this behaviour. For example, the cognitive effort (Shenhav, Botvinick, & Cohen, 2013) involved in maintaining a nosepoke, e.g. without checking for reward at the reward spigot, may have been very high. This would effectively increase the 'cost' of each successive tick during any given nosepoke. Alternatively, rats' shorter than expected nosepokes may reflect temporal discounting of future rewards. Delay discounting is a well-established characteristic of impulsive behaviour in humans (Green, Myerson, Oliveira, & Chang, 2014; Holt, Green, & Myerson, 2003), non-human primates (Hayden & Platt, 2007; Rajala, Jenison, & Populin, 2015) and rats (Calvert, Green, & Myerson, 2010; Valencia-Torres et al., 2012). If rats were steeply discounting delayed rewards, then a sooner reward would hold relatively more subjective value than a discounted delayed reward, which would result in shorter nosepoke durations. If either cognitive effort costs or delay discounting were meaningful factors contributing to shorter than expected nosepoke durations, then future studies could increase the rate of reward accrued at each tick in order to offset costs/discount rates.

A central aim of the study was to develop a task that would facilitate the study of loss on behaviour and within the nervous system of rats. Thus, it was important to demonstrate that rats differentiated between wins and losses within the task. Multiple behavioural measures indicated that rats were indeed sensitive to losing in the task. Specifically, we found that rats' behaviour was altered on trials immediately preceded by a win vs. those preceded by a loss (Figures 13 and 14a). Furthermore, it was clear from our analysis that rats

poked significantly shorter on trials with a previous loss compared to a previous gain. Together, significant differences in MT, error time, and poke duration suggest not only that rats differentiated between wins and losses in the task, but also that a loss outcome on the current trial affected behaviour on the subsequent trial.

Of note was also rats' unexpected development of a lose-stay/win-shift strategy, which is illustrated in Figure 14a. Although the task was not specifically designed to reinforce such behaviour, we found rats were more likely to return to a contingency after a loss compared to a gain. Lesion studies in rats have demonstrated that the acquisition of stimulus-response associations such as win-stay or win-shift strategies is critically dependent upon the nigrostriatal pathway (Da Cunha et al., 2003; McDonald & White, 1993), which becomes the focus of investigations in future chapters.

One possible explanation for loss-stay behaviour involves preference of the high contingency (Figure 14b) and could only be partially addressed here. Given that rats prefer the high contingency but that they are also more likely to lose in the high contingency, it follows that they are also more likely to return to a losing contingency on average. However, survival analysis indicated that rats are actually less likely to stay after a loss vs. win in the high contingency (Figure 16c). Whereas previous analysis (Figure 14a) suggested that there was a significant main effect of Previous Outcome on stay choices overall, previous outcome did not emerge as a significant predictor in the Cox proportional hazard model. This suggests that the addition of another factor (e.g. Previous Contingency or the PreviousContingency*PreviousOutcome interaction) may

have accounted for a large proportion of the variance previously attributed to Previous Outcome. Unfortunately, low numbers of loss-stay trials in the low contingency (Figure 16a) may have sufficiently reduced power to preclude any conclusion here. Future work might address this by increasing the lowest probability of losing to ensure more losses in the low contingency.

Another potential alternative explanation for the development of this behaviour is that rats may have treated losses as a varying response requirement (e.g. variable interval reinforcement schedule) rather than a lost trial. If this were the case, they would return to the previous contingency to respond again for reward. Alternatively, rats may have returned to a contingency after a loss due to the training protocol that was implemented before testing. Before testing, rats were gradually trained to sustain longer and longer nosepokes. If a rat failed to sustain a nosepoke for the required length of time, it received no reward. Thus, rats may have interpreted a loss in the task as a failure to sustain the requisite nosepoke length. Again, however, this would not explain why rats are less likely to stay after a loss in the high contingency, nor why rats poked shorter after a loss compared to a gain. Future studies should make the loss more explicit to avoid this potential confound.

Finally, we found that rats made significantly more choice errors (i.e. poking in an unlit hole) in the least-preferred low contingency (Figure 12c). Since this behaviour was stable across testing sessions, it is likely that such errors represent a lack of inhibition or an expression of frustration rather than an absence of understanding in the task. For example, this could have resulted from an animal initiating a nosepoke first in a more-preferred contingency

rather than the low contingency on forced-choice trials. This would be in line with work on frustrative nonreward (e.g. Amsel, 1958), whereby frustration with task demands elicits increases in non-instrumental responses (e.g. increased grooming or error responses). However, the inverse error rate could provide a key behavioural measure of preference (or dislike) in the event that this is otherwise ambiguous.

In conclusion, we have developed a rat gambling task that elicits risky decision-making behaviour that could be more readily translated to human behaviour than previous task designs have allowed. Decisions about whether to continue poking for more reward occur on the millisecond-timescale, which is conducive to future research using neuroscience methods with high temporal resolution such as in vivo electrophysiology. Alternatively, decisions about which contingency to select on free-choice trials lend themselves to behavioural researchers in fields such as neuroeconomics. Future work is necessary to establish the optimally effective levels of reward magnitude and probability associated with each contingency and to disambiguate the separate effects of risk and reward on choices within the task. Given the flexibility of the task design, however, future iterations of the task present promising opportunities for ascertaining fundamental insights into loss-related behaviour in the brain.

Chapter 3

Loss-stay behaviour in rats with more Substantia Nigra pars compacta neurons is mitigated by dopamine antagonism

Abstract

The nigrostriatal pathway has been implicated in reward-related motor learning, but it is uncertain whether it also plays a role in loss-related motor learning. Given the loss-related changes in behaviour observed in Chapter 2, we sought to identify whether pharmacological manipulation of dopamine availability could alter responding after a loss in the gambling task. A cohort of rats (N=16) was trained on the task presented in Chapter 2. Once baseline behaviour had become stable, we administered systemic injections of saline vehicle or three doses of *cis*-Flupenthixol, a nonspecific dopamine antagonist. Relative to vehicle, increasing doses of *cis*-Flupenthixol administration monotonically increased the survival time of a poke toward the maximum 2 sec on any given trial. This indicates that dopamine receptor blockade may lead to a reduction in the discounting of future rewards. Additionally, we reported a dose-dependent decline in the average number of errors made on forced-choice trials, suggesting dopamine antagonism also increased behavioural inhibition. In contrast, rats became more likely to choose the contingency associated with higher reward accrual rates and higher probability of losing as doses of *cis*-Flupenthixol increased, which implies increasing tolerance to probability of loss. We also found that systemic dopamine blockade disrupted baseline lose-stay/win-shift behaviour. Interestingly, rats with a greater number of putative dopamine neurons in the substantia nigra pars compacta (SNc) were more likely to 'stay' after a loss – and this effect was abolished by dopamine antagonism. These findings implicate nigrostriatal dopamine transmission as playing a role in the modulation of loss-related behaviour in rats. This may provide critical insight into the processes underlying diseases involving degeneration of the midbrain dopamine system such as Parkinson's disease.

Introduction

In Chapter 2, we demonstrated that rats' behaviour is sensitive to losses in a gambling task. We next sought to identify whether win- and loss-related behaviour were modulated by similar neural mechanisms. The nigrostriatal dopamine system plays a key role in motor control and reward-based learning, and has been implicated in the signalling of prediction errors (Romo & Schultz, 1990; Schultz et al., 1997), the attribution of incentive salience to predictors of reward (Wilson et al., 2006; Zhang, Berridge, Tindell, Smith, & Aldridge, 2009) and the formation of habits (Graybiel, 2008; Haber, 2003; Wise, 2009). This work has its foundations in the classical intracranial self-stimulation studies, which demonstrated that animals would learn to press a lever in order to receive a pulse of stimulation to the dopaminergic midbrain (Olds & Milner, 1954). More recent work has been able to utilize optogenetics techniques to break down the midbrain areas responsible for different aspects of learning and performance of instrumental actions. For example, work by Rossi and colleagues (2013) demonstrated that selective optogenetic activation of the mouse substantia nigra pars compacta (SNc) is sufficient to facilitate the acquisition of a new instrumental action. Others have found that optogenetic stimulation of SNc neurons elicits a positive affective state encouraging approach behaviour, while optogenetic inhibition of the area provokes avoidance (Ilango et al., 2014). Furthermore, dopamine-dependent plasticity from such learning can create long-term changes in nigrostriatal pathways (Wickens, Reynolds, & Hyland, 2003). Thus, the nigrostriatal dopamine system became a primary focus in the investigation of loss-related behaviour in the rat gambling task introduced in the previous chapter.

We hypothesized that the involvement of the dopamine system in risky decisions would be apparent through changes in poke duration and contingency choice after pharmacological blockade of dopamine. Specifically, we predicted that dopamine antagonism would increase poke durations by mitigating reward discounting (St Onge & Floresco, 2009) and reduce choice of the high contingency by attenuating rats' propensity to take risks (St. Onge, Chiu, & Floresco, 2010). Finally, we hypothesized that dopamine antagonism would significantly disrupt strategic control of stay-shift behaviour after a loss.

Methods

The Methods used here are detailed extensively elsewhere (see Chapter 2, 'Methods').

Pharmacological Manipulation

A cohort of 16 rats were trained on the risky decision-making task and submitted to the pharmacological challenge once baseline behaviour was considered to be stable. Intraperitoneal injections of the nonspecific dopamine antagonist *cis*-Flupenthixol (Sigma-Aldrich Co., U.K) or vehicle (saline) were administered 20 minutes prior to behavioural testing. 3 doses and vehicle (saline) were counterbalanced according to a modified Latin square design, with a minimum of 7 days between doses to minimize carry-over effects. *cis*-Flupenthixol dissolved in 0.9% w/v saline was injected intraperitoneally at a volume of 1.0ml/kg (molecular salt weight) at doses of 0.125, 0.25, and 0.5 mg/ml. The saline vehicle used as control was delivered at 1.0 ml/kg.

Histology

Following the pharmacological testing, rats were euthanized via overdose with 0.08 ml pentobarbital (Univet Ltd., Oxford, U.K.) and then perfused intracardially with 0.1% phosphate buffered saline followed with a 4% paraformaldehyde in 0.1M phosphate buffer fixative. Using a freezing microtome, the fixed brains were then cut into 50 µm sections and stored in 0.1M phosphate buffer. One out of every four sections were subsequently stained for tyrosine hydroxylase and examined under a conventional light microscope. Sections were mapped onto standardized brain areas following Paxinos and Watson (1997) as depicted in Figure 18, and the number of tyrosine hydroxylase stained cell bodies were counted in the ventral tegmental area and substantia nigra pars compacta at 3 levels: -5.3mm, -5.8mm, and -6.3mm posterior to bregma.

Rat atlas cytoarchitechtonic guide to SNc and VTA

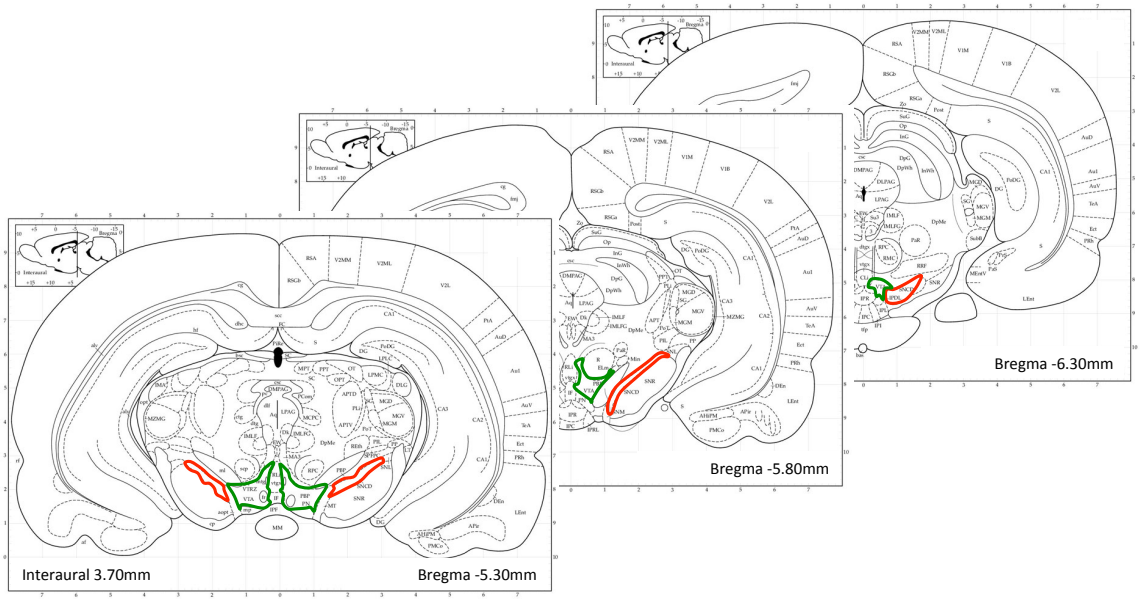


Figure 18: adapted from Paxinos and Watson (1997). Three sections (-5.30mm, -5.80mm, and -6.3mm behind bregma) of the rats' midbrains were stained for tyrosine hydroxylase (TH). TH-positive cell bodies were then counted under a conventional light microscope in the substantia nigra pars compacta (red) and the ventral tegmental area (green).

Count-recount correlations demonstrated highly reliable neuron counts (Pearson's $r = .953$, $p < .001$). A random sample of 10 sections (5 from animals presented in this chapter and 5 from animals presented in Chapter 4) was selected for blind recount using a random-number generator. The inter-rater reliability between the original counts and the counts conducted by a second blind counter was also high (Pearson's $r = .987$, $p < .001$).

Data Analysis

Session reconstruction with time-stamped data was performed using a self-written program in AWK programming language. Subsequent data analysis was carried out using Microsoft® Excel for Mac 2011 as well as R version 3.2.2 and SPSS® version 21 for Mac. Behavioural effects of Flupenthixol were analysed using repeated measures ANOVA's with contingency (three levels: high, medium, low) and dose (4 levels: vehicle, low, medium, high) as within-subject variables. In order to avoid potential ceiling-effects using proportion data, arcsine transformations were used on all variables expressed as a percentage (Zeeb et al., 2009). Missing data were replaced with series means. Greenhouse-Geisser adjusted degrees of freedom and Sidak-corrected *p*-values were applied where appropriate. All means are reported with standard errors and any significant main effects are reported with associated planned contrasts.

Given that average poke durations could be biased by truncated loss trials, descriptive statistics and ANOVA's including average poke duration are calculated based on successful trials only. All analyses measuring responses to a previous win or loss are defined as those trials immediately preceded by a

win or loss, omitting the first trial of a session. Due to insufficient trial completion rates, the high dose was omitted from regression analyses and Pearson's correlations performed in the association of neuron counts with stay-shift behaviour.

We also conducted survival analyses of poke duration and contingency choice by fitting a Cox proportional hazard model to the data. This analysis is carried out on a trial-by-trial basis, rather than averaging across subjects and sessions. The Cox proportional hazard model is a semi-parametric model that makes no assumption about the shape of the baseline curve (e.g. linear), and takes the following form:

$$h\big(t, x(t)\big) = h_0(t)\exp\big(\beta_1 x_1 + \cdots + \beta_p x_p\big)$$

where the hazard rate of *x* occurring, *h(t,x(t)),* on trial *t* is conditional on *p* predictors. The *β* coefficients are estimated from the data. While the model's primary assumption is that the hazard associated with any given covariate is proportional across time, it can be extended to incorporate time-varying covariates and stratified to accommodate within-subject designs. By stratifying across subjects, the models fit here include individual baseline hazards for each animal, which accounts for the variance in survival rates contributed by individual subjects.

Results

Effects of Flupenthixol on behaviour

We began by analysing any potential effects of Flupenthixol that may have emerged at the aggregate group level, followed by a more in depth trial-by-trial analysis of specific patterns of interest. We break down our primary analyses according to the two ways in which risky decisions are effected in the task: 1) length of a given poke and 2) choice among contingencies of risk/reward-accrual.

Based on session averages alone, we could not conclude that there was any significant change in the overall proportion of contingency chosen (Figure 19a) at any dose (Dose*Contingency: $F_{(6,90)}$=0.73, *p*=*NS*). Rats chose the high contingency option on 71.3% (±4.0%) of trials, while choice of the medium contingency (21.4 ± 3.1%,) and low contingency (7.3 ± 1.8%) remained low.

Flupenthixol increased animals' choice accuracy (Figure 19c) on forced-choice trials (Dose: $F_{(3,45)}$ = 7.82, $p < .001$, η_p^2 = .34). Post-hoc testing revealed that the medium ($p < .05$) and high ($p < .01$) doses significantly and linearly ($p < .01$) reduced error rates compared to vehicle. Although the error rates in the low-risk/low-reward contingency fell considerably, Flupenthixol did not abolish the main effect of contingency at any dose (Contingency: $F_{(2,30)}$ = 44.73, $p < .001$, η_p^2 = .75) and the interaction did not reach significance (Dose*Contingency: $F_{(6,90)}$ = 2.64, p = .68).

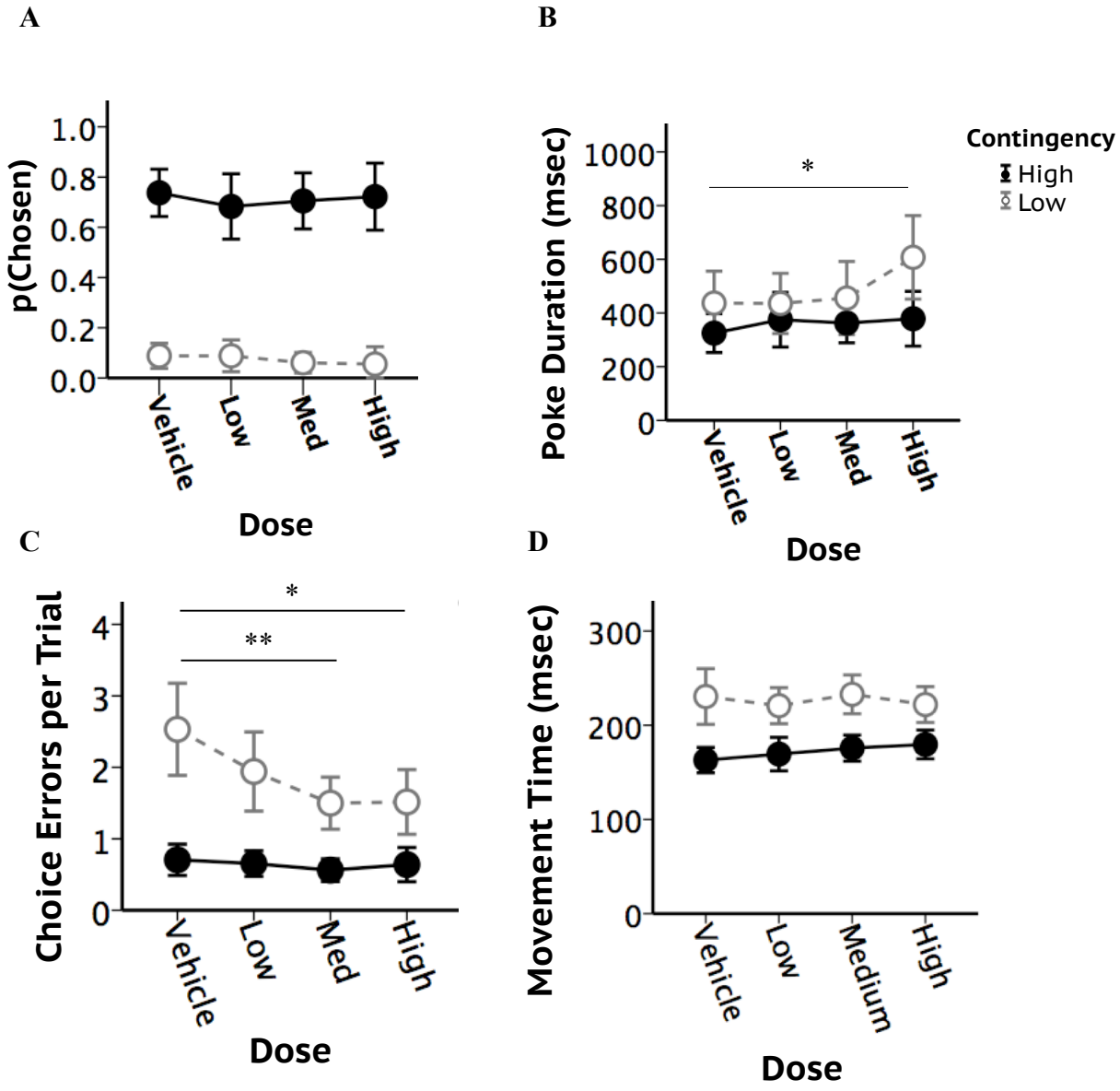


FIGURE 19: (**A)** Preference for the high-risk/high-reward contingency was not affected at any dose of dopamine antagonist Flupenthixol. (**B**) In the low-risk/low-reward contingency, rats poked on average 25% longer at the highest dose compared to the lowest dose (**C**) The number of errors made on forced-choice trials in the low-risk/low reward contingency decreased linearly with dose. (**D**) There was no significant change in movement time to reward across all doses. Error Bars are 95% CI's. **p*<.05, ***p*<.01, ****p*<.001.

Movement times were generally unaffected by dopamine antagonism, with animals moving more slowly in the low-risk/low-reward contingency (Figure 19d) at each dose (Contingency: $F_{(2,30)} = 46.27$, $p < .001$, $\eta_p^2 = .76$). Flupenthixol did not significantly affect rats' movement time to reward ($F_{(3,45)} = 0.22$, $p = NS$) or lick rate ($F_{(3,45)} = 0.51$, $p = NS$) during reward consumption.

Effects of DA antagonism on poke duration

We observed a main effect of both dose ($F_{(3,45)}$ = 5.14, p < .01, η_p^2 = .26) and contingency ($F_{(2,30)}$ = 5.18, p < .05, η_p^2 = .26) on poke duration (Figure 19b). Sidak-corrected pairwise comparisons revealed significant differences between the vehicle and high dose (p<.05) and between the high-risk/high-reward and low-risk/low-reward contingencies only (p < .01). The effect of Flupenthixol administration was most prominent in the low-risk/low-reward contingency, where animals poked on average 25% longer at the highest dose compared to the lowest dose. However, this interaction did not reach significance (Dose*Contingency: $F_{(6,90)}$ = 2.47, p = .065).

In order to further our examination of the effect of Flupenthixol on poke duration, we conducted a trial-by-trial analysis by fitting a Cox proportional hazard model to rats' poke durations at each dose. The model was stratified over subjects. The dose, contingency of the nosepoke hole, and the outcome of the previous trial (win/loss) were added as time-varying covariates to the model to adjust for non-proportional hazards. Table 5 of Appendix 2 provides particulars of model coefficients.

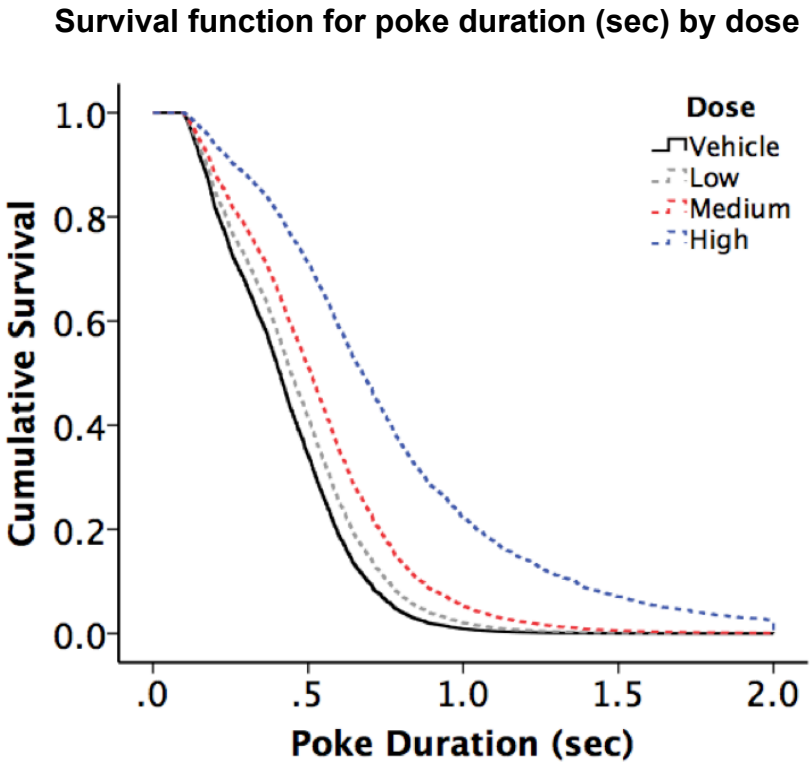


Figure 20: The 'survival time' of a poke decreased as a function of poke duration, and this effect was significantly attenuated by *cis*-Flupenthixol (main effect, $p<.001$). Compared to vehicle, the low ($p<.001$), medium ($p<.01$) and the high dose ($p<.001$) significantly increased the likelihood of continuing any given poke. CI's not shown; see Table 5 of Appendix 2.

The survival curves depicted in Figure 20 reflect a dose-dependant increase in poke durations compared to the saline vehicle, (main effect, $p < .001$). Dopamine antagonism significantly increased overall poke survival times by up to 27% at the highest dose ($p < .001$). Consistent with baseline, there was also a main effect of Contingency ($p < .001$), whereby rats poked longer in the low contingency ($p = .001$) and shorter in the high contingency ($p < .001$) in contrast to the medium contingency. Interestingly, previous outcome did not emerge as a significant factor in the model ($p = .11$).

Effects of DA antagonism on choice

We next fit separate Cox proportional hazard models to assess the likelihood that a rat would choose either the low or high contingency on free-choice trials. The model was stratified over subjects and previous trial outcome (win/loss)

was added as time-varying covariate to the model. Dose was also added as a covariate. The reader is referred to Table 6 and 7 of Appendix 2 for details of the model coefficients. The curves depicted Figure 21a and b represent the 'survival' time in trials until a rat chooses a given contingency. Steeper curves indicate that a rat was more likely to choose the contingency on an earlier trial than flatter curves.

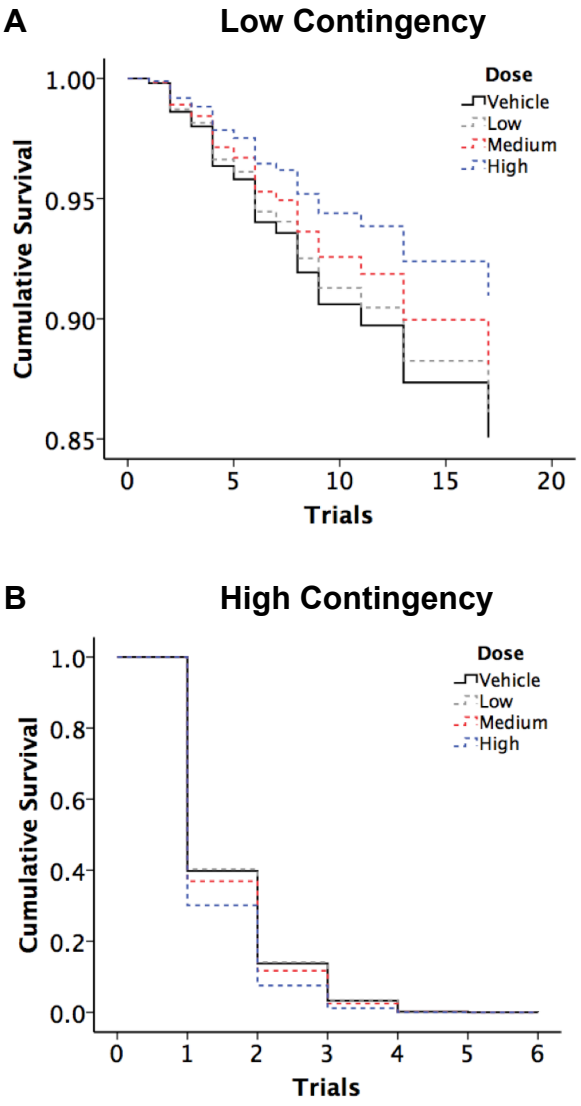


Figure 21: (**A**) **Low Contingency.** The survival function, which indicates the likelihood that a rat has *not* chosen the low contingency by a given number of free-choice trials, decreases as the sequence of free-choice trials increases. The limited range of the y-axis indicates that there is generally a low probability that a rat chooses the low contingency option on any free-choice trial. At vehicle, there is an 85% probability that rats will choose the low contingency by the 17th free-choice trial. Compared to vehicle, the likelihood of choosing the low contingency on a free-choice trial is reduced in a dose-dependent manner when DA agonist Flupenthixol is administered systemically.

(**B**) **High Contingency.** Note the axes differences from Figure 21a, which reflect rats' propensity to choose the high contingency over the low contingency on free choice trials. Here, the survival curve represents the likelihood that a rat did not choose the high contingency by the given free-choice trial. At all doses, rats do not last more than 6 free-choice trials without selecting the high contingency. The model indicates that 60% of rats chose the high contingency on the first free-choice trial in the saline condition and that this increases to 70% at the high dose (p<.05). However, the overall effect of dose in the model fell at 95% significance (p=.05). CI's not shown; see Tables 6-7 in Appendix 2.

These analyses indicate that, despite choosing the low contingency relatively infrequently on free-choice trials, there were significant dose-dependent decreases in choice of the low contingency (Figure 21a), main effect p < .05. In

the saline control condition, there was an 87% likelihood that rats would complete over 15 trials before selecting the low option on a free-choice trial. This probability increased to 93% at the high dose of Flupenthixol. In contrast, rats tended to choose the high contingency more at the highest dose (Figure 21b). Given that the high option was chosen within two free-choice trials about 90% of the time, modelling of any increases in high contingency choice was naturally rather limited. Despite this, we did find that the highest dose was a significant factor in the choice model ($p < .05$), although the overall effect of dose was only at significance ($p = .05$) so this effect should be interpreted with caution.

As demonstrated at baseline, choice of a given contingency may have also been influenced by a rats' strategy to stay or switch based on the previous outcome or contingency. In order to ascertain whether strategy was affected by dose, we conducted a survival analysis of rats' decision to stay/shift by fitting a Cox proportional hazards model with Dose, Previous Outcome, Previous Contingency, Current Contingency, and Contingency*PreviousOutcome as covariates (see Table 8 of Appendix 2). Previous Outcome and Contingency were entered into the model as time-varying covariates and the model was stratified by Subject.

Figure 22 illustrates the mediating effect of DA blockade on rats' propensity to return to the same contingency as the previous trial (main effect of Dose: $p < .01$). Compared to baseline, each the low ($p < .01$), medium ($p < .001$) and high ($p < .05$) doses significantly increased contingency choices associated

with a shift from the previous trial. This is consistent with the DA system being involved in the modulation of stay-shift strategy.

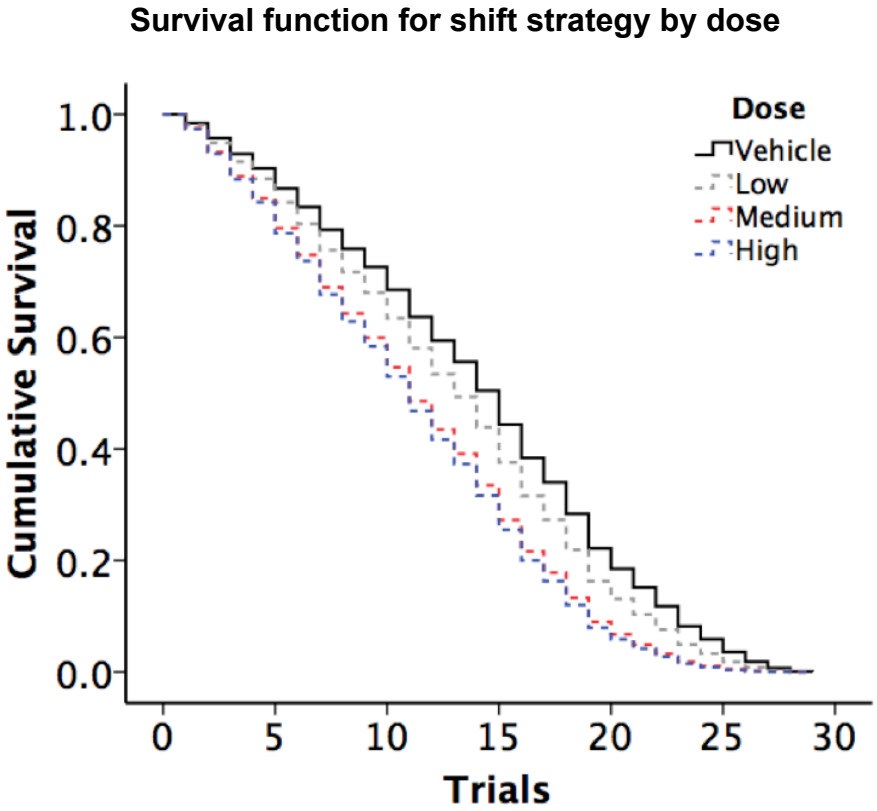


Figure 22: **Survival function for shift strategy by dose.** Flupenthixol administration significantly increased shift behaviour (thereby decreasing stay behaviour) overall compared to saline vehicle, main effect: $p<.01$. Each the low ($p<.01$), medium ($p<.001$) and high ($p<.05$) doses decreased the likelihood that a rat would return to the same contingency as the previous trial.

