Reward-related learning and individual differences

Francesca Pesola

Submitted to the University of London
for the degree of Doctor of Philosophy

Goldsmiths, University of London,
New Cross, London, SE14 6NW

Declaration

I, Francesca Pesola, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Acknowledgements

I am grateful to my supervisor Professor Alan Pickering for all the guidance, knowledge and enthusiasm that he has provided throughout the course of this work. I would like to thank the Psychology Department for the PhD bursary award which has allowed me to carry out this research.

I also need to thank the technical support staff, Maurice Douglas, Steve Yesson, Ian Hannent and Rob Davis, for all their help. I am also thankful to my family and friends for all their encouragement and, in particular, I would like to thank Dr. Felix DeBeaumont for his friendship and valuable advice. Last but not least, I would like to thank Dr. Ian Tharp for his involvement in the joint-recruitment and testing during the first two years of my PhD.

Dedication

This thesis is dedicated
to my parents, Antonio and Gladys Pesola,
and my sister Patrizia

Abstract

The thesis explored the impact of personality on performance during category learning (CL) tasks, following motivational and reinforcing manipulations. In particular, the thesis explored how motivation and reinforcement affect performance during category learning tasks and, concurrently, the research aspired to clarify how reward sensitivity is modulated by individual differences in personality. According to findings which indicate that dopamine (DA) plays an important role in reward-based learning (Schultz, 1998; 2002), personality traits which may have a DAergic basis were considered. The thesis makes broad reference to the Reinforcement Sensitivity Theory (RST; Gray and McNaughton, 2000) and, in particular, to the Behavioural Activation System (BAS). Indeed, the BAS is believed to involve DAergic midbrain projections and be sensitive to rewards (Pickering and Gray, 2001). Therefore, the personality traits underlying the BAS are believed to have a DAergic nature and, subsequently, determine inter-individual variations in reward sensitivity.

A series of behavioural experiments were conducted to explore the relationship between motivation and personality during CL tasks. Moreover, a biologically-inspired model was developed to simulate the behavioural data and capture individual differences. The model had a DAergic basis that represented some of the biological mechanisms that underlie procedural learning and that may occur within brain structures thought to be part of the BAS (Gray, 1987). The model was shown to be a useful tool to obtain further insights into the experimental data.

Impulsivity was found to mediate procedural learning in a series of studies. Thus according to RST (Gray, 1987), the present research shows that impulsivity might represent the underlying BAS trait. However, contrary to RST, the model indicated that inter-individual variations in procedural learning were dependent on individual's sensitivity to reward prediction error (RPE) signals rather than rewards *per se*. Finally, the model simulations suggest that category learning under asymmetric payoffs is mediated by both explicit and implicit (i.e. procedural) processes. The implications of these findings are discussed in light of personality theories and in relation to future studies.

Table of contents

List of tables

List of figures

List of Abbreviations

AIC	Akaike Information Criterion
EEG	Electroencephalogram
EMG	Electromyography
fMRI	Functional Magnetic Resonance Imaging
LI	Latent Inhibition
MTL	Medial Temporal Lobe
PFC	Prefrontal Cortex
DA	Dopamine
SNpc	Substantia nigra pars compact
NAc	Nucleus Accumbens
VTA	Ventral Tegmental Area
RPE	Reward Prediction Error
LTP	Long-Term Potentiation
LTD	Long-Term Depression
PET	Positron Emission Tomography
PD	Parkinson's Disease
RT	Reaction/Response Time
RST	Reinforcement Sensitivity Theory
BAS	Behavioural Activation/Approach System
BIS	Behavioural Inhibition System
FFFS	Fight, Flight and Freezing system
SR	Stimulus-Response
WM	Working-Memory
MID	Monetary Incentive Delay
DD	Delay Discounting
UCS	Unconditioned Stimulus
UCR	Unconditioned Response
CS	Conditioned Stimulus
CR	Conditioned Response
LR	Low Reward Probability
HR	High Reward Probability
COVIS	Competition between Verbal and Implicit Systems
CL	Category Learning
RB	Rule-Based
II	Information-Integration
SDT	Signal Detection Theory
PCA	Principal Component Analysis
ImpAss	Impulsive Antisocial Sensation Seeking
N	Neuroticism
E	Extraversion
DI	Dysfunctional Impulsivity
FI	Functional Impulsivity
PS	Positive Schizotypy
SZ	Schizophrenia

Chapter 1

Category Learning

1.1 Chapter aims

This chapter aims to describe the processes underlying category learning and the potential tool represented by category learning tasks to explore reward-related learning and to investigate the personality and motivation interplay.

This chapter presents both theoretical background and empirical evidence that indicate the basic principles and requirements of category learning performance. Moreover, it includes information on how both reward and motivation affect performance on category learning. This evidence indicates the existence of an association between cognition and motivation during category learning tasks.

1.2 Category learning - Introduction

Every day, individuals are faced with categorisation problems that they need to solve in order to produce adequate responses to deal with the circumstances. During categorisation, individuals learn to assign different objects to different categories (e.g. food vs. poison) and they also learn to produce different responses for each category (e.g. eat vs. leave; Ashby & Valentin, 2007). This process is referred to as category learning and it represents a fundamental skill for survival, although each task may vary in its survival importance (Maddox, Bohil & Dodd, 2003). In order to play such an important role in survival, category learning is involved in 'perception, thinking and language and is probably a significant factor in motor perception' (Harnad, 1987, p. 1). It can be concluded that the individual's ability to solve category problems plays a very important role in day-to-day effective human performance and survival (Ashby & Maddox, 2005).

On a daily basis individuals are presented with several and diverse categorisation problems. Some categorisation problems require solutions that cannot readily be expressed in terms of verbalisable rules, for example, the skills of the wine expert at identifying different types of wines. Other problems, however, can be more readily solved by using easily verbalised rules; these rules can be applied when identifying wine from grapefruit juice (Ashby, Alfonso-Reese, Turken & Waldron, 1998).

Considering the important role played by categorisation skills in every-day-life, it is not surprising that several studies have attempted to identify the psychological processes involved in solving categorisation problems and, in particular, during category learning (CL). CL is characterised by the need to lay down a memory trace that can be accessed subsequently to improve categorisation efficacy in similar instances (Maddox & Ashby, 2004).

In order to experimentally explore category learning, it is important that the main features of everyday categorisation problems are reproduced in the laboratory (Maddox, Bohil & Dodd, 2003). A main requirement to look at category learning is to have participants classify new, unfamiliar categories of objects. Otherwise, the experiment would really look at category representation rather than learning *per se* (Maddox, Bohil & Dodd, 2003; Maddox, Markman & Baldwin, 2007). Additionally, learning and expertise are mediated by different neural mechanisms (Ashby & Ell, 2002; Ashby, Ennis & Spiering, 2007). For example, Parkinson's disease (PD) patients are impaired at learning new categories but do not lose old categories acquired before they became ill. Similarly, patients affected by agnosia who show inability to recall learnt, old categories are able to learn new categories (Ashby & Spiering, 2004).

During a CL task, participants are presented with different stimuli and they are required to classify them as members of different categories. Categorisation will often first rely on guessing but it becomes more confident over trials thanks to the presentation of feedback on whether each response was correct or incorrect. In order to create the set of stimuli to be used in an experimental setting, it is essential to recreate stimuli that share the features typical of the category stimuli which are more commonly encountered in the real world. Maddox and colleagues (2003) have indicated that there are few main properties that are common to the diverse categories encountered in the real world.

Categories are characterised by several dimensions which are generally continuously valued rather than binary-valued, for example the range of sweetness in different types of wines. Categories usually have a graded structure so that the stimuli that constitute the category are symmetrically and uni-modally (normally) distributed around a prototype (i.e. the best representative of the category; e.g. typical flavour of red wines). Finally, everyday categories are probabilistic (i.e. they overlap), which implies that it is not possible to achieve perfect performance (i.e. 100% accuracy; Maddox, Bohil & Dodd, 2003). In other words, everyday categories are often constituted by a large (possibly infinite) number of stimuli whose dimensions are normally distributed and the categories overlap with each other rendering categorisation a difficult task.

1.3 Decision-bound theory

In order to understand the perceptual and cognitive processes involved in solving categorisation problems, researchers have opted to compare human performance with the performance of the optimal classifier, which is an ideal process that performs optimally. Optimality is defined as performance that maximises long-term rewards (Bohil & Maddox, 2001). The concept of the optimal classifier as well as the model framework is defined in terms of signal detection theory (SDT). Thus, the measures used in the decision-bound model are equivalent to the ones set by SDT, i.e. response sensitivity (d') and response bias (beta or C).

The decision-bound model suggests that people learn to assign different responses to different regions of perceptual space. They set a decision criterion that determines response regions related to the two categories and when categorising a stimulus, the individual needs to decide which region the stimulus belongs to. The optimal classifier is able to identify the relevant continuous dimension that determines category membership and it is also able to set a criterion along the relevant dimension to ensure maximal long-run winnings (Bohil & Maddox, 2001; Maddox & Dodd, 2001).

A typical circumstance where category learning is required occurs when an individual has to classify several stimuli into two categories (e.g. A and B). For example, when classifying red wine as a member of two different types (table vs. dessert wine) depending on its sweetness, sweetness represents the continuous relevant dimension (variable x). For any stimulus taken from variable x (xi), the optimal classifier estimates the probability of the stimulus given category A and B. The optimal classifier calculates the 'optimal decision function', which is computed from the likelihood ratio of the two category distributions and it is equal to:

$$\mathbf{L(x) = \text{function } (x|B)/\text{function}(x|A)} \qquad \textbf{Eq. 1.1}$$

where $\mathbf{f(x|A)}$ and $\mathbf{f(x|B)}$ indicate the likelihood of the stimulus (i.e. $\mathbf{x}$) given categories A and B (Maddox & Dodd, 2001; Maddox & Dodd, 2003). The optimal classifier possesses complete knowledge of the category distributions and can perfectly record all of the observed stimuli. According to the optimal decision function, the optimal classifier will identify the stimuli x as a member of category A if the likelihood ratio is less than 1 and as a member of category B if the ratio is greater than 1 (Bohil & Maddox, 2001).

Everyday categorisation is affected by the interplay of three factors, which are category discriminability, base-rates (prior probabilities) and costs and benefits to correct and incorrect responses (payoffs; Maddox, Bohil & Dodd, 2003).

Category discriminability, which is also indicated as dprime (d'), refers to the 'standardised distance between category means' (p .1175). The greater the d' value, the more distant the two category distributions and, therefore, the easier it is to assign stimuli to either category (higher accuracy). The d' value is represented in the optimal criterion function as the likelihood ratio, since the latter is affected by the two category distribution overlap.

Category base-rates indicate the probability of the occurrence of a particular category depending on previous instances. Asymmetric base-rates can lead the individual to over-classify new stimuli as members of the more frequently occurring category. For example, a doctor may be more likely to diagnose a patient with a disease (e.g. bird flu) if there is an epidemic.

Finally, asymmetric payoff matrices can also occur in the real world and they also produce a bias in the decision making process. For example, correctly diagnosing bird flu is more beneficial than diagnosing a simple flu; indeed, in this case the correct categorisation has a greater survival value.

The optimal classifier has perfect knowledge of these three factors which are then used to compute a decision criterion that allows reward maximisation (Bohil & Maddox, 2001; Maddox, Bohil & Dodd, 2003). The equation for the optimal decision criterion is represented below:

$$\boldsymbol{\beta = [P(A)/P(B)] * [(Va|A - Vb|A)/(Vb|B - Va|B)]} \qquad \textbf{Eq. 1. 2}$$

where **P(A)** and **P(B)** indicate the base-rates of the two categories, **Va|A** (i.e. an 'a' response given an 'A' category stimulus) and **Vb|B** represent the benefits of correct responses and **Vb|A** and **Va|B** represent the costs of incorrect responses.

After having formulated the decision function and the optimal criterion, the optimal classifier can formulate the optimal decision rule:

$$\textbf{If } \boldsymbol{L(x) > \beta}\textbf{, respond B; otherwise respond A} \qquad \textbf{Eq. 1. 3}$$

Research results indicate that human beings implement the same strategy used by the optimal classifier to solve categorisation problems, although their general performance is not optimal. This is due to the fact that perceptual and criterial noise impairs human performance (Maddox & Dodd, 2001; Maddox, Bohil & Dodd, 2003). Perceptual noise is due to the trial-by-trial variability in the way the individual perceives each stimulus whereas criterial noise is due to trial-by-trial variability in the placement of the decision bound (Maddox & Bohil, 1998).

Even though human beings do not perform as optimally as the optimal classifiers, because of trial-by-trial noise, it seems that they use the same strategy only less effectively. Hence, the optimal decision bound can be used to assess and describe individuals' performance on CL tasks. However, since the noise element is always active, they have to be included in the mathematical equation that describes the optimal decision function, and it is now rewritten as:

if $L(x) > \beta + e$, then respond B; otherwise respond A **Eq. 1. 4**

where β is the participants' average criterion decision over trials and e represents the error element due to perceptual and criterial noise.

Compared to the optimal classifiers, human beings tend to show a smaller alteration in response bias when presented with asymmetric payoff matrices. This is due to the fact, that human participants are less willing to sacrifice accuracy in favour of reward (i.e. they adopt a conservative cutoff placement). When base-rates are symmetric, the decision criterion that maximises accuracy corresponds to the one that maximises reward (Maddox & Bohil, 2005; Maddox, Bohil & Dodd, 2003).

1.4 COVIS - Explicit vs. implicit system

As previously mentioned category learning relies on laying down a memory trace for subsequent use. It is nowadays widely accepted that memory is mediated by multiple systems depending on the type of information to be encoded (Ashby & O'Brien, 2005). Therefore, it is also accepted that CL is mediated by different types of category systems, depending on the categorisation problem. One of the most comprehensive multiple systems model is the COmpetition between Verbal and Implicit Systems (COVIS) model (Ashby et al., 1998; a & Ashby, 2004). This theory of category learning is the first model that attempts to explain category learning at a neuropsychological level (Ashby et al., 1998). Indeed, the COVIS model aims to identify the underlying neural structures that account for the behavioural performance.

The COVIS model suggests the existence of two independent systems that compete during learning: an explicit, rule-based system and an implicit, learning-based one. The explicit system uses logical reasoning and requires working memory and executive attention for efficient performance. In contrast, the implicit system is automatic, closely related to motor activity and relies on procedural learning so that its activation does not require conscious awareness (Ashby et al., 1998; Keri, 2003). The explicit system is able to learn much faster than the implicit system, which instead learns in a slow and incremental way when, and only when, receiving trial-by-trial feedback (Ashby & Valentin, 2005). Evidence to support the existence of both the explicit system and the implicit system comes from the fact that they seem to be activated during performance on two distinct category learning tasks: rule-based and information-integration learning tasks, respectively (see below for details).

The two systems learn independently and they compete for response control throughout the task. Each system calculates a discrimination value every trial by estimating the distance between the stimulus and the criterion bound(s). The response is controlled by the system with the greatest discrimination value (Ashby & Maddox, 2005).