In order to further this investigation of the DA system's role in stay-shift strategies, we used histological preparations of TH-stained brain tissue to correlate putative dopamine cell counts from rats' midbrains with behaviour in the task. We found considerable variation between animals in the absolute number of TH-positive neurons from 3 representative sections of tissue (see Figure 23 for a photographic depiction) taken from the substantia nigra (SNc) of rats (N = 16 rats; *min* = 894; *max* = 1086; M = 984.93; SD = 66.26) and from the ventral tegmental area, or VTA (N = 16 rats; *min* = 757; *max* = 1242; M =

1010.16; *SD* = 137.26). SNc and VTA neuron counts were not correlated (Pearson's *r* = .35, *p* = .19).

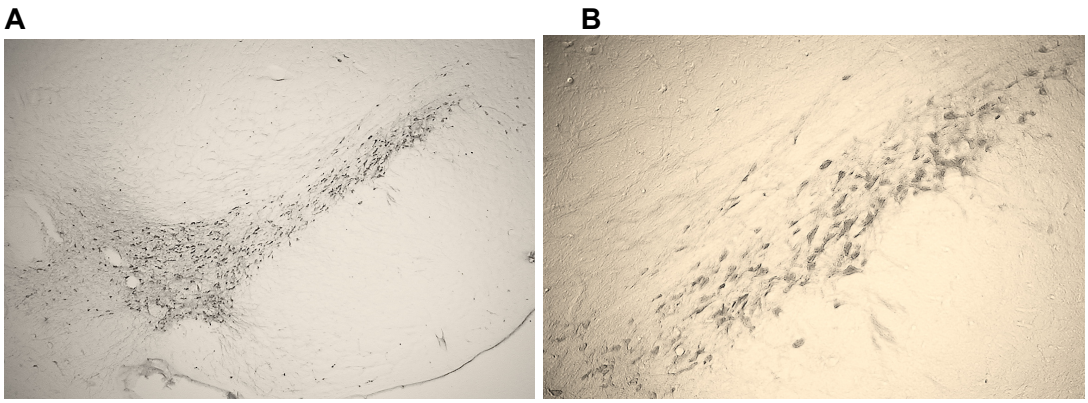


Figure 23: **A**) TH-positive dopamine cells of the rat SNc at ×4 magnification and (**B**) ×10 magnification.

Given that the nigrostriatal pathway is implicated in reward-related motor learning (for review, see Wickens et al., 2003), we hypothesized that variability in SNc neuron counts may also be reflected in variability in stay-shift behaviour. To establish the contribution of midbrain neuron counts to loss-stay behaviour, we carried out individual regression analyses for the proportion of loss-stay trials at each dose and baseline. Given that midbrain dopamine neurons have been shown to affect feeding behaviour (Hommel et al., 2006), we included rat weight as a predictor variable. We used the stepwise method to enter average weight, VTA, and SNc neuron counts as independent predictors of loss-stay trials in a linear regression. Details of model coefficients of each analysis can be found in Table 9-10 in Appendix 2. Unfortunately, the number of trials completed at the high dose was not sufficient to support a reliable comparison here, and future studies should take this reduction in trial completion at high doses of Flupenthixol into consideration.

Loss-stay strategy in rats with more SNc neurons is mitigated by DA antagonism

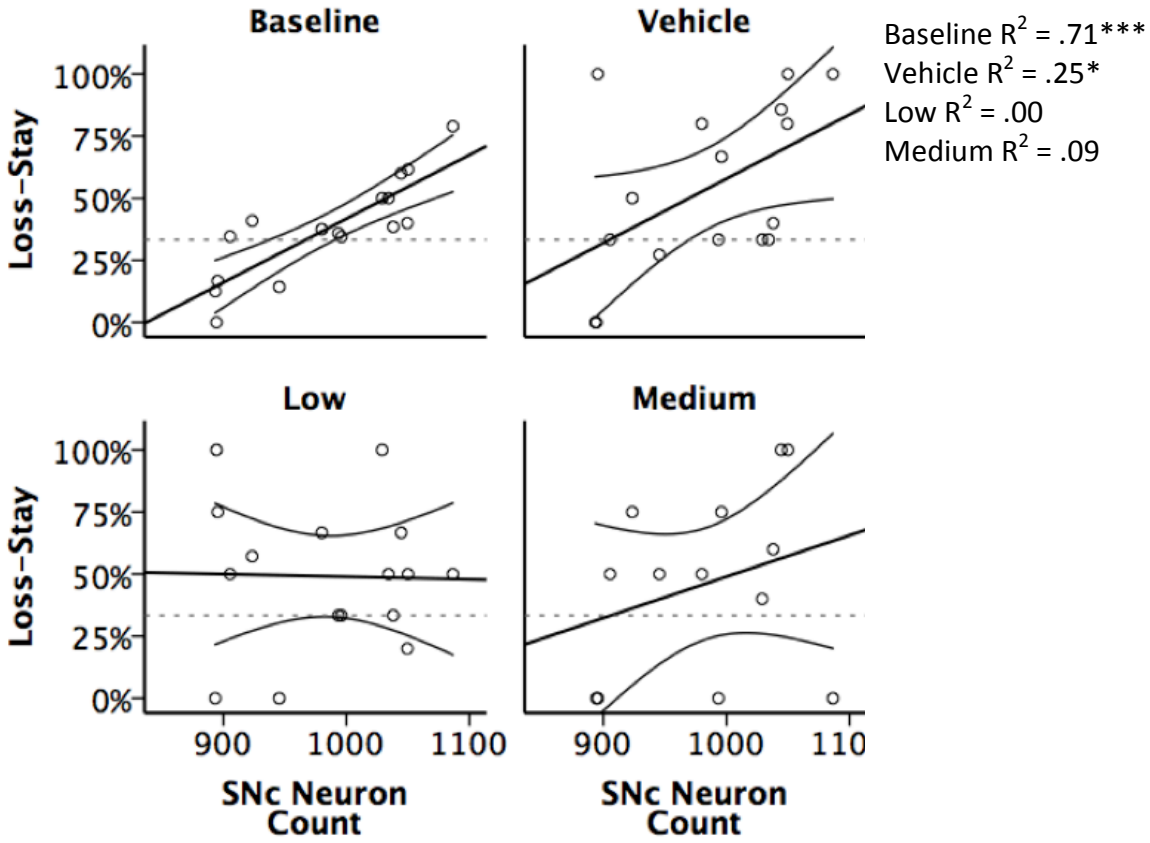


Figure 24: Loss-stay strategy in rats with more TH+ SNc neurons mitigated by DA antagonism. We found that rats with greater numbers TH-positive neurons in the SNc demonstrated a greater propensity to stay after a loss at baseline (R^2=.71) and at vehicle (R^2=.25), but that this effect was mitigated by administration of DA antagonist *cis*-Flupenthixol at both the low (R^2=.00) and medium doses (R^2=.09). Rats did not complete enough trials at the high dose to establish loss-stay proportions. Dotted lines represent chance at 33%. Error lines represent 95% CI's. ***p<.001, *p<.05.

As can be observed in Figure 24, rats with a greater number of SNc neurons were also more likely to stay after a loss. Furthermore, we found that this correlation was abolished by DA antagonism at the low and medium doses, but not by saline. At baseline, only the SNc count (β = .84, 95CI[.53,1.15], p < .001) was entered as a significant predictor in the model (R^2_{adj} = .69), while VTA count and weight were not. Similarly, at the saline vehicle, only SNc count (β = .50, 95CI[.01,1.00], p < .05) was entered as a significant predictor into the model (R^2_{adj} = .20). At both the low and medium doses, no predictors were significantly entered into the model. Together, the results shown in Figure 22

and Figure 24 provide evidence that loss-stay behaviour in our task is modulated by the dopamine system.

Discussion

In Chapter 2, we described the development of a rat gambling task and demonstrated that rats were not only sensitive to manipulations of reward volume and probability, but also that losses significantly affected subsequent behaviour in the task. Interestingly, we also found that rats develop a lose-stay/win-shift strategy that leads to greater efficiency in earning reward in the task. Here, we present research that established the effect of systemic dopamine antagonism on behaviour. We observed dose-dependent increases in poke durations and choice of the high contingency as well as decreases in 'stay' behaviour. These results not only suggest an important role for dopamine in guiding risky decisions in the task – but also implicate the substantia nigra directly based on correlations between behaviour and absolute cell counts. To the authors' knowledge, such a correlation between normal (i.e. healthy, non-lesioned) DA neuron counts and individual variability in behaviour has never been demonstrated. If such a relationship between neuron counts and behaviour were to stand up to further testing in the future, the implications for neuroscience are potentially quite wide – although we focus here on its prospective application to neuropsychological testing for early detection in Parkinson's Disease (PD).

Stay-shift pattern of behaviour

In the previous chapter we note rats' tendency toward lose-stay/win-shift behaviour. Although the task was not specifically designed to reinforce such behaviour, we found rats were more likely to return to a contingency after a loss compared to a gain. Lesion studies in rats have demonstrated that the

acquisition of stimulus-response associations such as win-stay or win-shift strategies is critically dependent upon the nigrostriatal pathway (Da Cunha et al., 2003; McDonald & White, 1993), which is characterized by afferent dopamine projections to the dorsal striatum from the substantia nigra (Haber, 2003). It is hypothesized that reward-based motor learning is driven by dopamine-dependent plasticity at the synapse and at the whole cell (for an excellent review, see Wickens et al., 2003). Therefore, we next counted TH-positive neurons in the SNc (Figure 23) and turned our investigation to the role of the midbrain dopamine system.

Dopamine Antagonism

Systemic injections of DA antagonist *cis*-Flupenthixol dose-dependently increased poke duration in the task. Since movement time to reward was unaffected by Flupenthixol administration, it is unlikely that increased poked durations reflect motor impairment. Given the probabilistic nature of accruing reward during a poke, potential future rewards are likely subject to probabilistic discounting during the decision process. Thus, this result would be consistent with previous work by St Onge and Floresco (2009), who found that amphetamine-induced probabilistic discounting was blocked by systemic dopamine D_1 and D_2 receptor antagonists (St Onge & Floresco, 2009). Therefore, these results support previous work demonstrating that Flupenthixol reduced rats' subjective overweighting of reward probabilities when evaluating gambles.

Our analyses suggest that systemic dopamine blockade decreased choice of the low contingency and increased choice of the high contingency at the

highest dose. This result is surprising given that a number of genetic, psychopharmacological, and imaging studies have implicated enhanced dopamine availability with risky decision-making (Kuhnen & Chiao, 2009; Mitchell et al., 2014; St Onge & Floresco, 2009). These findings also contrast with those of St Onge, Chiu, and Floresco (2010), who found that Flupenthixol decreased choice of the high-risk contingency on a risk discounting task in rats. However, this discrepancy may be attributable to differences in task designs. The previous study manipulated probability of the large/risky reward over four blocks of trials, whereas the task presented here maintained the same probabilities of reward throughout training and the task. Updating of reward probabilities and magnitude was necessary only while sustaining a nosepoke in the current task. It is possible that Flupenthixol may have impeded learning or updating of subjective reward probabilities (e.g. by blunting reward or risk prediction errors) over successive blocks of trials in the previous study. Thus, it is possible that choice of the high contingency in our task was maintained despite dopamine blockade solely due to preserved encoding of the larger magnitude of the reward (i.e. rats' choices were insensitive to risk). While rats' choices were highly sensitive to independent manipulations of risk (while reward magnitudes were constant) in the discrimination task (see Chapter 2), it remains possible that the discrepancy in reward accrual rate between the high and low contingencies was simply too large to motivate choice of the low contingency. This could reflect potential confounds arising from discounting of future rewards, or the tendency of individuals to preferentially weight rewards occurring sooner or with more certainty over those occurring later or with less certainty. This is a particularly relevant issue given that the dopamine system

has been heavily implicated in reward discounting, and that dysregulation of dopamine function often lead to exacerbated or abnormal discounting of rewards (Besson et al., 2010; Cardinal, Robbins, & Everitt, 2000; Dalley et al., 2007; Wade, de Wit, & Richards, 2000). Future studies could either disassociate risk- and reward-accrual rates or decrease the difference in rate of reward contingencies to further this line of enquiry.

Increasing doses of Flupenthixol reduced the average number of poke errors in a linear fashion. Interestingly, nosepokes appear to have become more deliberate at higher doses in the low contingency (Figure 19c). These results are consistent with previous work demonstrating that enhanced dopaminergic transmission potentiates premature responding while dopamine blockade reduces premature responding in the 5CSRTT (Passetti, Levita, & Robbins, 2003; van Gaalen, Brueggeman, Bronius, Schoffelmeer, & Vanderschuren, 2006). Choice errors in the current task may represent a form of frustrative behaviour in the face of a less-preferred option. Given that dopamine hyperactivity has been linked to aggression in both humans and animals (Brizer, 1988; Miczek, DeBold, & van Erp, 1994; for review, see Seo, Patrick, & Kennealy, 2008), the effect of Flupenthixol may be to diminish the negative emotion associated with an undesirable forced-choice trial.

We discovered a significant relationship between rats with more TH-positive neurons in the substantia nigra pars compacta (SNc) and a greater propensity for the win-shift/loss-stay strategy (Figure 24). This is bolstered by the finding that stay-shift strategies are significantly disrupted at all doses of Flupenthixol administration (Figure 22). While there is a well-established precedent for

linking the activity of SNc neurons to instrumental behaviour (for review, see Wickens et al., 2003), there is less evidence to suggest that absolute numbers of dopamine neurons can affect activity levels at the local or systems level.

A starting place may be the growing literature supporting a causal relationship between cognitive impairments and dopamine cell loss in Parkinson's disease (PD) patients. Impairments in cognitive function, such as deficits in spatial planning and attentional set shifting, as well as the prevalence of depression and anxiety are present in both the early and late stages of PD, and often predate diagnosis (Dubois & Pillon, 1996; Lees & Smith, 1983; Lewis, Dove, Robbins, Barker, & Owen, 2003). Indeed, given the subtle nature of the cognitive deficits in the early stages, non-motor symptoms are commonly unreported or overlooked by clinicians (Chaudhuri et al., 2010; Shulman, Taback, Rabinstein, & Weiner, 2002). In spite of this, research suggests that non-motor deficits typically precede the motor symptoms of PD by over a decade (Chaudhuri, Healy, & Schapira, 2006; Chaudhuri & Naidu, 2008). This is important to note because motor deficits in PD do not typically present until a substantial proportion (~70%) of dopamine neurons in the SNc have been lost (Truong, Allbutt, Kassiou, & Henderson, 2006; Zigmond, Berger, Grace, & Stricker, 1989). Surviving neurons exhibit a number of compensatory changes, such as: increasing dopamine release per terminal, decreasing dopamine reuptake, hemispheric inter-dependence, and increasing the proportion of active dopaminergic neurons (Blesa et al., 2011; Zigmond et al., 1989). Most researchers therefore expect that such compensatory mechanisms should preclude any observable differences in behaviour arising from natural variations in the absolute number of neurons in the substantia nigra (as demonstrated in

our findings). Given that we not only observe a moderate-to-strong correlation between neuron count and behaviour, but that this effect is also attenuated by systemic dopamine blockade, we believe these findings warrant further research. Replication of the study or direct manipulation of dopamine neurons in the SNc would be necessary to further substantiate such a claim.

In conclusion, we have found intriguing evidence to support a novel relationship between the number of neurons in a healthy system and behaviour. Given that clinical presentation typically occurs at very late stages of PD, identifying cognitive domains that are susceptible to impairment at lower rates of dopamine cell loss in the SNc presents an exciting area for future research. Establishing a link between the number of SNc neurons and a capacity for dopamine-dependent plasticity could potentially lead to new methods of neuropsychological testing for earlier detection of PD.

Chapter 4

Disassociating effects of dopamine neurons, probability and reward in a rat gambling task

Abstract

In this chapter, we extend the novel gambling task developed in Chapter 2 and seek to replicate the association between loss-stay behaviour and neurons in the substantia nigra pars compacta from Chapter 3. We successfully dissociated the effects of high probability of losing vs. high reward accrual rates on poke duration and choice. High reward accrual rates exhibited greater influence over choice while high probability of loss exhibited a greater effect on poke duration. However, we were unsuccessful in attempts to elicit longer poke durations by offsetting any potential hyperbolic discounting. Modifications to the task design changed the way rats responded after a loss. While rats did not exhibit a general tendency toward loss-stay behaviour as in Chapter 2, individual variability in loss-stay behaviour was still positively correlated with neuron counts in the substantia nigra. This suggests that risky or compulsive behaviour may be under the control of the number of neurons in the substantia nigra.

Introduction

In this chapter, we further develop the rat gambling task introduced in Chapter 2 – addressing a number of the limitations associated with the previous version of the task. Furthermore, we attempt to replicate any findings associated with the correlation between loss-stay behaviour and the number of neurons in the Substantia Nigra pars compacta (SNc) presented in Chapter 3. In terms of task design, we focused on addressing three key limitations identified in the discussion of the previous task:

1) The probability of losing and volume of reward both increased monotonically over the duration of a nosepoke, thus the distinct effects of either on behaviour were confounded.
2) A loss was not explicitly signalled in the original version of the task, which meant wins were more salient than losses. Furthermore, rats may have considered loss trials simply as a variable ratio schedule.
3) The volume of reward accrued in the low contingency was likely too low to make up for any effects of temporal discounting of delayed reward. Discounting of delayed rewards may have rendered longer poke durations in the low contingency less desirable than intended.

To implement these changes, we began by expanding the number of available nosepoke holes from 3 to 5. The five contingencies varied either by rate of reward accrual *or* by probability of losing the accrued reward. Therefore, the five contingencies were:

1. **Low reward:** low-reward/medium-probability
2. **High reward:** high-reward/medium-probability
3. **Medium-Medium:** medium-reward/medium-probability
4. **Low probability:** low-probability/medium-reward
5. **High probability:** high-probability/medium-reward

These separations allowed us to distinguish more clearly between the effects of reward and probability on choice and poke behaviour.

In the previous version of the task, losses were not signalled to the rats. Not only did this disproportionately increase the salience of wins to losses, it was also possible that the unanticipated development of loss-stay behaviour was a misperception of losing in the task. For example, rats may have perceived the high contingency as a variable ratio reinforcement schedule (Ferster & Skinner, 1957) requiring a varying number of responses to earn reward. To preclude this potential and to increase the salience of losses (our target outcome), we added a second auditory cue to the task. If the rat won, a win tone sounded and the reward magazine was illuminated as in the previous version of the task. In the event of a loss, the reward magazine was not illuminated and a second tone, distinct from the win tone, was paired with reward omission.

Finally, we found that rats rarely poked for to the max 2 sec in the low contingency, despite the fact that there was a 90% chance of success even at the full 2 seconds (compared to 50% in the medium and only 10% in the high contingency). This may have been due to temporal discounting of reward, whereby delayed rewards are discounted more heavily the further away they are in the future (Ainslie, 1975). The hyperbolic reduction of perceived reward can be captured by a factor of:

$$\frac{1}{1+kt}$$

where *t* is the duration of the delay until a reward is received and *k* is a constant discount rate per unit of time. Therefore to counteract any potential effect of reward discounting in the current version of the task, we included an extra reward 'buffer' at each successive 100msec tick.

Methods

Animals

Subjects were 16 male outbred Lister Hooded rats (Harlan U.K.) that were housed in groups of three (2 additional cage mates were not included in testing) in a climate-controlled colony room on a reverse 12-hour light: 12-hour dark cycle (6PM lights off). After three weeks of habituation to experimenter handling, rats were placed on restricted water access for the duration of behavioural training and testing with ad libitum food access in the home cage. Rats were tested 5 days a week. Water access was restricted to 1 hour on weekdays following testing but was available ad libitum on weekends from Friday at 4PM until Sunday afternoon (typically between 2-4PM). Rats' weights were monitored so that no animal dropped below 85% of its maximum body weight and showed growth throughout the experiment. All procedures conformed to the United Kingdom Animals (Scientific Procedures) Act (1986) under Project License 60/4040 and was approved by the Animal Welfare Ethics Committee of the University of St Andrews.

Apparatus

The reader is referred to the 'Methods' section of Chapter 2 for a detailed description of the apparatus employed here.

Training

After two weeks in which rats were habituated to human handling, the animals were placed on water restriction and submitted to 30 min training sessions in the testing chambers with no fixed trial limit. First, thirsty rats were trained over 2 days to associate a tone-light cue with delivery of 0.15 ml sweet liquid reward (sodium saccharin 0.3% w/v). Rats were subsequently trained over 23 sessions to sustain gradually longer nosepokes in any nosepoke hole of the five-hole array for up to 2 sec in order to receive the tone/light cue followed by the reward. Finally, rats were trained to sustain nosepokes for 2 sec in only lit nosepoke holes of the five-hole array. During these sessions rats were trained to poke in lit holes only, where pokes in unlit holes resulted in a 2 sec 'timeout.' Testing proceeded once all rats were above 90% accuracy in nosepoking (36 sessions).

Task Outline

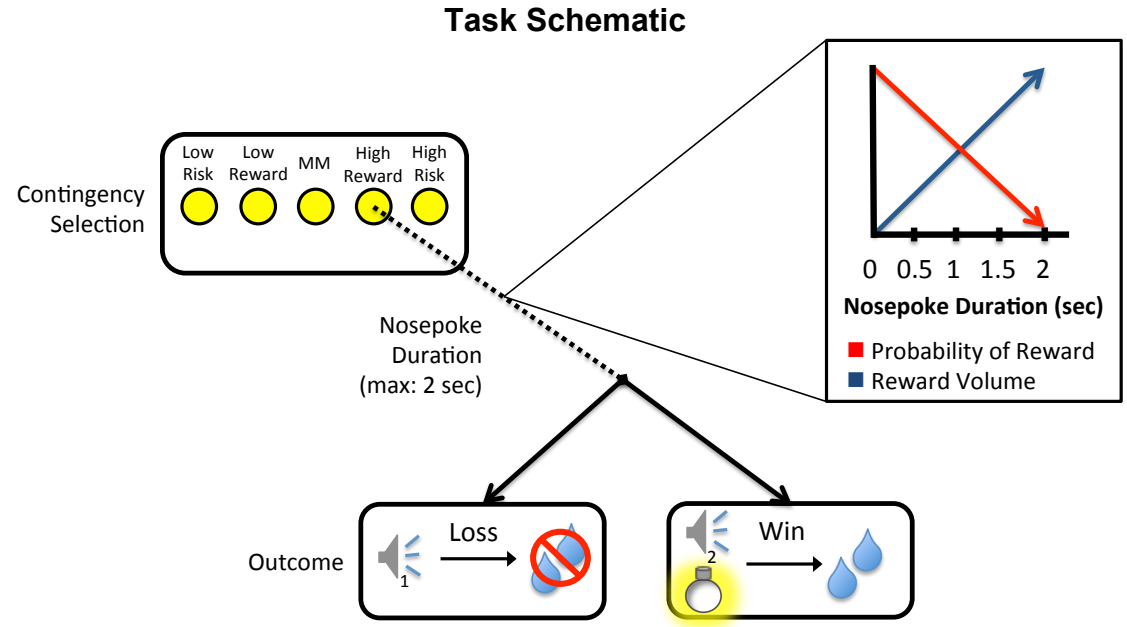
As discussed above, the task is modified from the original presented in Chapter 2. A brief description of the task is provided here. The reader is referred to Chapter 2 for a full description of the task and discussion of the associated behaviour.

A task schematic is depicted in Figure 25. Lit LEDs indicated the available hole(s) of a 5-hole array in which a rat could poke on any given trial, and the locations of the holes associated with each contingency were counterbalanced across testing chambers. All LED's in nosepoke holes were extinguished after the rat began poking into one of the lit nosepoke holes. If the animal either withdrew its snout from the nosepoke hole before a loss, or it successfully

reached the full 2 sec limit without losing, a tone-light cue emanating from the reward magazine would indicate the availability of reward for collection. If the rat lost the bet, a separate tone indicated the loss outcome and reward omission. Movement time (MT) to reward was measured from the onset of the conditioned stimuli to licking onset at the reward spigot. The amount of reward earned was a function of nosepoke length (refer to plot at top right of Figure 25) and varied by contingency. A 15 sec timer was activated from the onset of a nosepoke response, and a trial ended either when the timer elapsed or with the end of reward delivery, in which case a new trial began immediately. The animals were free to complete as many trials as possible over the course of the 30-minute session.

The major modifications to the previous task are illustrated in Figure 25. Here, one can see that the number of available nosepoke holes has been extended from three to five. The contingences associated with each hole vary from low to high either in probability of losing (depicted as ‘risk’ in Figure 25) or in the rate of reward accrual (depicted as ‘reward’ in Figure 25) with the ‘medium-medium’ contingency representing a medium rate of both. Furthermore, two distinct tones were incorporated as stimuli in the task. Tone 1 indicated to rats that a nosepoke had resulted in a loss and no reward would be available. Tone 2 (paired with a light at the reward spigot) indicated to rats that a nosepoke had resulted in a win and that liquid reward was available for consumption at the reward spigot.

A



B **Expected value of future reward as a function of poke duration**

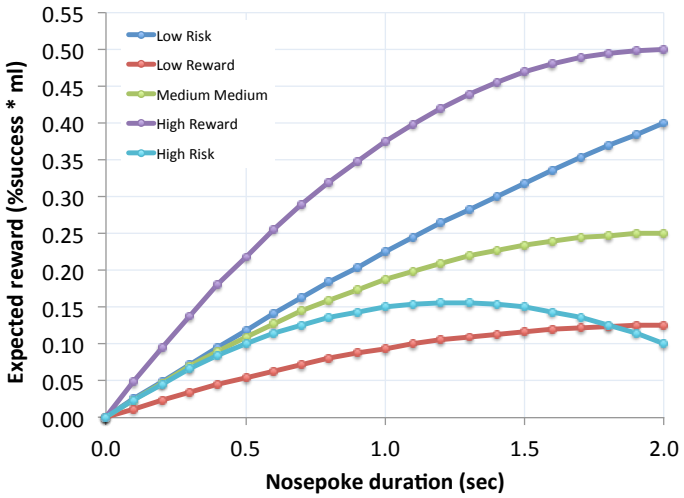


Figure 25: (A) Task Schematic. Note the major revisions from the previous version of the task (Chapter 2). On the top left of the figure, five nosepoke holes are available during the contingency selection stage instead of three. The probability of losing (denoted simply as 'risk') increased either at a low or high rate (with medium rate of reward accrual) in the 'Low Risk' and 'High Risk' contingencies, respectively. Alternatively, the rate of reward volume accrual (denoted simply as 'reward') increased at either a low or a high rate (with a medium rate of probability of losing) in the 'Low Reward' and 'High Reward' contingencies, respectively. Reward accrual and probability of losing both increased at medium rates in the 'Medium-Medium' (MM) contingency. To convey these changing rates of reward volume and probability more clearly, the box on the top right depicts the trade-off between the decreasing probability of winning (inverse of the probability of losing) vs. an increasing volume of reward. One should also note the two distinct tones (marked '1' and '2') that indicated either a win or a loss to the animal. **(B) Expected Reward.** The expected value of future reward as a function of poke duration is depicted for each of the five nosepoke contingencies. Apart from the High Risk contingency, the maximum expected reward can be achieved by maintaining a nosepoke for the maximum 2 seconds.

Behavioural measurements

Baseline data were from the last seven days of stable performance. Given that rats were not limited in the number of trials that they were able to complete in a session, we calculated percent choice (rather than absolute number of choices) using the number of trials chosen in a given contingency over the 7-day period, divided by the total number of free-choice trials over that period. The percentage of stay/shift trials was calculated as the total number of decisions to stay/switch after a previous trial across the 7 days, divided by the total number of free-choice trials across the 7 days. The following variables were also measured and analysed separately across conditions: error rate per forced-choice trial, time spent (sec) in incorrect nosepoke holes, lick rate (Hz), movement time to reward (sec).

Histology

Following testing, a subset of 9 rats (7 rats were chosen at random for a separate procedure not detailed here) were euthanized via overdose with 0.08 ml pentobarbital (Univet Ltd., Oxford, U.K.) and then perfused intracardially with 0.1% phosphate buffered saline followed with a 4% paraformaldehyde in 0.1M phosphate buffer fixative. Brains were postfixed in the cold (at a refrigerator temperature of ~ 1.6°C) for 24 hours in 20% sucrose solution and then washed for 30 minutes with buffer. Using a freezing microtome, 50 µm serial sections were taken through the midbrain and stored in 0.1M phosphate buffer. One out of every four sections were subsequently stained with antibody to tyrosine hydroxylase (TH) using avidin-biotin complex (ABC) immunohistochemical methods (S. M. Hsu & Raine, 1981) and 3’3-diaminobenzidine (DAB) for visualization of the antigen. Sections were mounted and examined under a

conventional light microscope at ×20 objective. Sections were mapped onto standardized brain areas following Paxinos and Watson (1997) as depicted in Figure 26, and the number of tyrosine hydroxylase stained cell bodies were counted in the substantia nigra pars compacta (SNc) at 3 sections: -5.3mm, -5.8mm, and -6.3mm behind bregma.

Rat atlas guide of SNc

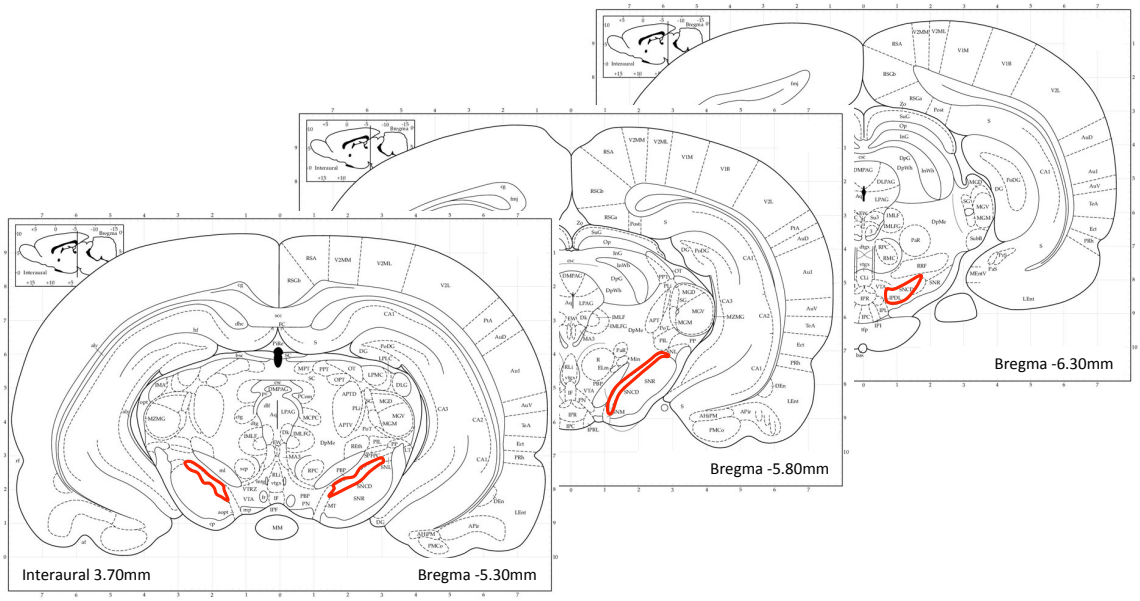


Figure 26: adapted from Paxinos and Watson (1997). Three sections (−5.30mm, −5.80mm, and −6.3mm behind bregma) of the rats' midbrains were stained for tyrosine hydroxylase (TH). TH-positive cell bodies were then counted under a conventional light microscope in the substantia nigra pars compacta (red) at ×20 magnification.

AH counted the number of midbrain dopamine neurons. AH's count-recount correlations demonstrated highly reliable neuron counts (Pearson's r = .953, p < .001). A random sample of 10 sections (5 from animals presented in this chapter and 5 from animals presented in Chapter 3) was selected for blind recount using a random-number generator for inter-rater reliability. The inter-rater reliability between the original (AH) counts and the counts conducted by a second blind counter (EMB) was also high (Pearson's r = .987, p < .001).

Data analysis

Session reconstruction with time-stamped data was performed using a program in the AWK programming language. Subsequent data analysis was carried out using Microsoft® Excel for Mac 2011 as well as R version 3.2.2 and SPSS® version 21 for Mac. Task behaviour was analysed using the average percentage each contingency was chosen (on free-choice trials) across rats and days during the last five testing sessions. Repeated-measures ANOVAs were performed with contingency (5 levels: high probability, low probability, medium probability/medium reward, high reward, and low reward) as the within-subject variable. Greenhouse-Geisser adjusted degrees of freedom and Sidak-corrected *p*-values were applied where appropriate. All means are reported with standard errors and any significant main effects are reported with associated planned contrasts. In order to avoid potential ceiling-effects associated with the use of proportion data, arcsine transformations were used on all variables expressed as a percentage (Zeeb et al., 2009). All analyses measuring responses to a previous win or loss are defined as those trials immediately preceded by a win or loss, omitting the first trial of a session.