One of the COVIS model's main assumptions is that individuals initially rely almost exclusively on the explicit system. Nonetheless, the explicit system is generally not effective at learning an information integration task; the implicit system, by contrast, is more effective when, and only when, it receives positive feedback (Ashby et al., 1998). Under these conditions, the implicit system takes over, although never completely. In fact, on some trials the explicit system will still take over and respond (Ashby and Maddox, 2005). Evidence also shows that this shift in use from the explicit to the implicit system can also be produced by asking participants to respond quickly and/or impulsively or when participants are required to perform a simultaneous secondary verbal task (e.g. simultaneously counting backwards orally; Smith & Kemler-Nelson, 1984).

1.4.1 RB vs. II tasks

At a psychological level, the evidence to support the existence of the two different systems comes from evidence showing that each system seems to be more effective at performing one of two different types of categorization tasks: rule-based and information-integration learning tasks (Ashby et al, 1998; Ashby & Ell, 2001).

These two types of CL tasks are qualitatively different. Rule-based tasks can be learnt by using some explicit reasoning through a process of hypothesis generation and testing. The optimal rule, which allows maximising accuracy, is usually verbalisable. A typical example of a rule-based task is the Wisconsin Sorting Card Test (WSCT) that requires participants to sort into different categories cards that vary on several binary-valued dimensions (colour, shape and number). In contrast, the optimal rule in information-integration tasks requires integrating information on several stimulus dimensions at a pre-decisional stage and the rule is not easily verbalized (Ashby & Maddox, 2005).

In order to be able to verbalise the rule during an RB task, the classifier must be able to assign a semantic label to each dimension and, subsequently, identify the relevant dimension. Once the relevant dimension has been identified it is necessary to place a decision criterion along this continuous dimension in order to determine the two categories. In the easiest type of RB task, even though the stimuli may vary on several dimensions only one of them is the relevant one (i.e. uni-dimensional rule). As an example of a potential rule-based (RB) task take the lines below:

/// // /// // \\\ \\ \\\ \\

The stimuli are groups of lines that vary on three binary dimensions, which are: direction of the lines, colour of the lines (black vs. grey) and numerosity of group (2 vs. 3 items). One of these three dimension could be the relevant one, which should be used to formulate the optimal rule, while the others should be ignored being irrelevant. For example, the relevant dimension could be the stimulus colour and the verbal, uni-dimensional rule could be as follows: "If the lines are black, respond A; otherwise respond B".

Optimal rules to solve RB tasks are not necessarily uni-dimensional. In fact, less frequently, the optimal rule may require combining information from several dimensions. In some cases it may be necessary to formulate a conjunctive, disjunctive or a exclusive-or ('xor') rule; all these rules can be verbally described (Ashby & Maddox, 2005; Ashby & Spiering, 2004). Formulating more complex verbalisable rules is more demanding on WM and individuals will stick to a simpler uni-dimensional rule, if it ensures effective performance (Ashby, et al., 1998; Gluck, Shohamy & Myers, 2002; Ashy & Maddox, 2005).

The optimal rule that maximises accuracy during an information-integration (II) task requires participants to integrate information from two or more stimulus dimensions at a pre-decisional stage. Therefore, the optimal rule is non-verbalisable. Even though both II tasks and conjunctive RB tasks require combining information from several dimensions, this is done differently in the two tasks. Indeed, during an RB task whose solution requires a conjunctive rule, information is combined only after decisions are made on the relevant dimensions (i.e. post-hoc rather than a-priori).

An example of an II task is also illustrated by the eight line stimuli presented above. In the case of an II task, one level of each dimension is assigned a value of +1, for example: black, number of items equal 2 and left direction. Then, the optimal rule could be set as follows: "if the sum of values on the relevant dimensions > 1.5, then respond A; otherwise respond B". In contrast, a conjunctive rule would be: 'if the stimulus is black and contains two lines, then respond A; otherwise respond B'.

Even though, the COVIS makes clear predictions of which system should be active depending on the CL task, the simple structure of the task does not determine the strategy that participants will actually employ (Gluck et al., 2002). In some cases participants stick to an incorrect RB rule even if it is suboptimal to solve the task (e.g. II task). This is expected to happen in situations where the implicit system is not sufficiently activated (i.e. no feedback is offered on a trial-by-trial fashion; Ashby Queller & Berrety, 1999). In other cases, an II rule may be employed to solve a more complex RB task (e.g. which may require a disjunctive rule; Maddox, Filoteo, Hejl & Ing, 2004; Ashby & Maddox, 2005). Hence, no prediction can be made about what strategy participants will adopt to maximise performance based on the type of task. Gluck and colleagues (2002) found that a high proportion of participants performing on the weather prediction task, a probabilistic II task, were using a singleton or a 1-dimensional, rather than a multi-dimensional, strategy to solve the task. This is easily explained by the fact that the use of a singleton or a uni-dimensional rule was sufficient to produce effective performance (70%-75% optimal responding). The reviewed evidence indicates that it is not possible to infer what strategy participants will implement based on the task's structure (Shohamy, Myers & Gluck, 2008). Hence, formal models, which allow fitting of the participants' responses, provide a greater insight into participants' performance. They are a great tool to identify the strategy used by each participant.

COVIS states that the main difference between the two systems is in the way they learn. In fact the explicit system is supposed to rely on logical reasoning which requires WM and attention whereas the implicit learning relies on implicit learning. Experimental data supports this assumption. Indeed, evidence shows that switching the location of the response keys (cf. switching hands position) impaired learning during an II task (Ashby, Ell & Waldron, 2003; Maddox, Bohil & Ing, 2004). However, performance on an RB task was not affected by the switch phase manipulations. These results suggest that while the explicit system learns response labels the implicit system learns response locations. These findings support the idea that the latter system relies on procedural learning.

Performance on an RB task was found to be impaired when participants had to simultaneously carry out a numerical analogue of the Stroop task, which requires both WM and attention effort (Waldron & Ashby, 2001; Zeithamova & Maddox, 2006). This impairment was not observed in participants performing on an II task. Similarly, it was found that increasing the number of potential categories in an RB task also impaired performance on an RB task but not on an II task. These results indicate that increasing the demand on WM and attention resources impairs performance on RB tasks, which are highly dependent on these processes, but not on II tasks, which rely on procedural learning (Maddox, Ashby, Ing &Pickering, 2004).

1.5 Neurobiological basis of the two systems

At a neurobiological level, support to the existence and independence of the explicit and the implicit systems comes from evidence indicating that different brain areas underlie the activation of each system. Moreover, evidence also supports that activation in these areas is necessary to ensure effective performance on the different tasks (RB and II). Neural evidence that shows the existence and the independence of these two learning systems comes from neuroimaging data and neuropsychological studies on different patient populations (Ashby & Ell, 2001).

As previously mentioned the two systems use different types of learning to perform CL tasks. The explicit system relies on logical reasoning whereas the implicit system relies on feedback (i.e. reward-related learning). More specifically, the explicit system applies hypothesis generation and testing to solve categorisation problems whereas the implicit system relies on procedural learning (Ashby & Valentin, 2005). Hence, it is expected that several brain structures would mediate performance of the two CL systems.

A series of brain imaging studies have been carried out to identify the brain areas that are involved during performance on rule-based and information-integration tasks. Evidence shows that the neural circuitry found to mediate performance on II tasks corresponds to the one identified as underlying reward processing and the behavioural activation system (BAS; chapter 3).

1.5.1 Explicit system

COVIS assumes that the explicit system solves CL problems by generating and testing hypotheses about category membership. fMRI studies have identified the brain areas that are active during RB tasks and that, therefore, mediate the activation of the explicit system. The main brain structures that were found to be active during the Wisconsin Card sorting test are the right dorso-lateral prefrontal cortex, the anterior cingulate and the head of the caudate (Rao et al., 1997; Filoteo et al., 2005a). These areas are involved in working memory and executive attention, which are involved in hypothesis generation and testing as well as rule switching (Goldman-Rakic, 1995; Ashby, et al., 1998; Ashby, Ell & Waldron, 2003; Ashby & Maddox, 2005, Ashby & O'Brien 2005).

During the hypothesis testing phase the relevant rule is kept active in working memory by the reverberating loop between the lateral units in the prefrontal cortex and nuclei in the thalamus. If the outcome of the response emitted is not satisfactory (i.e. negative feedback), it is necessary to generate an alterative rule. An fMRI study carried out by Filoteo and colleagues (2005a) indicated that the head of the caudate is involved in processing feedback and, in particular, error signals. The anterior cingulate is active during the hypothesis generation phase and it is, therefore, responsible for the rule selection among all the possible rules (Ashby et al., 1998; Ashby & Ell, 2001; Ashby & Spiering, 2004; Ashby & Ennis, 2006). Once a new rule has been selected, the system has to switch executive attention from the old to the new rule. COVIS assumes that the PFC sends a signal to the head of the caudate, which is responsible for this volitional switching between alternative rules (Ashby et al., 1998; Ashby & Ell, 2001; Filoteo et al., 2005a; Ashby & Valentin, 2005). The caudate activity is mediated by dopamine levels, with low levels of DA in the caudate being associated with a greater level of perseverative errors during RB tasks (Goldman-Rakic, 1995; Roberts, De Salvia, Wilkinson, Collins, Muir & Everitt, 1994; Ashby & Spiering, 2004; Ashby & Ennis, 2006).

Further evidence that supports the involvement of these brain areas during RB tasks comes from data collected on patients affected by Parkinson's disease (PD), who are impaired on RB tasks (Ashby, Noble, Filoteo, Waldron & Ell, 2003; Filoteo et al., 2005a; Price, 2006; Filoteo, Maddox, Ing, Zizak & Song, 2005b). PD is characterised by dopaminergic degeneration in the substantia nigra pars compacta (SNpc) and the ventral tegmental areas (VTA) whose projection sites include the PFC, the caudate and the anterior cingulate (Ashby et al., 1998; Pinel, 2002). PD patients were impaired when performing on RB tasks compared to an age-matched sample composed of healthy controls (Ashby, Noble, Filoteo, Waldron & Ell, 2003). Recent evidence shows that learning deficits in PD patients during RB tasks is associated with impaired selective attention processes, which are mediated by the striatum and, in particular, the head of the caudate (Ashby et al., 1998; Ashby, Noble, Filoteo, Waldron & Ell, 2003; Filoteo, Maddox, Ing & Song, 2007). These researchers observed that increasing the number of irrelevant dimensions (up to 3) impaired patients' performance compared to matched-old and younger controls, i.e. the patients required more trials to reach criterion and committed more errors (Filoteo et al., 2005b). These findings were replicated in a follow-up study where patients were found to be impaired in a 1-dimensional RB task when there was unrelated variation on the irrelevant dimension. However, patients did not differ from the control group when performing on an RB task that required a conjunctive or a disjunctive rule, which requires a greater WM effort. These results support the findings that lower performance of PD patients is due to impaired attentional functioning rather than WM activity (Filoteo et al., 2007).

Further evidence that supports the involvement of nigrostriatal DA activity in strategy switching and cognitive flexibility comes from a study by Cools and colleagues (2003). These researchers compared performance of PD patients on and off DA precursor treatment (L-dopa) to healthy controls during performance on a switching task. Patients who were off the medication showed impaired switching and, therefore, performance compared to those on medication, who showed performance levels comparable to those of the control group.

These results suggest that PD patients in the early stages of the disease are impaired in the rule-switching rather than the hypothesis generation and testing phase. In fact, rule-switching is mediated by the activity of the head of the caudate which is a structure affected by the disease at the early stages. Subsequent damage of DA projections to the PFC should impair WM capacity and, therefore, further impair performance on RB tasks by disrupting rule retention during the hypothesis generation and testing (Ashby et al., 1998; Ashby & Maddox, 2005; Filoteo et al., 2007). Hence, different forms of impairment on RB tasks should be observed across the disease spectrum.

1.5.2 Implicit system

An fMRI study by Poldrack and colleagues (1999) observed brain activation during a feedback-based learning and paired-associated learning on the weather prediction task, which is a variation of an II task (Pickering, 2004). Results showed that the tail of the caudate was significantly activated during feedback-related learning whereas the medial temporal lobe (MTL) was active during observational learning. These results indicate that memory (mediated by MTL) is essential when learning the association between exemplars and category labels; whereas feedback processing (mediated by the head of the caudate) is essential for the implicit system to perform on II tasks. Activation of the tail of the caudate during II tasks was confirmed in a more recent fMRI study conducted by Nomura and colleagues (2007). Further support to the double dissociation between memory functions and probabilistic classification comes from a study comparing CL performance on a weather prediction task between PD and amnesic patients (Knowlton, Mangels & Squire, 1996). PD patients were found to be impaired on the probabilistic CL task but not on a multiple-choice questionnaire on the details of the layout and stimuli presented in the CL task, which tested declarative memory. The opposite pattern was observed in the amnesic patients. These results suggest that effective performance on II tasks requires the activation of the neostriatum (caudate and putamen) and that dysfunction of these brain area, typical of PD patients, disrupts procedural learning while memory remains intact.

The caudate receives inputs from visual and auditory association areas and the PFC and it sends excitatory signals back into the cortex via the globus pallidus and the thalamus (Carlsson, 2002). The caudate projects primarily to the premotor areas (e.g. Area 8 and the supplementary eye fields), which are responsible for ocular orientation and orienting responses, and to the supplementary motor area, which selects motor programs for limb movements (Ashby & Valentin, 2005; Ashby & Ennis, 2006; Ashby, Ennis & Spiering, 2007).

The caudate not only receives glutamate projections from the visual cortex but it also receives DA projections from both the VTA and the SNpc. These synapses are supposed to play a critical role during implicit learning and, therefore, during performance on II tasks (Ashby & Ennis, 2006; Ashby et al., 2007). The effect of DA on the striatum has been found to play a key role in strengthening the cortical-striatal synapses (long term potentiation, LTP; Ashby et al., 2007). Evidence shows that three factors are necessary to ensure learning and strengthening of these synapses, they are: 1) strong pre-synaptic activation, 2) strong post-synaptic activation, and 3) DA firing (3-factor rule). Factors 1 and 2 depend on the intensity of the stimulus and the sensitivity of the glutamatergic NMDA receptor, which has a high threshold of excitation. Noise and weak signals will not produce LTP but rather long term depression (LTD) which weakens stimulus-response associations. An incorrect response would also suppress DA firing and therefore produce LTD.

In addition, evidence shows that for learning to occur DA has to be released shortly after the post-synaptic activation, i.e. before repolarisation has occurred. Several studies reported in the following section indicate that feedback (reward) timing is very important for implicit learning to occur.

The COVIS model assumes that through a procedural learning process mediated by the 3-factor rule, the caudate learns to associate a specific visual input (stimulus) to a motor response. In other words, procedural learning involved during II tasks produces strong associations between a stimulus and its corresponding response (i.e. correct categorisation response).

Evidence that supports the central role of the basal ganglia (especially the tail of the caudate) in the implicit system and during II tasks comes from data from patients with basal ganglia disorders (i.e. PD patients; Knowlton et al., 1996). Further studies suggest that PD patients are impaired on II tasks but only when the structure of the categories is complex. For example, a study carried out by Ashby and colleagues (2003) compared PD patients' performance to healthy controls not only on an RB task but also on an II one. Results showed that participants were not impaired on the II version, where the categories were linearly separable. However, studies observed that compared to a matched control sample, PD patients were impaired on more complex II tasks that require a quadratic decision bound for optimal performance (Maddox & Filoteo, 2001; Filoteo et al., 2005b).

PD is characterised by general dysfunction of the caudate. As previously mentioned, damage to the head of the caudate was found to be associated with impairment on RB tasks. Dysfunction in the tail of the caudate due to reduced DA activation in the area suggests that the caudate is not able to reinforce the correct SR association according to the 3-factor rule and, therefore, implicit learning is expected to be impaired in PD patients.

1.6 CL & feedback

Feedback facilitates the operations of both systems, although in different ways (Ashby et al., 1998). In particular, it has been suggested that in the explicit system feedback guides the process of hypothesis testing and generation, whereas in the implicit system positive feedback works as a reward signal that automatically reinforces the adequate response (3-factor rule). In the implicit system, positive feedback has been found to act like a reinforcement signal and elicit dopamine release from the substantia nigra pars compact (SNpc) into the caudate. These dopamine projections strengthen the synapses responsible for the correct stimulus-response association, thus producing learning (Maddox et al., 2004).

Data that support these claims come from patients' data (Ashby et al, 1998; Ashby & Ell, 2001). Empirical data further supports the dissociation between the two systems and their underlying neural mechanisms. In particular, the two systems seem to be differently affected by feedback manipulations. Ashby, Queller and Berrety (1999) observed that when participants received no trial-by-trial feedback, performance on an II task was impaired, whereas performance on an RB task was not affected. Additionally, participants were found to use an explicit rule to solve the II task when no feedback was offered (Ashby et al, 1999). Similarly, compared to feedback training, observational learning was found to impair learning on an II task but not performance on an RB task (Ashby, Maddox & Bohil, 2002)

Feedback timing was also found to differently affect performance on an RB and on an II task. For example, Maddox, Ashby and Bohil (2003) observed that performance on an II task was effective when feedback was presented immediately following the response. However, results also showed that performance on the II task was impaired when the feedback signal was presented after a delay equal or greater than 2.5 secs. In contrast, performance on the RB task was not impaired by varying the feedback delay. These results were replicated by Maddox and Ing (2005) who observed that during an II task accuracy levels were lower when feedback presentation was delayed by 5 seconds rather than immediate. Reduced accuracy was a result of the increased use of RB strategies in the delayed condition.

A recent study by Ashby and O'Brien (2007) assessed the impact of different feedback manipulations during an II task. Participants were randomly allocated to one of four feedback conditions; the four conditions consisted of full feedback, partial feedback (80% probability), gain-only or loss-only feedback. In the last two conditions, feedback was administered at a rate of 80% on relevant trials. Results showed that participants used sub-optimal, explicit rules when performing under a gain-only or a loss-only matrix whereas they implemented an II strategy when receiving both positive and negative feedback. In contrast, participants were found to learn to perform appropriately during an RB task under both gain- and loss- only matrices (Markman, Daldwin & Maddox, 2005).

However, performance on an RB task was affected by the available feedback processing interval (intertrial interval; ITI). In fact, performance was impaired when participants did not have time to process the feedback signal (ITI = 0) in an RB task compared to those who had 2.5 seconds to process the feedback. Performance on the II task was not affected by this ITI manipulation. These results are due to the fact that short ITI impairs hypothesis testing (Maddox, Ashby & Bohil, 2003; Maddox, Ashby, Ing & Pickering 2004).

These results highlight a main difference between the two systems, the explicit system relies on logical reasoning, which requires effort, attention and WM; whereas the implicit system seems to learn effortlessly by automatically processing the feedback signal.

According to the evidence reviewed in the two sections above, it seems that dopamine and feedback mediate performance under both types of CL tasks. Therefore, DA and feedback influence the functioning of the two systems by affecting the relevant processes that are essential for their performance. However, feedback and DA are only essential for the well-functioning of the implicit system which relies on procedural learning (see chapter 2).

1.7 CL & motivation

In order to understand the way individuals learn and decide to act, it is necessary to understand the interplay between motivational and cognitive factors (Higgins, 1997; Maddox, Markman &Baldwin, 2007). The motivation literature suggests that there are two types of goals: approach goals which have a positive end state which is desirable to attain; avoidance goals which have a negative end state, which is desirable to avoid (Maddox, Baldwin & Markman, 2006).

In particular, Higgins (1987; 1997) postulated that motivation can be influenced by regulatory focus, which affects decision making and cognitive processing. The regulatory focus theory suggests that there are 'different ways of achieving different types of desired end-states' (Higgins, 1997, p. 1281). In particular, Higgins suggested that individuals can have two types of regulatory focus: promotion and prevention. The promotion regulatory focus renders individuals sensitive to positive outcomes and it is characterised by a sensitivity to approach positive goals. On the contrary, prevention focus renders individuals sensitive to negative outcomes and it is characterised by a tendency to avoid possible mismatches to desired goals.

The model suggests that individuals with a promotion focus should be more willing to maximise hits and avoid misses (errors of omissions). In contrast, individuals with a prevention focus are expected to be more willing to make correct rejections and avoid making a mistake (errors of commissions). Higgins (1987; 1997) suggests that regulatory focus represents a trait variable (chronic regulatory focus) since individuals seem to have a predisposition towards regulatory focus. Individual differences in chronic regulatory focus have been found to influence the way individuals approach a task in relation to its incentives.

This account makes broad reference to the Reinforcement Sensitivity Theory (RST) by distinguishing between approach and avoidance stimuli. RST suggests that there are three systems that mediate responding to motivational factors; the behavioural activation system (BAS), Fight, Flight and Freezing system (FFFS) and the behavioural inhibition system (BIS; Corr, 2006; 2008). The BAS is more relevant to the regulatory focus theory and, in particular, its characteristics resemble the ones typical of the promotion focus. Indeed, the BAS is considered to be highly reactive to reward, leading to approach behaviour and positive affect (Smillie and Jackson, 2006; Smillie, Pickering & Jackson, 2006a).

Additionally, the model framework proposes that regulatory focus does not only represent a chronic state but it can also be induced experimentally by manipulating incentives (situational focus) and this has been shown to influence cognitive processing and decision-making during learning (Crowe & Higgins, 1997; Maddox et al., 2006; Maddox, et al., 2007). Finally, the model suggests that individuals show higher performance when the induced regulatory focus matches the chronic regulatory state and when the induced regulatory focus matches the reward structure offered in the task (Shah, Higgins & Friedman, 1998; Maddox et al., 2007). This match is known as 'regulatory fit' between the two foci and it is predictive of performance on category learning (Higgins, 1997; Maddox et al., 2006, 2007).

Shah and colleagues (1998) looked at the relationship between chronic regulatory focus and experimental regulatory focus manipulations. They had participants perform on an anagram solving task under either a prevention or a promotion regulatory focus. In the promotion framed condition, participants were told that they could win extra money (from $4 to $5) if they found 90% or more of the words. In contrast, during the prevention framed condition participants were told that they could avoid losing money (from $5 to $4) if they did not miss more than 10% of the words. Participants with a chronic promotion focus performed better under a promotion-framed condition (reward-oriented) than under a prevention-framed one. The opposite was true for individuals with a chronic prevention focus.

Shah and colleagues (1998) carried out a follow up study using the same paradigm and observed how different payoff matrices affected individuals' performance during the anagram solving task. They had participants solving two types of anagram tasks (green and red) which had different payoffs matrices. The green anagram task offered a more rewarding payoff structure than the red anagram's one. In fact, if participants completed correctly the green anagram set they could win one point; otherwise they would not lose or win any points. Incorrect performance on the red anagram set led to the loss of one point, while correct performance ensured maintaining the points. Participants with a chronic promotion focus were found to be more motivated to maximise their performance on the green anagrams rather than the red ones. The opposite pattern of behaviour was shown by individuals with a prevention focus.

Similarly, evidence from recent studies on RST has shown that individuals with a highly reactive BAS perform better when the motivational manipulation offers reward (i.e. gain-only matrix) rather than punishment (i.e. loss-only matrix). For example, Smillie and colleagues (2007) had participants perform on a yes/no category learning task. The stimuli were vignettes that indicated the characteristics of hypothetical job candidates and participants had to decide whether they were suitable (yes-response) to the job or not (no-response). High BAS participants were found to perform more accurately when performing under a gain-only payoff matrix rather than under a loss-only matrix.

These researchers carried out a further study that used the same design and aimed to assess the impact of asymmetric payoff matrices on learning. Participants performed on the same yes/no category learning task and 'yes' responses received either confirmatory feedback (gain-only matrix) or corrective feedback (loss-only matrix). In contrast, the probability of receiving confirmatory or corrective feedback following 'no' responses was equivalent across the two conditions (loss vs. gain payoffs). Results showed that high BAS individuals under the gain-only matrix developed a greater bias for responding 'yes' than those under the loss-only matrix (Smillie, Dalgleish & Jackson, 2007). Similar results were replicated in a go/no-go category learning task (Smillie & Jackson, 2006). Hence, these results show that BAS activation and chronic regulatory focus modulate the impact of experimental motivation manipulations and, therefore, performance on category learning tasks.

1.8 Summary

The present chapter offered a review of the theoretical background of category learning, particularly focusing on the COVIS model. Moreover, it presented evidence from both neuro-imaging and behavioural studies which underlined the processes involved in solving category learning problems.

The evidence also indicates that performance on CL tasks is affected by both motivation and cognition as indicated by studies manipulating reward, reward structures and cognitive demands. According to the COVIS model, different manipulations affect performance differently depending on the features and requirements of the task as well as the active system that moderates performance.

The review indicates that there exists extensive knowledge about the nature of different category learning problems and the skills and processes required for optimal performance. This knowledge renders these tasks ideal to explore the relationship between motivation and cognition during performance on CL tasks.

Chapter 2

Dopamine

2.1 Chapter aims

This chapter aims to offer an overview of the central role played by dopaminergic pathways in processing reward and mediating motivation as well as motivated approach behaviour. More specifically, the chapter explores how dopamine activity mediates reward-related learning. Evidence from both human and animal studies is examined in order to assess the relationship between reward processing, DA activation and learning.

2.2 DA and motivation

Experimental manipulation through electrical or chemical stimulation of the brain has helped identify the reward pathway and the neurotransmitters involved in reward processing. Olds and Milner (1954) observed that rats would learn to lever press at a high rate (over 6000 times per hour) in order to obtain direct electrical stimulation in the limbic system. Administration of electrical stimulation seems to be more rewarding than the administration of any natural rewards (e.g. food). In a study where rats were presented with a forced-choice between lever pressing for natural rewards (i.e. food and water) and for electrical stimulations, results indicated that rats would chose electrical stimulation over natural rewards (Routtenberg & Lindy, 1965). Electrical stimulation was administered in the limbic area in both studies. For ethical reasons, human studies on electrical stimulation are very limited but results have shown that electrical stimulation in the limbic system is rated as extremely pleasurable in humans (Heath, 1964)

Additionally, neurochemical studies have been conducted on animals to discover which neurotransmitters are involved in reward processing. Several studies observed that rats readily learnt to self-administer stimulants (e.g. cocaine) and opiates, which are highly addictive substances in humans for their rewarding effect. These drugs administrations were associated with increased levels of dopamine (DA) in the limbic system (Bozarth, 1990). In summary, these results and many others indicate that dopamine is the central neurotransmitter involved in emotional and cognitive processes and, especially, in processing reward signals (Bozarth, 1994).

The brain contains three main dopaminergic systems that originate in the midbrain, in the substantia nigra pars compact (SNpc) and in the ventral tegmental area (VTA; Carlsson, 2002; see figure 2.1). These systems consist of a limited number of cells with highly branched axons, which allow the neuron to project to several brain regions and, in particular, the prefrontal cortex and both the ventral and the dorsal striatum. These areas are involved in processing rewards and mediating approach behaviour amongst various functions (Schultz, 1998; Carlsson, 2002).

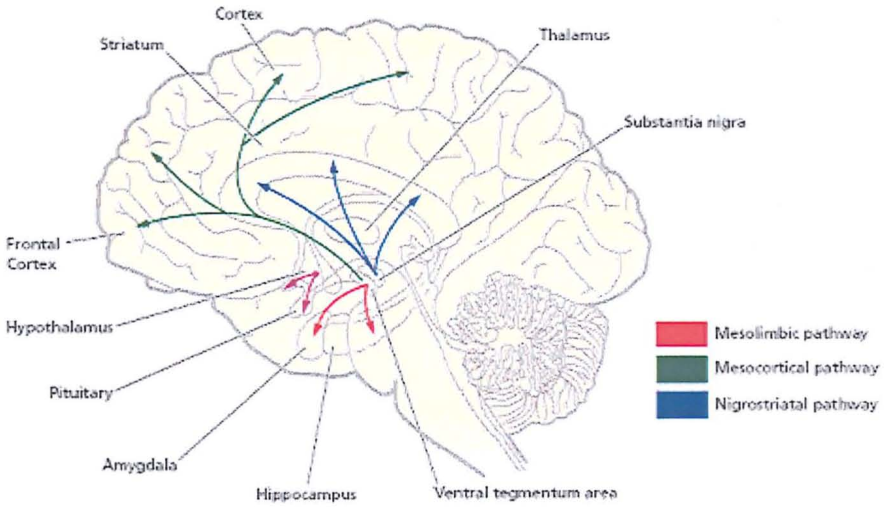


Figure 2.1 Three main dopaminergic systems. From: Townsend MC (2006). Psychiatric Mental Health Nursing: Concepts of Care in Evidence-Based Practice (5th Ed.) Philadelphia, PA. F. A. Davis Company.

Cells in the SNpc project to the neostriatum (caudate and putamen) that is located in the basal ganglia. These projections form the nigrostriatal system which is involved in movement control. Damage to this system is associated with symptoms typical of Parkinson's disease (PD). PD symptoms are generally reduced by increasing DA levels in the basal ganglia. This result can be obtained by DA replacement therapy (administering Ldopa), DA receptor (DA_2) agonist administration or inhibiting DA-breakdown (administration of MAO-B inhibiters; Julien, 2003).

Cells in the VTA area project to the cerebral cortex (mesocortical system) and the limbic system (mesolimbic system; Carlsson, 2002). The mesocortical system mediates executive functioning, e.g. strategy planning and execution. Schizophrenia's symptoms are associated with dysregulation of this system and, in particular, with high levels of DA in the midbrain (Heinz, 2002). Typical antipsychotic drugs used in the treatment of SZ symptoms (e.g. haloperidol) work as DA receptor blockers (especially D_2 receptors; Julien, 2003). In fact, schizophrenia is generally viewed to be characterised by dysregulation of the mesocortical (low DA) and the mesolimbic system (high DA) as well as their interaction (Julien, 2003).

The mesolimbic system, in particular the DA projections to the nucleus accumbens (NAc), has been identified as a key part of the reward system and, therefore, it is considered to be central to rewarded behaviour and motivation processing (Bozarth, 1994). Drugs of abuse have been found to increase DA activation in the mesolimbic system in the same fashion as natural rewards (e.g. food; Bozarth, 1990; DiChiara & North, 1992). For example, an fMRI study showed that amphetamine injections increased activity in the ventral tegmental area and nucleus accumbens, whereas DA depletion blocked DA-activity enhancing effect (Jenkins, Sanchez-Pernaute, Brownell, Chen & Isacson, 2004). Similarly, administration of antipsychotic drugs has been found to reduce cocaine administration in rats (de Wit & Wise, 1977; Childress & O'Brien, 2000).