Poke durations longer than the maximum allowed 2 seconds were truncated to 2 seconds for analysis. Furthermore, given that the average poke durations could be biased by truncated loss trials, descriptive statistics and ANOVA's including average poke duration are calculated based on successful trials only.

We also conducted survival analyses of poke duration and contingency choice by fitting a Cox proportional hazard model to the data. This analysis is carried out on a trial-by-trial basis, rather than averaging across subjects and sessions.

The Cox proportional hazard model is a semi-parametric model that makes no assumption about the shape of the baseline curve (e.g. linear), and takes the following form:

$$h\big(t, x(t)\big) = h_0(t)\exp\big(\beta_1 x_1 + \cdots + \beta_p x_p\big)$$

where the hazard rate of *x* occurring, *h(t,x(t)),* on trial *t* is conditional on *p* predictors. The *β* coefficients are estimated from the data. While the model's primary assumption is that the hazard associated with any given covariate is proportional across time, it can be extended to incorporate time-varying covariates and stratified to accommodate within-subject designs. By stratifying across subjects, the models fit here include individual baseline hazards for each animal, which accounts for the variance in survival rates contributed by individual subjects.

Results

Contingency

We began by determining how successful rats were at earning reward in each contingency. We used the reward won and the reward lost on each trial to calculate the average net reward for each subject in each of the 5 contingencies. The results are plotted in Figure 27. A repeated-measures ANOVA of net reward with Contingency (5 levels: high probability, low probability, high reward, low reward, and medium-medium) as a within-subjects factor confirmed that there was a main effect of contingency on net reward earned per trial ($F_{(4,60)} = 31.64$, $\eta_p^2 = .68$, $p < .001$). This suggests that our manipulations were successful.

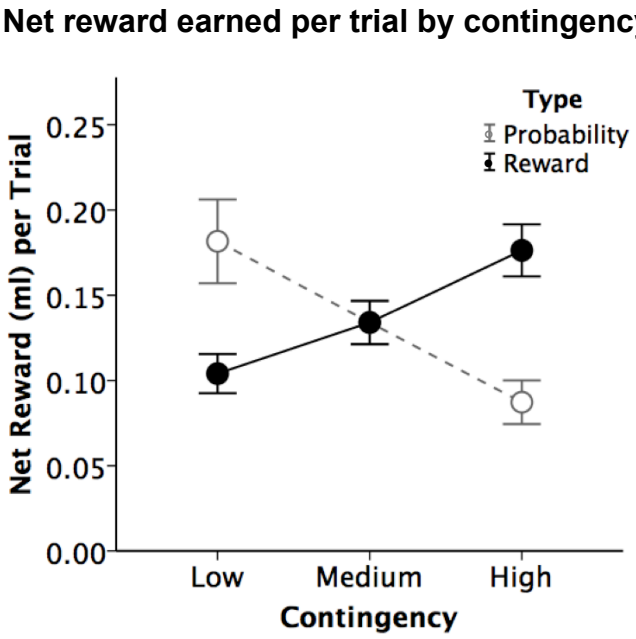


Figure 27: Net reward per contingency. On average, the net reward earned per trial was highest in the contingencies associated with a high rate of reward accrual and a low probability of losing. In contrast, rats netted the least reward in the low reward and high probability contingencies. Net reward earned in the medium-medium contingency fell in between these, which simple planned contrasts revealed was significantly different than all other contingencies (low reward: $p<.001$, low probability: $p<.01$, high reward: $p=.001$, high probability: $p<.001$). This suggests that our manipulations were successful. Error bars represent 95% CI's.

We next investigated whether the 5 contingencies of reward volume and probability had any effect on the various behavioural measures in the task. To establish the effect of contingency on hole choice during free choice trials, we conducted a repeated measures ANOVA for the percentage of trials each contingency was chosen on free choice trials, with Contingency (5 levels: high probability, low probability, high reward, low reward, and medium-medium) as a within-subjects factor. As depicted in Figure 28a, we found a significant main effect of contingency on rats' choice of nosepoke holes on free-choice trials ($F_{(4,60)} = 14.50$, $\eta_p^2 = .49$, Greenhouse-Geisser corrected $p < .001$). Although there appeared to be a linearly decreasing trend as the probability of losing increased, planned contrasts revealed that contingency choice did not significantly change from the low probability to the high probability contingency

(p = .54). There was a trend in the opposite direction as the reward accrual rates increased, with significant contrasts between choice of the low and high reward contingencies (p = .001) and medium-medium and high reward contingencies (p = .001). Overall, rats chose the high reward contingency the most (M = 42.9%, SEM = 4.4%), while they chose the high probability of losing contingency the least (M = 10.0%, SEM = 1.7%). One-sample t-tests revealed that only choice of the low probability contingency failed to differ significantly from chance ($t(15)$ = 0.07, p = NS).

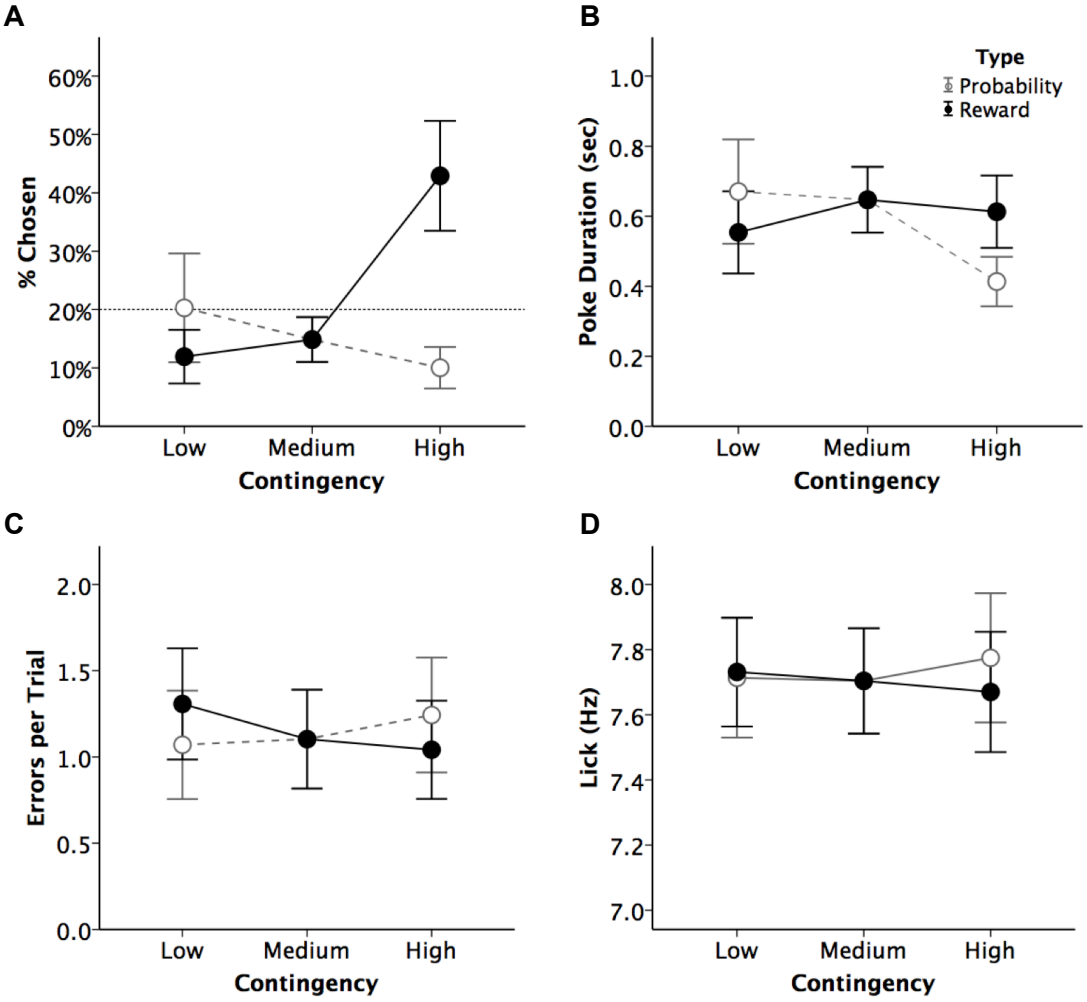


Figure 28: (**A**) Choices appear to have been more affected by reward-type manipulations than probability-type manipulations. Despite earning the same net reward as the contingency associated with a low probability of losing, rats choose the high reward contingency 22.6% (*SEM* = 8.1%) more often on free-choice trials. The low reward, high probability, and medium-medium contingencies were all significantly below chance (20% represented by the black

dotted line). (**B**) Poke durations appear to have been more affected by manipulations of probability than reward-type manipulations. On average, rats poked the longest in the low probability contingency (*M*=0.67sec, *SEM*=0.07sec) and the shortest in the high probability contingency (*M*=0.41sec, *SEM*=0.03sec). (**C**) There was a significant main effect of contingency on error rates (p<.01), which appear to be the inverse of net reward (see **Figure 27**). (**D**) There is a marginal trend toward faster lick rates in the high probability contingency and slower lick rates high reward contingency (Main effect: p=.05). Error bars represent 95% CI's.

Figure 28b illustrates the effect of contingency on the average time that rats were willing to sustain a nosepoke. We performed a repeated-measures ANOVA on average poke duration with Contingency (5 levels: high probability, low probability, high reward, low reward, and medium-medium) as a within-subjects factor. The results indicate that there was a strong main effect of contingency on the duration a rat was willing to sustain a nosepoke ($F_{(4,60)}$ = 12.85, η_p^2 = .46, Greenhouse-Geisser corrected p < .001). In contrast to the medium-medium contingency, rats poked significantly shorter in the high probability of loss contingency ($M_{\text{Difference}}$ = 0.23 sec, *SEM* = 0.04 sec, p < .001) but not significantly longer or shorter in the low probability of loss contingency ($M_{\text{Difference}}$ = 0.02 sec, *SEM* = 0.04 sec, p = *NS*). Although rats did tend to exhibit quicker nosepokes in the low reward contingency (*M* = 0.55 sec, *SEM* = 0.06 sec) compared to the high reward contingency (*M* = 0.61 sec, *SEM* = 0.05 sec), the manipulations of reward accrual rate did not significantly affect poke durations.

In the previous version of the task, we demonstrated that rats made more errors on forced-choice trials involving less preferred contingencies. To investigate this pattern, we performed a repeated-measures ANOVA on the average number of errors per trial with Contingency (5 levels: high probability, low probability, high reward, low reward, and medium-medium) as a within-

subjects factor. As depicted in Figure 28c, we again found that error rates were significantly affected by contingency ($F_{(4,60)}$ = 4.39, η_p^2 = .23, p < .01). On average, rats made the greatest amount of errors per trial in the low reward contingency (*M* = 1.31 errors/trial, *SEM* = 0.15 errors/trial) and the fewest number of errors per trial in the high reward contingency (*M* = 1.04 errors/trial, *SEM* = 0.13 errors/trial).

Based on the results from the previous iteration of the task, we also expected to find a significant effect of contingency on movement time to reward (MT). We performed a repeated-measures ANOVA on the average MT (sec) with Contingency (5 levels: high probability, low probability, high reward, low reward, and medium-medium) as a within-subjects factor. While rats did move somewhat more quickly to collect reward in the high reward contingency (*M* = 1.60 sec, *SEM* = 0.08 sec) compared to the slowest MT in the low probability contingency (*M* = 1.74 sec, *SEM* = 0.13 sec), we did not find a significant main effect of contingency on MT ($F_{(4,60)}$ = 1.04, p = .39).

Similarly, we performed a repeated-measures ANOVA on the average lick rate (Hz) with Contingency (5 levels: high probability, low probability, high reward, low reward, and medium-medium) as a within-subjects factor. We had previously found a weak effect of contingency on lick rate, and this was again the case ($F_{(4,60)}$ = 3.31, η_p^2 = .18, Greenhouse-Geisser adjusted p = .05), although it should be noted that this effect just reached our criterion for statistical significance. As can be seen in Figure 28d, planned contrasts revealed significantly faster lick rates when the probability of loss was the

highest (*M* = 7.78 Hz, *SEM* = 0.09 Hz, $p < .05$). Lick rates were slowest in the high reward contingency (*M* = .67 Hz, *SEM* = 0.09 Hz, *p* = .057).

In order to gain a more detailed understanding of the effects of contingency on behaviour, we next evaluated poke duration on a trial-by-trial basis. To achieve this, we conducted a survival analysis by fitting a Cox proportional hazard model to the data. This analysis has the potential for greater power given that that: 1) it can account for censored poke durations (i.e. premature unpokes due to losses) and, 2) it is computed on a trial-by-trial basis. Previous Outcome (2 levels: win and loss) and Contingency (5 levels: high probability, low probability, high reward, low reward, and medium-medium) were added as time-varying covariates and the model was stratified over subjects.

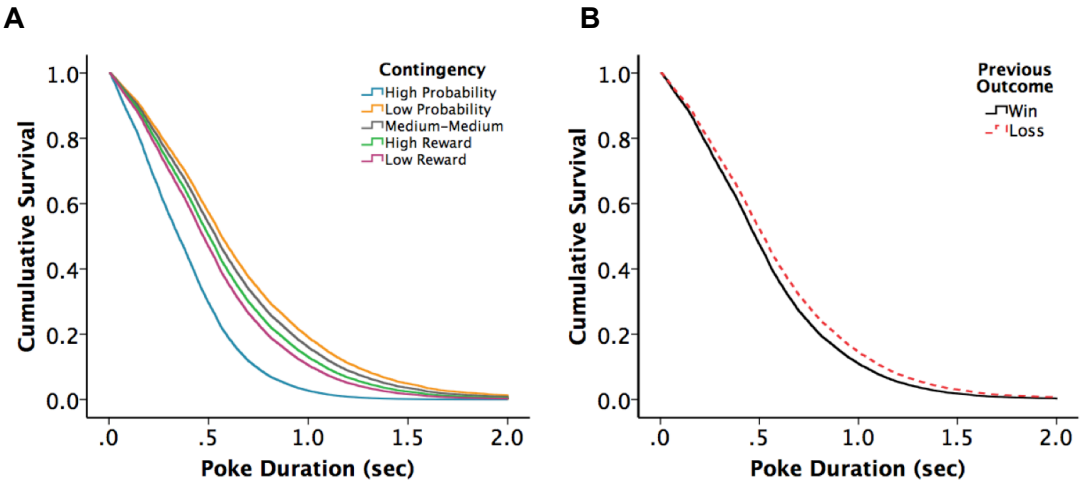


Figure 29: (A) The 'survival' time of a nosepoke varies significantly as a function of contingency ($p<.001$). Poke durations are the shortest when the probability of losing is the highest ($p<.001$) and longest when the probability of losing is the lowest ($p<.001$). Variations in rate of reward accrual had less of an apparent effect on poke duration, with the low reward contingency eliciting the shortest pokes ($p<.05$) while the high reward contingency did not significantly differ from the medium-medium contingency ($p=.524$). (**B**) Previous outcome was a significant factor in the model, but this effect was reversed and smaller in comparison to the previous task. Here, rats poked longer after a loss compared to a gain ($p<.001$). CI's not depicted, for details the reader is referred to Appendix 3, Table 11.

The model survival curves for each contingency are illustrated in Figure 29a. Contingency was found to be a highly significant factor in the model ($p < .001$).

The high probability contingency ($p < .001$) and the low probability contingency ($p < .001$) represented the shortest and the longest poke durations, respectively. In contrast to the medium-medium contingency, poke durations in the low reward contingency ($p < .05$) were shorter while those in the high reward contingency did not significantly differ ($p = .524$). Previous outcome was also entered as a significant factor in the model ($p < .001$), as depicted in Figure 29b. Interestingly, rats generally poked longer after a loss in the current version of the task, which contrasted with the previous 3-hole version (see Figure 17b). A table detailing further particulars of the model coefficients can be found in Table 11 of Appendix 3.

Stay-Shift Strategy

Given the surprising lose-stay/win-shift strategy in the previous version of the task, we were interested in determining whether this pattern was again present in rats' behaviour in the current task. We began our analysis by conducting a paired-samples *t*-test on the proportion of trials a rat chose to stay after a win *vs.* a loss. This analysis revealed that rats were more likely to stay after a win rather than after a loss, $t(15) = 2.21$, $p < .05$). See Figure 30 for a graph depicting this win-stay/lose-shift behaviour.

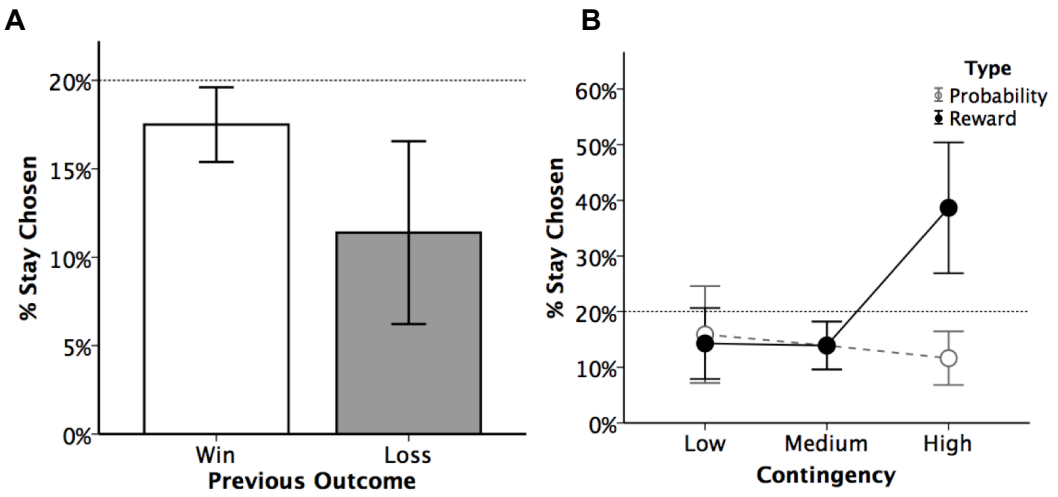


Figure 30: (**A**) Unlike in the previous version of the task, we found that rats developed a win-stay/lose-shift strategy in the current task. After a win, rats stayed on about 5% more trials than after a loss (p<.05). Note that the percentage of stay trials was significantly below chance on trials with both a previous win and a previous loss. (**B**) There was a main effect of contingency (p=.001), whereby rats returned to the high reward contingency significantly more often than any other contingency on free-choice trials.

We also wanted to establish whether rats were more likely to stay/shift after trials with some contingencies compared to others. We therefore conducted a repeated-measures ANOVA on the average percentage of stay trials per session with Contingency (5 levels: high probability, low probability, high reward, low reward, and medium-medium) as a within-subjects factor. As illustrated in Figure 30b, rats were significantly more likely to return to the high reward contingency (M = 38.7%, SEM = 5.5%) than the other contingencies, where averages all fell below 20% chance, main effect: $F_{(4,60)}$ = 8.49, η_p^2 = .36, Greenhouse-Geisser adjusted p = .001. The contingency associated with a high probability of loss also had the lowest stay percentages, with rats choosing to return to it on only 11.6% (SEM = 2.3%) of free-choice trials.

Neuron counts and behaviour

In Chapter 3, we provided evidence of a significant relationship between loss-stay behaviour and rats with more putative dopamine neurons in the substantia nigra pars compacta (SNc). Here, we sought to determine whether a similar

association was present in the current cohort of rats performing the new variation of the task. The number and variation of TH-positive neurons in the SNc (Figure 31) was first ascertained (N = 9 rats; *min* = 820; *max* = 1045; M = 936.94; SD = 61.37). These values were similar to counts from the previous study, and an independent t-test revealed that the differences between groups were not significant, $t(23) = 1.78$, Cohen's $d = 0.74$, $p = .09$.

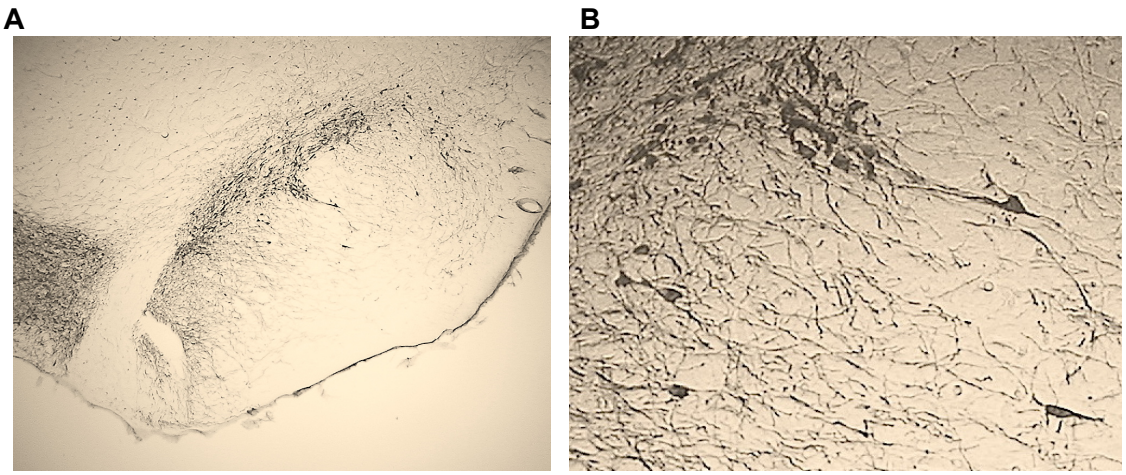


Figure 31: (**A**) TH-positive dopamine cells of the rat SNc at ×4 magnification and (**B**) ×10 magnification.

We implemented this analysis by regressing the percentage of lose-stay trials on neuron counts in the SNc. Given implications of midbrain dopamine neurons in feeding behaviour (Hommel et al., 2006), rat weight was also included as a predictor variable. We used the stepwise method to enter average weight and SNc neuron counts as independent predictors of lose-stay trials in a linear regression. The results of a linear regression (refer to Figure 32 and to Table 12 in Appendix 3 for full details of the model coefficients) suggest a strong relationship between neuron count and behaviour ($\beta = .83$, 95CI[.33,1.33], $p < .01$, $R^2_{adj} = .64$). Rat weight was not entered as a significant predictor into the model nor did it raise any potential issues with collinearity ($\beta_{in} = .004$, $p = .99$,

VIF = 1.28). These findings are in line with the positive relationship between greater numbers of SNc neurons and a greater propensity to shift after a loss identified in Chapter 3.

Loss-stay strategy increases in rats with more SNc neurons

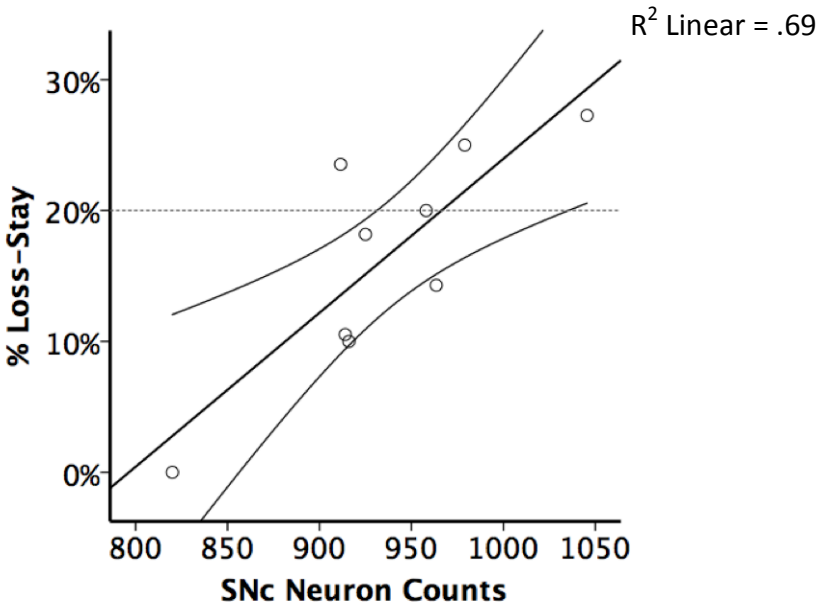


Figure 32: The figure above depicts the relationship between counts of TH-positive neurons in the SNc and the proportion of trials on which a rat (N=9) chose to 'stay,' or return to the same contingency, after a loss. We similarly observed a positive relationship in Chapter 3 with a separate group of rats on the previous version of the task. The dotted line represents chance at 20%. Error lines represent 95% CI's.

Discussion

This research was intended to extend and replicate previous work (presented in Chapters 2 and 3), which focused on the development of a novel rat gambling task. Here, we implemented three major changes to the task design: 1) Disassociated contingencies of varying probability and reward, 2) cued losses, and 3) 'extra' reward to offset hyperbolic discounting and incentivize longer poke durations.

By varying the rate of magnitude and probability accrual separately between nosepoke holes, we were able to determine that rats' choices (Figure 28a) reflect the expected reward rates in each hole (Figure 27). This finding was again in line with the matching law (Chung & Herrnstein, 1967), whereby an individual's choices among alternatives reflect the relative rate of reward (rather than net reward) of those alternatives. Given two contingencies of differing expected reward (refer to Figure 25b) but comparable net reward (high reward vs. low probability of loss, Figure 27) rats demonstrably preferred the high reward contingency to the low probability contingency (and all other contingencies).

Rats' average nosepoke durations varied between 400 and 700 msec. This suggests that our previous assertion – namely that delayed rewards are discounted – remains true despite the additional reward 'buffer.' It stands to reason in the first instance that our assumption about the function of the discount factor was incorrect. For example, rats may discount reward exponentially as opposed to hyperbolically, which would render later rewards

once more less desirable. Although such a result would stand in opposition to a number of studies documenting hyperbolic discount rates in animals and humans (e.g. Mazur & Biondi, 2009; Rajala et al., 2015; Rodriguez & Logue, 1988; Sopher & Sheth, 2006), future studies should test this assumption explicitly. On the other hand, this may be a factor arising from the timed task design. In other words, with an unlimited number of trials to complete the task, rats may have developed the strategy of completing many short nosepokes rather than a few longer nosepokes. While we did restrict the minimum trial duration to 7 seconds, either integrating pseudorandom inter-trial intervals or limiting the number of trials per task to a set amount could potentially resolve this in future versions of the task.

On average, poke durations did not increase greatly from the previous version of the task as intended. However, we were able to determine that high probability of losing was the most effective manipulation at eliciting different nosepoke durations (Figure 28b and Figure 29a). Survival analysis revealed that the low probability contingency did result in significantly longer poke durations than the medium-medium contingency – but this effect was not sufficiently strong enough to come through in the summary analysis of average behaviour. In contrast, the manipulations of reward accrual had little apparent effect on nosepoke behaviour. This suggests that rats do not have some predetermined threshold of either reward expectancy or effort expenditure. If, for example, a rat wished to achieve an arbitrary amount across all contingencies, it would stop poking sooner in the high reward contingency than in the low reward contingency. We observed the opposite effect (Figure 29a) –

nosepokes were the shorter in the low reward contingency than either high reward or medium-medium contingencies.

This offers much room for speculation. It is possible, for example, that the small reward was (still) not worth the cognitive effort of sustaining a nosepoke nor the time lost from starting a new trial in a new hole. Alternatively, it may be that losing a large reward elicits more negative affect than losing a medium-sized reward – thus conditioning rats to reduce poke durations in the high reward contingency. Future iterations of the task could take advantage of individual differences in tolerance of increasing probabilities / rates of reward accrual to maximize desired behaviour by establishing subject-specific contingencies before training.

We also integrated a second auditory cue for losses that was distinct from the win tone into the task. This was intended to increase salience of a loss, which was the target outcome of the task design and study. This was also implemented in order to minimize the potential that rats developed lose-stay behaviour because they incorrectly associated the high probability of loss contingency with a variable ratio schedule of reinforcement. Encouragingly, we found that rats performing in the current version of the task did not employ a lose-stay/win-shift strategy as in Chapters 2 and 3. Instead, rats were significantly more likely to stay after a win (Figure 30a). It is possible that the inclusion of an auditory loss cue effectively changed behaviour. However, this change may also be attributable to the fact that rats stayed on nearly 40% of trials in the high reward contingency (Figure 30b), which was well above 20% chance. Unfortunately, the large number of contingencies prevented us from

employing a Cox proportional hazard model to differentiate between any potential effects of a PreviousOutcome*PreviousContingency interaction, as was done in Chapter 2. We can, however, definitively state that rats returned to the contingency associated with a high probability of loss the least often of all five contingencies and that this was significantly less than chance. Thus, while it is unclear whether the switch from loss-stay to win-stay behaviour was motivated by addition of a loss tone or the change in contingencies (or both), it is unlikely that rats associate the high probability contingency with a variable reinforcement schedule.

Given that we did not observe a pattern of loss-stay behaviour in this version of the task, it is perhaps rather surprising that we once more identified a significant positive relationship between rats with more TH-positive SNc neurons and a greater propensity to stay after a loss (Figure 32). This finding was intriguing in that it rules out the possibility that the correlation was only present in animals that have performed a task in which lose-stay is the preferred response. Furthermore, the current cohort of rats was drug-naïve – which indicates that the relationship was not the spurious by-product of altered neural tissue potentially arising from *cis*-Flupenthixol administration. We also did not observe any significant differences in neuron counts (Figure 31) between cohorts, although the mean count in the current cohort (M = 936.94 cells) was somewhat lower than the previous cohort (M = 984.93 cells).

Previous work by Baker, Joh, and Reis (1980) used a strain of inbred mice with 20% greater midbrain DA neuron counts to establish that variation in midbrain TH activity is wholly accounted for by the number of neurons containing the

enzyme. The authors (Reis, Baker, Fink, & Joh, 1979) also performed a series of experiments and were also able to draw the conclusion that variability in both drug reactivity and behaviour was reducible to differences in neuron counts in the mouse nigrostriatal, mesolimbic, and mesocortical systems. Indeed, inbred mice with more midbrain dopamine neurons not only exhibited more exploratory behaviour and spontaneous motor activity, but their behaviour was also more sensitive to d-amphetamine administration. As a caveat, however, Sved, Baker, and Reis (1984) later found that DA neuron counts could not be used to predict overall neurotransmitter levels.

The results of our study bear striking resemblance to the work done in an inbred strain of mice as detailed above (Baker et al., 1980; Reis et al., 1979; Reis, Fink, & Baker, 1982; Sved et al., 1984). The current study is comparable despite no known strain differences (rats from both studies were outbred by the same breeder), with the minimum count (820 cells) equal to ~22% of the maximum count (1045). This would suggest that either rats have greater natural variations in midbrain DA numbers than the mice did, that this genetic variant is also present in the subjects used here, or that some other potential factor affecting neuron counts exists.

If variations in neuron count reflected faster instrumental learning in general, a plausible hypothesis would have been to predict a correlation between neuron counts and win-stay behaviour in the current task. In contrast, we found that neuron counts once again predicted loss-stay behaviour (Figure 32). This suggests that any putative dopaminergic control exerted on behaviour is related to either losses or a lack of spatial exploration. Given that Baker et al. (1980)

observed *more* spontaneous exploratory behaviour in mice with more midbrain numbers, it is more likely that this association is linked to losses. Repeating a loss can be seen as a risky, compulsive behaviour that putatively arises from the failure of a reward prediction error. If loss-related learning fell within the purview of SNc dopamine neuron populations, this could largely explain increased risk taking in, for example: susceptibility to addiction (Dalley et al., 2007), variation in life financial outcomes (Brian Knutson, Samanez-Larkin, & Kuhnen, 2011), Parkinson's Disease (Jee-Young et al., 2010), and those with genetic variations in the dopamine receptor D4 gene (Kuhnen & Chiao, 2009). Alternatively, loss-stay behaviour may be interpreted as compulsivity, which also has a substantial body of research linking it to the DA system (Eagle et al., 2011; Evans, Lawrence, Potts, Appel, & Lees, 2005; Evans et al., 2006; McKeon et al., 2007; Voon et al., 2010).

With the present chapter, we endeavoured to further advance the development of a novel rat gambling task. By replicating the correlation between loss-stay behaviour and neuron counts, but not an overall pattern of lose-stay behaviour in the task, our understanding of task behaviour and the neurobiological mechanisms governing that behaviour has also been critically expanded. This study provides evidence for a relationship between DA neuron counts and loss-stay behaviour that both complements and extends beyond the work by Reis and colleagues (1979). These results seem to indicate that the self-regulating nature of the dopamine system still leaves enough variability to allow for quantifiable individual differences in behaviour. Future work should focus on determining whether SNc neuron counts can also distinguish between individual differences in compulsivity and risk-taking. A number of viable

methods could further elucidate this potential relationship in future studies, including: psychopharmacology, targeted lesions, reversible inactivation, and optogenetics.