Animal studies have indicated that DA neurons show short, phasic activation when animals receive natural rewards and this activation is associated with motivated responses to approach the reward (Schultz, Dayan & Montague, 1997; Schultz, 2002). DA neurons respond in the same way to different types of primary rewards (e.g. food and liquids) as long as they are not fully predicted

(Schultz, et al., 1997). DA can discriminate between rewarding and aversive stimuli, in fact DA shows little or no phasic activation in response to aversive stimuli (Schultz et al., 1997). For example, DA phasic activation was observed in monkeys when they found hidden food resulting from their self-initiated exploratory movement into a box (i.e. unpredicted rewards). In contrast, DA neurons did not show any activation in response to punishments or non-reward stimuli (Romo & Schultz, 1990).

In a PET study conducted by Pappata and colleagues (2002) human participants performed on a number comparison task where they had to decide whether the number on the screen was greater or smaller than 5. As monetary feedback, participants experienced unpredicted sequences of winning and losing money. The study aimed to assess the relationship between monetary feedback and DA release, indexed by decreased [^{11}C]raclopride concentration. Results showed that DA release in the mesolimbic system (especially, NAc) increased when participants experienced unpredicted rewards but no DA release was observed in response to losses (Pappata, et al., 2002).

Incentive motivation mediates behaviour which leads the individual to approach the (appropriate) primary rewards (e.g. water when thirsty) that are necessary for wellbeing, survival and reproduction (Bozarth, 1990, 1994). As a direct consequence, rewards induce positive motivational states in the individual. They also have direct behavioural consequences. Rewards and reward-related stimuli lead to immediate approach behaviour and they also enhance response learning for future circumstances. In fact, motivated behaviours lead to experiencing reward and this experience increases the chance of reproducing the same reward-related behaviour (response) when presented with similar circumstances in the future (Corr, 2006). This process will be discussed further later in this chapter.

fMRI techniques have been used in order to assess the role played by the striatum during reward-related learning during a gambling task (e.g. Delgado et al., 2000; 2003). On each trial, participants were presented with a card, whose value could vary between 1 and 9, and they had to guess whether the value of the card was higher or lower than 5. During the first 0.5 second of the trial (i.e. the probabilistic cue period), participants saw one of five cues (star, circle, square, triangle or diamond). Participants were told to pay attention to the cues since each of them had a predicting value on the type of card that would follow (high vs. low). Participants had 2.5 seconds to make a guess and they received auditory, visual and monetary feedback, for each response they made; when the card was equal to five they did not receive any feedback (neutral condition). Participants performed under a predetermined pseudorandom reinforcement schedule. They received reward on 40% of their responses, symmetrically they also received punishment for 40% of their responses and they received neutral feedback on 20% of their overall responses. Results showed that the basal ganglia were activated when processing reward-related information, but not the other type of feedback. In particular, the fMRI data showed that the areas most active during reward processing were the ventral striatum (nucleus accumbens) as well as the dorsal striatum (caudate).

As mentioned earlier, addictive drugs activate the DA midbrain reward system and, therefore, studies exploring the effects of these substances can offer a useful insight into normal appetitive motivation and motivated behaviour, which are characterised by a sensitivity to, and a tendency to approach, reward-related stimuli (e.g. food). Goal-directed behaviours are also mediated by the salience of the stimulus, which is associated with its rewarding value (Volkow, Fowler & Wang, 2003). The salience of a stimulus is relative to the circumstances encountered by the individuals and the salience value of potential alternative stimuli (Volkow et al., 2003). For example, when experiencing thirst, water would have a high salience value and it would have a higher value than food-related stimuli. The higher the salience value associated with the stimulus, the greater the activation of the motivational system and, therefore, the motivational response. Evidence indicates that drugs of abuse are three-to-five times more rewarding than natural rewards, which would explain their high addictive potential (DiChiara, 2002). Prolonged chronic use (typical of addiction) leads to decreased sensitivity to the rewarding effect of natural rewards and, hence, drugs become the main motivational drive (Volkow et al., 2003). For example, a

PET study showed that monetary rewards failed to activate midbrain DA activity in smokers but not in non-smokers (Martin-Sölch et al., 2000).

As mentioned earlier, drugs of abuse mimic the effect of natural rewards and lead to increased activation in the VTA DA cells which project to the limbic system and, in particular, the nucleus accumbens. This activation thus leads to higher DA levels in the limbic system. A PET study by Volkow and colleagues (1999) showed that intravenous administration of methylphenidate (MP; a psychostimulant drug) produced increased release of DA (phasic firing) from the VTA into the NAc. Increased DA activity was also found to be associated with participants' self-reports of experiencing 'high', with greater 'high' ratings reported by those administered the higher dose of MP. According to these results reward processing is mediated by phasic DA firing.

Robinson and Berridge (2001; 2003) suggest that repeated use of drugs of abuse leads to alterations in the mesocortical limbic system and, by doing so, render the system hypersensitive to drugs and drug-related stimuli, which is responsible for further drug-seeking and consuming behaviours. In other words, extensive use of these substances increases the salience of drug-related cues, because of their positive incentive value, and it also leads to consummatory motivated behaviour. For example, Mendrek and colleagues (1998) observed that compared to naïve rats, those pre-treated with amphetamine showed a greater break point (number of drug self-administrations) under a progressive ratio of administration (i.e. higher motivation). In other words, amphetamine pre-treatment sensitises the system to drugs, possibly, by modifying the mesolimbic system and its reactivity to drugs and rewards in general.

A study by Fiorino and Phillips (1999) shows that intraperitoneal (IP) injections of psychostimulants should also increase the incentive properties of natural rewards and, therefore, enhance goal-directed behaviour to obtain them. More specifically, the study showed that amphetamine, as opposed to saline, administration facilitated sexual behaviour in male rats. Similar results were obtained by Wyvell and Berridge (2000) who observed that DA micro-injections into the NAc led to increased lever pressing for sucrose administration in the presence of reward-related cues. An fMRI study by Knutson and colleagues (2004) explored the effect of one oral administration of amphetamine on performance during a monetary incentive delay (MID) task. During this task, human participants had to decide whether to respond to a stimulus or withhold their responses (go-nogo task), basing their decisions on the preceding cue. The cues indicated whether responding to the stimulus had a potential gain or a potential loss outcome. Results showed that amphetamine, compared to placebo, administration was associated with a blunted but prolonged activation of the ventral striatum during expectation of gains. Since ventral striatal activation is supposed to be an index of DA activation (Schultz, 2002), the results suggest that amphetamine administration increased tonic but reduced phasic DA activity.

2.3 Learning and reward – DA mediation

DA release is not only associated with natural rewards and consumption of addictive drugs. In fact, phasic DA release in the VTA has been found to be activated by visual and auditory conditioned stimuli that are predictive of rewards. This association occurs through learning and, more specifically, via classical and operant conditioning (Schultz, 1998). Hence, DA activation is related not only to the presentation of reward but also to reward prediction that can be elicited by conditioned cues (Schultz, et al., 1997; Schultz, 1998; Knutson et al., 2004). In fact, animal studies have indicated that DA neurons show phasic firing in response to visual and auditory stimuli that, following conditioning, act as reward predictors and are able to trigger approach behaviour (Schultz, et al., 1997; Schultz, Tremblay & Hollerman, 2003; Schultz, 2006). Hence, the reward predicting cues adopt the values which are characteristic of the natural reward, according to classical conditioning processes.

Classical and operant conditioning reflect associations between two events. In the case of classical conditioning the association made is between two stimuli (stimulus-stimulus association) whereas in the case of instrumental conditioning the association learnt is between a stimulus and the adequate response (stimulus-response association), which leads to reward (Corr, 2006).

On the one hand, the main assumption of the classical conditioning model is that after pairing a neutral stimulus with a stimulus with primary reward properties, the neutral stimulus acquires the rewarding properties of the natural reward and produces the same response (Corr, 2006). A few rewards seem to elicit innate vegetative or muscular responses but the majority of rewards are learnt throughout life (Schultz, 1998).

On the other hand, the main assumption made by the instrumental conditioning model is that each response made by an individual is followed by feedback from the environment. Depending on the kind of feedback received, individuals will determine whether such actions should be re-enacted in the future when presented with similar circumstances. Thus, reward acts as a positive reinforcer by enhancing learning of a stimulus-response association which is of value to the individual (Schultz, 2006).

The main difference between classical conditioning and instrumental learning is that in classical conditioning the outcome follows the stimulus without the need of any voluntary action to be produced, while in instrumental conditioning the outcome is obtained as a result of a voluntary response to the present stimulus or circumstance (Schultz, 2006). Rewards and DA activation reinforce both types of associative learning. In the case of classical conditioning, the stimulus conditioned becomes more salient and acquire properties typical of primary rewards (Schultz, et al., 2003). In operant learning the correct response becomes more salient and, therefore, more likely to be adopted in the future (O'Doherty, 2004).

Animal studies have observed that after learning has occurred, the conditioned cue acts as a predictor of the reward probability, magnitude and timing (Abler, Walter, Erk, Kammerer & Spitzer, 2006). Moreover, the cue alone is able to elicit DA firing and approach behaviour (Schultz, et al., 1997). DA firing to conditioned rewards shares the main characteristic of firing to natural rewards, although the phasic response is slightly weaker (Schultz, 2002). Indeed, the magnitude of the activation is smaller and the number of neurons responding is lower than when firing is activated by primary rewards (Scultz, 1998; 2002). The timing of the firing is also modified. Indeed, prior to learning DA fires in response to unpredicted reward delivery (figure 2.2a) whereas once task performance is established, DA fires in response to the reward-predicting cue but not to the actual reward (figure 2.2b; Schultz, et al., 1997; Mirenowicz and Schultz, 1994; Schultz, 2002). Observations have indicated that, following cue-reward learning, if no reward is delivered following the cue presentation, DA firing is depressed below basal firing rate at the time when reward would usually be delivered (Schultz, et al., 1997; Schultz, 1998; figure 2.2c).

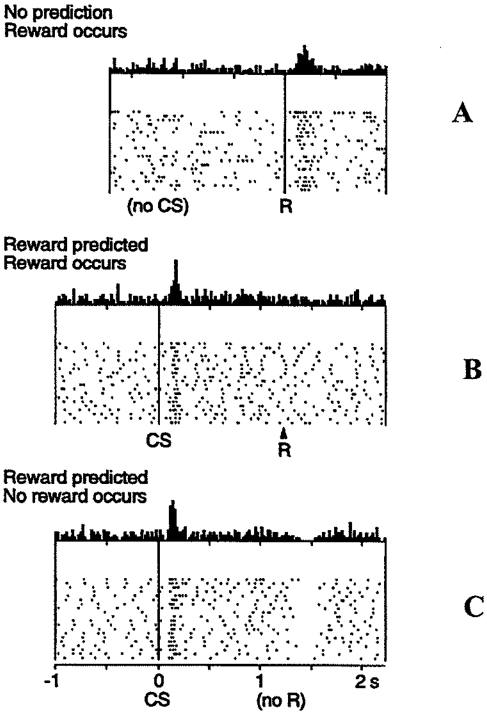


Figure 2.2. DA cell recording of the relationship between DA activity and reward, evidence for reward prediction error (CS = conditioned stimulus; R = reward). From: Schultz W. (1998), Predictive reward signal of dopamine neurons, Journal of Neurophysiology, 80(1), pp. 1-27.

Several animal studies have indicated that following both classical and operant learning, conditioned stimuli can elicit DA phasic firing similar to activation caused by natural rewards. For example, in a study conducted by Ljungberg and colleagues (1991) monkeys learned to lever press when instructed by a light to press the lever in order to receive liquid reward. DA neurons were found to respond both to reward administration and the conditioned stimulus. Similar results were obtained when monkeys were trained to reach out in order to obtain food morsels when a box door opened. In this study DA firing was associated both to reward presentation and the conditioned stimulus (i.e. door opening).

In a study by Schultz and colleagues (1993) monkeys were gradually trained to perform on a delayed response task via two transitional spatial tasks. In the three tasks, in order to receive a liquid reward monkeys had to learn to press the correct lever (right vs. left) when instructed by a light. Performance was dependent on two stimuli (lights), a trigger light and an instruction light. The former indicated when the response had to be made while the instruction light indicated which lever had to be pressed to receive the reward. DA activity was recorded at different stages of learning: the initial training stages and later stages when performance is established. Results indicated that during the training stage, DA cells fired (especially in the VTA) in response to reward presentation and to both the instruction and trigger stimuli (CS). In the later stages of learning reward delivery did not lead to DA firing whereas DA fired in response to the two conditioned stimuli. These observations applied to the three different tasks. Additionally, data indicated that once monkeys reached stable performance, DA firing in response to both primary rewards and conditioned cues was significantly reduced (Schultz, 2003)

These results were replicated in a study by Mirenowicz and Schultz (1994) where monkeys were trained to lever-press to obtain liquid reward when prompted by a sound. Results showed that during the initial stages, DA fired in response to the presence of the conditioned stimulus (i.e. sound) as well as in response to reward delivery. In contrast, in the later stages of training, when conditioning had occurred, DA fired only in response to the conditioned stimulus but not in response to the actual reward. Similar results were obtained by Hollerman and Schultz (1998) who observed that DA activity was greater during the early phase of learning while rewards were unpredictable but DA firing stopped once learning had established. Results from these studies support the assumption that the conditioned stimulus is also able to elicit the approach response that would have been elicited by primary rewards (Schultz, 1998).

An fMRI study was conducted on humans to explore DA levels in the NAc during performance on a monetary incentive delay (MID) task with varying reward probability (0%, 25%, 50%, 75% and 100%; Abler et al., 2006). Results indicated that during the reward prediction period, the phasic activity of the NAc increased linearly as a function of reward probability. Similarly, during the outcome period there was a linear relationship between the NAcc activity and RPE, so that the NAc activity was found to be higher with the most positive RPE. All these results show that DA activation is associated to reward unpredictability. Therefore, they support the claim that DA fires to primary reward stimuli only during learning but this activity stops once learning has established.

Pessiglione and colleagues (2006) conducted an MRI study to explore the effects of L-DOPA (i.e. DA agonist) and haloperidol (i.e. DA antagonist) administration during performance on an instrumental learning task where participants had to respond to one of two stimuli, in order to earn money. Each stimulus was associated with a potential loss, gain or neutral monetary outcome. L-DOPA administration, compared with haloperidol, was found to enhance responding to the gain-related stimuli and it was also associated with increased activity in the ventral striatum, which reflected greater RPE signals. The opposite pattern was observed in participants who had been administered haloperidol, which inhibits DA activation. The results indicate that reward-related decision-making is modulated by DA function (especially, reward prediction error; RPE).

Further evidence that indicates the important role played by DA in reward processing comes from patients' data. Indeed, never-medicated schizophrenics showed reduced activation in the ventral striatum (esp. NAc) during anticipation of reward while performing on a MID task compared to healthy controls (Juckel et al., 2006a). Reduced striatal activation was positively correlated with both negative and (as a positive trend) with positive symptoms. The authors suggested that reduced reward sensitivity resulted from high tonic DA activity, which decreases the signal-to-noise ratio and, therefore, interferes with normal processing of reward and reward-related cues (Juckel et al., 2006a). In a follow-up study, schizophrenic patients treated with typical neuroleptics, showed a blunted activation of the ventral striatum, similar to that observed in the earlier study of unmedicated SZ patients (Juckel, et al., 2006b). The researchers suggested that these effects are due to the primary effects of typical neuroleptics to block D2 receptors and, subsequently, blunt the reward system and worsen negative symptoms. These results suggest that higher tonic DA levels hinder DA phasic response to reward (prediction error). This assumption is supported by results showing that SZ patients treated with atypical neuroleptics, which do not blunt the DA system, showed the same levels of ventral striatal activation as healthy controls (Juckel et al., 2006b).

Finally, these observations are supported by the study by Knutson and colleagues (2004), which showed that during an MID task, amphetamine administration increased tonic DA levels but blunted DA phasic responses in response to potential gains (SZ-like activations). Murray and colleagues (2007) found that patients with positive psychotic symptoms showed lower DA activation in the ventral striatum, which was associated with weaker RPE signals, compared to healthy controls during performance on a reward learning task. During the task, participants had to respond to one of two stimuli to earn money. Each pair of stimuli was constituted by one high reward probability (HR) stimulus and one low reward probability (LR) stimulus. Results showed that patients tended to respond to the HR stimulus less frequently than the healthy controls (although this difference was non-significant).

Evidence from human opiates addicts indicated that drug-related cues elicit drug-like positive motivational and physiological states (i.e. ‘opiate-like euphoria’; Stewart, et al., 1984, p.260). This phenomenon is observed in needle freak behaviour amongst former opiate addicts and it is associated with increased DA activation in the midbrain (Steward, 1984). These findings further indicate that conditioned cues acquire the same appetitive properties of primary rewards through associative learning.

As already noted, DA activation only occurs when reward incidence is unpredictable and it does not occur when reward delivery is fully predicted by a stimulus or a cue that has been previously experienced to precede rewards (Schultz, 1998; Schultz, 2002). Once a reward becomes fully predictable, DA phasic firing and, subsequently, learning terminates (Hollerman & Schultz, 1998). DA activity in response to reward is affected by the unpredictability of the reward timing and the actual occurrence of reward (Schultz, 2002; Schultz, et al., 2003).

Animal studies show that DA activation varies throughout learning and it also indicates that once learning has fully occurred DA firing ceases (Mirenowicz & Schultz, 1994; Hollerman & Schultz, 1998; Schultz, 1998; Schultz et al., 2003). In an fMRI study conducted by Delgado and colleagues (2005), human participants had to guess whether the value of a card they were about to see was high or low (card values ranged between 1 and 9). They were instructed to base their decision on the preceding cue. There were five different cues and each contained probability information of the next card value. Results showed that DA activity (in the caudate) was greater in the initial stages of performance and decreased as learning progressed and the cues became effective predictors of the card value and, therefore, reward.

The loss of DA firing cannot be attributed to reduced reward sensitivity, as in fact DA activity is reinstated when rewards are administered outside the task (Mirenowicz & Schultz, 1994; Hollerman & Schultz, 1998; Apicella, Ljungberg, Scarnati & Schultz, 2001). DA firing seems to be related to reward unpredictability rather than reward presentation *per se*. Indeed, the main difference between learning and acquired knowledge lies in reward unpredictability. Once learning has terminated the reward occurrence can be predicted by the conditioned stimuli and DA does not fire.

2.3.1 RPE

The ability to make reward predictions is adaptive for the organism since it allows the animal (including humans) to produce adequate behavioural responses that will increase the chance of earning reward. In particular, animals seem able to predict the magnitude, probability and timing of the future reward based on their previous experience (Schultz et al., 1997; Schultz, 1998).

The data reported above indicate that DA fires in response to both unexpected and underpredicted rewards. Therefore, the data suggest that DA is responsive to reward unpredictability rather than to reward *per se* and that DA-driven learning only occurs when rewards are unpredictable (Schultz, 1998).

During the early stages of learning during a go-nogo task, performance relies on guessing and reward is predicted to occur at a chance which represents a reward expectation (i.e. prediction). The reward prediction is then weighted against the actual reward outcome. The difference between the actual and the predicted reward represents the reward prediction error (RPE). If the RPE is different from zero, the reward is partially unpredictable and learning is required. Moreover, a correct response would lead to a reward, which represents a positive prediction error whereas an incorrect response would lead to a punishment or a non-reward, which represents a negative prediction error (Schultz, 2002). As learning progresses, RPE approaches zero, which represents total predictability of the reward.

A study by Tobler and colleagues (2005a) indicated that the unpredictability of reward magnitude also leads to positive or negative RPE, if the reward obtained was greater or smaller than the predicted one, respectively. These results taken together suggest that, owing to the unpredictability of reward magnitude, probability or timing, animals formulate outcome expectations that can either match or mismatch the actual outcome. Thus, RPEs guide future behaviour and expectations (Schultz, 1998; Cromwell & Schultz, 2003).

Since DA neurons fire in response to reward unpredictability, they may be sensitive to reward prediction errors and their phasic activation can be considered to reflect an RPE signal, which can be formalised in the equation below:

$$\textbf{Phasic DA response} = \textbf{f(Reward occurring} - \textbf{Reward expected)} \qquad \textbf{Eq.2. 1}$$

where **f()** means "a function of".

A positive prediction error is associated with increased DA firing whereas a negative prediction error is associated with DA depression below baseline levels (Schultz et al., 1993; Schultz, 1998; 2002). In a study by Schultz and colleagues (2003) when a monkey made a wrong response, reward delivery failed to occur, which led to DA depression at the time when reward delivery was expected. Since feedback was administered 500 msec after the response had been made, it is possible to affirm that DA depression was not determined by the incorrect response *per se* but it was a direct result of mismatch between the reward prediction and the actual outcome (i.e. it reflects a negative RPE).

Once learning has terminated and the individual has mastered performance on the task, the reward prediction will be equal to the actual reward. Therefore, the RPE value will be equal to zero and according to the equation above there will be no phasic DA activation (Schultz, 1998; 2006), and thus no DA-dependent learning.

The assumption that the RPE is crucial to learning is supported by studies that explored Kamin's 'blocking' paradigm (Schultz, 2002; Corr, 2006). These studies indicate that once associative learning between a conditioned stimulus (CS) and a reward has occurred, a new CS cannot become conditioned to predict the reward in spite of being paired with the predictor.

An fMRI study conducted by Tobler and colleagues (2005b) used the blocking paradigm to observe how blocking affects appetitive learning. They trained participants to respond to different visual stimuli. During the pre-training phase one stimulus was followed by reward (A+) whereas the other was not followed by reward (B-). Participants had to indicate on what side of the screen the stimulus appeared by pressing the corresponding button. In the training phase the stimulus X was associated with A+ to create the AX+ compound which predicted reward; while the stimulus Y was associated with B- to create the BY+ compound which also predicted reward. In the test phase, the X and Y stimuli were tested alone to assess appetitive learning. As expected, participants showed shorter RT in response to A and Y than to the neutral B stimulus and the blocked X stimulus. Behavioural results were associated with phasic activation in the putamen and the orbitofrontal cortex. Results showed that phasic responses were greater for the non-blocked stimulus (Y) than for the blocked one (X), which suggest that the blocked stimulus may have not led to phasic DA firing in the striatum, which is a critical brain area involved in reward prediction signals (Schultz, 1998).

The fMRI study conducted by Tobler and colleagues (2005b) observed that the activation of the medial orbitofrontal cortex (OFC) was greater for non-blocked stimuli compared to the blocked ones. Moreover, this activation was positively associated with learning levels and it ceased when learning terminated. Similar observations were made for the ventral putamen. In fact, activation in the ventral putamen was associated with learning and it was only observed for non-blocked stimuli but not for the blocked ones. The OFC seems to be responsible for reward processing in relation to reward magnitude (Cromwell and Schultz, 2003).

These results suggest that the medial orbitofrontal cortex (OFC) and the putamen are also involved in processing reward prediction errors. O'Doherty (2004) reviewed several fMRI studies looking at conditioning of reward and results show that OFC and the ventral putamen are active during RPE processing. O'Doherty and colleagues (2001) also found that medial OFC showed increased activation that was proportional to the monetary reward magnitude.

Moreover, these researchers have suggested that coding of the RPE in the OFC may be mediated by DA projections into these regions (Tobler et al., 2005b; O'Doherty, 2004). Contrary to DA responses that are sensitive only to reward occurrence/prediction, responses from these brain areas are capable of discriminating the different nature (e.g. liquid or visual) and magnitude of rewards and, therefore, can modulate motivated approach behaviour (Schultz 1998; 2002). Subsequently, the striatum may use RPE information to learn the adequate motor response (Cromwell & Schultz, 2003). In fact, the striatum is involved in the control of movement as indicated by patients' data that indicate that damage to midbrain projections into the striatum is associated with PD (Carlsson, 2002).

2.4 The three-factor neural model of learning

The striatum receives projections from the midbrain DA system and sensory inputs from several areas of the cortex; for example the putamen receives inputs from somatosensory and motor areas while the caudate receives visual and auditory inputs from the associative areas. In contrast, the NAc receives input from the OFC. The striatum indirectly projects back into the prefrontal and premotor cortical areas (Ashby and Ennis, 2006). It is assumed that the striatum develops an abstract motor plan to deal with different stimuli and circumstances. The motor plan is subsequently sent to the premotor cortical areas, which will produce the actual motor response. In other words, the striatum is involved in creating adequate stimulus-response associations via procedural (i.e. implicit) learning (Schultz, 1998; Ashby et al., 2007).

Dopamine RPE signals have both immediate and long-term effects. The immediate effects of DA signals are to increase the signal-to-noise ratio, which filters incoming information into the striatum. In particular, strong rather than weak inputs are processed and this selection affects the behavioural response. Indeed, DA activity may act as a selective filter for reward-related stimuli which lead to a shift in favour of processing and approaching these stimuli (Schultz, 1998; 2002; Franken, Booij & van den Brink, 2005a). Long-term effects of DA signals are associated with reward-related learning as a result of synaptic changes. In fact, DA firing is thought to act as a teaching or reinforcing signal at the synapse and, therefore, it produces synaptic plasticity according to Hebb's theory (Sutton a& Barto, 1981; Schultz, 2002).

The main assumption of Hebb's theory is that 'cells that fire together wire together' (Corr, 2006, p. 223). Indeed, Hebb suggested that an active synapse between a group of cells ('cell assembly') may lead to long-lasting structural changes that strengthen the synapse. A result of these structural changes is the increased effectiveness of the synapse. This is knows as a Hebbian synapse and neural evidence that supports this theory come from studies on long-term potentiation and long-term depression (LTP and LTD; Corr, 2006). LTP is associated with strengthening of synapses whereas LTD results in a weakening of synapses (Carlsson, 2002).

Several experiments have shown that synaptic plasticity occurs when glutamate NMDA receptors are activated. Once the post-synaptic cell is partially depolarised by glutamate binding onto the NMDA receptor, the calcium ion channel, usually closed by a magnesium ion, opens allowing further depolarisation of the post-synaptic membrane (Carlsson, 2002). A strong pre-synaptic glutamate signal is necessary to obtain sufficient depolarisation in the post-synaptic cell to open the calcium channel (Carlsson, 2002; Ashby and Ennis, 2006).

DA firing signals have been identified as the teaching signals that mediate implicit learning that requires the activation of the striatum (Schultz, 1998; Grace, Floresco, Goto & Lodge, 2007). Therefore, DA is also involved in long-term potentiation and log-term depression, which mediate procedural learning (e.g. during performance on an II CL task; see chapter 1). Many authors have suggested that three factors are involved in procedural learning, the three factors are: 1) pre-synaptic activation (glutamate signal), 2) post-synaptic activation (NMDA depolarisation) and 3) phasic DA signal (RPE; Ashby et al., 1998; Ashby & Ennis, 2006; Ashby et al., 2007; Seger, 2008).

Before neuroscientific data implicated DA as one of the three factors discussed above, a closely related three-factor model was formalised in the Rescorla-Wagner (1972) learning rule, and this can be interpreted as a rule that accounts for synaptic changes (Schultz, 2002). The Rescorla-Wagner's rule captures associative learning by assuming that over-repeated pairings of the unconditioned stimulus (UCS) and the reward learning progresses according to the rule in equation 2.2 below:

$$\Delta V = \alpha\beta(\lambda - V) \qquad \textbf{Eq.2. 2}$$

where **ΔV** represents the change in the associative strength (**V**) of the stimulus, **α** and **β** are constants indicating the saliency value of the unconditioned stimulus and the reward stimulus, respectively. Additionally, **λ** represents the maximum associative strength that the reward can support (Schultz, 1998). The difference (**λ-V**) indicates the unpredictability of the reward and, therefore, it represents the reward prediction error term. Indeed, when **λ=V** the UCS fully predicts the reward and no further (associative) learning occurs.

The three-factor rule of synaptic weight changes, which characterise operant conditioning (procedural learning), is expressed in the equation 2.3 below:

$$\Delta\omega = \varepsilon * r * i * o \qquad \textbf{Eq.2. 3}$$

where **$\Delta\omega$** indicates the change in a synaptic weight (strength), **ε** is a learning rate constant, **r** represents the DA RPE signal (equation 2.1), **i** indicates the strength of the presynaptic input signal and **o** indicates the strength of the postsynaptic activation or output (Schultz, 2002). A positive RPE will determine LTP (**$\Delta\omega$** will be positive) whereas a negative RPE will determine LTD (**$\Delta\omega$** will be negative), i.e. a weakening of the stimulus-response activation which was not associated with the expected reward. Moreover, DA signals strengthen or weaken only the active synapses that are responsible for the response that led to the reward (or non-reward) and so can train appropriate stimulus-response associations (Ashby et al., 2007).

In addition, evidence shows that for procedural learning to occur all three factors described by the model need to be present. Additionally, the timing of the DA signal is important for learning to occur. Indeed, DA has to be released shortly after the post-synaptic activation has occurred so that the DA signal can strengthen the active stimulus-response synapses that led to reward (Schultz, 1998; Maddox et al, 2004). In other words, the DA release needs to occur while the synapse is still depolarised otherwise the reinforcing DA signal will not be able to strengthen the correct stimulus-response association once it is no longer active.

A study by Maddox and colleagues (2003) explored the impact of delayed feedback during performance on a procedural (II) CL task. They observed that delays as short as 2.5 seconds impaired performance compared to immediate reward delivery. These results indicate that the timing of reward administration is crucial to procedural learning. In order to ensure learning, rewards need to be administered within a few seconds of responding while the active synapse is still depolarised (Ashby et al., 2007).

According to the three-factor model, then, all three factors should be present to observe synaptic changes and procedural learning to occur. The stimulus input should be strong enough to elicit post-synaptic activation and each response should receive a reward. Indeed, if the stimulus is not sufficiently salient the post-synaptic cell will not be depolarised. Similarly, evidence indicates that procedural learning does not occur when trial-by-trial reward is not provided (Ashby et al., 1999).