Chapter 5

Rats exhibit anchoring, loss aversion and the disposition effect in an experimental stock market task

Abstract

Empirical research suggests that there are a number of behavioural biases that characterize systematic deviations from optimal trading behaviour in the stock market. Biases such as ‘anchoring’, ‘loss aversion’, and the ‘disposition effect’ all describe suboptimal behavioural patterns exhibited by human investors. Explanations for these biases, such as the ‘realization utility’ hypothesis of the disposition effect, are supported by theories of model-free reinforcement learning and dopaminergic reward prediction errors. This opens up the potential for exploring the aforementioned biases using more primitive models of behaviour. However, to date no research has explicitly tested whether or not anchoring, loss-aversion, or the disposition effect could be observed in rat behaviour. To this end, we have developed a stock market task in rats that simulates key aspects of investor decision-making. Using the notion of reference dependence from Prospect Theory, we first trained thirsty rats (N=24) to develop a reference point set at 0.15 ml of sweet liquid reward. Thereafter, cohorts of four rats drove a virtual stock market by nosepoking first to select an asset, followed by a second nosepoke to subsequently buy, sell, or hold the selected asset. If a rat chose the buy or sell option, the reward earned on that trial was equal to the reference point plus (minus) the liquid equivalent of the gain (loss) incurred by the trade. Choice of the hold option always resulted in a gain, albeit much smaller in volume relative to a potential gain from either the buy or sell options. Analysis of rats’ choices relative to changes in price of the selected stock revealed that rats learn either to buy, hold or sell optimally (but not all three). Our results indicate that rats move much more slowly to collect reward after a loss than after a gain. Furthermore, rats choose the riskier buy and sell options more often than the safer hold option on trials immediately preceded by a loss. These findings suggest that rats’ behaviour reflects both reference-dependence and loss aversion. Our results also indicate that rats – like humans – demonstrate a significant disposition toward selling at a gain relative to selling at a loss (i.e. the disposition effect). Together, these results suggest that behavioural biases such as anchoring, loss aversion, and the disposition effect can be elicited in a simulated rat stock market task.

Introduction

One of the most robust empirical findings in behavioural finance is the tendency of investors to hold on to losing stocks too long and to sell winning stocks too quickly. Shefrin and Statman (1985) characterize this bias as the 'disposition effect.' Evidence for the disposition effect, along with a number of other behavioural biases, represents a mounting challenge to normative economic theories predicated on 'rational,' utility-maximizing Bayesian updaters. Irrespective of the prescriptions for portfolio management laid out by leading standards such as the Capital Asset Pricing Model (Markowitz, 1959; Sharpe, 1964), individual investors dependably exhibit a reluctance to realize their losses. While such systematic deviations from rationality appear to be ubiquitous across cultures (Grinblatt & Keloharju, 2001), professional money managers (Shapira & Venezia, 2001), and even primate species (M. K. Chen et al., 2006; Lakshminaryanan et al., 2008), the origins of such behaviour remain unclear.

Although a number of potential theories have been posited to account for the disposition effect (e.g. Hirshleifer, 2001; Kaustia, 2010; Weber & Camerer, 1998), the 'realization utility' hypothesis is unique in its ability to explain a number of behavioural biases (Barberis & Xiong, 2012; Shefrin & Statman, 1985) while also garnering support from neurobiological and psychological research (Barber & Odean, 2011; Frydman et al., 2014; Frydman & Rangel, 2014). The hypothesis posits that in addition to the utility derived from consumption, an investor receives a 'burst' of realization utility at the moment of sale that is proportional to the amount gained or lost in the trade. Frydman et

al. (2014) have recently used human fMRI BOLD activity to provide particularly compelling neural evidence for this hypothesis. The observed 'bursts' of utility closely align with learning signals described by reinforcement learning models (Erev & Roth, 1998; Sutton & Barto, 1998) in their potential to increase the likelihood of (not) repeating an (un)successful action in the future. Indeed, Charness and Levin (2005) demonstrate that utility maximization via Bayesian updating often fails when Bayes' rule clashes with reinforcement. The authors find that participants violated Bayes' updating rule on nearly 50% of trials where Bayesian updating and reinforcement were in conflict, yet hardly ever when the two rules were aligned. Thus, naïve reinforcement learning may lie at the heart of investor bias in situations where the immediate disutility of realizing a loss conflicts with an individual's ability to update priors to achieve the optimal outcome. This assertion has gained traction in a growing number of studies (Choi, Laibson, Madrian, & Metrick, 2009; Fuster, Laibson, & Mendel, 2010; Kaustia & Knüpfer, 2008) and highlights a role for more primitive models of investor behaviour.

Given the evidence implicating the potential contribution of reinforcement learning rules to investor bias, it may be possible to test such assertions using animal models. On the face it may seem a rather unusual notion to simulate financial decision-making behaviour in animals, but such models constitute an important basis for research in reinforcement learning (Lee & Dorris, 2014; Tolman, 1932). Indeed, animal models represent a critical means of interrogating the more primitive subcortical reward circuitry of neural systems without confounding 'human' factors, such as preconceived notions of how a stock market works or individual differences in numeracy and education.

This assertion is exemplified by previous research investigating economic decision-making in capuchin monkeys (M. K. Chen et al., 2006), pigeons (Kagel et al., 1975), and rodents (Kagel & Battalio, 1980), which has established that the foundations of rational economic choice extend further into humanity's evolutionary past than one might expect. A group of researchers (Kagel & Battalio, 1980; Kagel, Battalio, & Green, 1995; Kagel et al., 1975) have demonstrated in numerous studies that when rats are given the opportunity to work (e.g. press a bar) to earn food and water in their home cage, they adjust their consumption patterns in a way that is consistent with rational economic pricing models. For example, when the 'price' of an item of food is increased from one bar press to three bar presses, rats naturally adjust their consumption patterns to reflect the new higher prices, revealing expected elasticities in demand. van Wingerden, Marx, and Kalenscher (2015) have recently extended this work, altering the experimental design of the previous authors to reflect an open economy (i.e. consumption was outside the home cage and not essential for homeostasis). By changing animals' budgets (i.e. the number of trials in a session), the authors demonstrate that corresponding changes in consumption patterns imply that the subjective value attributed to a particular good is relative not only to price and preference, but also to the total budget. Together, these studies overcome the limitations of human experimental research and offer a valid model of behaviour that allows researchers to probe the neural representation of valuation in changing economic conditions.

To a similar end, we have developed a stock market task for rats that simulates key aspects of investor decision-making. To facilitate such behaviour, we utilize

the notion of reference-dependence from Kahneman and Tversky's (1979) Prospect Theory. In Prospect Theory, the subjective value of a prospect can be plotted on axes with an origin at some reference point (often the *status quo*), rather than at the objective value of zero. This implies that, given a gamble where most other players received an outcome of $100 and you receive an outcome of $5, a positive outcome can be viewed as a loss when it falls short of the reference point. We incorporate this concept in the rat stock market task by first establishing a reference point of reward, and subsequently by signalling losses and gains with respect to that reference point. Given that this is fundamental to the task design, our first hypothesis is therefore that rats are capable of reference-dependent behaviour. This directly challenges assertions that rodents do not possess the cognitive sophistication required for an organism to exhibit reference point effects (M. K. Chen et al., 2006).

Based on the research outlined above, we further hypothesize that rats will display rational trading behaviour with regard to changes in prices. This would manifest as subjects buying stocks that are undervalued (i.e. at low prices) and selling stocks that are overvalued. Alternatively, we expect to see 'irrational' behaviour with respect to losses (e.g. loss aversion) and to the realization of negative returns (e.g. the disposition effect).

Methods

Animals

Subjects were 24 male Lister Hooded rats, 8 of which were bred in house and 16 of which were outbred (Harlan U.K.), with initial weights between 125g and 250g. Animals were housed in groups of two or three on a 12-hour light: 12-

hour dark cycle (6PM lights off). All testing was carried out in the light part of the cycle. Rats were habituated to human handling for two weeks and then placed on restricted water access. Rats received water *ad libitum* from Friday afternoon to Sunday afternoon and for one hour each weekday after testing. Rats weights were monitored daily before testing so that no animal was allowed to drop below 85% of its free-drinking body weight. All procedures were carried out under Project License number 60/4040, conformed to the United Kingdom Animals (Scientific Procedures) Act (1986), and were approved by the Animal Ethics Committee at the University of St Andrews.

Apparatus

The apparatus used here are detailed extensively elsewhere (see Chapter 2, 'Methods').

Behavioural testing was interfaced by the MED-PC™ data experimental control system (Med Associates Inc., St Albans, VT) with an HP® computer running Windows7™ at a temporal resolution of 2 msec. Summary measures were also available in an online display on the computer screen along side real-time video feeds. Behavioural events were also time-stamped and recorded for offline data analysis and session reconstruction using a self-written program in AWK (Thompson Toolkit, Thompson Automation) programming language. Subsequent data analysis was carried out using Microsoft® Excel for Mac 2011 as well as SPSS® version 21 for Mac and R version 3.2.2 for Mac.

Training

Thirsty rats completed three stages of 30 min training sessions in the testing chambers. In the first stage of training, rats were trained to pair a tone-light cue

with the availability of sweet liquid reward at the reward spigot. Thereafter, rats were trained to nosepoke in a lit nosepoke hole of the five-hole array in order to receive the cued reward. Finally, the animals were trained to complete a sequence of two nosepokes in the lit holes in order to receive the cued reward. Response accuracy was measured as the number of nosepokes into incorrect (unlit) holes versus the number of nosepokes into correct (lit) holes. In total, 18 training sessions were required for rats to reach >90% nosepoke accuracy.

Testing

Task Outline

In our task, four freely moving adult rats drive a virtual stock market by nosepoking in holes to select and subsequently buy, sell, or hold assets. At the beginning and end of a testing session, each cohort of four rats was carried together in one transport cage between the colony room and the testing room. Rats were then placed in one of four separate standard operant boxes (described above) and the outer sound-attenuating chamber doors were closed to indicate to the rat that the session had begun. On each trial, a rat was required to make two distinct nosepokes into a lit nosepoke hole within the 5-hole array. Free-choice (three stock choices), paired-choice (two stock choices), and forced-choice (one stock choice) trials were randomly interleaved throughout the session. Each rat completed two blocks of trials: 1) a reference-point establishment block lasting 15 trials and 2) a trading block with a 45-minute duration from block onset.

Block 1

At the start of a free-choice trial, the recessed nosepoke-hole LED lights of the three middle holes were illuminated to indicate that all three stocks were available for selection. The rat then made a 'stock selection' by poking its nose into one of the lit holes. The lights were then immediately extinguished for 2 seconds. After this pause, the lights on both the left and right of the chosen hole, as well as the chosen hole itself, were re-illuminated. Rats then made an 'option selection' by completing a second nosepoke into one of the three lit holes. Counter-balanced across subjects, rats poked into the left lit hole to select a 'buy' option and into the right lit hole to select a 'sell' option. The centre lit hole was always a 'hold' option. After an option was selected a reward tone and light indicated the availability of reward at the reward spigot. In order to establish a reference point, all trials in block 1 resulted in 0.15 ml of reward. Block 1 consisted of 15 trials, and all rats were required to complete all 15 trials before the cohort could progress to block 2.

Block 2

The trial structure of block 2 was similar to block 1, with the exception that rats received information about the market volume of each stock and could earn more or less reward than the 0.15 ml reference point depending on their stock and option selections. At the onset of a new trial, available stocks were indicated via blinking LED lights recessed in the nosepoke holes. The blink rate was proportional to the number of shares currently being held in that stock across all four rats (i.e. market volume). Rats were then able to select a stock by poking in one of the blinking holes. Once a 'stock selection' nosepoke was

made, the hole-lights were immediately extinguished and a 2-second pause ensued. After 2 seconds, the selected hole and the two adjacent holes (to the left and right) were once more illuminated, but no longer blinking. Rats nosepoked a second time into one of the illuminated holes to indicate either a buy, hold, or sell option. One of two tones immediately indicated whether this choice resulted in a gain or loss, and rats were free to lick at the reward spigot for reward. Movement time (MT) was measured from the onset of the tone to the onset of licking at the reward spigot. The volume of reward a rat received on a given trial was proportional to the amount gained or lost, added to the reference point.

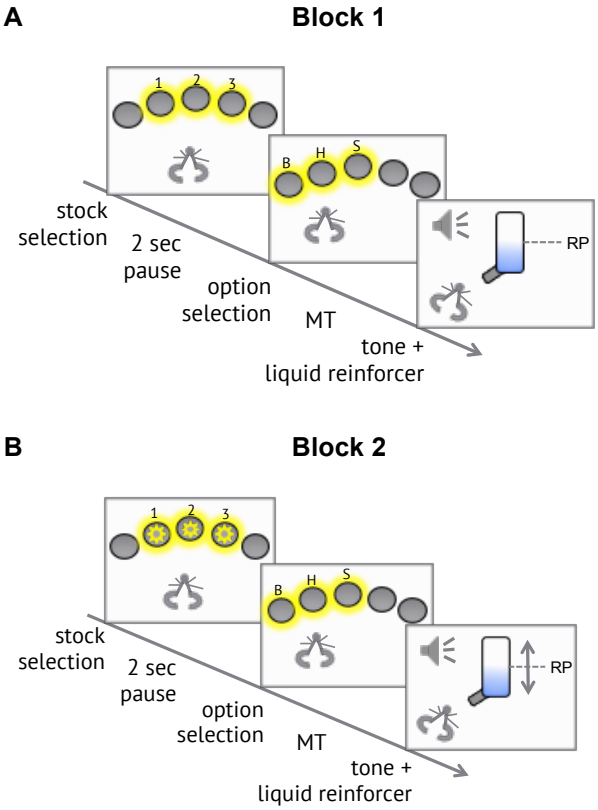


Figure 33: Task Schematic:

(**A**) In Block 1, a (free-choice) trial began with three centre nosepoke holes indicating the availability of stocks 1, 2, and 3. The rat then nosepoked in a lit hole to select a stock (here, stock 1), and the lights were immediately extinguished. After 2 sec, the selected hole and the adjacent holes on either side were illuminated, indicating that the rat could select an option to buy (B), hold (H), or sell (S). Once the rat poked to select an option, a tone immediately indicated that reward was available at the reward spigot. In block 1, rats always received 0.15 ml of reward in order to establish a reference point (RP).

(**B**) In Block 2, hole lights blinked to indicate the market volume of a stock, whereby stocks with more total shares had faster blinking rates. As before, the rat nosepoked first to select a stock (1, 2, or 3) and again to subsequently select an option (buy, hold, or sell). After the second nosepoke, a gain or a loss tone immediately indicated the trial outcome. Movement time (MT) was measured from tone onset to lick onset at the reward spigot. The volume of reward delivered was greater than the RP for gains and less than the RP for losses.

Pricing

When encountering a stock for the first time in a session, prices were arbitrarily set at 80, 140, and 200 for stock 1, 2 and 3, respectively. All rats were initially endowed with 100 shares of each stock. Prices in all boxes would update dynamically from that initial price point as rats bought or sold shares of a stock. A stock's price depended on its total number of shares across all four rats (market volume), with greater market volumes leading to higher prices and lower market volumes resulting in lower prices. The price of a share at any given moment was equal to:

$$\textit{Share Price} = (\textit{Initial Price} \times \textit{Total \# of Market Shares})/400$$

The blink rate was a 50/50 on/off cycle (i.e. the time between each flash was equal to the flash length). The on/off time period is the reciprocal of blink rate (Hz):

$$\textit{On/Off Period} = [2/(\textit{Total \# of Market Shares}/100)] \times 60 \text{ seconds}$$

Selling

Take, for example, the investor that buys a stock at $100 and subsequently sells it at $110 for a $10 profit. Likewise, rats could choose the 'sell' option, which decremented the number of shares held by that rat by 10. In the task, if the price of a stock had increased from the last time that a rat selected that stock (or from the arbitrarily set price on an initial encounter), then the rat gained reward on that trial. The volume of reward received was calculated as the reference point (0.15 ml) *plus* the liquid equivalent of the profit. On the other hand, if the stock had decreased in price since it was last selected, the rat lost reward on that trial. The volume of reward received was calculated as the

reference point (0.15 ml) *minus* the liquid equivalent of the loss. In the event that a rat lost 0.15 ml of reward or more, it received nothing upon licking at the reward spigot. Rats were not allowed to 'short-sell,' i.e. in the event that a rat no longer held any shares of a selected stock, the 'sell' option was not illuminated and only the 'buy' or 'hold' option could be chosen.

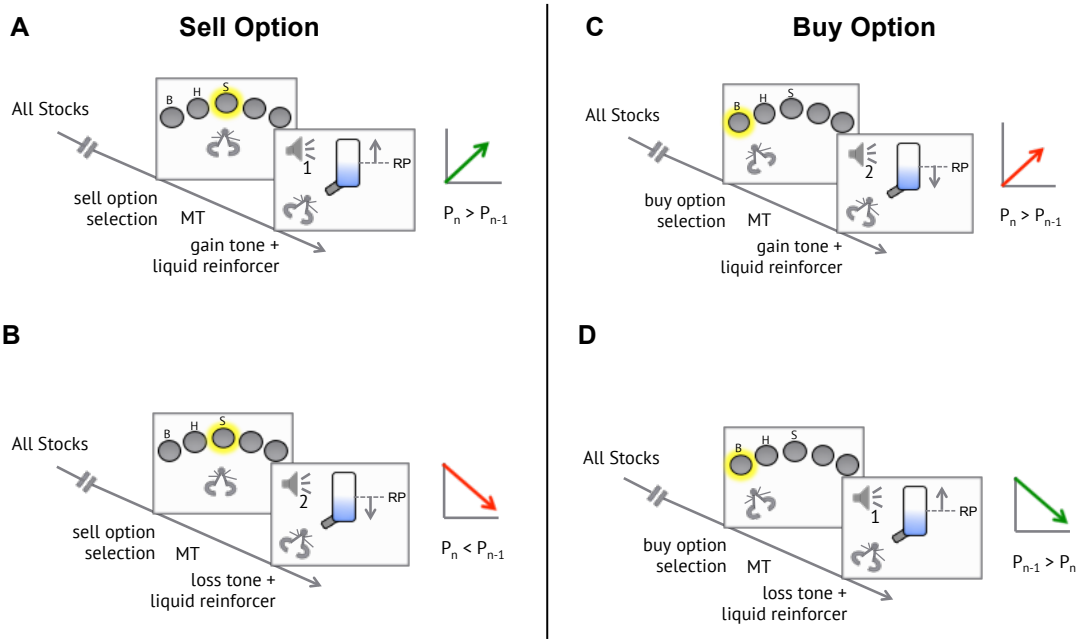


Figure 34: Trading (A) The rat chooses the sell option and the price of a stock on the current trial (P_n) has increased since the previous trial (P_{n-1}), resulting in a gain. The rat receives the reference point (RP) plus the liquid equivalent of the profit gained from the sale. (B) The rat chooses the sell option but the price has decreased since the previous trial, resulting in a loss. The rat receives the RP minus the liquid equivalent of the loss incurred from the sale. (C) The rat chooses the buy option but the price has gone up since the previous encounter with that stock, which represents a loss. The rat receives the RP minus the price differential between the two time points. (D) The rat chooses the buy option and the price has gone down since it was previously encountered, which represents a gain. The rat receives the RP plus the price differential between the two time points.

Buying

Take, for example, the investor that has a chance to buy a stock at $100 but waits until the next day, only to find that it had gone up in price to $110. Or perhaps that the stock price had gone down to $90 and the waiting had paid off. Likewise, rats could choose the 'buy' option, which incremented the number of shares held by that rat by 10. In the current task, if the price of a stock had

increased in price since the last encounter (or from the arbitrarily set price on an initial encounter) with that stock, then the rat lost reward on that trial. However, if the price of a stock had decreased since the previous selection, then the rat gained reward on that trial. Similar to the selling option, the amount that the rat received was equal to the reference point plus or minus the liquid equivalent of the gain or loss, respectively.

Holding

Rats also have the option of 'holding' on any given trial, which results in a 'dividend' payment and no change in the number of shares being held in the selected stock. The dividend amount was based on the individual subject's current share holdings of that stock, and had a 2/3 probability of being low (e.g. 2% of holdings) and a 1/3 probability of being high (e.g. 6% of holdings). The dividend gains were the 'safe' option, but on average resulted in a smaller reward than could have been earned with either the buy or sell options. In the event that a rat no longer held any shares of a stock, the hold option delivered only the reference point 0.15 ml of reward.

Data Analysis

We compare individual choices and behaviour recorded over seven testing sessions. Since Block 1 was intended to set a reference point only, all analyses are performed on trials from Block 2 unless otherwise specified. We use repeated-measure ANOVAs in order to evaluate basic behaviour in the task, as well as the effects of our reference-point manipulation and any potential effects of loss aversion. In order facilitate comparison of rat behaviour with human behaviour, we also adapt the methodology used in the behavioural finance

literature (Barber & Odean, 2011; Odean, 1998) to evaluate any potential for the disposition effect.

Summary Measures

For summary measures, data were averaged per subject across the last seven testing sessions. Behavioural measures were: percentage choice of each stock, percentage choice of each option, movement time to collect reward (MT), choice errors (pokes in unlit holes), lick rate (Hz), inter-poke interval (IPI), and post-pump licking (PPL). PPL can be thought of as rats 'savouring' reward, and has been identified as a putative measure of 'liking' vs. 'wanting.' PPL is defined as the amount of time spent licking at the reward spigot after mechanical cessation of reward delivery. Greenhouse-Geisser corrections to *p*-values have been applied where appropriate, although uncorrected *p*-values are reported for ease of comprehension. Trial durations > 40 seconds were omitted from analysis.

Proportion of Realized Gains & Losses

Adapted from Odean (1998), we calculated the proportion of gains realized (PGR) and the proportion of losses realized (PLR) in order to establish whether rats exhibited the disposition effect (PGR > PLR) in our stock market task. Unlike the previous author, we perform these calculations on a subject-wise level. Odean (1998) computed PGR and PLR as:

$$PGR = \frac{\#\ of\ realized\ gains}{\#\ of\ realized\ gains + \#\ of\ paper\ gains}$$

$$PLR = \frac{\#\ of\ realized\ losses}{\#\ of\ realized\ losses + \#\ of\ paper\ losses}$$

Above, the denominator represents the number of opportunities to realize a gain (loss). In the rat task, this is slightly less straightforward. Given the forced-choice and paired-choice trials within a session, rats did not have the opportunity to realize a gain (or loss) on every stock on every trial. Therefore, the PGR (PLR) denominator was calculated on a stock-by-stock basis as any trial on which the rat had the *opportunity* to select and sell a stock that had gone up (down) in price since the previous purchase.

Cox Proportional-Hazard Modelling

The cox proportional-hazard model is a semi-parametric analysis that makes no assumption about the shape (e.g. linear) of the baseline hazard rate. This model has been employed in a number of behavioural finance studies (Barber & Odean, 2011; Feng & Seasholes, 2005; Shumway & Wu, 2005; Strahilevitz, Odean, & Barber, 2011) to characterize the likelihood of selling a stock in a time-series conditional on some factor (e.g. return magnitude and valence). The estimated model takes the following form:

$$h\big(t, x(t)\big) = h_0(t)\exp\left(\beta_1 x_1 + \cdots + \beta_p x_p\right)$$

where the hazard rate, *h(t,x(t)),* on trial *t* is conditional on *p* predictors. The *β* coefficients are estimated from the data. The main assumption of the model is that the hazards are proportionally dispersed at all time-points, but this model can be extended to include time-varying covariates (e.g. blink-rate). This model can also be stratified to incorporate repeated-measures designs, which has

been done here. From this model, one can predict the hazard ratio of a subject choosing to sell a given stock at time *t* for each covariate *k* as:

$$\exp(\beta_k) = \frac{h_0(t)\exp\left(\beta_1 x_1 + \cdots + \beta_k(x_k + 1) + \cdots + \beta_p x_p\right)}{h_0(t)\exp\left(\beta_1 x_1 + \cdots + \beta_k x_k + \cdots + \beta_p x_p\right)}$$

Here, the hazard ratio, $exp(\beta_k)$, is the ratio of two stocks with the same *k* covariates and where the numerator stock has an x_k that is one unit greater than the denominator (Barber & Odean, 2011). To maximize the potential of the model, a continuous variable (such as return on sale) can be transformed into dummy variables that represent 4% wide bins, taking on the value of 1 on trials that fall into that range and 0 otherwise. This allows the model to isolate the marginal hazard contributed by each bin when all other bins are zero. The reader is referred to Cox and Oakes (1984) for further details on the Cox Proportional Hazard analysis. For any sale trial x, return on sale was calculated as:

$$Return(x) = \frac{Price_{sale} - Price_{purchase}}{Price_{purchase}}$$

where the difference in current sale price and previous purchase price was averaged relative to the previous purchase price. Trials were included in the analysis only if the selected stock had been purchased at least once previously. The model was stratified over subject, stock, and session to account for the effects of the repeated-measures design. Blink rate, counterfactual reward, and return on sale were included as time-varying factors.

Results

Descriptive Statistics

Cumulatively, rats (N=24) completed 19,323 trials over the course of 7 testing sessions. On average, rats completed 115.02 (*SD* = 27.96) trials and earned 18.78 ml (*SD* = 4.75 ml) of reward per session. The mean trial duration was 19.62 (*SD* = 1.91) seconds. This equated to an average rate of reward of 3.36 (*SD* = 0.39) µl per trial second over the 45-minute duration of Block 2. During trading (i.e. when a buy or sell option was selected), rats received a profit on nearly 2/3rd of trials (63.53%) and a loss on 36.47% of trials. Although rats profited on a greater proportion of trials, rats lost an average of 0.052 (*SD* = 0.005) ml per trial while profiting only 0.026 (*SD* = 0.006) ml of reward on average ($M_{Difference}$ = 0.027 ml, *SEM* = 0.001 ml). The results of a paired-sample *t*-test indicated that rats lost significantly more reward than they gained on average, $t(23) = 21.04$, $p < .001$. Rats that lost more reward on average did not also gain more reward on average, Pearson's $r = .37$, $p = .08$.

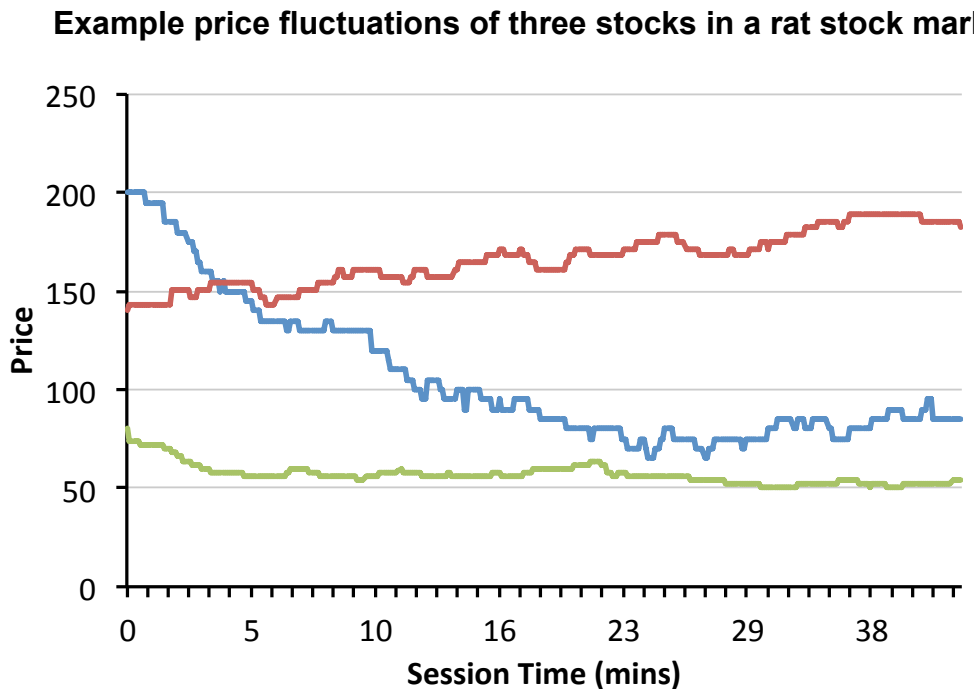


Figure 35 depicts share price fluctuations during a single testing session from one cohort of four rats. Each coloured line represents the evolution of price (in arbitrary units) of the three stocks over the course of the 45-minute session. From the graph, it is apparent that stock prices move both up and down over time. Since price is a function of market volume, the steep decreasing price of the blue line represents continuous sales of shares in that stock (i.e. lower market volume) up until about minute 24. After that point, sales of the blue stock level out and rats begin buying shares of it again, causing share price to begin a slow recovery.

Rats nosepoked in error an average of 1.88 (*SD* = 0.29) times per trial, with the error occurring either during stock/option selection or during the 2 second pause between stock and option selection. To ensure that these errors did not reflect a lack of understanding about the task demands (i.e. poking twice in a lit nosepoke hole, but not in an unlit nosepoke hole) that was subject to learning over further training sessions, a repeated-measures ANOVA was performed on error rate with Session (7 levels) as the within-subject variable. A significant effect would potentially suggest that rats were learning over the course of the sessions. However, the results indicate that there were no significant differences in error rates across the seven testing sessions ($F_{(6,138)}$ = 1.45, p = .20), which implies that these errors resulted from any number of

affective/cognitive factors (e.g. frustrative non-reward, impulsivity, exploration, habit, etc.) rather than a misunderstanding of the task.

At the beginning of any given testing session, each of the three centre nosepoke holes was randomly assigned stock 1, 2 or 3. This remained constant throughout the session. It is therefore possible that over many sessions, rats could develop a preference for a stock based on its original starting price. However, we did not find any indication that rats developed preferences for a particular stock across testing sessions. On free-choice trials, the average proportion of trials that subjects chose Stock 1, 2, and 3 were distributed tightly around chance at .33, .35, and .32, respectively. There was no significant effect of Stock on choice ($F_{(2,46)} = 0.51$, $p = NS$).

Manipulation check: Trading payoffs are greater than dividend payoffs

When a rat chose to hold a selected stock instead of a trade option, it received a dividend payout. Dividends were calculated as a percentage of the rats' current holdings of the selected stock. Theoretically, the percentage paid was sufficiently small that the expected value of a hold option was on average distinctly lower than that of either a buy or a sell option. As a manipulation check to ensure that this played out in reality, we analysed the average reward earned per trial second under each option using a repeated-measures ANOVA with Option (3 levels: Buy, Hold and Sell) as the within-subject variable.

Rats earn more reward when trading as compared to holding

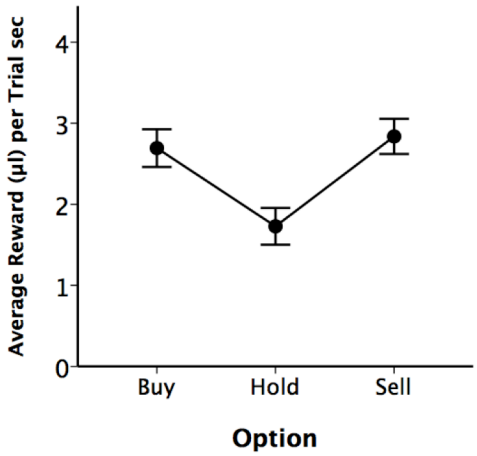


Figure 36: On average, rats earned almost twice as much reward per trial second when selecting a trade option (buy or sell) as compared to a hold option.

As theorized, we found that the proportion of trials on which each option was selected had a highly significant effect on rate of reward, $F_{(2,46)} = 48.30$, $\eta_p^2 = .68$, $p < .001$. Either trading option was nearly 50% more profitable than the hold option ($p < .001$ for both contrasts), with the hold option averaging 1.73 µl/trial sec and the buy and sell options averaging 2.69 µl/trial sec and 2.84 µl/trial sec, respectively. Neither buying nor selling more resulted in significantly greater payoff rates than the other ($M_{Difference} = 0.14$ µl/trial sec, $p = .33$). Using the variability of outcome volume as a measure of an option's risk, we found that the sell option was the most risky ($\sigma^2 = .0028$) although not significantly more so than the buy option ($\sigma^2 = .0026$, p=*NS*). As intended, the hold option was a much safer option ($\sigma^2 = .0006$), leading to the average reward more than four times more reliably than either of the trade options.

The transaction cost of trading within the task

Within the operant boxes, the spatial location of the buy and sell options on either side of the hold option may have added a temporal disadvantage to

trading. Although such a disadvantage could potentially increase the face validity of the task when translating to human behaviour (e.g. due to transaction costs), we endeavoured to identify and quantify such a cost, should it exist. Since the hold option involved poking for a second time in the same nosepoke hole, the inter-poke interval (IPI) was shorter on hold trials (M = 2.42 sec, SEM = 0.05 sec) compared to buy (M = 2.76 sec, SEM = 0.06 sec) and sell (M = 2.66 sec, SEM = 0.07 sec) trials. A repeated measures ANOVA with Option as a within-subjects factor revealed that there was a significant main effect of Option on IPI, $F_{(2,46)} = 7.81$, $\eta_p^2 = .25$, Greenhouse-Geisser adjusted $p < .01$. Planned contrasts determined that this effect was primarily due to the significant differences between hold trials and each buy ($p < .001$) and sell ($p < .01$) trials. We found no significant difference between buy and sell IPI's. Thus, while the added cost of trading compared to holding may be small ($M_{\text{Difference}}$ = 285.5 msec), it is nonetheless significant.