2.5 Summary

The aim of the chapter was to explore the role played by DA pathways in reward processing and assess how DA activation mediates reward-related learning.

Evidence from electrical and chemical stimulation studies and the addiction literature have identified the midbrain DA projections as the central system that is involved in reward processing and motivation. Moreover, animal and human studies have indicated that phasic DA firing is associated with reward processing and motivated responding. Hence, these results represent the main source of evidence that supports the central role played by DA in directing approach behaviour.

These studies have allowed us to identify the basic principles underlying reward-processing, motivation and reward-related learning. Finally, the insight offered by animal and human research has allowed us to develop computational models that can formalise learning and facilitate further explorations of the topic of motivation and learning (see chapter 7).

Chapter 3

Reinforcement sensitivity theory

3.1 Chapter aims

The present chapter aims to introduce the reader to personality and in particular to a major biological personality theory developed in the 1970s by Jeffrey Gray. This theory, now known as Reinforcement Sensitivity Theory (RST), states that individual differences in personality are determined by inter-individual variation in emotional-motivational systems. The theory offers a biological explanation for individual differences and personality. The theory postulates the existence of three systems which interact with each other for behavioural control. This chapter will introduce a brief history and overview of the model and it will present the three motivational-emotional systems. However, it will mainly focus on the system involved in reward processing and motivation in order to show the hypothesised links between reward-processing and personality.

3.2 Introduction

Eysenck's (1967) arousal theory is one of the main biological theories of personality and it considers personality and psychopathology as a continuum where clinical disorders lie at one extreme of normal behaviour (Eysenk, 1997). In 1944, Eysenck gathered individual data based on a medical checklist from 700 neurotic individuals and factor-analysed the data. By doing so, he extracted two dimensions of personality: extraversion (E) and neuroticism (N; Carver and Scheier, 2003; Corr, 2006). These two factors are considered to be orthogonal and each factor is continuously distributed. In 1952, a third personality trait was added to the theory; this trait was psychoticism, which was thought of as a measure related to psychotic conditions such as schizophrenia (Corr, 2006).

Eysenck's theory aimed to build a link between personality and biological functions, especially focusing on cortical excitation and inhibition. In the 1960s, a particular structure in the brainstem had been discovered and it was held responsible for the regulation of cortical arousal, which is responsible for alertness, concentration and sleep-waking cycles (Bullock and Gilliland, 1993). This structure was known as the 'Ascending Reticular Activating System' (ARAS) and it was supposed to activate or deactivate the cerebral cortex. Eysenck suggested that extraverts and introverts differed in their level of activation in the ARAS (i.e. cortical arousal). Moreover, differences in arousal were expected to determine individual differences in every day life. Low arousal characterised high extraverts who are defined as sociable, outgoing, and enthusiastic individuals (Hogan, 1997). In contrast, high arousal was considered typical of introverts who are defined as withdrawn, shy and reserved (Hogan, 1997). Neuroticism was postulated to mediate sensitivity to emotional stimuli and situations (i.e. emotionality) and to regulate the activation of the autonomic nervous system (Gray, 1970; Corr, 2004).

Starting from a critique to Eysenck's theory, Jeffrey Gray developed an alternative biological personality theory, the reinforcement sensitivity theory (RST; Gray, 1970; Corr, 2006). Indeed, after reviewing evidence from animal and human studies, Gray (1970) suggested that Eysenck's theory should be modified. This evidence indicated that introverts are more responsive to stimuli related to punishment and non-reward whereas extraverts are more sensitive to stimuli associated with reward and punishment relief. Moreover, neuroticism was considered to be responsive to reinforcing stimuli which are predictive of either punishment or reward. Hence, varying degrees of neuroticism reflect the overall levels of sensitivity to either reward or punishment.

Following these observations, Gray (1970; 1987) suggested that the relationship between N and E traits determines the sensitivity to reward or punishment and, therefore, the axes representing E and N should be rotated by 30° to obtain the personality components, which align more directly with the corresponding biological system. The two traits proposed by Gray are: Impulsivity and Anxiety. The impulsivity personality trait (Imp) ranges from E+/N+ (high Imp) to E-/N- (low Imp) while the Anxiety personality trait (Anx) ranges from E-/N+ (high Anx) to E+/N- (low Anx; Corr, 2002). In this scheme, impulsivity reflects sensitivity to reward whereas Anxiety reflects sensitivity to punishment. As impulsivity levels increase, reward sensitivity also augments whereas sensitivity to punishment is constant (Smillie, Pickering & Jackson, 2006a). The opposite pattern was argued to be true for anxiety.

RST is a theory-driven model of personality that relies on the assumption that individual differences derive from the variation in the activation of emotional and motivational systems across individuals (Smillie et al., 2006a). Initially, the theory was interested in explaining animal learning and, therefore, it aimed to explore the interplay between motivation, emotion and learning. In a subsequent stage, emotional and motivational biological systems were proposed to be related to personality.

Over the years, some of the main RST principles have been reviewed in light of findings obtained in animal and human studies. These modifications have finally led to a revised version of the theory (Gray & McNaughton, 2000; Corr, 2004a; McNaughton & Corr, 2004; Smillie, et al., 2006a). The revised theory offers a more detailed description of the motivational-emotional systems, their functions as well as their neural basis (Corr, 2004a).

3.3 Revised RST

According to the revised version of RST model there are three main systems of emotion that mediate motivation and learning, which are the Behavioural Activation System (or Behavioural Approach System; BAS), the Fight, Flight and Freezing system (FFFS) and the Behavioural Inhibition system (BIS; McNaughton & Corr, 2004; Smillie et al., 2006a, Corr, 2006). The BAS is sensitive to reward and reward-related stimuli (also punishment relief stimuli; Gray, 1970; Schultz, 1998) and it directs motivated behaviour towards positive goals. This system is associated with positive emotions/affect (Corr, 2006). Behaviour mediated by the BAS to obtain reward is known as 'Approach Behaviour' whereas BAS-directed behaviour which aims to obtain punishment relief is known as 'Active Avoidance' (Smillie and Jackson, 2005).

The BAS system is supposed to lie in neural regions which are involved in reward processing and motivation, i.e. the midbrain DA projections from the VTA and the SNpc into the striatum and the prefrontal cortex (Schultz, 1998; 2002; Gray, 1987; Pickering, 2004; Corr, 2006). The FFFS represents the complement to the BAS as it is sensitive and responsive to both natural and conditioned stimuli associated with punishment (Corr, 2006; Smillie et al., 2006a). In the original version of RST this function used to be performed by the BIS, but in the revised theory the BIS is a conflict resolution system (see details below). The emotional states associated with activation of the FFFS is fear and panic which are states experienced in the presence of threats and aversive stimuli. Thus, this system is considered to be associated with negative affect (Corr, 2006; 2008). The neural substrate of this system has been identified to be the amygdala, anterior cingulate, medial hypothalamus and the periaqueductal grey (McNaughton & Corr, 2004).

The potential behavioural responses in the presence of threats and aversive stimuli are, as suggested by the name of the system, fight or freezing and flight. Which one of the three behaviours will be more dominant depends on the distance between the actor and the threat ('defensive distance'; Corr, 2006, p. 540). The freezing and flight responses are viable only when there is sufficient distance from the threat; however flight occurs when escape is possible otherwise the animal will freeze (e.g. if caught in a trap). When the distance is small, the only possible response available is 'fight'. The personality correlates of the FFFS are unclear; however, since it is a fear system the high-end may represent traits, such as fearfulness, that are components of neuroticism (Corr, 2008; Pickering, 2008).

The BIS is a goal conflict detection and resolution system which becomes active when both the BAS and the FFFS system are concurrently and similarly active and, therefore, they are competing for control. In this type of situation, the individual is presented with a conflict of interest where both rewards and actual or potential punishments are present (e.g. drinking an extra glass of wine knowing one has to drive home and may be stopped by the police). Conflict can also arise in the presence of two competing rewards or two competing punishments (McNaughton & Corr, 2004).

In order to resolve the conflict the BIS shows cautious approach behaviour which requires termination of any ongoing behaviour and it involves risk assessing and behavioural caution, increased arousal and attention in response to external stimuli (Corr, 2006). In other words, the BIS interrupts any ongoing behaviour to assess what strategy is optimal and it does so by evaluating both the circumstances and the reinforcement value (Smillie, et al., 2006a). Hence, the BIS is supposed to be associated with the experience of anxiety states and, therefore, the relationship with the FFFS system (which is linked to fear) is unclear. The BIS resolves the situation by favouring the activation of either the BAS or the FFFS (Corr, 2006; 2008). However, it is generally organised so as to favour the FFFS by inhibiting the BAS and, subsequently, stimulating the FFFS. This is achieved by increasing the negative valence of the stimulus or situations (Smillie, et al., 2006a; Corr, 2008).

In line with the theory which stresses the important role played by the BIS in evaluating the emotional valence of stimuli and situations, the BIS is supposed to lie in neural areas such as the hypothalamus and the limbic system that are involved in activation of the autonomic nervous system and in mediating motivated behaviour and emotion, respectively (Gray, 1970; Carlsson, 2002; Corr, 2004a; Corr 2006). The main brain structures that underlie the BIS are the amygdala and the septo-hippocampal system (Gray, 1987; McNaughton & Corr, 2004; Fowles, 2006).

3.4 The interacting nature of the systems

The original RST theory postulated that the three systems and their underlying personality traits were orthogonal and independent from each other (Gray, 1987). However, more recent considerations have led to the conclusion that these systems (and possibly their associated traits) may actually be functionally interdependent (Corr, 2002a, 2004a; Smillie, et al., 2006a). Indeed, it has been indicated that the BAS and the FFFS systems compete for behavioural control by mutually inhibiting each other and, eventually, the conflict is resolved by the BIS which inhibits one of the two systems and indirectly enhances the other. Therefore, in order to observe a pure BAS or FFFS activation it is necessary to use a reward- or a punishment-only payoff matrix (Pickering and Gray, 2001) or if using mixed payoff matrices, it may be necessary to test only individuals who are hyper-sensitive to rewards or punishments, respectively (Corr, 2002b). Otherwise, it is necessary to render the rewards or the punishment highly salient (Corr, 2002b). If neither of these requirements is fulfilled, then all the three systems are active at the same time and, therefore, personality measures underlying the three systems may be found to mediate performance and participants' choices (i.e. joint subsystem hypothesis; Corr, 2002b). This is true in most situations experienced in every-day life as opposed to experimental settings. Moreover, this consideration may account for the mixed results obtained in studies exploring RST. Indeed, evidence in support of the revised RST comes from both psychometric and experimental observations and studies.

For example, Pickering and Gray (2001) reviewed several studies that attempted to identify the BAS underlying personality traits but showed inconclusive results. They indicated that most of these studies used mixed incentives feedback which would account for the results showing an activation of either the FFFS alone or an interaction between FFFS and BAS in mediating reward processing. For example, in a study carried out by Ball and Zuckerman where participants saw a pair of stimuli on each trial, one being the target and the other the distractor, participants had to learn to discriminate the target stimulus from the distractor. The visual stimuli varied on eight bivariate dimensions, but only two of the eight dimensions were relevant to determine whether the stimulus was a target or not. Learning was reinforced by feedback (verbal vs. monetary) which was delivered either in a reward-only or a punishment-only matrix. Results showed that participants who scored high on the sensation seeking scale (a measure of impulsivity) and those who scored high on neuroticism performed better (i.e. required fewer trials to reach criterion) than their low scoring counterparts. These results show that potential FFFS- and BAS-related measures mediated performance, regardless of payoff matrix. Moreover, a study carried out by Zinbarg and Revelle (1989) showed that both the BAS and the FFFS systems mediated learning during a discrimination go/no-go learning task where both punishment and reward cues were presented. Participants had to learn to discriminate between cues that predicted whether a key press would lead to monetary winning or to monetary loss (25 cents). Indeed, results showed that performance on the task (which was a measure of learning over trials) was mediated by an interaction between impulsivity and anxiety (Imp x Anx). In fact, individuals with low scores on impulsivity and high on anxiety and those with high scores on impulsivity and low on anxiety showed better performance than the other two combinations. A similar discrimination task was used by Zinbarg and Mohlman (1998) and the same personality measures were taken into consideration. Results showed that performance (i.e. avoidance learning) was mediated by an interaction between anxiety and impulsivity (Imp x Anx). Contrary to the original results, they observed that individuals who scored high on both impulsivity and anxiety learned to withhold responses to the punishment cues. However, when feedback was ego-relevant (i.e. winning or losing "IQ points") anxiety alone was found to be associated with performance. Hence, these results show that feedback manipulations can modulate and identify individual differences during learning.

Contrary to the assumptions made by the original RST, impulsivity was found to mediate processing of punishments in a study that observed the impact of acoustic startle reflex during the presentation of emotionally valenced pictures (Corr, 2002b). According to RST predictions, results showed that anxiety increased the eyeblink reflex in response to slides showing negative stimuli but anxiety was also found to interact with impulsivity. Individuals who scored high on anxiety and low on impulsivity showed a greater startle response than those individuals who scored high on anxiety and impulsivity. These results suggest that Imp and Anx interact with each other and that Imp has antagonising effects on Anx.

Overall, these studies can be considered to be inconclusive as tests of the original version of RST due to their mixed findings. However, in the light of the revised version of RST they simply seem to indicate that the feedback manipulation was ambivalent since it offered both rewards and punishments (mixed payoff) and therefore it would activate both the reward and the punishment systems. Therefore, these results are probably unhelpful in the quest to identify the personality traits that underlie each individual system. These results, however, highlight the importance of employing the correct payoff matrix (reward-only or punishment-only matrices) in order to avoid ambivalent experimental conditions that lead to activation of the reward as well as the punishment systems.

Further evidence in support of the claim that the three emotional systems and their related personality traits are not fully independent, but may be related, comes from evidence indicating that psychometric measures of the reward and the punishment systems overlap to a certain degree. An example of RST measures is represented by the scale developed by Carver and White (1994) to measure sensitivity of the reward and punishment systems, the BAS and the BIS (FFFS in the revised version of RST) systems, respectively. The scale contains 4 subscales, three of these subscales measure the BAS activation [Drive (BAS-D), Reward Responsiveness (BAS-RR) and Fun Seeking(BAS-FS)] and one subscale measures BIS/FFFS activation (punishment system). The three BAS scales were found to be positively correlated (as expected) amongst themselves [correlation between: BAS-D and BAS-RR: r = 0.34; BAS-RR and BAS-FS: r = 0.41; BAS-D and BAS-FS: r = 0.36]. However, they were also significantly correlated with the BIS/FFFS measure (r = -0.12 with Drive, r = 0.28 with Reward Responsiveness and r = -0.08 with Fun Seeking). The results obtained from these correlations indicate that the reward and punishment systems overlap to a certain degree. Hence, these results suggest that the reward and punishment systems (and their measures) are not entirely orthogonal but somewhat oblique to each other.

Finally, formal modelling evidence also showed that successful simulation of self-reported trait scores expressing the BAS output required the inclusion of the BAS, BIS and FFFS activation values (i.e. weights). These simulations indicate that the BAS output is not only predicted by the activation of the BAS alone but by the activation of the FFFS and, if conflict is present, the BIS sensitivity and reactivity (Pickering, 2008). In particular, BAS activation was positively correlated with the reward system level of sensitivity (reactivity value) but also negatively associated with the reactivity values of the two other systems. Neural model simulations also indicated that there was a negative relationship between the reward (BAS) system and the punishment (FFFS) system. Additionally, the simulation dealt with the case of conflict (when the reward and punishment systems had roughly equal activation levels) by controlling for the BIS. These simulations indicate that measures of the BAS-related trait(s), which could be interpreted as self-reported measures of general BAS output levels, are influenced by the activation of all three systems rather than the BAS reactivity alone (Pickering, 2008; Smillie et al., 2006a). Therefore, the reward and the punishment systems interact with each other rather than being orthogonal to each other, as suggested by the original RST.

In conclusion, evidence from the studies reviewed above indicates that in contexts offering mixed incentive values, and with individuals who are not reward or punishment hypersensitive, the three systems interact. In particular, the FFFS system is activated in the presence of punishments and it inhibits the BAS activation; in contrast, the BAS is activated by reward stimuli and it inhibits the FFFS. However, in cases where both rewards and punishments are present and they are roughly equivalent (i.e. goal conflict), the BIS also becomes activated to solve the conflict. The BIS decides whether to produce an approach or an avoidance response depending on the value of reward and punishment. It generally shows a bias towards the FFFS by augmenting the aversive value of the ambivalent stimuli and decreasing their positive value (Corr, 2004a; McNaughton& Corr, 2004).

Following these observations, for the purpose of this chapter, in the next few sections attention will be focused on the BAS and its underlying neural substrate and personality traits. This choice is driven by the fact that the BAS is responsible for reward processing and motivation which is the main focus of the empirical studies reported in the following chapters. Nonetheless, according to the observations made in the section above it will be necessary to include potential measures of both the BIS and FFFS when conducting empirical observations.

3.5 BAS and DA

As mentioned in the section above, the behavioural activation system is a motivational system that is sensitive to reward and once activated it mediates approach behaviour towards reward stimuli (Corr, 2006). Considering the assumption that the BAS is reactive to reward, it was suggested that the BAS lies within the brain areas that are involved in reward processing and motivation, i.e. midbrain DA projections (Schultz, 1998; Carver & White, 1994; Pickering & Gray; 2001). Midbrain DA has been found to fire phasically in response to unpredicted rewards and reward-related cues (Schultz, 1998; 2002). Therefore, midbrain DA projections have been identified as the main structures involved in reward processing and motivation (Gray, 1987; Fowles, 2006).

RST suggests that BAS-related traits are a by-product of inter-individual differences in the level of sensitivity of the emotional system to rewards (Smillie, et al., 2006a). Variations on the BAS-related trait measures correspond to BAS reactivity and, therefore, reactivity to rewards. Therefore, individuals scoring high on BAS-related traits are expected to show greater sensitivity to rewards than those with lower scores on the BAS-traits. Moreover, DA firing in response to rewards is expected to be greater in high BAS individuals. Since personality variations derive from differences in the sensitivity of the emotional system, the personality traits underlying the system are also expected to have the same biological basis. Thus, BAS-like traits are expected to be dopaminergic (Pickering & Gray, 2001). A study by Lee and colleagues (2007) has indicated that scores on the BAS-RR (Carver & White, 1994) were associated with density levels of the D2 dopaminergic receptor. However, these researchers observed that high scores were observed in individuals with reduced D2 density which could be associated with a greater synthesis and production of DA (Klein et al., 2007; Lee, Ham, Cho, Lee & Shim, 2007). Indeed, receptor down-regulation is considered to represent a compensatory mechanism that maintains DAergic activity in equilibrium (Koob & LeMoal, 1997).

3.6 BAS and personality

The up-to-date literature is not conclusive in pinpointing which are the BAS-related traits and the debate is on-going. Even though Gray originally suggested that impulsivity represented the BAS underlying trait, empirical results do not fully support the claim. Moreover, recently it has been suggested that extraversion may be the actual BAS trait due to the fact that extraversion scores are associated with DA activity (Depue & Collins, 1999). There is an open and on-going debate on whether it is impulsivity or extraversion that underlies the BAS trait. Some of the evidence in favour of the two sides of the debate is reviewed below.

The studies reviewed below contain evidence from behavioural, psychometric and neurobiological evidence to support either the extraversion or the impulsivity theory. The behavioural tasks fulfil the requirements, set by Pickering (2004), which determine whether the experiment classifies as an RST study effective at eliciting BAS activation and, hence, exploring BAS-related traits. As mentioned earlier, only gains payoff matrices should be used to be able to assess a pure BAS activation. Additionally, the task employed should index learning, especially reward-related learning, since it is affected by motivational factors which are mediated by DA activation and, therefore, BAS activity. In fact, DA firing enhances associative and procedural learning and it has been assumed to modulate BAS reactivity (Smillie, et al., 2006a).

3.6.1 Impulsivity as the BAS trait

In the original RST, Gray (1970) suggested that impulsivity is determined by an additive combination of extraversion and neuroticism. However, psychometric observations indicated that impulsivity measures correlated with measures of sensation seeking, antisocial and non-conformity behaviour (Pickering and

Gray, 2001, p.120). Additionally, measures on these scales are associated with antisocial and delinquent behaviour, the use of alcohol and illegal drugs as well as psychopathy (Zuckerman, 1993). These measures have been found to be partly accounted for by Eysenck's psychoticism scale (Zuckerman, 1993). Hence, these authors suggested that measures from all these scales should be included in order to obtain reliable measures of the BAS trait. Therefore, impulsivity should be re-labelled as impulsive antisocial sensation seeking (ImpAss). The ImpAss component suggested by Pickering and Gray (2001) makes a broad reference to the Impulsive Unsocialised Sensation Seeking (ImpUSS) personality component proposed by Zuckerman (1993). The switch from the 'unsocialised' to the 'antisocial' definition was made to avoid any possible interpretation that failed socialisation must be due entirely to environmental or developmental mechanisms.

Zuckerman (1993) pointed out that existing evidence showed that the ImpUSS component was associated with DA activity. In particular, high levels of DA activity were expected to be characterised by high scores on the ImpUSS scale. Zuckerman reviewed genetic and neuroanatomical evidence that indicated that ImpUSS (hence, ImpAss) had a dopaminergic basis and, therefore, it also supports RST claim that impulsivity (and impulsivity related traits) was the underlying trait of the BAS.

Theoretical evidence that supports the view that impulsivity (or ImpAss) is the underlying personality component of the BAS system also comes from studies exploring substance use. Indeed, substances of abuse (e.g. nicotine, alcohol and opiates) have rewarding properties and they are able to activate DA firing in the mesolimbic pathway (Bozarth, 1990; Julien, 2003). Moreover, studies that investigated the impact of individual differences on substance use (cf. addiction) have found that personality traits such as impulsivity, venturesome, novelty- and sensation-seeking (i.e. ImpAss sub-factors) are good predictors of drugs and alcohol use (e.g. quantity and frequency) amongst college students (Waldeck & Miller, 1997; Ham & Hope, 2003). Similarly, smokers, compared to non-smokers, were found to score higher on Impulsivity (Mitchell, 1999; 2004) and similarly ecstasy users were found to score higher on Impulsivity than drug-naïve controls (Morgan, 1998). According to RST assumptions, this evidence seems to confirm that the ImpAss component is a good candidate as a BAS trait. In fact, RST suggests that the BAS is responsive to rewards and high BAS individuals should show greater reward sensitivity and approach behaviour to reward stimuli (i.e. as seen in drug use).

Further empirical evidence from the field of substance use comes from a study conducted by Franken and Muris (2006). They assessed the relationship between scores on the BAS scale (Carver and White, 1994) and alcohol as well as drug use (e.g. cannabis, amphetamine, and opiates) among non-dependent college students. Alcohol use was measured in terms of quantity of alcohol use and frequency of binge drinking, whereas drug use was determined by the number of illegal substances ever used. Results showed that the BAS Fun-Seeking subscale was positively correlated with all three measures of use, while the BAS-Drive subscale was correlated only with the number of substances ever used. Hence, BAS Fun-Seeking seems to be a better predictor of drug use. This scale was found to be positively correlated with Novelty Seeking ($r = .51$, $p <.001$; Carver and White, 1994), which is a sub-factor of the ImpAss component. Smillie, Jackson and Dalgleish (2006) also reported that the BAS Fun Seeking subscale was a good predictor of the Impulsivity construct ($r = .64$, $p < .001$) that consisted of measures from the EPQ-P, Impulsiveness, Risk Taking and Sensation Seeking, which are subscales of the Eysenck Personality Profiler (EPP) Psychoticism trait. However, results also showed that the BAS Fun Seeking scale was also a good predictor of the BAS construct ($r = .51$, $p <.001$), which consisted of EPQ-E, PANAS positive affect and reward sensitivity (Smillie, Jackson & Dalgleish, 2006). Even though, the researchers considered these results ambivalent, they seem to indicate that both impulsivity and extraversion may represent the underlying BAS traits and represent different aspects of the BAS. Pardo and colleagues (2007) replicated the results obtained by Franken and Muris (2006) as they observed that alcohol use and frequency were predicted by BAS scores, assessed by the Sensitivity to Reward scale (SPSRQ; Torrubia, Ávila, Moltó, & Caseras, 2001).

Similar observations to the ones in the Franken and Muris' study were made on a clinical sample of addicted drug (heroin and/or cocaine) and alcohol users who reported higher scores on the BAS Fun Seeking as well as the BAS Drive scales compared to controls (Franken and Muris, 2006). These results indicate that high impulsivity scores are associated with drugs use, which is mediated by the rewarding effects of these substances. Therefore, these results suggest that impulsivity is a personality trait highly reactive to rewards and, therefore, they suggest that it may be the BAS trait. However, when considering individuals who show chronic use of the substances it is difficult to disentangle whether impulsivity score are the cause or the result of the use, or vice-versa. Nonetheless, results from occasional users seem to suggest that the personality trait is the precursor of chronic use.

In a study by Stuettgen and colleagues (2005), a neuroendocrine challenge paradigm was used to investigate what personality trait was associated with DA variations and, therefore, BAS responsivity. Participants either received a capsule containing placebo or mazindol, which acts as a DA reuptake inhibitor and, consequently, reduces prolactin secretion by the hypothalamus. Results showed that individuals who score high on the Novelty Seeking (NS) scale showed general low levels of prolactin compared to those with low scores on the scale regardless of the drug manipulation. Moreover, high NS individuals showed a blunted response to mazindol. Overall, these results show that high NS scores are associated with higher DA levels and activity. Additionally, DA levels were not affected by administration of the DA agonist for individuals with high NS scores (due to a possible ceiling effect).

A study conducted by Abler and colleagues (2006) offers further neurobiological evidence in favour of the ImpAss components as the BAS trait. These researchers carried out an fMRI study that revealed that midbrain DA activity in the NAc was involved in reward processing during a probabilistic monetary incentive task. Additionally, the results showed that NAc activity was positively correlated with personality scores on the Sensation Seeking (SS) and the Novelty Seeking (NS) scales. In particular, it was the Exploratory Excitability (NS1) subscale the Thrill and Adventure Seeking (TAS) subscale of the NS and SS scales, respectively, that were found to be associated with NAc activity and, therefore, reward processing. These two subscales are measures of the ImpAss compound as NS1 measures attraction to novel stimuli and TAS is a measure of the tendency to approach novel and risky stimuli (Abler, et al., 2006).

Finally, Cools and colleagues (2007) found that the administration of a dopamine D2 receptor agonist (i.e. bromocriptine) improved performance during a match-to-sample task in individuals who scored high (cf. low) on the Barratt Impulsiveness Scale. These findings, thus, show a link between dopamine (D2 receptor) activity and impulsivity.

3.6.2 Extraversion as the BAS trait

Depue and Collins (1999) postulated the existence of a motivational system which is responsive to rewards and modulates motivated behaviour. This system was labelled as the 'Behavioural Facilitation System' (BFS) and it closely resembles the Behavioural Activation System postulated by RST. In line with RST, Depue and Collins state that the motivational system has a dopaminergic nature and, in particular, it involves the mesolimbic DA system. This suggestion implies that inter-individual differences are related to the activation of the projections from the VTA to the NAc. Contrary to RST they suggested that the underlying personality trait of the motivational system was extraversion rather than impulsivity, and indeed, they identified the agency sub-factor of extraversion as critical. Studies from their laboratory found that extraversion had a dopaminergic nature. For example, extraversion scores were found to be positively correlated with psychophysiological correlates of DA levels. Indeed, they suggested that variations in midbrain DA and, therefore, in the system sensitivity to rewards were also reflected in individual differences, e.g. extraversion. After reviewing several animal and the few available human studies, they developed a 'psychological threshold model', which states that individual differences in reward sensitivity and in reward responsivity (i.e. threshold) are mediated by midbrain DA transmission. In other words, extraverts were postulated to have a lower threshold and so be more responsive to rewards. As a result, extraverts show more motivated behaviour in response to rewards than introverted individuals.

Several researchers have found evidence that supports the claim that extraversion may be sensitive to rewards and reward-related cues and, therefore, that variations in extraversion may be determined by inter-individual differences in the reward system (BAS; Cohen, et al., 2005). These results support the claim that extraversion represents the underlying personality trait of the Behavioural Activation System (BAS) proposed by Gray (1970; 1987) and that it mediates performance on reward-related learning tasks. For example, a study carried out by Pickering and colleagues found that during performance on a 'weather categorisation task' (i.e. a probabilistic information-integration, II, task[1]), performance was positively correlated with extraversion when learning was guided by reward (reward-dependent learning) whereas performance was mediated by psychoticism scores (EPQ-P, a measure of ImpAss) when learning was based on pure observation (i.e. paired-associative training; Pickering, 2004).

Similar observations were obtained in a study by Smillie, Dalgleish and Jackson (2007) where participants performed on a yes/no category learning task. The stimuli were vignettes that indicated the characteristics of hypothetical job candidates and participants had to decide whether they were suitable (yes-response) to the job or not (no-response). BAS reactivity was measured using the I^7 Impulsiveness questionnaire (Eysenck, Pearson, Easting and Allsopp, 1985) and the BAS scales developed by Carver and White (1994). Results showed that individuals who scored high on the BAS scales by Carver and White performed better under a gain-only payoff matrix than under a loss-only matrix. In contrast, the impulsivity measure was found to be a predictor of poorer performance on the task, regardless of feedback manipulation.

---

[1] Optimal solution of II tasks requires participants to integrate information from several dimensions at a pre-decisional stage (see chapter 1 for a review)

These researchers carried out a further study that used the same design and aimed to assess the impact of asymmetric payoff matrices on learning. Participants performed on the same yes/no category learning task under either a gain only or a loss only payoff matrix. However, the internal payoff matrices for the two possible responses were asymmetric. In fact participants received predominantly more feedback for 'yes' responses (80%) than 'no' responses (10%). The same personality measures observed in the first study were collected in this second one. Results showed that participants with high scores on the BAS scale and under the gain-only matrix developed a greater bias for responding 'yes' than those under the loss-only matrix. Impulsivity was not found to predict performance on the task.