Given that the 2 sec pause also allows ample time to move between spatial locations, it may also be possible that this difference represents deliberation time. Although the task design does not allow us to conclude with certainty one way or the other, evidence of this lies in the discrepancy between IPI after a gain (M = 2.49 sec, SEM = 0.08) *vs.* after a loss (M = 2.79, SEM = 0.11). A 2-way repeated-measures ANOVA of IPI with Previous Outcome and Current Option indicated that both Previous Outcome ($F_{(1,23)} = 4.54$, $\eta_p^2 = .17$, $p < .05$) and Current Option ($F_{(2,46)} = 6.06$, $\eta_p^2 = .21$, Greenhouse-Geisser adjusted $p = .01$) had significant main effects on IPI. The interaction term was not significant ($F_{(2,46)} = 1.50$, $p = .23$). Therefore, rats take longer to deliberate between

options on trials preceded by a loss, and this effect is independent of option choice.

Rats exhibit reference-dependent behaviour

Within the task, a rat could incur losses and gains based on its stock and option selection. After the rat had made its selections, the profit (loss) was translated into a liquid equivalent and added to (subtracted from) a 0.15 ml reference point. To determine whether or not rats' behaviour demonstrated reference-dependence, we evaluated the movement time (MT) to collect reward after the tone onset. Given that even loss trials resulted in reward, there should be no observable difference in MT after a loss tone vs. gain tone if rats did not form a reference point at 0.15 ml. The exception to this would be trials where a rat incurred a very large loss (greater than the liquid equivalent of 0.15 ml), resulting in a payoff of zero reward. However, such trials represent on average only 0.7% of all trials encountered by the animals and are therefore unlikely to strongly bias such behaviour.

We performed a repeated-measures ANOVA on rats' mean MT with Current Outcome (3 levels: loss, gain and dividend) as a within-subjects factor. The results indicate that the outcome signalled by the gain/loss tone has a large significant effect on movement time to reward ($F_{(2,46)} = 43.88$, $\eta_p^2 = .66$, Greenhouse-Geisser corrected $p < .001$). Rats were 1.73 sec (*SEM* = 0.25 sec) faster to collect a dividend reward compared to a trading loss reward ($p < .001$), and 1.51 sec (*SEM* = 0.22 sec) faster to collect a trading gain reward compared to a trading loss reward ($p < .001$). There was no significant difference between MT after a trading gain and a dividend gain ($p = .13$).

Rats are faster to collect a gain reward than a loss reward

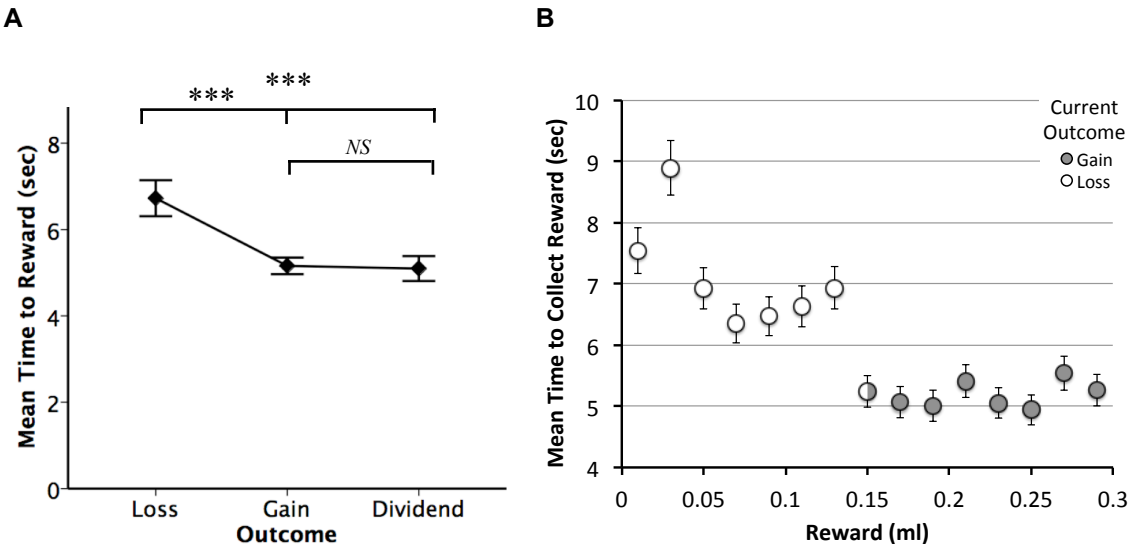


Figure 37: (A) Upon hearing a gain tone (as the outcome of either a trading profit or a dividend payout), rats moved ~2 sec faster from the tone onset to collect reward from the spigot on the opposing wall. This is contrasted with rats' average movement time to reward after a loss tone. (B) The trend is alternatively displayed in 0.025 ml reward bins, with grey circles indicating movement time to collect reward after a gain and white circles indicating movement time after a loss. ***p<.001, *NS* = Not Significant.

In support of these analyses, we also found other behavioural measures suggestive of reference-dependency. For instance, we observed that rats spent on average 250 msec longer licking at the reward spigot (PPL) on a gain trial compared to a loss trial ($t(23) = 4.68$, $p < .001$), suggesting that rats 'savoured' the gain rewards more (Wilson et al., 2006). Animals also increased the number of erroneous pokes into unlit holes on trials with a loss outcome (M = 2.60 pokes/trial, SEM = 0.14) with respect to a gain (M = 2.06 pokes/trial, SEM = 0.13), which was supported by a paired-sample t-test $t(23) = 5.63$, $p < .001$. This suggests that the erroneous pokes observed in our task are in part an affective response to losses (e.g. frustrative nonreward). Together, these results provide strong evidence that the reference-point manipulation was successful.

Rats are more risk-seeking after a loss than after a gain (loss-aversion)

We next analysed rats' choices on a trial-by-trial basis to determine whether or not individual animals' behaviour was significantly affected by a previous loss compared to a previous gain. Under Prospect Theory (Kahneman & Tversky, 1979), loss aversion manifests in repeated gambles as an individual's increased propensity to take risks following a loss. Thus, we analysed whether there were any differences in choices of the risky trade options vs. the safe hold option on trials with a previous loss vs. previous gain. We performed a 2-way repeated-measures ANOVA on the mean proportion of trials that each option was chosen. Within-subject variables were the Current Option (3 levels: Buy, Hold and Sell) and Previous Outcome (2 levels: Gain and Loss). The analysis indicates that rats chose the hold option on a greater proportion of trials overall (Effect of Current Option: $F_{(2,46)} = 37.81$, $\eta_p^2 = .62$, $p < .001$), but that this effect was attenuated by a previous loss trial compared with a previous gain trial (Interaction Effect of CurrentOption*PreviousOutcome: $F_{(2,46)} = 10.63$, $\eta_p^2 = .32$, $p < .001$). We subsequently performed paired-sample t-tests to assess individual contrasts. On trials with a previous loss, rats reduced their choice of the safer hold option by nearly 10% compared to trials with a previous gain (Loss: 43.1% *vs.* Gain: 51.7%, $t(23) = 3.99$, $p = .001$). Instead, rats were increasingly likely to choose the riskier buy (Loss: 29.6% *vs.* Gain: 24.6%, $t(23) = 2.64$, $p < .05$) or sell options (Loss: 27.4% *vs.* Gain: 23.6%, $t(23) = 2.38$, $p < .05$).

Loss-averse behaviour in rats

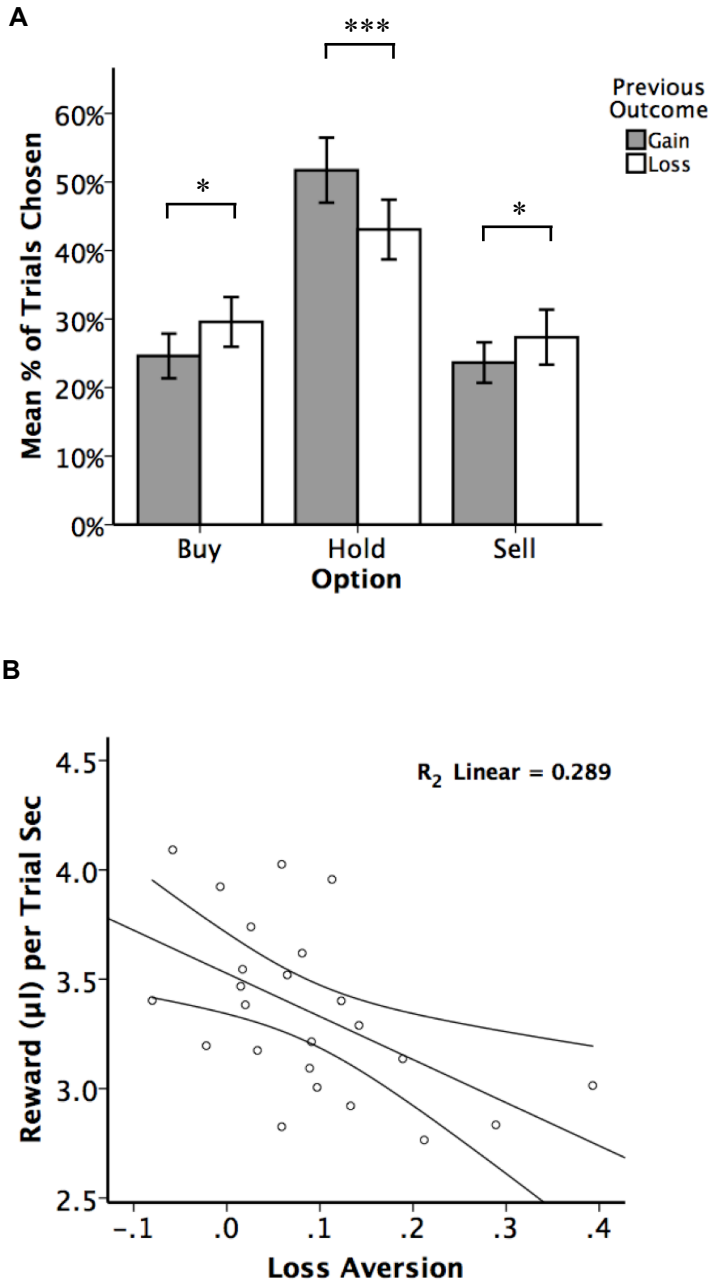


Figure 38: (A) On trials immediately preceded by a gain, rats select the 'safe' hold option on 51% of trials. The remaining proportions of choices were distributed evenly between the buy (25%) and sell (24%) trading options. When the current trial was preceded by a loss, rats chose the hold option nearly 10% less overall. Instead, rats increased selection of the two trade options. Since the trade options are associated with more variability in reward volume, this redistribution of choices represents a general shift toward more risk-seeking behaviour. Post-hoc paired-sample t-tests revealed significant differences between previous gains and loss at each the buy ($t(23)=2.64$, $p<.05$), hold ($t(23)=3.99$, $p=.001$) and sell ($t(23)=2.38$, $p<.05$) options. *$p<.05$, ***$p<.001$

(B) As an individual measure of loss aversion, we calculated the difference in proportion of choices between gains and losses in the hold option shown in Figure 5a above. Loss aversion was negatively correlated with average rate of reward earned per unit of time across sessions (Pearson's $r = -.54$, $p<.01$). In a linear regression analysis, individual loss aversion significantly accounted for nearly 30% of average reward (µl) earned per trial second.

We were therefore able to use the difference between the proportions of trials on which the hold option was chosen after a gain vs. after a loss as a proxy measure of loss aversion. This measure also correlated with a general bias against selection of the buy and sell options (Pearson's $r = -.54$, $p < .01$). We found that this individual measure of loss aversion was also negatively

correlated with the average rate of reward that a given rat earned per session (Pearson's r = -.54, p < .01). A simple linear regression analysis revealed that loss aversion was a significant predictor of lower volumes of reward per trial second, R^2_{adj} = .26, β = -.54, and where the equation was significant: $F_{(1,22)}$ = 8.95, p < .01. This indicates that loss aversion was disadvantageous in the task.

Rats perform suboptimally when trading

In the task, rats were given information about the market volume of a stock (i.e. the cumulative number of shares being held by all four rats) via the relative frequency of blinking LED lights. Lights blinked faster as rats bought more shares of a stock, and slower as rats sold shares of a stock. It was therefore necessary to investigate whether rats were capable of using memory of blink rates from previous trials to guide behaviour on a current trial (i.e. contrasting the previous blink rate with the current blink rate). Within the task, an optimal strategy would be to sell when a stock price has gone up since the previous trial, to buy when the stock has gone down since the previous trial, and to hold when the stock has made no change in price. Note that due to the interleaved forced-choice trials, rats did not always see the price of a given stock on every trial. Thus, Price Change was calculated as the difference between the price of the selected stock on the current trial and the price of the selected stock on last trial that it had been selected, and not necessarily the trial directly preceding it.

We carried out a 2-way repeated-measures ANOVA with direction of Price Change (3 levels: No Change, Up, Down) and average proportion of trials on which an Option (3 levels: Buy, Hold, Sell) was chosen as within-subject

factors. The average count of each the buy, hold and sell options per rat were included as covariates to control for potential effects of individual rates of trading vs. holding.

As previously discussed (see Figure 38a), rats chose the hold option on average on about half of all trials. This serves as an immediate indicator that rats are neither performing optimally nor randomly (33%) in the task. The results of this analysis indicated that the strong preference for the hold option did not significantly change between conditions, although it was slightly higher in the no change condition (M = 54.78%, SE = 0.76%) than conditions where the price increased (M = 51.68%, SE = 0.73%) or decreased (M = 50.73%, SE = 0.96%). Overall, there was no significant interaction between Price Change and the Option Choice ($F_{(4,80)}$ = 0.58, p = .68).

We also classified option selection decisions based on whether the decision was optimal or suboptimal given the change of price from the previous trial (illustrated in the right panel of Figure 39b). A repeated measures ANOVA was carried out on the proportion of optimal decisions resulting from choice of each the buy, hold and sell options, with Option (3 levels: Buy, Hold, Sell) as a within-subject factor. Post-hoc one-sample t-tests were carried out to establish significance of sample means from chance (33%). Although we did observe an increase in the proportion of optimal hold decisions (M = 36.95%, SE = 1.22%) compared to buy and sell decisions, the main effect fell short of significance (Option: $F_{(2,46)}$ = 2.95, p = .06). On average, rats' optimal decisions to buy (M = 31.85%, SEM = 1.47%) and to sell (M = 32.30%, SEM = 1.48%) did not differ from chance (33%). However, the increased selection of the hold option on

trials with no change led to optimal responding that significantly exceeded chance (*M* 37.00%, *SEM* = 1.23%, *t*(23) = 2.95, *p* < .01), although this should be interpreted with caution given the lack of significance in the ANOVA results. Overall, these group-level analyses suggest that rats did not learn to perform optimally according to the task contingencies.

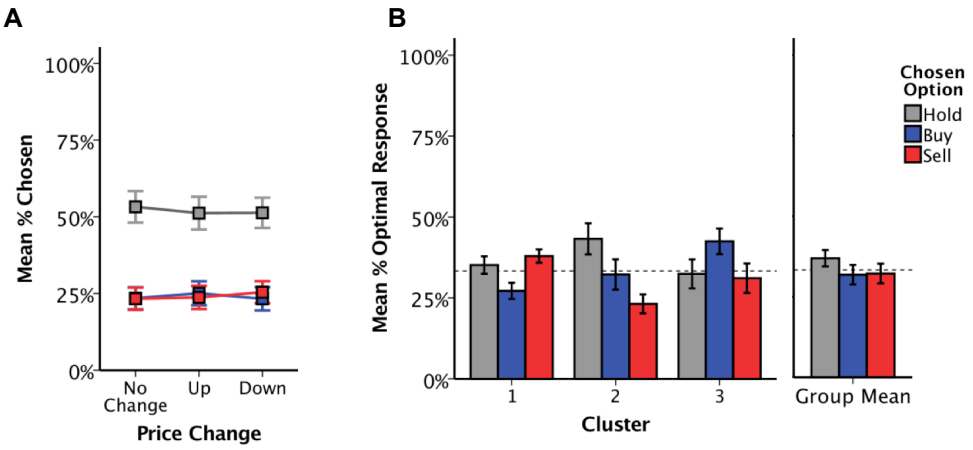


Figure 39: Optimal performance in the task would reflect three strategies: 1) to hold on trials where there has been no price change, 2) to choose the buy option when price had gone down, and 3) the sell option when price had gone up. (**A**) At the group level, there were no significant differences in the mean proportion of choices of each the buy/hold/sell options relative to the selected stock's price change from the previous trial. Overall, rats chose the hold option on 52.4% of trials. (**B**) At the group level (right panel, N=24 rats), rats performed around chance when choosing the buy and sell options, with a modest increase in optimal hold responding that exceeded chance (*t*(23)=2.95, *p*<.01). Cluster analysis (left panel) classified three subsets of rats based on optimal responding in either the sell (Cluster 1, N=12), hold (Cluster 2, N=7) or buy (Cluster 3, N=5) option. Thus, it appears that different individuals learned one of the three optimal strategies, often at the expense of another. Dotted lines at 33% represent chance. Error bars represent 95% CI's.

We then investigated whether there were any subsets of rats that had learned a partial strategy. Given that there were three optimal learning rules (buy when down, sell when up, and hold when there is no change), it stood to reason that individuals learned one or even two of the strategies without learning all three. We performed a hierarchical cluster analysis with the squared Euclidean distance of the average linkage (between groups) method on the percentage of optimal responses when selecting each the buy, hold, and sell options.

We identified three clusters (see left hand panel of Figure 39b) based on these three indices. Rats in cluster 1 (N=12) exhibited increased optimal selling behaviour, but also suboptimal sell behaviour and chance hold performance. Rats in cluster 2 (N=7) learned to hold optimally, but sell behaviour suffered and choice of the buy option did not differ from random behaviour. Cluster 3 (N=5) exhibited high optimal performance when choosing the buy option, while optimal choice of the sell and hold options was at chance. We performed a mixed 2-way ANOVA with Option Choice (3 levels: Buy, Hold, Sell) as a within-subjects factor and Cluster as a between-subjects factor. A significant interaction between Option Choice and Cluster ($F_{(2,42)} = 25.92$, $\eta_p^2 = .71$, $p < .001$) confirmed that the cluster analysis was able to account for over 70% of the variance in optimal responding between options. Furthermore, planned contrasts revealed a strong interaction between optimal responding in the buy and sell options within each cluster ($F_{(2,42)} = 48.99$, $\eta_p^2 = .82$, $p < .001$), which suggests that rats learned either to buy or sell at the expense of the other. Together, these analyses suggest that rats were capable of both perceiving and learning from changes in blink rate, but that individual rats learned only one of three optimal strategies. Rats had particular difficulties learning both the optimal buy and sell strategies together, which makes sense given that the optimal buy and sell strategies require the same operational response, but the opposite instrumental response, to the other.

We next sought to investigate the possibility that optimal responding could be driven by responses to market information from the current trial alone. In other words, we sought to clarify whether rats' decisions to buy/hold/sell were based on the change in volume between trials or rather on the current volume alone.

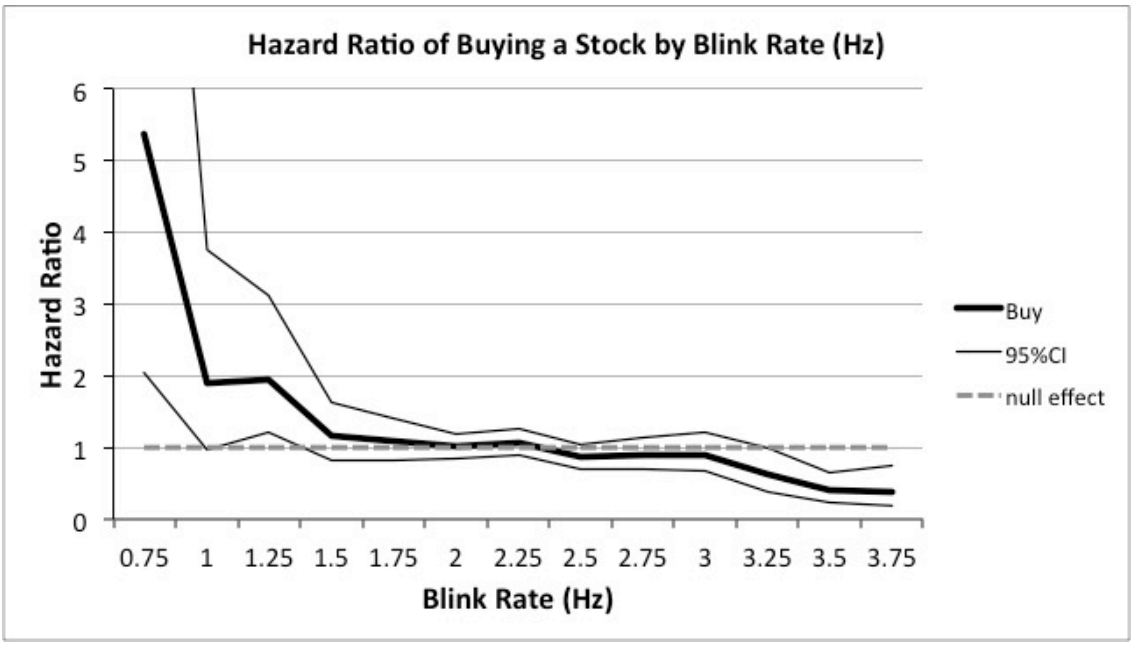
We began by investigating whether rats chose stocks with greater market volumes on average, which could reflect a potential confound in stimulus saliency and stock selection (i.e. faster blink rates lead to a greater proportion of choices). We established the mean difference in chosen vs. unchosen stock volumes by subtracting the average unchosen stock volumes from the average chosen stock volume on a subject-wise basis. On average, the chosen option had a total volume that was 3.00 stocks (*SEM* = 4.11 stocks) greater than the average of the two unchosen options. A one-sample *t*-test revealed that this was not significantly different from zero, $t(23) = 0.73$, $p = .47$. Thus, we could conclude that rats' initial stock choice was not a function of its current volume. We then carried out a Univariate ANOVA with mean difference in chosen and unchosen stock volumes as the dependent variable and cluster as a between subjects variable. The results indicated that there was no significant effect of cluster on the mean difference of chosen vs. unchosen stocks ($F_{(2,21)} = 0.50$, $p = .95$). This suggests that there was no significant effect of stimulus saliency (i.e. blink rate)/market information (i.e. market volume) on the initial poke (i.e. stock choice) in isolation.

In order to investigate whether rats' reacted to changes in market information in their combined first and second pokes (i.e. the choice to buy, hold, or sell the selected stock), we began by performing a mixed ANOVA on the market volume of the selected stocks at each Chosen Option (within-subjects, 3 levels: buy, hold, sell) with Cluster as a between-subjects factor. The results indicate that there was a main effect of option ($F_{(2,42)} = 3.32$, $\eta_p^2 = .14$, $p < .05$), and planned contrasts revealed that this effect was primarily driven by significant differences in chosen volume between the buy and sell options ($F_{(2,21)} = 7.17$,

η_p^2 = .26, p < .05). We found that the average market volume of a selected stock was significantly lower when a rat chose to sell compared to when it chose to buy ($M_{Difference}$ = 8.71 shares, SE = 3.25 shares), and that this did not significantly differ between clusters (Option*Cluster: $F_{(4,42)}$ = .85, p = .50). If optimal responding was the product of decisions on a single trial, we would expect rats to chose the sell option more often at higher market volumes and the buy option more often and lower market volumes. Given that we observe the opposite effect here, we can conclude that optimal responding in the task was not the result of rats learning a general rule to buy at low blink rates and to sell at high blink rates.

Given that we utilized blinking lights to convey market information, it is not possible to decouple changes in information content from changes in information saliency in the current task. Acknowledging this point, we modelled the likelihood that a rat would buy or sell based on the market volume and information saliency (i.e. Blinking Hz) on any given trial. We employed the Cox Proportional Hazard model to identify the hazard ratio of buying or selling (separately) at varying blink rates. Here, we create dummy variables for each stock on every free-choice trial according to its current blink rate. The dummy variables span from <0.75 to ≥4.00 Hz in 0.25 Hz bins. A bin with a hazard ratio of 1 corresponds to a null effect on the trading choice. Hazard ratios above 1 indicate that the likelihood of buying/selling is higher in that blink rate range, while hazard ratios below 1 denote a reduced likelihood of trading in that range of blink rates. Stratifying over subjects, we also included either buy count (for the model) or sell count (for the sell model) as a time-varying covariate in the model.

A



B

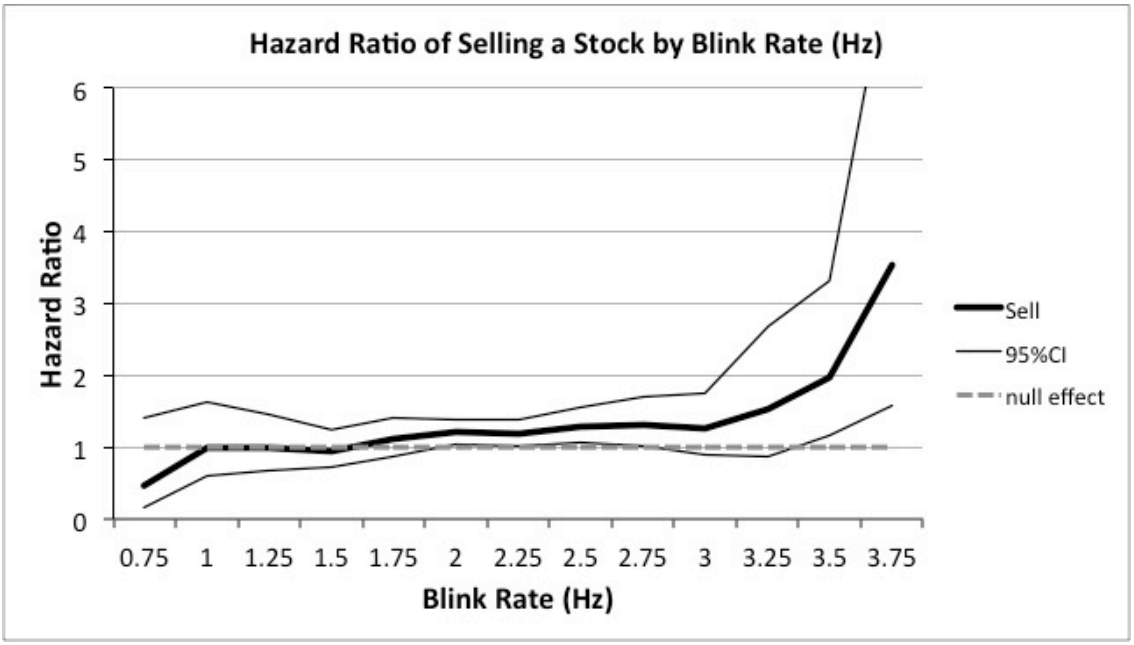


Figure 40: (A) The blink rate of any given stock significantly increased the hazard of purchase at rates below 1.5 Hz. Increased purchase rates at slower blink Hz translates to rats buying stocks with lower market volumes and lower prices more frequently. When blink rates exceed 3.25 Hz, subjects were significantly less likely to select the buy option. See Table 14 in Appendix 4 for further particulars regarding model coefficients and *p*-values. (B) The hazard of selling a stock was significantly increased at average blink rates between 1.75 Hz and 2.75 Hz and at high blink rates above 3.25 Hz. There was no significant effect of market information/salience on selection of the sale option at blink rates below 1.75 Hz. See Table 15 in Appendix 4 for further details regarding model coefficients and *p*-values.

In Figures 39a-b, it is apparent that market information/saliency has less of an effect on trading at average values, but has a high impact on rats' choices at the extremes (either blinking very fast or very slow). Rats are more likely to purchase a stock at very low market volume (slow blink rates) and less likely to buy a stock that has much greater market volume (fast blink rates). This is equivalent to buying an undervalued stock and not buying an overvalued stock. The opposite is true for sales, whereby rats are significantly more likely to sell a stock when it has either an average market value between 1.75 – 2.75 Hz, or a particularly large market value above 3.25 Hz. From these analyses, we conclude that rats' trading behaviour is sensitive to large changes in market information, and that this behaviour conforms to both explanations of rational trading (e.g. buying stocks that are undervalued) and stimulus saliency. However, we cannot conclude that rats respond optimally within the blink rates centred around the mean. Given these results and those of our previous ANOVA of average chosen market volume, it is likely that rats did not dissociate between small differences in blink rate, but did dissociate between particularly large differences in blink rate.

Rats exhibit the disposition effect

In actual stock markets, empirical research suggests that humans behave suboptimally in sell decisions depending on whether the potential return is positive or negative (i.e. the disposition effect). In the simulated rat stock market presented here, the return on an 'investment' can be determined by evaluating price changes over the series of trials between a choice of the buy option and a subsequent sell option. Over the 7 testing sessions, returns from selling a stock ranged from -57.1% to 26.5% (*SD* = 8.5%), with an average

return of 0.3%. Here, we have a first indication that rats may exhibit something akin to the disposition effect based on the negative bias in the range of experienced outcomes. If rats held losses for too long – or rather, persistently avoided the sell option over successive trials – this would increase the potential for more negative returns once a loss was finally realized. Alternatively, if rats were selling gains too quickly, they would not get the chance to experience an equivalent level of positive returns. To further this line of inquiry, we follow the methodology of behavioural finance studies based on human subjects (e.g. Barber & Odean, 2011; Frydman et al., 2014; Odean, 1998).

We begin by determining PGR and PLR for each rat across each stock and session (see Methods). We found that rats had an average PGR of .12 (*SEM* = .02) and an average PLR of .09 (*SEM* = .01). We then calculated the average difference between PGR and PLR per subject. There was no significant correlation (Figure 41a) between PGR and PLR, (Pearson's $r = .33$, $p = .12$). The results of a paired-sample t-test reveal that individual rats realize gains significantly more often than they realize losses, $t(23) = 2.22$, $p < .05$.

As often observed in human studies, we found that individual subjects exhibited a range of disposition effect strengths. Effect sizes ranged from -.09 to .23 (*M* = .04, *SEM* = .02), with higher positive values indicating stronger tendencies to hold losses too long and sell gains too quickly. This effect (Figure 41b) was not significantly correlated with greater average volumes of reward per trial second (Pearson's $r = .10$, $p = .65$). There were also no correlations between disposition effect strength and subjects' average proportion of sales per

session (Pearson's r = .13, p = .53) or overall proportion of trades (*vs.* holds) per session (Pearson's r = .30, p = .15).

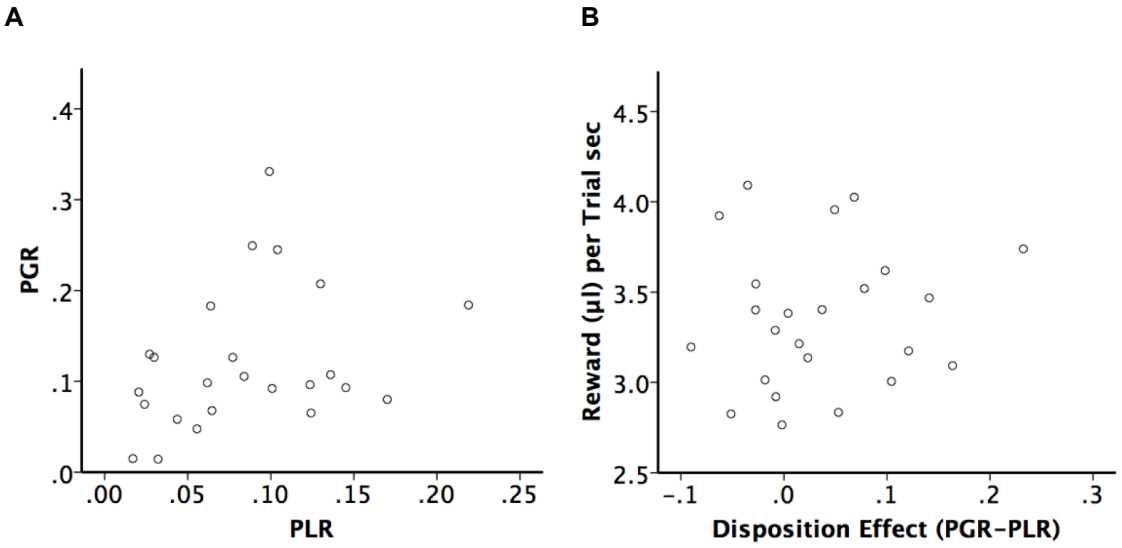


Figure 41: (A) The proportion of gains realized (PGR) and the proportion of losses realized (PLR) were not significantly correlated within individual subjects (Pearson's r = .33, p = .12). (B) The disposition effect, calculated as the difference between PGR and PLR, was not significantly correlated with increased or decreased rates of reward (µl/trial sec), Pearson's r = .10, p = .65.

Our analysis of the disposition effect proceeded by fitting a Cox Proportional Hazards model (see Methods). Here, we create dummy variables for each stock on every trial according to its current return (in the case of a realized sale) or its current potential for return (in the case of an unrealized sale). The dummy variables span from <-22% to ≥18% return in bins spanning 4% each. A bin with a hazard ratio of 1 corresponds to a null effect on the trading choice. Hazard ratios lying significantly above 1 indicate that the likelihood of buying/selling is higher in that range of returns, while hazard ratios below 1 denote a reduced likelihood of trading in that range of returns. Stratifying over subject, stock, and session, we also included the subjects' sales count and counterfactual reward as time-varying covariates in the model. Counterfactual rewards represent the volume of reward a rat *could* have earned had it chosen

a different option. Interestingly, the addition of counterfactual reward significantly increased the model's R^2 from .57 to .65 ($\chi^2(1) = 635.23$, $p < .001$). This suggests that rats' choices were influenced by the potential reward outcome of non-selected actions.

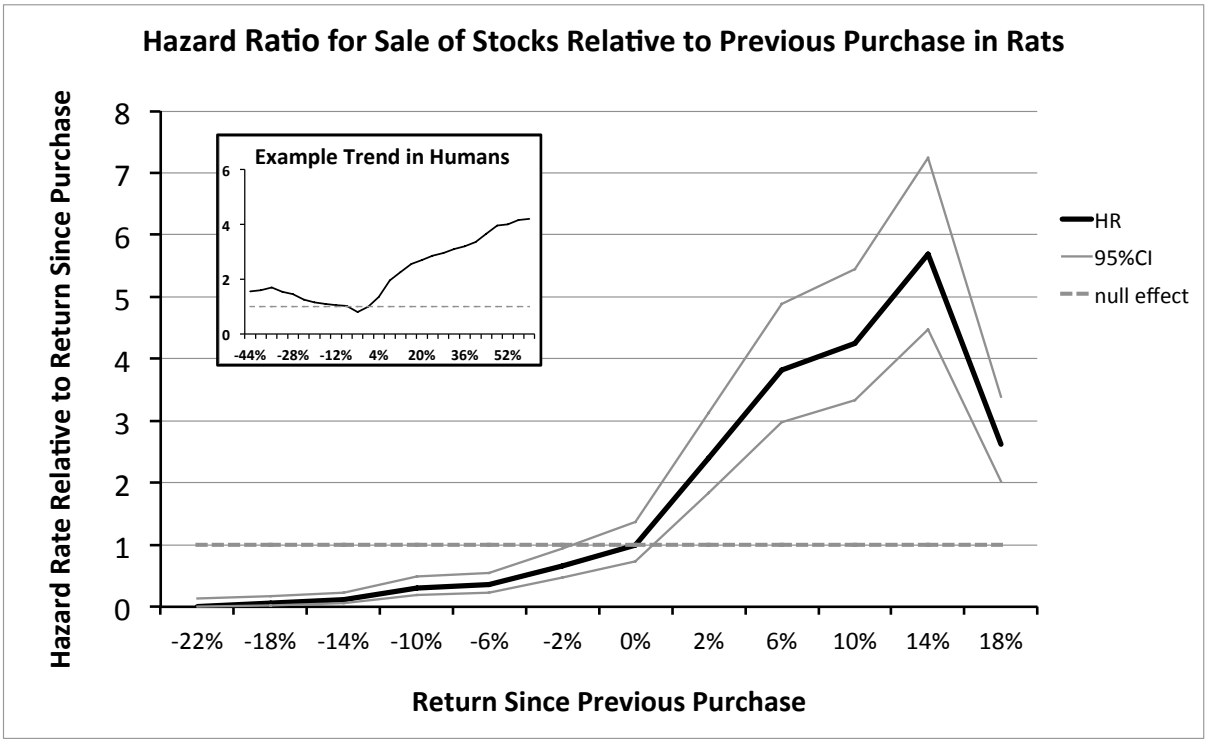


Figure 42: The hazard ratio that a rat selects the sale option on any given trial as a function of return. The hazard rate for each return bin is calculated relative to zero return. Peaking around 14% returns, rats are up to 500% more likely to sell a stock when returns are positive. The opposite is true in the case of negative returns, whereby subjects demonstrate an increasing disposition against realizing losses as returns become more negative. To facilitate comparison, the inset at the top-left depicts typical human behaviour (e.g. Barber & Odean, 2011). See table 13 in Appendix 4 for further particulars regarding model coefficients and p-values.

Figure 42 illustrates the 'hazard' of an animal selling a stock on a given trial based on the potential returns. Similar to human behaviour, rats are more likely to sell a stock at a gain than at a loss. An interesting deviation from human behaviour however, is the observation that rats are far less willing to sell at a loss. While this rather robust effect may represent a learned avoidance of immediate losses in the task, it is also probable that rats lack the information

and overview to ascertain that selecting the sale option at a loss will reduce the potential for larger future losses.

In order to determine whether loss aversion might be driving the disposition effect, we analysed the partial correlations between individual measures of loss aversion and the disposition effect while controlling for trading frequency. We found no significant correlation (Pearson's r =.15, p = .51). This suggests that loss aversion does not lead to the disposition effect in our task. We speculate as to the potential cause of disposition effect-related behaviour further in the discussion.

Finally, we investigated whether the three different optimal strategies learned by each subset of rats also represented valid clusters with regard to the strength of their individual behavioural biases (i.e. the disposition effect, loss aversion, and anchoring). We conducted separate Univariate ANOVA's with individual measures of each of the three behavioural biases as dependent variables and Cluster membership as a between-subjects variable. To attain a single individual measure of the anchoring effect, we computed the difference between average latency to collect reward after a trading loss vs. gain (MT_{Loss} – MT_{Gain}) per rat. The results of our analyses are illustrated in Figures 43a-c.

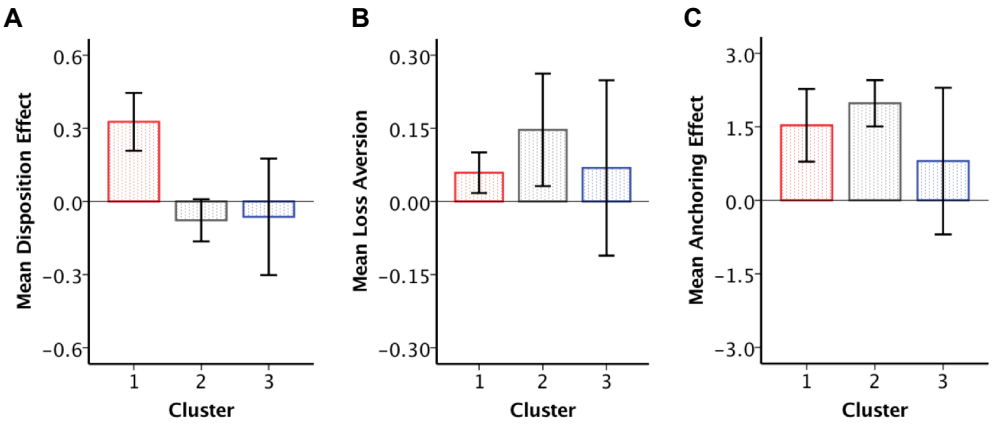


Figure 43: The composition of individual measures of three behavioural biases was examined across clusters, where: Cluster 1 (N=12) responded optimally when selling, Cluster 2 (N=7) responded optimally while holding and Cluster 3 (N=5) responded optimally while buying. (**A**) There was a significant main effect of cluster membership on individual disposition effect measures, $F_{(2,21)}$=17.13, η_p^2=.62, p<.001. Rats that sold optimally (Cluster 1) exhibited a significantly greater disposition effect than those that learned to hold (p<.001) or sell (p<.01) optimally. There was no significant difference in the disposition effect between rats in Clusters 2 and 3 (p=*NS*). (**B**) There was no significant main effect of cluster membership on individual loss aversion ($F_{(2,23)}$=1.69, p=NS) or (**C**) on anchoring ($F_{(2,21)}$=1.91, p=NS), but these effects may have been obscured by the relatively low number of animals in Cluster 3 (N=5). Error bars represent 95% CI's.

We found that individual measures of the disposition effect varied significantly between clusters, $F_{(2,21)}$=17.13, η_p^2=.62, p<.001, (Figure 43a). Interestingly, planned contrasts revealed that rats with the highest proportion of optimal responding during selling also demonstrated the highest strengths of the disposition effect (M = 0.33, SE = 0.05) compared to clusters 2 (p < .001) and 3 (p < .01). Therefore, we cannot argue that rats exhibit the disposition effect due to a lack of understanding about how to sell optimally in the task. In fact, those rats that learned to hold optimally (Cluster 2) and buy optimally (Cluster 3) exhibited negative disposition effects on average (M = -0.08, SE = .06 and M = -0.06, SE = .08, respectively). This indicates that rats in clusters 2 and 3 realized losses slightly more quickly than gains on average, although not significantly so.

Loss aversion was expressed most strongly in rats from Cluster 2 (M = 0.15, SE = .04). Rats in clusters 1 and 3 exhibited relatively lower levels of loss

aversion (Cluster 1: *M* = .06, *SE* = .03; Cluster 2: *M* = .07, *SE* = .05). While this trend was interesting, the results of the Univariate ANOVA indicated that individual measures of loss aversion did not significantly vary between clusters, $F_{(2,23)}$=1.69, *p* = .21. Similarly, we found that the anchoring effect was strongest in Cluster 2 (*M* = 1.98, *SEM* = 0.39) compared to clusters 1 and 2 (*M* =1.53, *SEM* = 0.30; *M* = 0.80, *SEM* = 0.46, respectively). However, these differences were not significant, ($F_{(2,21)}$=1.91, *p*=NS). It should be noted that the relatively small number of rats that learned to choose the buy option optimally (Cluster 3, N = 5) might have obscured potential significant effects.

Discussion

In this paper we develop a reinforcement-learning task that creates a simulated stock market between cohorts of four rats 'trading' in networked operant boxes. Our results suggest that rodents successfully establish a reference point in Block 1, which has never before been demonstrated in rats. In Block 2, we found that certain subsets of rats learned an optimal buy, hold, or sell strategy (but not all three together). Our analyses indicate that rats exhibited two further suboptimal patterns of behaviour that are well-established in humans (and to some extent in non-human primates as well, see Santos and Platt (2014) for a review), but have never before been explicitly demonstrated in rodents: 1) loss-aversion, and 2) the disposition effect. Thus, rats not only demonstrated reference-dependency (otherwise known as anchoring) to experimenter-determined expectations in the current task, but the animals also became more risk seeking on trials immediately preceded by a loss, as predicted by loss aversion in Prospect Theory (Kahneman & Tversky, 1979). Furthermore, we also found that rats had a tendency to hold on to their losses for too long while selling their winners too soon – behaviour that is hallmark of the disposition effect in humans. This effect was primarily exhibited by rats that had learned the optimal sell strategy, which suggests that it was unlikely the result of random selling behaviour in the task, but rather a natural consequence of selling behaviour that was focused on short-term profits rather than long-term returns. Together, these results suggest that behavioural biases such as anchoring, loss aversion and the disposition effect are more deeply rooted in humanity's evolutionary past than previously considered.

Reference Dependence (Anchoring)

At the start of every testing session, rats completed a block of 15 trials that always resulted in a set volume of 0.15 ml sweet liquid reward. Block 1 was intended to establish a reference point for the subsequent trading block. To facilitate this, rewards less than 0.15 ml were paired with a loss tone, while rewards at or above 0.15 ml were paired with a gain tone during Block 2. Although the average loss trial still resulted in an average of 0.10 ml reward, rats took about 2 sec longer (~10% of the average trial time) to collect reward after hearing a loss tone as opposed to a gain tone. Before even experiencing the amount of reward, rats' lugubrious approach after hearing the loss tone implies that the stimulus reshaped expectations about the desirability of the reward. Thus, rats' behaviour is consistent with the idea that the loss tone conceptually reframed the 0.10 ml reward as a 0.05 ml loss.

Aside from moving more slowly to collect reward, rats react to trading losses in other quantifiable ways as well. This is evinced by the observation that rats significantly increase the number of erroneous pokes into unlit holes on losing trials. Such responses may represent a frustrative reaction to the aversive stimulus (Amsel, 1958; Rescorla, 1992), despite the fact that animals are still receiving positive reward on (nearly all) loss trials. While it may be argued that animals find the tone itself aversive, this would not explain why rats move more slowly to collect reward. Furthermore, rats spend less time licking at the reward spigot (PPL) after a loss reward delivery compared to a gain reward delivery. PPL is thought to reflect the notion of 'savouring' in rats (Wilson et al., 2006), which would imply that rats savour losses less so than they do gains. All of these measures suggest that rats not only have the cognitive capacity to exhibit

reference-dependent behaviour, but that reference-dependent behaviour has been selected for throughout mammalian evolution.

Loss Aversion

In repeated gambles, loss aversion biases an individual toward risk-seeking behaviour on events preceded by a loss. After establishing that the hold option was associated with a lower variance in rate of reward (i.e. less risky) compared to either of the trade options, we conducted a trial-by-trial analysis to determine whether rats increased risk-seeking behaviour subsequent to a loss trial. On trials immediately preceded by a loss trial, we found that rats significantly decreased the average proportion of hold trials, choosing instead one of the riskier trade options (i.e. buy or sell). This serves as direct evidence that rats exhibit loss aversion.

Furthermore, individual measures of loss aversion were negatively correlated with the average rate of reward that a rat earned per session and accounted for over 25% of the variability in observed reward rates. While we found no significant group-level preferences between the three stocks as differentiated by their starting prices, future studies could also manipulate the inherent riskiness of each stock in order to determine whether loss aversion can also be observed in stock selection choices.

Disposition Effect

In this study, we employed analyses from the existing behavioural finance literature (Barber & Odean, 2011; Frydman et al., 2014; Kaustia, 2010; Odean, 1998) in order to: 1) determine whether rats exhibited the disposition effect, and

2) to increase the potential for translational links in behaviour between rats and humans. By calculating the difference between the proportion of gains realized (PGR) and the proportion of losses realized (PLR), we found that rats exhibited a significant disposition toward realizing gains more often than realizing losses. In line with Kaustia (2010) and Barberis and Xiong (2009), who argue that loss aversion does not account for the disposition effect, we find no correlation between loss aversion and the disposition effect. Intriguingly, we found that rats that demonstrated the highest proportion of optimal sell behaviour (Figure 30b) also exhibited the disposition effect most strongly (Figure 43a). This suggests that the disposition effect arises from a focus on short-term profits rather than long-term returns. These results are in line with the realization utility hypothesis (Barberis & Xiong, 2012; Frydman et al., 2014), whereby an individual is postulated to receive an instantaneous neural reinforcement (dis)utility signal at the moment of the sale, which biases behaviour towards a suboptimal short-term focus on gains/losses.

By fitting an extended Cox Proportional Hazard Model to the data, we demonstrate that the 'hazard' of selling is increased over positive returns and decreased over negative returns. Intriguingly, we find that counterfactual reward is a significant predictor in the model. This suggests that rats are considering the potential value of all available options when making a trading decision. This result has also been identified in a simulated stock market task using human participants (Lohrenz, McCabe, Camerer, & Montague, 2007). Using fMRI, the authors found that the difference between experienced outcome and the counterfactual reward – or fictive error – was associated with BOLD activity in the ventral caudate nucleus. Furthermore, dysfunctional

processing of counterfactual rewards has been implicated in disorders such as addiction and depression (Chiu, Lohrenz, & Montague, 2008; Platt & Hayden, 2011). Thus, this area represents a potentially exciting target for future work.

In general, the hazard model for sales as a function of returns bears a striking resemblance to the human data (Barber & Odean, 2011). It diverges in that rats are notably unwilling to sell at negative returns. It is possible that this reflects a species-level difference in the general willingness or ability to incur a smaller immediate loss in order to avoid a larger future loss. However, it may also be the case that the rats in our task simply did not have enough information/experience to come to such a conclusion.

We strictly limited the availability of information in the task to the current market volume of each stock and the outcome of a trade/hold option. This was done in the first instance in order to establish the contribution of reinforcement learning to the most basic elements of financial decision-making behaviour. In theory, restricting the kinds of available information to profit/loss and market volume afforded rats the ability to develop a personal memory-based trading history and to respond to other rats' responses (albeit without explicit knowledge of doing so). On the other hand, such a restrictive model did not allow rats to directly associate changing blink rates with other rats' actions or to ascertain an overview of current portfolio holdings. Although such capacities represent interesting additions to the task design and clearly constitute intriguing directions for future work, it is nonetheless striking that such behavioural biases can arise in a context where the potential for both theory of mind and portfolio optimization has been precluded. This suggests that behavioural biases such

as the disposition effect do not arise from mechanisms supporting higher levels of cognition, such as those implicated in the primate prefrontal cortex. In support of this notion, studies by Grinblatt, Keloharju, and Linnainmaa (2012) and G. Chen et al. (2007) demonstrate that the disposition effect is diminished in individuals with higher IQs and with greater trading experience, respectively.

Information salience & content

In the task, rats could glean information about market volume and respective changes in market volume from the blink rate of each stock's associated nosepoke hole. While the proportion of choices of the buy/hold/sell options did not approximate random choice (i.e. 33% distribution between the three choices), they also did not reflect optimal performance. Therefore it remains unclear precisely what kind of strategy the rats were employing in the task. Aside from trials with very fast blink rates and very slow blink rates, our results suggest that rats did not use changes in blink rate to perform optimally in the task. This may have occurred due to limitations in working memory and/or to an inability to dissociate between small changes in blink rate. This would therefore be a particularly relevant area for improvement in future task designs. Given that the visual acuity of the rat is quite poor, future studies might employ a different stimulus to convey market information, such as a tone with varying pitch. Strain on working memory may be ameliorated by changing the contingencies so that an optimal response is relative to a stock's historical average rather than the previous encounter with a given stock. Furthermore, it may be beneficial to carry out a related study in humans. Human participants may be able to more easily communicate any potential strategies that they may

have employed to maximize reward in the task, and this could in turn shed light on rats' behaviour.

Another important caveat to consider in the current task design is the correlation between information content and information salience. While we failed to demonstrate that trading behaviour was significantly influenced by blink rate on average, we did find some indication that rats' choices were affected by market information when blink rates were either very fast or very slow. Both salience-based and content-based accounts present plausible explanations of the observed behaviour. Namely, it is equally reasonable to believe that rats are responding to the information contained within the blink rate (e.g. low price = better value) as it is to believe that they are responding to the saliency of the information (e.g. faster blink rate = more attention). However, while the notion that highly salient stimuli attract more responses cannot be disentangled from high market volumes eliciting higher sell rates, the salience hypothesis cannot explain the observation that purchase rates increase at low blink rates. A rational investor will buy stocks that she deems to be undervalued, while selling stocks that she believes to be overvalued. Translated to the task, consider the example where only a small number of shares are currently being invested in Stock 1 between all four rats. In this situation, the market volume will be low, which represents an opportunity to purchase the stock at a bargain. Thus, rats using information content will have an increased propensity to select the buy option when blink rates are low and the sell option when blink rates are high, as was observed (see Figures 39a-b). While this does not exclude potential saliency effects at high blink rates, the information content explanation fits more closely with the data overall. Future

studies could control for such effects by inverting the blink rates (e.g. convey higher market volumes with slower blink rates).

The incorporation of two tones to indicate a loss or gain outcome before reward collection likely plays an important role not only in the presence of loss aversion as previously discussed, but also in eliciting the disposition effect. Frydman and Rangel (2014) demonstrate that making an investment's original purchase price less salient can attenuate the strength of the disposition effect. Therefore, drawing attention to the amount to be gained or lost in a potential transaction enhances the disposition effect. In a similar vein, our task ensures that information regarding a gain or loss is made explicitly salient via auditory stimuli, albeit after choice. One might therefore predict that by either omitting the tones or by making them less salient, one might also reduce the disposition effect in rats.

Future directions

We believe this task represents an exciting opportunity for future neuroeconomic research to explore the neural correlates of financial decision-making and its associated behavioural biases. For example, this task could be employed to further investigate the 'realization utility' hypothesis (Barber & Odean, 2011; Barberis & Xiong, 2012; Frydman et al., 2014). It would be possible to temporally separate the option selection nosepoke and the sound of the outcome tone in order to determine whether the 'realization utility' came from the action itself or from the signal of the gain/loss on a single neuron level. Furthermore, future studies could explore a number of manipulations of the task contingencies, such as higher transaction cost on trades (e.g. longer

required nosepokes), discrepancies between risk/reward between stocks, asymmetrical initial endowments or information (insiders), and visible competitors.

While future prospects are many, it is also important to bear in mind that there are a number of limitations associated with the current task. For example, we do not require that rats liquidate at the end of a session, so there is little consequence to continuously buying stocks, or never selling bad stocks. Rats are also paid a 'dividend' each time they select the hold option, which is not consistent with real life. It may be possible to employ a delayed reinforcement schedule on hold options, although this would almost certainly lead to much higher proportions of trading relative to holding. Another critical obstacle that we faced was the fact that rats did not learn the full optimal strategy with regard to price changes from the previous trial (i.e. buy when price had gone down, sell when price had gone up, and hold when there was no change). Instead, we found that rats learned one strategy well, at the expense of one or both of the other strategies. This was especially problematic between the optimal buy and sell options, where rats that learned to respond to changes in one direction (e.g. nosepoke to the left sell option when price increased) did not learn the reverse instrumental response (e.g. nosepoke to the right buy option when price decreased). This may be a critical cognitive limitation, or it may be a matter that could be resolved through more targeted training paradigms. Future studies might ensure that rats are trained to respond optimally to each option (buy, hold and sell) independently before testing commences.

As previously discussed, another limitation to the study is that market volume, information salience, and stock price are all correlated under the current task design. Thus, disentangling the distinct effects of each proves challenging. However, one could either invert blink rates or change the stimulus that signals information in the task to ameliorate this confound. Finally, despite the fact that rats are transported together in one transporter cage from the colony room to the testing room, it is very improbable that they have any understanding of social competition within the task. Even communication through high-frequency vocalizations is unlikely to lead a rat to associate the presence of another rat with changing blink rates. A future study in which rats perform the task in high visibility chambers may resolve this issue.

To conclude, the research presented here represents an initial effort toward modelling investor behaviour in rodents. Thus, this task forms a platform from which neuroscientists, psychologists, and financial economists alike may probe the neural underpinnings of financial decision-making. Future iterations of the task can be easily adjusted to address divergent research questions (e.g. asymmetry in initial resource endowment or insider information) as well as current shortcomings (e.g. optimal response training and correlation between information content and salience). However, the simple notion that the behavioural biases observed here instantiate from reinforcement learning provides key insight into the mechanisms that may be governing investor biases in the brain. Areas commonly implicated in reinforcement learning, such as the dopaminergic midbrain, represent an obvious target for primary investigations in the future.

Chapter 6

General Discussion

Summary

The psychology and neuroscience of economics, and specifically the notion of financial loss, represented a central theme of this thesis. I argued that animals, like humans, perceive, encode, and react to a loss of resources in a categorically different way than to a gain of the same magnitude. Specifically, this work extends ideas first proposed by Kahneman and Tversky (1979) from humans to rats. We began in Chapter 1 with a detailed introduction into normative models of economic valuation in order to demonstrate that prescriptions of 'rational' decision-making take objective valence as a presupposition. Put simply, economic models struggle to accommodate the subjective nature of losses and gains. At its very core, the definition of optimal, or 'rational,' behaviour is necessarily dependent on the basic notions of what constitutes a loss and what constitutes a gain. Given that an individual's observable behaviour is often an unreliable proxy for internal preferences, and that we know relatively little about the neural underpinnings of resource loss, we posited that a better understanding of how neural mechanisms instantiate behaviour may lead to more reliable models of economic decision making. Furthermore, we argue that rats represent a viable animal model of risky decision-making for neuroeconomic research.

The original research presented in Chapters 2 – 5 represents my endeavours to address this knowledge gap. By employing insight from psychology and economics, I developed models of rat behaviour that can be used to study the neural substrates of loss valuation in risky decision-making. I presented two behavioural paradigms (Chapters 2 and 5), while demonstrating novel loss-

related correlations between the midbrain dopamine system and loss behaviour in Chapters 3 and 4. The results presented in Chapter 5 demonstrate that rats are capable of producing behavioural patterns, such as loss aversion and the disposition effect, that economists studying human behaviour would recognize. While addressing this critical gap in the existing literature, this work has also highlighted a number of areas for future research.

Operationalizing Resource Loss

Resource loss is inherently difficult to operationalize in animal decision-making tasks, which represents a critical obstacle in translating loss-related research outcomes between species. The use of rats to study decision-making allows one to avoid many of the potential confounds arising during the study of human participants, such as numeracy or a pre-existing notion of how a stock market works. However, in rat tasks it is not possible to create a token economy, nor is retracting a previously consumed reward an effective means of simulating loss. To overcome this problem, previous rat models of risky decision-making have often incorporated either punishment (e.g. footshock) or opportunity costs (e.g. time-out). Unfortunately, it is difficult to ascertain whether these substitutes are subserved by the same neural mechanisms as those resulting from human representations of loss. For example, consider the purchasing strategy of 'in-app purchases' increasingly employed by game developers, where gamers have to choose between a time-out from game progression (i.e. opportunity cost) or a small monetary cost (i.e. resource loss). The strategy is extremely successful, because gamers are often willing to pay the small sum to forgo the time-out and resume game play immediately. This serves as anecdotal

evidence that different types of costs are not necessarily equivalently perceived. To address this, we developed two novel rat decision-making tasks, each operationalizing 'loss' in ways that we believe constitute closer representations of resource loss.

The gambling task (Chapters 2 – 4) was designed with the specific intention of eliciting a mental representation of potential losses. Thus, a loss in that task is represented by the loss of a resource that could have been consumed had the rat chosen differently just a moment earlier. In the stock market task (Chapter 5), rats were primed to perceive rewards that were smaller than a reference point as losses and those that were larger than the reference point as gains. Losses in this task were based on the prediction that rats, like humans, behave in a qualitatively different way when expectations are violated compared to when they are exceeded. On the single-neuron level, the gambling task provides a means for researchers to identify cells that putatively encode losses. On the behavioural level, the stock market task opens up the possibility of characterizing the transformation from loss-related learning to behavioural bias. Together, these two approaches to modelling resource loss in rats offer researchers the opportunity to tease apart the effects of loss on both brain and behaviour. Each adds a unique thread to the repertoire of the rapidly expanding neuroeconomics literature.

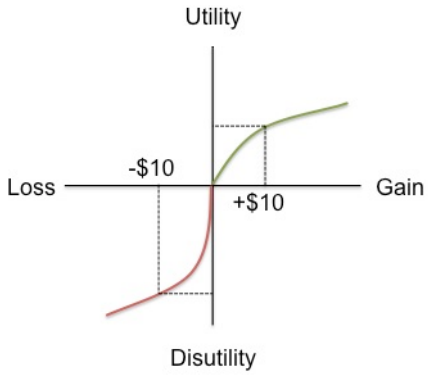
Loss-dependent behaviour

Humans exhibit a number of suboptimal behaviours in the wake of a loss. In an attempt to break even, gamblers often 'chase' their losses. Similarly, investors tend to hold on to losing stocks too long in the hope that the declining share

price might make a recovery. Both of these behaviours rely on the notion that some reference point has been set, and that this reference point is commonly *not* at absolute zero. For example, a professional trader trying to meet a 10% return target might perceive a 5% return as a loss, despite the fact that it is still a positive return. Thus, it is imperative that any study of the neural mechanisms subserving losses vs. gains takes reference-dependence into account.

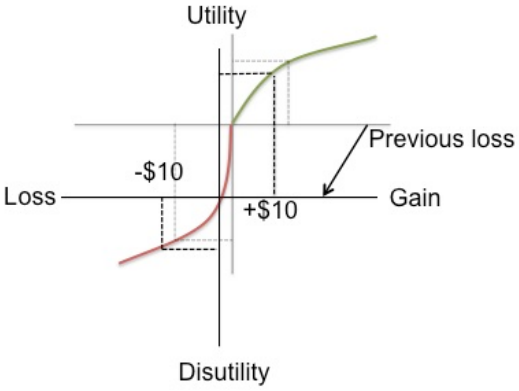
Given a single 'one-off' decision (see Figure 42a), Prospect Theory (Kahneman & Tversky, 1979) predicts that an individual will be risk-seeking over losses and risk-averse over gains. In other words, the loss-averse individual is willing to accept more risk in order to avoid a loss than to win the same amount, because she is more sensitive to decreases in wealth than to increases in wealth. Extended to a series of investment decisions, the effects of loss aversion depend on prior experiences of losses and gains. In general, Prospect Theory predicts that an individual will take more risk after a loss and less risk after a gain. It is thought that this occurs because the individual will either be in the convex (risk-seeking) domain after a loss, or in the concave (risk-averse) domain after a gain. This is exemplified by a study in which Coval and Moskowitz (2000) observed the trading activity of professional futures traders and found that individuals with daily losses at the middle of the day took more risks in the second half of the day. In contrast, individuals trading at a gain at midday took significantly less risks during the afternoon.

A: Single Gamble



B: Curve Shift after Loss/Gain

C: Reference Point Shift after Loss



D: Reference Point Shift after Gain

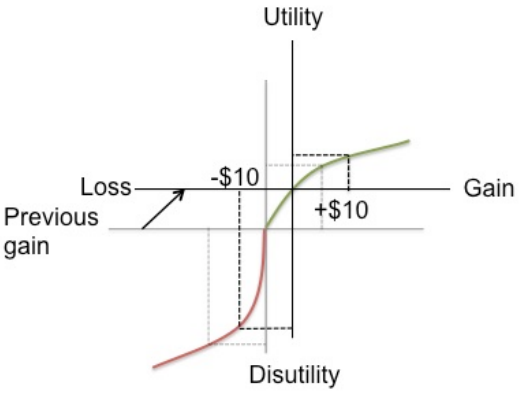


Figure 44: (**A**) given a single 50-50 gamble between losing $10 or gaining $10, an individual's choice typically reflects the 'S' shaped value curve illustrated here. The steep, convex red curve in the loss domain demonstrates that a loss of $10 would bring about twice as much disutility as positive utility (flatter green line) from a $10 gain would. (**B**) Given a series of gambles, empirical data suggests that a previous loss causes an upward shift red tail and an associated reduction in the discrepancy between loss disutility *vs.* gain utility. Conversely, a previous win leads to a downward shift in the green tail and a relative increase in the discrepancy between gain utility and loss disutility. Theoretically, these curve shifts (i.e. changes in risk attitude) result from a (**C**) downward shift in reference point after a previous loss, and (**D**) an upward shift in the reference point after a previous gain.

Given the example above, imagine the scenario in which an individual is trading at a loss at midday. As the outcome of each trade becomes integrated into the trader's payoff expectation, the reference point is updated to a lower position.

The seeming shift in the value curve (Figure 42b) that reflects more risk-seeking preferences after a loss is likely the result of a downward reference point shift along the 's-curve' (Figure 42c). The opposite upward reference point shift (Figure 42d) would also correspond with the apparent downward shift of the green curve in Figure 42b after a previous gain.

The explanation for this behaviour during a series of decisions highlights a key area for input from psychology and neuroscience. Critically, the reference point shifts in Figures 42c-d can be seen as the result of reinforcement learning as a decision maker uses previous outcomes to update predictions about future rewards. Thus, it is likely that one's willingness to accept uncertainty (i.e. risk attitude) and one's perception of valence (i.e. what constitutes a loss *vs*. what constitutes a gain) are both at least partially modulated by reward prediction errors generated in the midbrain dopamine system (Schultz et al., 1997). Despite the general evolutionary preservation of the midbrain dopamine system between rats and humans, it was unclear up until this point whether rats were capable of exhibiting reference-dependent behaviour. While we observed some indication that this might be the case in the gambling task, we were able to explicitly support this argument with the results of the stock market task in Chapter 5. By setting all reward volumes with respect to an arbitrary amount (i.e. 0.15 ml) and delivering only that amount on the initial 15 trials of every session, we were able to observe reference-dependent changes in movement time to collect reward for relative gains and losses. Furthermore, rats' increased risk-seeking behaviour after receiving a volume of reward that was less than the reference point was indicative of loss aversion. These results

indicate that future studies can use this paradigm to explicitly test how reference points are represented and updated in the rat brain.

In conclusion, the novel tasks developed here allowed us to conclude that rats often exhibit behavioural biases in a way that is remarkably similar to humans when faced with a loss compared to a gain. This suggests that the neural mechanisms governing such behaviour are evolutionarily conserved, which corroborates previous animal and human research implicating the midbrain dopamine system. Given the substantial evidence suggesting that reward expectations are encoded via phasic midbrain dopamine signals, one might speculate that this activity also acts as a switch or gating mechanism for the separable systems purported to be involved in valuation. For example, when expectations are not met, suppression of tonic dopamine levels via negative reward prediction errors may prime the system to selectively attend to the mechanisms subserving losses while inhibiting those associated with gains. Results compatible with precisely such a mechanism were highlighted in the histological characterization of brain and behaviour in our rat gambling task (Chapters 3 and 4).

Brain and behaviour

The notion that there are separable systems in the nervous system associated with losses and gains also introduces the potential for asymmetrical learning about losses and gains. The decision to stay or switch hole contingencies after a loss compared to a gain emerged as a significant factor in Chapter 2. Instead of adopting a win-stay/lose-shift strategy, rats tended to stay more after a loss than after a gain. Adding to a growing body of evidence suggesting that the

midbrain dopamine system is involved in the modulation of instrumental behaviour (for review, see Wickens et al., 2003), we also found evidence that this behaviour was under dopaminergic control. For example, the administration of dopamine antagonist *cis*-Flupenthixol dose-dependently decreased the likelihood that a rat would return to the same contingency as the immediately preceding trial. When this behaviour was analysed with respect to previous wins and losses, we found that rats with more putative dopamine neurons in the substantia nigra pars compacta (SNc) were also more likely to stay after a loss, which was replicated in a separate version of the task. Intriguingly, dopamine receptor blockade with *cis*-Flupenthixol mitigated this effect. These results suggest that in the rat, decision-making after a loss is critically modulated by dopamine neurons in the SNc.

These results complement work done in Parkinson's disease patients, who show improved learning from negative outcomes when off medication compared to controls (Frank, Seeberger, & O'Reilly, 2004). Conversely, patients show greater sensitivity to positive outcomes than negative outcomes when on medication. Cohen and Frank (2009) developed a neurocomputational model in support of these findings (see Figure 43), wherein efferent projections from the SNc play a central role.

Dopamine Subreceptor-Mediated Direct and Indirect Pathways

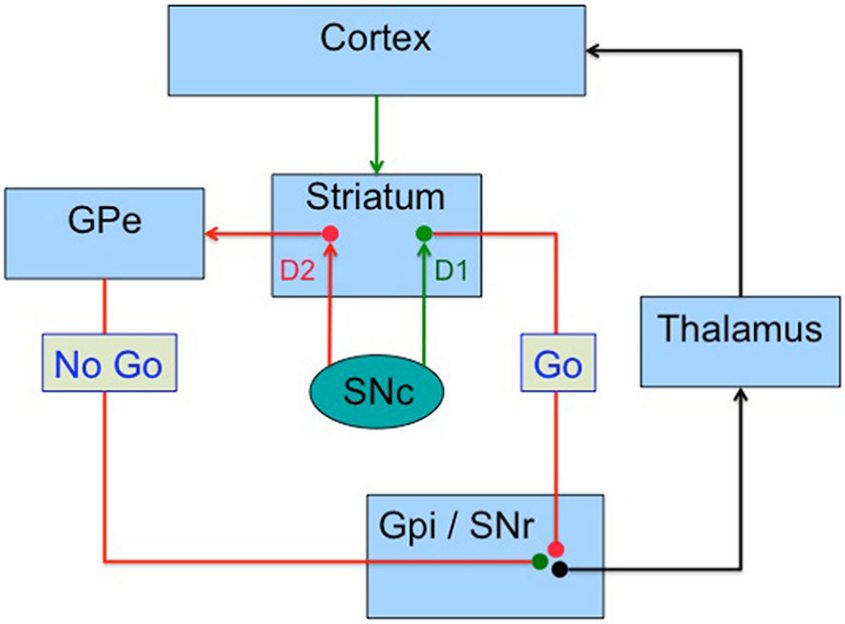


Figure 45 reproduced* from (Clark & Dagher, 2014) and based on computational theory by (Cohen & Frank, 2009): Output from the substantia nigra pars compacta (SNc) influences choice via two distinct pathways (direct and indirect). Postsynaptic neurons of the direct pathway primarily express dopamine D1 receptors at the striatum, which then projects a 'Go' response through the internal global pallidus (GPi) and substantia nigra pars reticulate (SNr), and on to the thalamus and cortex. In the indirect pathway, the dopamine D2 receptors facilitate a 'No Go' response by disinhibiting the GPi (i.e. suppressing action selection) via the external global pallidus (GPe). Excitatory projections are illustrated in green and inhibitory projections are shown in red. *Reproduced under the open-access terms of the Creative Commons Attribution License (CC BY).

Dopamine D_1 and D_2 receptors have different affinities for dopamine, and indeed often have opposing functions in the brain (Beaulieu & Gainetdinov, 2011). D_1 receptors have a low affinity for dopamine, making them sensitive only to larger phasic changes, whereas high-affinity D_2 receptors are responsive to lower tonic changes in dopamine concentrations. The Cohen and Frank (2009) model asserts that information about the utility of a gain is relayed via phasic dopamine activity and nigrostriatal D_1 receptor signalling in the direct 'Go' pathway, which elicits an approach response. Conversely, information about the disutility of a loss is conveyed via the indirect 'No Go' pathway,

whereby changes in tonic dopamine levels prompt D_2-expressing striatal neurons to disinhibit the 'Go' response from the inner globus pallidus/substantia nigra pars reticulata. Pharmacological manipulations of these receptors in humans support the model's predictions in reinforcement learning (Frank & O'Reilly, 2006; Pizzagalli et al., 2007).

Our findings are compatible with this model. Similar to Parkinson's disease patients off of medication, rats with fewer SNc neurons were less likely to repeat a choice after a loss and therefore demonstrated greater responsiveness to negative outcomes. By contrast, rats with more SNc neurons were more likely to repeat a choice after a loss, indicating that they were relatively less sensitive to negative outcomes. After administering dopamine antagonist *cis*-Flupenthixol, the association between DA neuron number and repeating a choice after a loss was abolished. The primary mechanism of action is hypothesized to reflect suppression of the indirect 'No Go' pathway by cis-Flupenthixol, which has been shown to preferentially interact with D_2 receptors (Hess, Norman, & Creese, 1988). Some parallel D_1 receptor blockade of phasic responses in the direct 'Go' pathway may have also counteracted any straightforward effect reversals. Relative differences in receptor densities could have subsequently led to an elimination of the loss-stay bias by asymmetrically interrupting these two pathways. Furthermore, due to fast-acting autoreceptor feedback loops, dopamine antagonists effectively increase dopamine cell firing due to increased dopamine availability at autoreceptors after target receptor blockade (Bunney, Walters, Roth, & Aghajanian, 1973). This would further amplify the asymmetrical disruption to

the ‘No Go’ signalling pathway that putatively conveys information about negative outcomes.

A number of cellular mechanisms are purported to compensate for system-level differences in neural cell counts, especially in the midbrain dopamine system (Blesa et al., 2011; Zigmond et al., 1989). It is believed that compensatory mechanisms, such as autoreceptor and dendritic spine densities, adaptively constrain dopamine levels and effectively preclude any meaningful effects of individual differences in healthy neuron counts on behaviour. Evidence to the contrary (i.e. in support of the notion that the number of neurons in a substrate can indeed lead to behavioural biases in healthy individuals) can be found in both animal and human literatures. For example, a number of studies investigating a genetic strain of mice with naturally increased midbrain dopamine cell numbers have also reliably linked variability in dopamine cell counts with variations in exploratory behaviour and drug reactivity (Baker et al., 1980; Reis et al., 1979; Reis et al., 1982; Sved et al., 1984). Furthermore, studies in humans have indicated that patients suffering from neurodegenerative diseases of the dopamine system (e.g. Parkinson’s disease and Huntington’s disease) demonstrate cognitive deficits very early on in disease progression (Chaudhuri et al., 2006; Chaudhuri & Naidu, 2008; Chaudhuri et al., 2010), which denotes a putative link between neuron count and behaviour. These findings raise the general question of how much control a single neuron can have over activity at the local and systems level in any species.[10] In conclusion, future research is warranted to establish whether such

[10] *While beyond the scope of this discussion, the reader is referred to recent work by Pitkow, Liu, Angelaki, DeAngelis, and Pouget (2015) for a computational*

a relationship between neural cell counts and behaviour extends to healthy human populations.

Conclusions

In this thesis, the research presented from the development of two experimental rat decision-making tasks brings together theoretical accounts of loss-related behaviour from economics, psychology and neuroscience. The experimental results are compatible with the notion that separable neural substrates and receptor pathways in the brain parallel the asymmetrical sensitivity that both rats and humans exhibit towards losses and gains. Individual differences in neuron densities in key substrates such as the SNc may underlie differences in reinforcement learning, which may in turn reflect reference point updating, macro-level biases and attitudes towards risk. In conclusion, rat models of risky decision-making, which are currently underutilized in neuroeconomic research, offer a critical link between the microscopic and macroscopic levels of behavioural analysis.

---

theory accounting for correlations between choice behaviour and single sensory neurons.

Appendices

Appendix 1

Table 3: Cox Proportional Hazard model coefficients for nosepoke duration

					95.0% CI for Exp(B)	
	B	SE	Sig.	Exp(B)	Lower	Upper
$Log(t)_{PrevOutcome}$	.059	.020	.003	1.060	1.020	1.103
$Log(t)_{Contingency}$	-.024	.009	.008	.977	.960	.994
PrevOutcome	.080	.030	.008	1.083	1.021	1.149
$Contingency_{MainEffect}$			.000			
$Contingency_{Low}$	-.039	.015	.008	.962	.935	.990
$Contingency_{High}$	.075	.017	.000	1.078	1.042	1.115
Contingency* $PrevOutcome_{MainEffect}$			.136			
$Contingency_{Low}$* PrevOutcome	.063	.032	.047	1.065	1.001	1.133
$Contingency_{High}$* PrevOutcome	-.038	.032	.230	.962	.904	1.025

Table 4: Cox Proportional Hazard model coefficients for stay behaviour on free-choice trials

					95.0% CI for Exp(B)	
	B	SE	Sig.	Exp(B)	Lower	Upper
$Log(t)_{PrevOutcome}$	-.122	.135	.367	.885	.680	1.153
$Log(t)_{Contingency}$	.078	.059	.184	1.082	.964	1.214
PrevOutcome	.138	.098	.160	1.148	.947	1.393
$Contingency_{MainEffect}$			.000			
$Contingency_{Low}$	-.657	.120	.000	.518	.409	.656
$Contingency_{High}$	.433	.075	.000	1.542	1.331	1.786
Contingency* $PrevOutcome_{MainEffect}$			.014			
$Contingency_{Low}$* PrevOutcome	.284	.170	.095	1.328	.952	1.853
$Contingency_{High}$* PrevOutcome	-.285	.101	.005	.752	.617	.916

Appendix 2

Table 5: Cox proportional hazard model coefficients for poke duration (sec) at each dose of *cis*-Flupenthixol

					95.0% CI for Exp(B)	
	B	SE	Sig.	Exp(B)	Lower	Upper
Log(t)PrevOutcome	.013	.062	.838	1.013	.898	1.143
Log(t)Contingency	-.031	.031	.319	.970	.912	1.030
Log(t)Dose	.020	.024	.398	1.021	.974	1.070
DoseMainEffect			.000			
DoseLow	-.188	.050	.000	.829	.751	.914
DoseMedium	-.203	.066	.002	.816	.718	.928
DoseHigh	-.309	.078	.000	.734	.631	.855
PrevOutcomeLossContrast	-.104	.065	.109	.901	.794	1.024
ContingencyMainEffect			.004			
ContingencyLL	.153	.049	.002	1.166	1.059	1.284
ContingencyHH	.046	.034	.183	1.047	.979	1.119

Table 6: Cox proportional hazard model coefficients for choice of the Low contingency at each dose

					95.0% CI for Exp(B)	
	B	SE	Sig.	Exp(B)	Lower	Upper
Log(t)PrevError	.464	.389	.233	1.591	.742	3.410
DoseMainEffect			.044			
DoseLow	.423	.233	.069	1.526	.967	2.409
DoseMedium	-.080	.271	.769	.923	.543	1.571
DoseHigh	-.830	.376	.027	.436	.209	.912
PrevOutcome			.553			

Table 7: Cox proportional hazard model coefficients for choice of the High contingency at each dose

					95.0% CI for Exp(B)	
	B	SE	Sig.	Exp(B)	Lower	Upper
Log(t)PrevError	-.163	.259	.530	.850	.512	1.411
DoseMainEffect			.053			
DoseLow	-.113	.064	.076	.893	.788	1.012
DoseMedium	-.103	.073	.157	.902	.781	1.041
DoseHigh	.174	.077	.025	1.190	1.023	1.385
PrevOutcome			.844			

Table 8: Cox Proportional Hazard model coefficients for shifting contingency choice from previous trial

					95.0% CI for Exp(B)	
	B	SE	Sig.	Exp(B)	Lower	Upper
Log(t)PrevError	.098	.143	.491	1.103	.834	1.460
Log(t)Contingency	.091	.059	.126	1.095	.975	1.230
Log(t)Dose	-.072	.051	.160	.931	.842	1.029
Log(t)PrevContingency	-.072	.048	.131	.931	.848	1.022
DoseMainEffect			.002			
DoseLow	-.220	.084	.009	.803	.681	.946
DoseMedium	.332	.095	.001	1.393	1.156	1.679
DoseHigh	.396	.174	.023	1.486	1.056	2.092
PrevOutcome	.019	.320	.953	1.019	.545	1.907
Contingency			.000			
ContingencyLL	.917	.197	.000	2.502	1.700	3.682
ContingencyHH	-1.176	.193	.000	.309	.211	.451
PrevContingency			.000			
PrevContingencyLL	.380	.066	.000	1.463	1.286	1.664
PrevContingencyHH	-.472	.172	.006	.624	.446	.873
PrevOutcome* Contingency			.002			
PrevOutcome* ContingencyLL	-.626	.239	.009	.535	.335	.854
PrevOutcome* ContingencyHH	.601	.170	.000	1.823	1.308	2.542

Table 9: Linear regression coefficients for baseline loss-stay behaviour as a dependent variable

							95% CI for B	
Model		B	SE	Beta	t	Sig.	Lower	Upper
1	(Constant)	8.42E-17	.140		.000	1.000	-.301	.301
	SNc Count	.841	.145	.841	5.804	.000	.530	1.151

* Average weight was excluded from the model via stepwise entry method (β_{in} = -.03, t = -0.20, p = .85, VIF = 1.03).

** VTA count was excluded from the model via stepwise entry method (β_{in} = .04, t = 1.20, p = .25, VIF = 1.14).

Table 10: Linear regression coefficients for saline vehicle loss-stay behaviour as a dependent variable

Model		B	SE	Beta	t	Sig.	95% CI for B Lower	95% CI for B Upper
1	(Constant)	-4.6E-17	.223		.000	1.000	-.479	.479
	SNc Count	.504	.231	.504	2.186	.046	.010	.999

* Average weight was excluded from the model via stepwise entry method (β_{in} = .14, t = 0.60, p = .56, VIF = 1.03).

** VTA count was excluded from the model via stepwise entry method (β_{in} = -.11, t = -0.42, p = .68, VIF = 1.14).

Appendix 3

Table 11: Coefficients of Cox hazard model for poke duration

	B	SE	Sig.	Exp(B)	95% CI for Exp(B) Lower	Upper
$Log(t)_{PrevOutcome}$	-.038	.008	.000	.963	.949	.978
$Log(t)_{Contingency}$	-.029	.033	.387	.972	.910	1.037
PrevOutcome	-.160	.040	.000	.852	.788	.921
$Contingency_{MainEffect}$			.000			
$Contingency_{HighProb}$	.668	.032	.000	1.950	1.834	2.075
$Contingency_{LowProb}$	-.156	.033	.000	.856	.803	.912
$Contingency_{HighRew}$	.022	.034	.524	1.022	.956	1.093
$Contingency_{LowRew}$	.087	.040	.028	1.091	1.009	1.179

Table 12: Linear regression coefficients with Lose-shift as a dependent variable

Model		B	Std. Error	Beta	t	Sig.	95% CI for B Lower	Upper
1	(Constant)	2.1E-16	.20		.00	1.00	-.47	.47
	SNc Neuron Ct	.828	.212	.828	3.914	.006	.33	1.33

* Average weight was excluded from the model via stepwise entry method (β_{in} = .004, t = 0.02, p = .99, VIF = 1.28).

Appendix 4

Table 13: Coefficients of a Cox Proportional Hazard Model of sales as a function of return

	Coef.	SE	Lower CI	Upper CI	Sig.
< -22%	-1.16	0.48	-9.01	-2.01	*
-22% – -18%	1.65	1.01	-3.59	-1.82	
-18% – -14%	2.13	0.88	-2.94	-1.52	*
-14% – -10%	3.16	0.90	-1.65	-0.73	***
-10% – -6%	3.31	0.90	-1.48	-0.61	***
-6%– -2%	3.94	0.90	-0.77	-0.06	***
-2%– 2%	4.35	0.89	-0.31	0.31	***
2%– 6%	5.22	0.90	0.61	1.14	***
6%– 10%	5.69	0.91	1.09	1.59	***
10%– 14%	5.80	0.95	1.20	1.70	***
14%– 18%	6.09	1.02	1.50	1.98	***
≥18%	5.32	1.22	0.70	1.22	***
Sales Count	-0.15	0.01	0.01	-30.04	***
Counterfactual Reward	-6.19	1.23	0.27	-22.86	***

Likelihood ratio test(14) = 3352, p <.001

*p<.05, **p<.01, ***p<.001

Table 14: Coefficients of a Cox Proportional Hazard Model of purchases as a function of blink rate

	Coef.	SE	Lower CI	Upper CI	Sig.
< 0.75 Hz	-0.77	0.46	-1.88	0.35	***
0.75 Hz − 1.0 Hz	-0.01	0.99	-0.49	0.49	*
1.0 Hz − 1.25 Hz	-0.01	0.99	-0.40	0.38	**
1.25 Hz − 1.5 Hz	-0.06	0.94	-0.34	0.22	
1.5 Hz − 1.75 Hz	0.10	1.11	-0.13	0.34	
1.75 Hz − 2.0 Hz	0.19	1.20	0.05	0.33	
2.0 Hz − 2.25 Hz	0.17	1.18	0.01	0.32	
2.25 Hz − 2.5 Hz	0.25	1.28	0.06	0.43	
2.5 Hz − 2.75 Hz	0.27	1.31	0.02	0.52	
2.75 Hz − 3.0 Hz	0.22	1.25	-0.12	0.56	
3.0 Hz − 3.25 Hz	0.42	1.52	-0.14	0.98	*
3.25 Hz − 3.5 Hz	0.67	1.96	0.15	1.20	***
≥ 3.5 Hz	1.26	3.54	0.45	2.08	**
Buys Count	0.04	1.04	0.04	0.05	***
$\text{Log}_{\text{s1blinkhz}}$	0.03	1.03	0.01	0.07	*
$\text{Log}_{\text{s2blinkhz}}$	-0.02	0.98	-0.05	0.01	
$\text{Log}_{\text{s3blinkhz}}$	0.04	1.04	0.01	0.07	*

Likelihood Ratio Test(17) = 490.8, *P* <.001

*p<.05, **p<.01, ***p<.001

Table 15: Coefficients of a Cox Proportional Hazard Model of sales as a function of blink rate

	Coef.	SE	Lower CI	Upper CI	Sig.
< 0.75 Hz	**-0.77**	**0.46**	**-1.88**	**0.35**	
0.75 Hz − 1.0 Hz	-0.01	0.99	-0.49	0.49	
1.0 Hz − 1.25 HZ	-0.01	0.99	-0.40	0.38	
1.25 Hz − 1.5 Hz	-0.06	0.94	-0.34	0.22	
1.5 Hz − 1.75 Hz	0.10	1.11	-0.13	0.34	
1.75 Hz − 2.0 Hz	0.19	1.20	0.05	0.33	**
2.0 Hz − 2.25 Hz	0.17	1.18	0.01	0.32	*
2.25 Hz − 2.5 Hz	0.25	1.28	0.06	0.43	*
2.5 Hz − 2.75 Hz	0.27	1.31	0.02	0.52	*
2.75 Hz − 3.0 Hz	0.22	1.25	-0.12	0.56	
3.0 Hz − 3.25 Hz	0.42	1.52	-0.14	0.98	
3.25 Hz − 3.5 Hz	0.67	1.96	0.15	1.20	*
≥ 3.5 Hz	1.26	3.54	0.45	2.08	**
Sales Count	0.04	1.04	0.04	0.05	***
$\text{Log}_{\text{S1BlinkHz}}$	0.06	1.07	0.03	0.10	**
$\text{Log}_{\text{S2BlinkHz}}$	0.02	1.02	-0.01	0.06	
$\text{Log}_{\text{S3BlinkHz}}$	-0.02	0.98	-0.05	0.01	

Likelihood ratio test(16) = 489.9, p <.001

*p<.05, **p<.01, ***p<.001

References

Aarts, E. (2012). Aberrant reward processing in Parkinson's disease is associated with dopamine cell loss. *NeuroImage, 59*(4), 3339-3346.

Adolphs, R., Tranel, D., & Damasio, A. R. (1995). Fear and the human amygdala. *J. Neurosci., 15*, 5879-5891.

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin, 82*(4), 463-496.

Allais, M. (1953). Le Comportement de l'Homme Rationnel devant le Risque, Critique des Postulats et Adiomes de l'Ecole Americaine. *Econometrica*(21), 503-546.

Amsel, A. (1958). The role of frustrative nonreward in noncontinuous reward situations. *Psychological Bulletin, 55*, 102-119.

Ashraf, N., Camerer, C. F., & Loewenstein, G. F. (2005). Adam Smith, behavioral economist. *Journal of Economic Perspectives, 19*(3), 131-138.

Bachelier, L. (1900). Theorie de la speculation. *Annales Scientifiques de l'Ecole Normale Superieure Ser., 3*(17), 21-86.

Baker, H., Joh, T. H., & Reis, D. J. (1980). Genetic control of number of midbrain dopaminergic neurons in inbred strains of mice: relationship to size and neuronal density of the striatum. *Proceedings of the National Academy of Sciences of the United States of America, 77*(7), 4369-4373.

Barber, B. M., & Odean, T. (2011). The behavior of individual investors. Retrieved 14.04.2015 http://ssrn.com/abstract=187221

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. In G. M. Constantinides, M. Harris, & R. M. Stultz (Eds.), *Handbook of Economics and Finance* (1st ed., Vol. 1, pp. 1053-1128): Elsevier.

Barberis, N., & Xiong, W. (2012). Realization utility. *Journal of Financial Economics, 104*(2), 251-271.

Barberis, N., & Xiong, W. E. I. (2009). What drives the disposition effect? An analysis of a long-standing preference-based explanation. *The Journal of Finance, 64*(2), 751-784. doi:10.1111/j.1540-6261.2009.01448.x

Basten, U., Biele, G., Heekeren, H. R., & Fiebach, C. J. (2009). How the brain integrates costs and benefits during decision making. *PNAS, 107*(50), 21767-21772.

Beaulieu, J.-M., & Gainetdinov, R. R. (2011). The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacological Reviews, 63*(1), 182-217. doi:10.1124/pr.110.002642

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior, 52*, 336-372.

Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. ["Exposition of a new theory on the measurement of risk." *Econometrica* 22 (1954): 23 - 26.]. *Commentari Academiae Scientiatatum Imperalas Petropolitanae, 5*, 175-192.

Berns, G. S., Laibson, D., & Loewenstein, G. (2007). Intertemporal choice–toward an integrative framework. *Trends in cognitive sciences, 11*(11), 482-488.

Berridge, K. C. (1996). Food reward: Brain substrates of wanting and liking. *Neuroscience & Biobehavioral Reviews, 20*(1), 1-25. doi:http://dx.doi.org/10.1016/0149-7634(95)00033-B

Besson, M., Belin, D., McNamara, R., Theobald, D. E. H., Castel, A., Beckett, V. L., . . . Dalley, J. W. (2010). Dissociable control of impulsivity in rats by dopamine D2/3 receptors in the core and shell subregions of the nucleus accumbens. *Neuropsychopharmacology, 35*(2), 560-569. doi:10.1038/npp.2009.162

Björklund, A., & Dunnett, S. B. Dopamine neuron systems in the brain: an update. *Trends Neurosci, 30*(5), 194-202. doi:10.1016/j.tins.2007.03.006

Blesa, J., Juri, C., García-Cabezas, M. Á., Adánez, R., Sánchez-González, M. Á., Cavada, C., & Obeso, J. A. (2011). Inter-hemispheric asymmetry of nigrostriatal dopaminergic lesion. A possible compensatory mechanism in Parkinson's disease. *Frontiers in Systems Neuroscience, 5*. doi:10.3389/fnsys.2011.00092

Bondt, W. F. M. D., & Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance, 40*(3), 793-805. doi:10.2307/2327804

Boorman, E. D., Behrens, T. E., & Rushworth, M. F. (2011). Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS Biol, 9*(6), e1001093. doi:10.1371/journal.pbio.1001093

Bornovalova, M. A., Daughters, S. B., Hernandez, G. D., Richards, J. B., & Lejuez, C. W. (2005). Differences in impulsivity and risk-taking propensity between primary users of crack cocaine and primary users of heroin in a residential substance-use program. *Exp Clin Psychopharmacol, 13*(4), 311-318. doi:10.1037/1064-1297.13.4.311

Bossaerts, P. (2010). Risk and risk prediction error signals in anterior insula. *Brain Struct Funct, 214*, 645-653.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity: perceptual versus cognitive mechanisms in chimpanzees (Pan troglodytes). *Journal of experimental psychology. Animal behavior processes, 21*(1), 82-86. doi:10.1037//0097-7403.21.1.82

Brizer, D. A. (1988). Psychopharmacology and the management of violent patients. *Psychiatr Clin North Am, 11*(4), 551-568.

Brown, P., Chappel, N., da Silva Rosa, R., & Walter, T. (2006). The reach of the disposition effect: Large sample evidence across investor classes. *International Review of Finance, 6*, 42-78.

Brown, R. (1828). A brief account of microscopical observations made in the months of June, July and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Edinburgh New Philosophical Journal, 4*, 358-371.

Buchel, C., & Dolan, R. J. (2000). Classical fear conditioning in functional neuroimaging. *Curr Opin Neurobiol, 10*, 219-223.

Bunney, B. S., Walters, J. R., Roth, R. H., & Aghajanian, G. K. (1973). Dopaminergic neurons: effect of antipsychotic drugs and amphetamine on single cell activity. *J Pharmacol Exp Ther, 185*, 560-571.

Buzasaki, G. (2002). Theta oscillations in the hippocampus. *Neuron, 33*, 325-340.

Calvert, A. L., Green, L., & Myerson, J. (2010). Delay Discounting of Qualitatively Different Reinforcers in Rats. *Journal of the Experimental Analysis of Behavior, 93*(2), 171-184.

Camerer, C. (2005). Three cheers - Psychological, Theoretical, Empircal - for loss aversion. *Journal of Marketing Research, 42*(2), 129-133.

Campbell, J. Y., Lo, A. W. C., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Caplin, A., & Glimcher, P. W. (2014). Chapter 1 - Basic methods from neoclassical economics. In P. W. G. Fehr (Ed.), *Neuroeconomics (Second Edition)* (pp. 3-17). San Diego: Academic Press.

Cardinal, R. N., Robbins, T. W., & Everitt, B. J. (2000). The effects of d-amphetamine, chlordiazepoxide, α-flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology, 152*(4), 362-375. doi:10.1007/s002130000536

Carlson, S. M., Davis, A. C., & Leach, J. G. (2005). Less is more: Executive function and symbolic representation in preschool children. *Psychological Science, 16*(8), 609-616. doi:10.1111/j.1467-9280.2005.01583.x

Charness, G., & Levin, D. (2005). When optimal choices feel wrong: A laboratory study of bayesian updating, complexity, and affect. *The American economic review, 95*(4), 1300-1309. doi:10.2307/4132717

Chaudhuri, K. R., Healy, D. G., & Schapira, A. H. V. (2006). Non-motor symptoms of Parkinson's disease: diagnosis and management. *The Lancet Neurology, 5*(3), 235-245. doi:http://dx.doi.org/10.1016/S1474-4422(06)70373-8

Chaudhuri, K. R., & Naidu, Y. (2008). Early Parkinson's disease and non-motor issues. *Journal of Neurology, 255*(5), 33-38. doi:10.1007/s00415-008-5006-1

Chaudhuri, K. R., Prieto-Jurcynska, C., Naidu, Y., Mitra, T., Frades-Payo, B., Tluk, S., . . . Martinez-Martin, P. (2010). The nondeclaration of nonmotor symptoms of Parkinson's disease to health care professionals: An international study using the nonmotor symptoms questionnaire. *Movement Disorders, 25*(6), 704-709. doi:10.1002/mds.22868

Chen, G., Kim, K., Nofsinger, J. R., & Rui, O. M. (2007). Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors. Retrieved 05.05.2015 SSRN: http://ssrn.com/abstract=957504

Chen, M. K., Lakshminarayanan, V., & Santos, L. R. (2006). How basic are behavioral biases? Evidence from capuchin monkey trading behavior. *Journal of Political Economy, 114*(3), 517-537.

Chiu, P., Lohrenz, T., & Montague, P. R. (2008). Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. *Nat Neurosci, 11*, 514-520.

Choi, J. J., Laibson, D., Madrian, B. C., & Metrick, A. (2009). Reinforcement learning and savings behavior. *The Journal of Finance, 64*(6), 2515-2534. doi:10.1111/j.1540-6261.2009.01509.x

Chua, P., Krams, M., Toni, I., Passingham, R. E., & Dolan, R. J. (1999). A functional anatomy of anticipatory anxiety. *NeuroImage, 9*, 563-571.

Chung, S., & Herrnstein, R. J. (1967). Choice and delay of reinforcement. *Journal of Analysis of Behavior, 10*, 67-74.

Clark, C. A., & Dagher, A. (2014). The role of dopamine in risk taking: A specific look at Parkinson's disease and gambling. *Frontiers in Behavioral Neuroscience, 8*. doi:10.3389/fnbeh.2014.00196

Cohen, M. X., & Frank, M. J. (2009). Neurocomputational models of basal ganglia function in learning, memory and choice. *Behavioural brain research, 199*(1), 141-156. doi:http://dx.doi.org/10.1016/j.bbr.2008.09.029

Collier, G., Hirsch, E., & Hamlin, P. H. (1972). The ecological determinants of reinforcement in the rat. *Physiology and Behavior*(9), 705 - 716.
Coval, J., & Moskowitz, T. (2000). *Do behavioral biases affect prices?* University of Michigan. Ann Arbor, MI.
Cowie, R. J. (1977). Optimal foraging in great tits (Parus major). *Nature, 268*(5616), 137-139.
Cox, D. R., & Oakes, D. (1984). *Analysis of Survival Data*. New York, NY: Chapman and Hall.
Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat Rev Neurosci, 3*, 655-666.
Da Cunha, C., Wietzikoski, S., Wietzikoski, E. C., Miyoshi, E., Ferro, M. M., Anselmo-Franci, J. A., & Canteras, N. S. (2003). Evidence for the substantia nigra pars compacta as an essential component of a memory system independent of the hippocampal memory system. *Neurobiology of Learning and Memory, 79*(3), 236-242. doi:http://dx.doi.org/10.1016/S1074-7427(03)00008-X
Dalley, J. W., Fryer, T. D., Brichard, L., Robinson, E. S. J., Theobald, D. E. H., Lääne, K., . . . Robbins, T. W. (2007). Nucleus accumbens D2/3 receptors predict trait impulsivity and cocaine reinforcement. *Science, 315*(5816), 1267-1270. doi:10.1126/science.1137073
Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, NY: Putnam.
De Martino, B., Camerer, C., & Adolphs, R. (2010). Amygdala damage eliminates monetary loss aversion. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 3788-3792.
De, S., Gondhi, N. R., & Pochiraju, B. (2010). *Does sign matter more than size? An investigation into the source of investor overconfidence*.
Dubois, B., & Pillon, B. (1996). Cognitive deficits in Parkinson's disease. *Journal of Neurology, 244*(1), 2-8. doi:10.1007/PL00007725
Eagle, D. M., Wong, J. C. K., Allan, M. E., Mar, A. C., Theobald, D. E., & Robbins, T. W. (2011). Contrasting roles for dopamine D1 and D2 receptor subtypes in the dorsomedial striatum but not the nucleus accumbens core during behavioral inhibition in the Stop-Signal Task in rats. *The Journal of neuroscience, 31*(20), 7349-7356. doi:10.1523/jneurosci.6182-10.2011
Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik, 322*(8), 549-560.
Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics, 75*(4), 643-669. doi:10.2307/1884324
Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review, 88*, 848-881.
Evans, A. H., Lawrence, A. D., Potts, J., Appel, S., & Lees, A. J. (2005). Factors influencing susceptibility to compulsive dopaminergic drug use in Parkinson disease. *Neurology, 65*, 1570-1574.
Evans, A. H., Pavese, N., Lawrence, A. D., Tai, Y. F., Appel, S., Doder, M., . . . Piccini, P. (2006). Compulsive drug use linked to sensitized ventral striatal dopamine transmission. *Annals of Neurology, 59*(5), 852-858. doi:10.1002/ana.20822
Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal, 21*(5), 55-59.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383-417. doi:10.1111/j.1540-6261.1970.tb00518.x
Feng, L., & Seasholes, M. (2005). Do investor sophistication and trading experience eliminate behavioral biases in financial markets? *Review of Finance, 9*, 305-351.
Feng, L., & Seasholes, M. (2008). Individual investors and gender similarities in an emerging stock market. *Pasific-Basin Finance Journal, 16*, 44-60.
Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.
Festinger, L. (1957). *A Theory of Cognitive Dissonance*: Stanford University Press.
Fibiger, H. C., & Phillips, A. G. (1986). The Nervous System. In F. E. Bloom (Ed.), *Handbook of Physiology* (Vol. 4, pp. 647-675). Baltimore, MD: Williams & Wilkins.
Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science, 299*, 1898-1902.
Fishburn, P. C. (1986). The axioms of subjective probability. *Statistical Science, 1*(3), 335-345. doi:10.2307/2245466
Fox, C. R., & Poldrack, R. A. (2009). Prospect Theory and the Brain. In P. W. Glimcher (Ed.), *Neuroeconomics: Decision Making and the Brain* (pp. 145-171). London, UK: Elsevier.
Fox, C. R., & Poldrack, R. A. (2014). Appendix - Prospect Theory and the Brain. In P. W. G. Fehr (Ed.), *Neuroeconomics (Second Edition)* (pp. 533-567). San Diego: Academic Press.
Frank, M. J., & O'Reilly, R. C. (2006). A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. *Behav Neurosci, 120*(3), 497-517. doi:10.1037/0735-7044.120.3.497
Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism. *Science, 306*(5703), 1940-1943. doi:10.1126/science.1102941
Friedman, M. (1953). *Essays in Positive Economics*. Chicago, IL: University of Chicago Press.
Frydman, C., Barberis, N., Camerer, C., Bossaerts, P., & Rangel, A. (2014). Using neural data to test a theory of investor behavior: An application to realization utility. *The Journal of Finance, 69*(2), 907-946.
Frydman, C., & Rangel, A. (2014). Debiasing the disposition effect by reducing the saliency of information about a stock's purchase price. *Journal of Economic Behavior & Organization*.
Fuster, A., Laibson, D., & Mendel, B. (2010). Natural expectations and macroeconomic fluctuations. *The journal of economic perspectives : a journal of the American Economic Association, 24*(4), 67-84. doi:10.1257/jep.24.4.67
Gächter, S., Johnson, E. J., & Herrmann, A. (2010). Individual-level loss aversion in riskless and risky choices. (No. 2010-20). Retrieved 07.06.2015, from CeDEx discussion paper series http://hdl.handle.net/10419/49656
Gan, J. O., Walton, M. E., & Philips, P. E. M. (2010). Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nat Neurosci, 13*, 25-27.

Gigerenzer, G., & Selten, R. (Eds.). (2001). *Bounded Rationality : The Adaptive Toolbox*. Cambridge, MA: MIT Press.
Gilovich, T., & Medvec, V. H. (1995). The experience of regret: What, when, and why. *Psychological Review, 102*(2), 379-395. doi:10.1037/0033-295X.102.2.379
Gläscher, J., Hampton, A. N., & O'Doherty, J. P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cerebral Cortex, 19*, 483-495.
Glimcher, P. W. (2008). Understanding risk: A guide for the perplexed. *Cognitive, Affective, & Behavioral Neuroscience, 8*(4), 348-354.
Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annu Rev Neurosci, 31*, 359-387. doi:10.1146/annurev.neuro.29.051605.112851
Green, L., Myerson, J., Oliveira, L., & Chang, S. E. (2014). Discounting of delayed and probabilistic losses over a wide range of amounts. *Journal of the Experimental Analysis of Behavior, 101*(2), 186-200. doi:10.1002/jeab.56
Grinblatt, M., & Keloharju, M. (2001). What makes investors trade? *The Journal of Finance, 56*(2), 589-616. doi:10.2307/222575
Grinblatt, M., Keloharju, M., & Linnainmaa, J. T. (2012). IQ, trading behavior, and performance. *Journal of Financial Economics, 104*(2), 339-362. doi:http://dx.doi.org/10.1016/j.jfineco.2011.05.016
Haber, S. N. (2003). The primate basal ganglia: parallel and integrative networks. *Journal of Chemical Neuroanatomy, 26*, 317-330.
Haigh, M. S., & List, J. A. (2005). Do professional traders exhibit myopic loss aversion? An experimental analysis. *The Journal of Finance, 60*(1), 523-534. doi:10.1111/j.1540-6261.2005.00737.x
Hayden, B., & Platt, M. L. (2007). Temporal discounting predicts risk sensitivity in rhesus macaques. *Curr Biol, 17*, 49-53.
Heilbronner, S. R., & Hayden, B. Y. (2013). Contextual factors explain risk-seeking preferences in Rhesus monkeys. *Frontiers in Neuroscience, 7*, 7. doi:10.3389/fnins.2013.00007
Herrnstein, R. J. (1990). Rational choice theory: Necessary but not sufficient. *American Psychologist, 45*(3), 356-367.
Hershberger, W. (1986). An approach through the looking-glass. *Animal Learning & Behavior, 14*(4), 443-451. doi:10.3758/BF03200092
Hess, E. J., Norman, A. B., & Creese, I. (1988). Chronic treatment with dopamine receptor antagonists: behavioral and pharmacologic effects on D1 and D2 dopamine receptors. *J Neurosci, 8*(7), 2361-2370.
Hirshleifer, D. (2001). Investor psychology and asset pricing. *Journal of Finance, 56*, 1533-1597.
Holt, D. D., Green, L., & Myerson, J. (2003). Is discounting impulsive?: Evidence from temporal and probability discounting in gambling and non-gambling college students. *Behavioural Processes, 64*(3), 355-367. doi:http://dx.doi.org/10.1016/S0376-6357(03)00141-4
Hommel, J. D., Trinko, R., Sears, R. M., Georgescu, D., Liu, Z.-W., Gao, X.-B., . . . DiLeone, R. J. (2006). Leptin receptor signaling in midbrain dopamine neurons regulates feeding. *Neuron, 51*(6), 801-810. doi:http://dx.doi.org/10.1016/j.neuron.2006.08.023
Houthakker, H. S. (1950). Revealed preference and the utility function. *Economica, 17*(66), 159-174. doi:10.2307/2549382

Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M., & Graybiel, A. M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature, 500*(7464), 575-579. doi:10.1038/nature12475

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*(310), 1680-1683.

Hsu, S. M., & Raine, L. (1981). Protein A, avidin, and biotin in immunohistochemistry. *J Histochem Cytochem, 29*(11), 1349-1353.

Ilango, A., Kesner, A. J., Keller, K. L., Stuber, G. D., Bonci, A., & Ikemoto, S. (2014). Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion. *The Journal of neuroscience, 34*(3), 817-822. doi:10.1523/JNEUROSCI.1703-13.2014

Jastrow, J. (1899). The mind's eye. *Popular Science Monthly*(54), 299-312.

Jee-Young, L., Jong-Min, K., Jae Woo, K., Jinwhan, C., Won Yong, L., Han-Joon, K., & Beom, S. J. (2010). Association between the dose of dopaminergic medication and the behavioral disturbances in Parkinson disease. *Parkinsonism and Related Disorders, 16*, 202-207. doi:10.1016/j.parkreldis.2009.12.002

Jentsch, J. D., Woods, J. A., Groman, S. M., & Seu, E. (2010). Behavioral characteristics and neural mechanisms mediating performance in a rodent version of the Balloon Analogue Risk Task. *Neuropsychopharmacology, 35*(8), 1797-1806.

Johnson, E. J., Gächter, S., & Herrmann, A. (2006). Exploring the nature of loss aversion. (No. 2015). Retrieved 07.06.2015, from IZA Discussion Paper SSRN: http://ssrn.com/abstract=892336

Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty, 7*(1), 35-51. doi:10.1007/BF01065313

Kable, J. W. (2013). Valuation, intertemporal choice, and self-control. In W. G. Paul, Ernst Fehr (Ed.), *Neuroeconomics: Decision Making and the Brain* (pp. 173). London, U.K.: Elsevier.

Kable, J. W., & Glimcher, P. W. (2009). The neurobiology of decision: Consensus and controversy. *Neuron, 63*(6), 733-745. doi:10.1016/j.neuron.2009.09.003

Kagel, J. H., & Battalio, R. C. (1980). Token economy and animal models for the experimental analysis of economic behavior. In J. Kmenta & J. B. Ramsey (Eds.), *Evaluation of econometric models*: Academic Press.

Kagel, J. H., Battalio, R. C., & Green, L. (1995). *Economic choice theory: An experimental analysis of animal behavior*: Cambridge University Press.

Kagel, J. H., Battalio, R. C., Rachlin, H., & Green, L. (1981). Demand curves for animal consumers. *The Quarterly Journal of Economics, 96*(1), 1-15. doi:10.2307/2936137

Kagel, J. H., Battalio, R. C., Rachlin, H., Green, L., Basmann, R. L., & Klemm, W. R. (1975). Experimental studies of consumer demand behavior using laboratory animals. *Economic Inquiry, 13*(1), 22-38. doi:10.1111/j.1465-7295.1975.tb01101.x

Kahn, I., Yeshurun, Y., Rotshtein, P., Fried, I., Ben-Bashat, D., & Hendler, T. (2002). The role of the amygdala in signaling prospective outcome of choice. *Neuron, 33*, 983-994.
Kahneman, D., & Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-291.
Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: Selective attention to information. *Journal of Risk and Uncertainty, 38*(2), 95-115. doi:10.1007/s11166-009-9060-6
Kaustia, M. (2010). Prospect Theory and the disposition effect. *Journal of Financial and Quantitative Analysis, 45*(3), 791-812. doi:doi:10.1017/S0022109010000244
Kaustia, M., & Knüpfer, S. (2008). Do investors overweight personal experience? Evidence from IPO subscriptions. *Jounral of Finance, 63*, 2679-2702.
Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Current Opinion in Neurology, 18*(4), 411-417.
Knutson, B., Samanez-Larkin, G. R., & Kuhnen, C. M. (2011). Gain and loss learning differentially contribute to life financial outcomes. *Plos One, 6*(9), e24390-e24390.
Knutson, B., Wimmer, G. E., Kuhnen, C. M., & Winkielman, P. (2008). Nucleus accumbens activation mediates the influence of reward cues on financial risk taking. *Neuroreport, 19*(5), 509-513.
Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural mechanisms of foraging. *Science, 336*(6077), 95-98. doi:10.1126/science.1216930
Krupenye, C., Rosati, A. G., & Hare, B. (2015). Bonobos and chimpanzees exhibit human-like framing effects. *Biology letters, 11*(2), 20140527.
Kuhnen, C. M., & Chiao, J. Y. (2009). Genetic determinants of financial risk taking. *Plos One, 4*(2), e4362. doi:10.1371/journal.pone.0004362
Kuhnen, C. M., & Knutson, B. (2005). The neural basis of financial risk taking. *Neuron, 47*(5), 763-770.
Kuhnen, C. M., & Knutson, B. (2011). The influence of affect on beliefs, preferences, and financial decisions. *Journal of Financial and Quantitative Analysis, 46*(03), 605-626. doi:10.1017/S0022109011000123
Lakshminaryanan, V., Chen, M. K., & Santos, L. R. (2008). Endowment effect in capuchin monkeys. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 363*(1511), 3837-3844. doi:10.1098/rstb.2008.0149
Lebreton, M., Jorge, S., Michel, V., Thirion, B., & Pessiglione, M. (2009). An automatic valuation system in the human brain: Evidence from functional neuroimaging. *Neuron, 64*, 431-439.
Lee, D., & Dorris, M. C. (2014). Brain circuitry for social decision making in non-human primates. In E. a. P. W. G. Fehr (Ed.), *Neuroeconomics (Second Edition)* (pp. 493-511). San Diego: Academic Press.
Lees, A. J., & Smith, E. (1983). *Cognitive deficits in the early stages of Parkinson's Disease* (Vol. 106).
Lejuez, C. W., Aklin, W., Daughters, S., Zvolensky, M., Kahler, C., & Gwadz, M. (2007). Reliability and validity of the youth version of the Balloon Analogue Risk Task (BART-Y) in the assessment of risk-taking behavior among inner-

city adolescents. *J Clin Child Adolesc Psychol, 36*(1), 106-111. doi:10.1080/15374410709336573
Lejuez, C. W., Aklin, W. M., Jones, H. A., Richards, J. B., Strong, D. R., Kahler, C. W., & Read, J. P. (2003). The Balloon Analogue Risk Task (BART) differentiates smokers and nonsmokers. *Exp Clin Psychopharmacol, 11*(1), 26-33.
Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., & Stuart, G. L. (2002). Evaluation of a behavioural measure of risk taking: the Balloon Analogue Risk Task (BART). *J. Exp. Psychol. Appl., 8*, 75-84.
Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes, 76*(2), 149-188. doi:http://dx.doi.org/10.1006/obhd.1998.2804
Lewis, S. J. G., Dove, A., Robbins, T. W., Barker, R. A., & Owen, A. M. (2003). Cognitive impairments in early Parkinson's disease are accompanied by reductions in activity in frontostriatal neural circuitry. *The Journal of neuroscience, 23*(15), 6351-6356.
Lohrenz, T., McCabe, K., Camerer, C., & Montague, P. R. (2007). Neural signature of fictive learning signals in a sequential learning task. *Proceedings of the National Academy of Sciences, 104*, 9493-9498.
Lowry, J. P., Griffin, K., McHugh, S. B., Lowe, A. S., Tricklebank, M., & Sibson, N. R. (2010). Real-time electrochemical monitoring of brain tissue oxygen: a surrogate for functional magnetic resonance imaging in rodents. *NeuroImage, 52*(2), 549-555.
Mäki, U. (2009). Economics imperialism: Concept and constraints. *Philosophy of the Social Sciences, 39*(3), 351-380. doi:10.1177/0048393108319023
Malkiel, B. G. (1973). *A Random Walk Down Wall Street*. New York, NY: Norton.
Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. New York, NY: Jon Wiley & Sons.
Mason, W. A., Capitanio, J. P., Machado, C. J., Mendoza, S. P., & Amaral, D. G. (2006). Amygdalectomy and responsiveness to novelty in rhesus monkeys (Macaca mulatta): Generality and individual consistency of effects. *Emotion, 6*, 73-81.
Matthews, S. C., Simmons, A. N., Lane, S. D., & Paulus, M. P. (2004). Selective activation of the nucleus accumbens during risk-taking decision making. *Neuroreport, 15*, 2123-2127.
Mazur, J. E., & Biondi, D. R. (2009). Delay-amount tradeoffs in choices by pigeons and rats: hyperbolic versus exponential discounting *Journal of the Experimental Analysis of Behavior, 91*(2), 197-211. doi:10.1901/jeab.2009.91-197
McDonald, R. J., & White, N. M. (1993). A triple dissociation of memory systems: Hippocampus, amygdala, and dorsal striatum. *Behavioral Neuroscience, 107*(1), 3-22.
McHugh, S. B., Barkus, C., Huber, A., Capitão, L., Lima, J., Lowry, J. P., & Bannerman, D. M. (2014). Aversive prediction error signals in the amygdala. *The Journal of neuroscience, 34*(27), 9024-9033. doi:10.1523/jneurosci.4465-13.2014
McKeon, A., Josephs, K. A., Klos, K. J., Hecksel, K., Bower, J. H., Michael Bostwick, J., & Eric Ahlskog, J. (2007). Unusual compulsive behaviors primarily related to dopamine agonist therapy in Parkinson's disease and multiple system atrophy. *Parkinsonism and Related Disorders, 13*(8), 516-519.

Michaely, R., Thaler, R., & Womack, K. (1995). Price reactions to dividend initiations and omissions. *Journal of Finance, 50*, 573-608.
Miczek, K. A., DeBold, J. F., & van Erp, A. M. (1994). Neuropharmacological characteristics of individual differences in alcohol effects on aggression in rodents and primates. *Behav Pharmacol, 5*(4 And 5), 407-421.
Mitchell, M. R., Weiss, V. G., Beas, B. S., Morgan, D., Bizon, J. L., & Setlow, B. (2014). Adolescent risk taking, cocaine self-administration, and striatal dopamine signaling. *Neuropsychopharmacology, 39*(4), 955-962. doi:10.1038/npp.2013.295
Mohr, P. N. C., Biele, G., Krugel, L. K., Li, S.-C., & Heekeren, H. R. (2010). Neural foundations of risk–return trade-off in investment decisions. *NeuroImage, 49*(3), 2556-2563. doi:http://dx.doi.org/10.1016/j.neuroimage.2009.10.060
Mukherji, S. (2011). Are stock returns still mean-reverting? *Review of Financial Economics, 20*(1), 22-27. doi:http://dx.doi.org/10.1016/j.rfe.2010.08.001
Neumann, N., & Böckenholt, U. (2014). A meta-analysis of loss aversion in product choice. *Journal of Retailing, 90*(2), 182-197. doi:http://dx.doi.org/10.1016/j.jretai.2014.02.002
Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance, 53*, 1775-1798.
Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology, 47*(419-427).
Pascal, B. (1670). *Pensées* (A. J. Krailsheimer, Trans.). New York, NY: Penguin.
Passetti, F., Levita, L., & Robbins, T. W. (2003). Sulpiride alleviates the attentional impairments of rats with medial prefrontal cortex lesions. *Behavioural brain research, 138*(1), 59-69. doi:http://dx.doi.org/10.1016/S0166-4328(02)00229-2
Paulus, M. P., Rogalsky, C., Simmons, A., Feinstein, J. S., & Stein, M. B. (2003). Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *NeuroImage, 19*, 1439-1448.
Paxinos, G., & Watson, C. (1997). *The Rat Brain in Stereotaxic Coordinates* (4th ed.). San Diego: Academic Press.
Pitkow, X., Liu, S., Angelaki, Dora E., DeAngelis, Gregory C., & Pouget, A. (2015). How can single sensory neurons predict behavior? *Neuron, 87*(2), 411-423. doi:10.1016/j.neuron.2015.06.033
Pizzagalli, D. A., Evins, A. E., Schetter, E. C., Frank, M. J., Pajtas, P. E., Santesso, D. L., & Culhane, M. (2007). Single dose of a dopamine agonist impairs reinforcement learning in humans: Behavioral evidence from a laboratory-based measure of reward responsiveness. *Psychopharmacology, 196*(2), 221-232. doi:10.1007/s00213-007-0957-y
Platt, M. L., & Hayden, B. (2011). Learning: Not just the facts, Ma'am, but the counterfactuals as well. *PLoS Biol, 9*(6), e1001092. doi:10.1371/journal.pbio.1001092
Platt, M. L., & Plassmann, H. (2014). Multistage valuation signals and common neural currencies. In E. Fehr & P. W. Glimcher (Eds.), *Neuroeconomics (Second Edition)* (pp. 237-258). San Diego: Academic Press.

Poterba, J. M., & Summers, L. H. (1988). Mean reversion in stock prices: Evidence and Implications. *Journal of Financial Economics, 22*(1), 27-59. doi:http://dx.doi.org/10.1016/0304-405X(88)90021-9

Prelec, D., & Loewenstein, G. (1998). The red and the black: Mental accounting of savings and debt. *Marketing Science, 17*(1), 4-28. doi:doi:10.1287/mksc.17.1.4

Preuschoff, K., & Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. *Ann NY Acad Sci, 1104*, 135-146.

Preuschoff, K., Bossaerts, P., & Quartz, S. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron, 51*, 381-390.

Puig, M. V., Rose, J., Schmidt, R., & Freund, N. (2014). Dopamine modulation of learning and memory in the prefrontal cortex: insights from studies in primates, rodents, and birds. *Frontiers in Neural Circuits, 8*. doi:10.3389/fncir.2014.00093

Rajala, A. Z., Jenison, R. L., & Populin, L. C. (2015). Decision making: effects of methylphenidate on temporal discounting in nonhuman primates. *J Neurophysiol, 114*(1), 70-79. doi:10.1152/jn.00278.2015

Rao, H., Korczykowski, M., Pluta, J., Hoang, A., & Detre, J. A. (2008). Neural correlates of voluntary and involuntary risk taking in the human brain: An fMRI study of the Balloon Analog Risk Task (BART). *NeuroImage, 42*, 902-910.

Reis, D. J., Baker, H., Fink, J. S., & Joh, T. H. (1979). A genetic control of the number of central dopamine neurons in relationship to brain organization, drug responses, and behavior. In E. U. J. K. Barchas (Ed.), *Catecholamines: Basic and Clinical Frontiers* (pp. 23-33): Pergamon.

Reis, D. J., Fink, J. S., & Baker, H. (1982). A genetic control of number of dopamine neurons in brain: Relationship to behavior and responses to psychoactive drugs. In S. S. Kety, L. P. Rowland, R. L. Sidman, & S. W. Matthysse (Eds.), *Genetics of Neurological and Psychiatric Disorders* (pp. 55-75). New York: Raven Press.

Rescorla, R. A. (1992). Response-independent outcome presentation can leave instrumental R-O associations intact. *Animal Learning & Behavior, 20*(2), 104-111. doi:10.3758/BF03200407

Rodriguez, M. L., & Logue, A. W. (1988). Adjusting delay to reinforcement: comparing choice in pigeons and humans. *Journal of experimental psychology. Animal behavior processes, 14*(1), 105-117.

Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009). Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *Journal of Neuroscience, 29*, 13365-13376.

Rolls, E. T., Sienkiewicz, Z. J., & Yaxley, S. (1989). Hunger modulates the responses to gustatory stimuli of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. *European Journal of Neuroscience, 1*(1), 53-60.

Romo, R., & Schultz, W. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *J Neurophysiol, 63*(3), 592-606.

Rossi, M. A., Sukharnikova, T., Hayrapetyan, V. Y., Yang, L., & Yin, H. H. (2013). Operant self-stimulation of dopamine neurons in the substantia nigra. *Plos One, 8*(6), e65799. doi:10.1371/journal.pone.0065799

Rubenstein, M. (2001). Rational markets: yes or no? The affirmative case. *Financial Analysts Journal*(May-June), 15-29.

Rushworth, M. F. S., Kolling, N., Sallet, J., & Mars, R. B. (2012). Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Current Opinion in Neurobiology, 22*(6), 946-955. doi:http://dx.doi.org/10.1016/j.conb.2012.04.011

Rutishauser, U., Mamelak, A. N., & Schuman, E. M. (2006). Single-trial learning of novel stimuli by individual neurons of the human hippocampus-amygdala complex. *Neuron, 49*, 805-813.

Santana, N., Mengod, G., & Artigas, F. (2009). Quantitative analysis of the expression of dopamine D1 and D2 receptors in pyramidal and GABAergic neurons of the rat prefrontal cortex. *Cerebral Cortex, 19*(4), 849-860. doi:10.1093/cercor/bhn134

Santos, L. R., & Platt, M. L. (2014). Evolutionary anthropological insights into neuroeconomics: What non-human primates can tell us about human decision-making strategies. In E. a. P. W. G. Fehr (Ed.), *Neuroeconomics (Second Edition)* (pp. 109-122). San Diego: Academic Press.

Savage, L. J. (1954). *The Foundations of Statistics*. NY, New York: Wiley.

Sayman, S., & Öncüler, A. (2005). Effects of study design characteristics on the WTA–WTP disparity: A meta analytical framework. *Journal of Economic Psychology, 26*(2), 289-312. doi:http://dx.doi.org/10.1016/j.joep.2004.07.002

Schonberg, T., Fox, C. R., & Poldrack, R. A. (2011). Mind the gap: Bridging economic and naturalistic risk-taking with cognitive neuroscience. *Trends in cognitive sciences, 15*(1), 11-19. doi:10.1016/j.tics.2010.10.002

Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behav Brain Funct, 6*(24).

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *The Journal of neuroscience, 13*(3), 900-913.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593-1599. doi:10.1126/science.275.5306.1593

Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience, 23*(1), 473-500. doi:doi:10.1146/annurev.neuro.23.1.473

Sejnowski, T. J., & Paulsen, O. (2006). Network oscillations: emerging computational principles. *J. Neurosci., 26*(1673-1676).

Seo, D., Patrick, C. J., & Kennealy, P. J. (2008). Role of serotonin and dopamine system interactions in the neurobiology of impulsive aggression and its comorbidity with other clinical disorders. *Aggress Violent Behav, 13*(5), 383-395. doi:10.1016/j.avb.2008.06.003

Sewell, M. (2011). *History of the efficient market hypothesis*. Retrieved from http://www.cs.ucl.ac.uk/fileadmin/UCL-CS/images/Research_Student_Information/RN_11_04.pdf

Shafir, S. (1994). Intransitivity of preferences in honey bees: support for 'comparative' evaluation of foraging options. *Animal Behaviour, 48*(1), 55-67. doi:http://dx.doi.org/10.1006/anbe.1994.1211
Shapira, Z., & Venezia, I. (2001). Patterns of behavior of professionally managed and independent investors. *Journal of Banking & Finance, 25*(8), 1573 - 1587.
Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*(3), 425-442.
Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance, 40*(3), 777-790. doi:10.1111/j.1540-6261.1985.tb05002.x
Shenhav, A., Botvinick, Matthew M., & Cohen, Jonathan D. (2013). The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. *Neuron, 79*(2), 217-240. doi:10.1016/j.neuron.2013.07.007
Shiller, R. J. (2003). From efficient markets theory to behavioral finance. *The Journal of Economic Perspectives, 17*(1), 83-104. doi:10.2307/3216841
Shulman, L. M., Taback, R. L., Rabinstein, A. A., & Weiner, W. J. (2002). Non-recognition of depression and other non-motor symptoms in Parkinson's disease. *Parkinsonism & Related Disorders, 8*(3), 193-197. doi:http://dx.doi.org/10.1016/S1353-8020(01)00015-3
Shumway, T., & Wu, G. (2005). *Does the disposition effect drive momentum?* Paper presented at the AFA 2006 Boston Meeting.
Sigurdsson, V., Saevarsson, H., & Foxall, G. (2009). Brand placement and consumer choice: An in-store experiment. *Journal of Applied Behavior Analysis, 42*(3), 741-745. doi:10.1901/jaba.2009.42-741
Simmons, A., Matthews, S. C., Stein, M. B., & Paulus, M. P. (2004). Anticipation of emotionally aversive visual stimuli activates right insula. *Neuroreport, 15*, 2261-2265.
Simon, H. A. (1959). Theories of decision-making in economics and behavioral science. *The American economic review*, 253-283.
Simon, N. W., Gilbert, R. J., Mayse, J. D., Bizon, J. L., & Setlow, B. (2009). Balancing risk and reward: a rat model of risky decision making. *Neuropsychopharmacology: Official Publication Of The American College Of Neuropsychopharmacology, 34*(10), 2208-2217.
Simonson, I., Nowlis, S., & Lemon, K. (1993). The rffect of local consideration dets on global choice between lower price and higher quality. *Marketing Science, 12*(4), 357-377. doi:10.1287/mksc.12.4.357
Skinner, B. F. (1953). *Science and Human Behavior*. Oxford, England: Macmillan.
Sokol-Hessner, P., Camerer, C. F., & Phelps, E. A. (2013). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Social Cognitive and Affective Neuroscience, 8*(3), 341-350. doi:10.1093/scan/nss002
Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences, 106*(13), 5035-5040. doi:10.1073/pnas.0806761106
Sopher, B., & Sheth, A. (2006). A deeper look at hyperbolic discounting. *Theory and Decision, 60*(2-3), 219-255. doi:10.1007/s11238-005-4596-7

St Onge, J. R., & Floresco, S. B. (2009). Dopaminergic modulation of risk-based decision making. *Neuropsychopharmacology, 34*(3), 681-697. doi:10.1038/npp.2008.121
St. Onge, J., Chiu, Y., & Floresco, S. (2010). Differential effects of dopaminergic manipulations on risky choice. *Psychopharmacology, 211*(2), 209-221. doi:10.1007/s00213-010-1883-y
Steiner, A. P., & Redish, A. D. (2012). The road not taken: neural correlates of decision making in orbitofrontal cortex. *Frontiers in Neuroscience, 6*. doi:10.3389/fnins.2012.00131
Stephens, D. W., & Krebs, J. R. (1986). *Foraging Theory*: Princeton University Press.
Stopper, C., & Floresco, S. (2011). Contributions of the nucleus accumbens and its subregions to different aspects of risk-based decision making. *Cognitive, Affective, & Behavioral Neuroscience, 11*(1), 97-112. doi:10.3758/s13415-010-0015-9
Strahilevitz, M. A., Odean, T., & Barber, B. M. (2011). Once burned, twice shy: How naive learning, counterfactuals, and regret affect the repurchase of stocks previously sold. *Journal of Marketing Research, 48*(SPL), S102-S120. doi:doi:10.1509/jmkr.48.SPL.S102
Suri, R., & Schultz, W. (1999). A neural network with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience, 91*, 871-890.
Sutton, R., & Barto, A. (1998). *Reinforcement Learning*. Cambridge, MA: MIT Press.
Sved, A. F., Baker, H. A., & Reis, D. J. (1984). Dopamine synthesis in inbred mouse strains which differ in numbers of dopamine neurons. *Brain Res, 303*(2), 261-266.
Symmonds, M., & Dolan, R. J. (2012). The neurobiology of preferences. In R. D. Sharot (Ed.), *Neuroscience of Preference and Choice* (pp. 3-31). San Diego: Academic Press.
Thaler, R. (1998). Mental accounting matters. *Journal of Behavioral Decision Making, 12*, 183-206.
Tindell, A. J., Smith, K. S., Berridge, K. C., & Aldridge, J. W. (2009). Dynamic computation of incentive salience: "Wanting" what was never "liked". *The Journal of neuroscience, 29*(39), 12220-12228. doi:10.1523/jneurosci.2499-09.2009
Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2003). Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci., 23*, 10402-10410.
Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science, 307*, 1642-1645.
Tobler, P. N., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *Neurophysiol, 97*, 1621-1632.
Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Century.
Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science, 315*, 515-518.
Truong, L., Allbutt, H., Kassiou, M., & Henderson, J. M. (2006). Developing a preclinical model of Parkinson's disease: A study of behaviour in rats with graded 6-OHDA lesions. *Behavioural brain research, 169*(1), 1-9. doi:http://dx.doi.org/10.1016/j.bbr.2005.11.026

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.
Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*(211), 453-458.
Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297-323.
Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological Science, 3*(5), 305-309. doi:10.1111/j.1467-9280.1992.tb00678.x
Valencia-Torres, L., Olarte-Sanchez, C. M., da Costa Araujo, S., Body, S., Bradshaw, C. M., & Szabadi, E. (2012). Nucleus accumbens and delay discounting in rats: Evidence from a new quantitative protocol for analysing inter-temporal choice. *Psychopharmacology, 219*(2), 271-283. doi:10.1007/s00213-011-2459-1
van den Bos, R. (2006). Toward a rodent model of the Iowa gambling task. *Behavior research methods, 38*(3), 470-478.
van Gaalen, M. M., Brueggeman, R. J., Bronius, P. F. C., Schoffelmeer, A. N. M., & Vanderschuren, L. J. M. J. (2006). Behavioral disinhibition requires dopamine receptor activation. *Psychopharmacology, 187*(1), 73-85. doi:10.1007/s00213-006-0396-1
van Wingerden, M., Marx, C., & Kalenscher, T. (2015). Budget constraints affect male rats' choices between differently priced commodities. *Plos One, 10*(6), e0129581.
Voon, V., Pessiglione, M., Brezing, C., Gallea, C., Fernandez, H. H., Dolan, R. J., & Hallett, M. (2010). Mechanisms underlying dopamine-mediated reward bias in compulsive behaviors. *Neuron, 65*(1), 135-142.
Wade, T. R., de Wit, H., & Richards, J. B. (2000). Effects of dopaminergic drugs on delayed reward as a measure of impulsive behavior in rats. *Psychopharmacology (Berl), 150*(1), 90-101.
Waite, T. (2001). Intransitive preferences in hoarding gray jays (Perisoreus canadensis). *Behavioral Ecology and Sociobiology, 50*(2), 116-121. doi:10.1007/s002650100346
Weber, M., & Camerer, C. (1998). The disposition effect in securities trading: An eperimental analysis. *Journal of Economic Behavior & Organization, 33*, 167-184.
Wickens, J. R., Reynolds, J. N. J., & Hyland, B. I. (2003). Neural mechanisms of reward-related motor learning. *Current Opinion in Neurobiology, 13*(6), 685-690. doi:http://dx.doi.org/10.1016/j.conb.2003.10.013
Wilson, D. I., Laidlaw, A., Butler, E., Hofmann, D., & Bowman, E. M. (2006). Development of a behavioral task measuring reward "wanting" and "liking" in rats. *Physiol Behav, 87*(1), 154-161. doi:10.1016/j.physbeh.2005.09.016
Wise, R. A. (2009). Roles for nigrostriatal--not just mesocorticolimbic--dopamine in reward and addiction. *Trends Neurosci, 32*(10), 517-524. doi:10.1016/j.tins.2009.06.004
Yacubian, J., & al, e. (2006). Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *Journal of Neuroscience, 26*, 9530-9537.

Zaghloul, K. A., Blanco, J. A., Weidemann, C. T., McGill, K., Jaggi, J. L., Baltuch, G. H., & Kahana, M. J. (2009). Human substantia nigra neurons encode unexpected financial rewards. *Science, 323*(5920), 1496-1499. doi:10.1126/science.1167342

Zeeb, F. D., Robbins, T. W., & Winstanley, C. A. (2009). Serotonergic and dopaminergic modulation of gambling behavior as assessed using a novel rat gambling task. *Neuropsychopharmacology (New York, N.Y.), 34*(10), 2329-2343.

Zhang, J., Berridge, K. C., Tindell, A. J., Smith, K. S., & Aldridge, J. W. (2009). A neural computational model of incentive salience. *PLoS Comput Biol, 5*(7), e1000437. doi:10.1371/journal.pcbi.1000437

Zigmond, M., Berger, T., Grace, A., & Stricker, E. (1989). Compensatory responses to nigrostriatal bundle injury. *Molecular and Chemical Neuropathology, 10*(3), 185-200. doi:10.1007/BF03159728