Overall, the three BAS scales were found to be highly positively correlated with extraversion (p smaller than .001; Carver & White, 1994) and, therefore, these results indicate that extraversion (or extraversion-related measures) is sensitive to reward manipulations. On the contrary, the present results suggest that impulsivity is not particularly sensitive to rewards. Therefore, these results support the claim that extraversion, rather than impulsivity, represents the best candidate to be the underlying BAS personality trait.

Further evidence in support of the extraversion theory comes from neurobiological studies. For example, Cohen and colleagues (2005) ran an fMRI study to assess how DA variations affected both extraversion scores and reward sensitivity. In the study participants performed on a gambling task. The task had a go/no-go format where participants had to decide whether to respond or withhold their response based on evaluation of a prediction cue. In fact, one cue predicted that responding would lead to a low risk gamble whereas response withholding would lead to a high risk gamble. Results showed that extraversion variation was not related to different strategy in response to the cues and gamble magnitude. As expected, the medial orbitofrontal cortex (mOFC), the amygdala and the nucleus accumbens (NAc) were found to be active when participants received rewards. In particular, these areas were significantly more reactive in extraverts than introverts, regardless of reward magnitudes. A follow-up study explored how extraversion mediated activation of these brain areas in the reward anticipation (prior to feedback) and the reward evaluation (following feedback) phase. Results showed that high extraversion scores were associated with greater DA activation in the mOFC, the amygdala and the NAc but only in the reward evaluation phases. Moreover, extraverts showed greater response to bigger compared to smaller rewards ($1.25 vs. $2.50). These results support the claim that extraversion has a dopaminergic basis and it mediates reward processing. Therefore, the further support the possibility that extraversion is the underlying personality trait of the BAS.

Franken (2002) observed that scores on the BAS-Drive scale were positively correlated with positive craving and intention to drink alcohol following presentation of alcohol related cues. BAS-drive scores were found to be positively correlated with extraversion ($r = 0.41$, $p<.001$: Carver and White, 1994) and it is considered to be a measure fairly closely related to extraversion. Hence, the present results suggest that extraversion may be the BAS-underlying personality trait and, therefore, scores on this scale mediate incentive sensitisation as described by Robinson and Berridge (2001, 2003). Robinson and Berridge suggest that prolonged drug use leads to a sensitisation of the mesolimbic system that renders the individual more responsive to the drug and drug-related cues. Together, these results suggest that following prolonged drug use, extraversion is associated with the changes experienced by the dopaminergic system (the neural basis of the BAS). Moreover, they indicate that extraversion mediates reward sensitivity to relevant stimuli. Therefore, they support the claim made by Depue and Collins (1999) that extraversion has a dopaminergic basis and, therefore, that it is represents the BAS-related trait.

Finally, similar neurobiological evidence comes from an EEG study conducted by Wacker and colleagues (2006). In this study, extraverts and introverts performed an n-back task under the effects of either a placebo or a DA antagonist (sulpiride, which act as a blocker on D_2 sites). The n-back task consisted of four levels of difficulty (0 to 3 back). During the task participants received trial-by-trial accuracy feedback and, depending on their performance, they could earn 1-Euro worth of goodies. This manipulation aimed to ensure an incentive experimental setting, i.e. a promotion regulatory focus (Higgins, 1997). The study is based on literature and previous studies indicating that EEG measures of activity in the frontal cortex and, in particular the anterior cingulate, are an index of Extraversion and they are modulated by dopaminergic activity. Indeed, midbrain DA projects to the cortex via the mesocortical DA pathway and DA mediates its functioning (Depue & Collins, 1999; Wacker, Chavanon & Stemmler, 2006).

Results indicated that under the placebo condition, extraverts had lower RT scores than introverts, especially in the hardest (3-back) version of the task. Effective performance on the task requires WM activation, which has been found to be enhanced by DA (Ashby et al, 1998). Hence, superior performance of extraverts over introverts could be due to higher DA levels and DA modulating effect on the PFC. These assumptions are supported by the results obtained in the study following administration of the DA antagonist. In fact, following sulpiride administration, introverts showed shorter response latencies than extraverts. As expected, these behavioural observations were matched by activations of the frontal cortex.

3.7 Conclusions

The studies reviewed above indicate that there is some evidence in favour of both extraversion and impulsivity as candidate BAS traits. The results of the studies, however, are not conclusive on the issue of whether it is extraversion or impulsivity (ImpAss) that constitutes the BAS-related trait. In fact, both personality traits seem to mediate reward-related performance and scores on the two scales seem to be modulated by DA activation levels. The results reviewed in this chapter indicate that different BAS scales seem to be correlated with extraversion or impulsivity. Indeed, recent evidence shows that the three BAS subscales may describe conceptually different constructs in relation to the EPQ scales. Smillie, Jackson and Dalgleish (2006) found that BAS-D and BAS-RR reflect BAS variation and, therefore, reward sensitivity whereas the BAS-FS scale was found to reflect both BAS and impulsivity. Overall, these results suggest that the BAS scales may capture two different constructs; in particular, BAS-RR and BAS-D may be purer measures of reward sensitivity whereas BAS-FS may be a broader term that captures both reward sensitivity and impulsivity. Heym, Ferguson and Lawrence (in press) have further suggested a three-factor model where each BAS subscales may capture a different constructs. They observed that EPQ-E was positively correlated with the three BAS subscales; however, the EPQ-P scale (i.e. an impulsivity measure) was found to be positively correlated with the BAS-FS scale and negatively with the BAS-D scale. These observations supports the conclusions drawn by Smillie and colleagues (2006) that BAS-D and BAS-RR are pure measures of the BAS. Additionally, Heym and colleagues suggested a further distinction between BAS-D and BAS-RR due to the positive relationship obtained between BAS-RR and the BIS scale. The researchers suggested that BAS-RR is a measure of future oriented reward sensitivity associated with worry and uncertainty (cf. BAS-D which is more responsive to immediate rewards).

These results suggest that the different BAS subscales describe reward sensitivity in a slightly different fashion. Therefore, these results render it harder to determine whether impulsivity or extraversion is the underlying BAS trait and, therefore, solve the ongoing debate. However, Smillie and colleagues (2006) state that the fact that the three BAS scales may reflect different constructs does not take away the fact that they still index BAS activation. In conclusion, it may be possible that both personality components may represent BAS-related traits and influence reward-processing and learning.

3.8 Summary

The present chapter introduced the Reinforcement Sensitivity Theory that states that individual differences are determined by inter-individual differences in the reactivity of emotional-motivational systems. There are three systems that mediate the actions of the individual in every-day life: the behavioural activation system (BAS), the behavioural inhibition system (BIS) and the Fear, Fight and Flight system (FFFS). The BAS is responsive to rewards, whereas the FFFS is responsive to punishment. The BIS is a conflict resolution system which becomes active only in ambivalent situations where both the BAS and the

FFFS are active. Hence, the activation of the three systems interacts with each other and affects the final behavioural output. The chapter mainly focused on the BAS and reviewed the studies that attempted to identify the BAS-related trait. The debate is still open, and both extraversion and impulsivity seem to be plausible BAS-traits. Some of the studies reported in the following chapter attempt to shed more light on the debate.

Chapter 4

Personality data

4.1 Chapter aims

The present chapter reviews the personality measures that were collected during the empirical studies. In each study, participants filled in a battery of questionnaires that included scales relevant to the main aim of the research. Several measures were included in order to fully capture individual variations in BAS reactivity. In this chapter, the different personality measures collected will be introduced and additionally a principal component analysis will be run in order to summarise the different measures and obtain fewer super-factors.

4.2 Introduction

The studies reported in the following sections aimed to explore the impact of individual differences on reward-related learning. According to RST, the emotional-motivation system sensitive to reward and responsible for approach behaviour to reward is the Behavioural Activation (Approach) System (BAS; Gray, 1970; Corr, 2004a; Corr, 2006; Smillie et al., 2006a). RST also suggests that variations in the BAS characterise inter-individual differences in reward sensitivity and goal approach, which are reflected in personality traits. High BAS individuals are expected to show greater sensitivity and more frequent reward approach behaviours than their low BAS counterparts. Gray (1987) originally identified impulsivity as the personality trait underlying the BAS. He postulated that impulsivity was composed by Eysenck's extraversion and neuroticism and, in particular, it ranged from high E and high N (high impulsivity) to low E and low N (low impulsivity). Therefore it was necessary to include trait impulsivity measures in the battery administered to participants.

However, as reviewed in the RST chapter there is evidence from empirical and psychometric studies that supports the original claim made by RST that impulsivity is the BAS-related trait. Indeed, the results obtained in these studies indicate that impulsivity is associated with reward sensitivity and reward-related learning. Moreover, there is evidence that indicates that impulsivity levels are positively correlated with DA levels, which is in line with RST assumption that the BAS lies on midbrain DA system and that the BAS-related trait should also have a DAergic nature (Pickering & Gray, 2001). However, as reviewed in that same chapter, there is also some contrasting evidence that indicates that extraversion, and not impulsivity, is related to reward sensitivity and learning (Pickering, 2004; Smillie et al., 2007). These data indicate that there is no conclusive empirical evidence in favour of one over the other theory (impulsivity vs. extraversion) and, therefore, the debate on which one is the BAS-related trait is still open. It is possible that both personality components may be independently associated with DA and, hence, represent different aspects of the BAS. For this reason it was decided to include measures of both the impulsivity and the extraversion components.

There are several scales that measure different elements of these two main personality traits. In order to obtain a more complete representation of each trait it is necessary to include a wide range of these scales and questionnaires.

Pickering and Gray (2001) have suggested that impulsivity is a higher-order personality component and, therefore, it is composed of several personality traits. They labelled this component Impulsivity Antisocial sensation seeking (ImpAss), which is a variation of the Impulsive Unsocialized Sensation Seeking (ImpUSS) component proposed by Zuckerman (1991; 1993)[2]. Pickering and Gray have also suggested that this component might be a better descriptor of the BAS and, hence, a better candidate as BAS-related trait. In fact, impulsivity is a complex trait component which is characterised by varied impulsive behaviours and different types of impulsivity measures (Evenden, 1999; Leshem & Glicksohn, 2007). Hence, in order to capture all the facets of impulsivity, it is necessary to consider several personality measures which are central to its definition, such as novelty-, sensation-seeking and psychoticism scales (Pickering & Gray, 2001). Zuckerman originally suggested that ImpUSS was

---

[2] Pickering (2004) preferred the term antisocial over unsocialised because the former terms does not imply that the behaviour is exclusively determined by environmental factors

strongly related to the psychoticism (EPQ-P) scale developed by Eysenck. Indeed, the ImpUSS (revised ImpAss) component contains measures of the sensation seeking, impulsivity, inhibition and EPQ-P scales (Zuckerman, 1993; Gibbons & Rammsayer, 1999). The psychoticism scale is constituted by several elements, such as 'hostility, lack of empathy and non-conformity' (Eysenck, Eysenck & Barrett, 1985, p.25).

Measures of positive schizotypy were also collected since this personality factor has been found to correlate with ImpAss scores and it plausibly also has a dopaminergic nature (Pickering, 2004; Kumari et al., 1999). Evidence to support the former claim comes from psychometric considerations. One of the best known questionnaires that measures schizotypy is the Oxford-Liverpool Inventory of Feelings and Experiences (OLIFE) scale developed by Mason and colleagues (1995). This scale contains four subscales, which are Positive Schizotypy, Negative Schizotypy, Cognitive Disorganisation and Impulsive Nonconformity. In contrast to the first three scales, as previously mentioned, the Impulsivity Nonconformity (ImpNon) one does not represent behavioural and cognitive symptoms typical of schizophrenia. This scale rather describes behaviours more typical of personality disorders (e.g. reckless behaviour), such as borderline and antisocial personality disorder (Evenden, 1999; Gray, Fernandez, Williams, Ruddle & Snowden, 2002; Pickering, 2004). Moreover, the ImpNon subscale contains seven items of the EPQ-P that is also a main measure of the ImpAss component (Gray et al., 2002). Therefore, it is not surprising that ImpNon represents a good measure of antisocial and sensation seeking attitudes (hence, ImpAss).

These observations highlight the importance of including measures of schizotypal traits (as measured by the OLIFE scale) as well as ImpAss measures since they seem to be inter-correlated. Indeed, including both measures should capture a greater variation of personality and, possibly, a more complete picture of the ImpAss component. Subsequently, this should lead to a greater understanding of the impact of personality on reward-related learning.

Evidence that indicates that positive schizotypy has a DAergic nature comes from evidence that identifies a close parallel in cognitive and attentional processes between positive schizotypy and schizophrenia. Schizotypal personality traits capture psychological and biological symptoms typical of schizophrenia across the healthy population (Mohanty, et al., 2005; Koo, et al., 2006). Schizophrenia is characterised by DAergic dysregulation and, in particular, by high levels of DA in the midbrain (Julien, 2003; Mohanty et al., 2005). Empirical evidence has indicated that both schizophrenic patients displaying psychotic positive symptoms and normal individuals who score high on positive schizotypy measures show cognitive and attentional impairment (Koo et al., 2006). Attentional impairment is characterised by the inability to 'filter out' irrelevant information to facilitate processing of relevant information and it may account for the positive symptoms typical of schizophrenia (i.e. hallucinations and delusions; Mohanty, et al., 2005; Corr, 2006). For example, studies using the latent inhibition paradigm indicate that compared to control healthy individuals, schizophrenic patients show attentional deficit (Corr, 2006). Latent inhibition is observed during stimulus-reward association learning is impaired when one of the stimuli had been previously presented with no direct consequences (Gray et al., 1992; Pickering & Gray, 2001). Following this pre-exposure, the stimulus loses salience and the individual finds it harder to learn the newly acquired (reward-predicting) valence of the stimulus in a second phase. As a direct effect of the pre-exposure phase, learning of the stimulus-stimulus association is retarded in the second (test) phase. This retardation is the phenomenon typical of LI and it is reduced in patients affected by positive symptoms of schizophrenia (Pickering and Gray, 2001; Gray et al., 2002; Corr, 2006). Blunted LI has also been observed among healthy individuals who score high on schizotypal personality trait as well as ImpAss measures (Pickering & Gray, 2001; Gray et al., 2002; Weiner, Schiller & Gaisler-Salomon, 2003). This is in line with the idea that psychotic symptoms are represented across a continuum (Gray et al., 2002). Indeed, during the test phase these individuals learn the new association faster than their low counterparts and this suggests that they perceive the stimulus as salient despite the pre-exposure phase (Pickering & Gray, 2001; Pickering, 2004).

Both animal and human studies have indicated that LI is modulated by DA activity (Kumari, et al., 1999; Corr, 2006). In particular, reduced LI in schizophrenia has been found to be associated with lower levels of selective attention which are related to higher DA levels in the mesolimbic system (Gray, Pickering, Hemsley, Dawling & Gray, 1992; Pickering & Gray, 2001; Pickering, 2004). Indeed, administration to healthy individuals of low doses of amphetamine was found to decrease the LI phenomenon compared to placebo administration and higher amphetamine doses (Gray et al., 1992). The latter results may be due to the fact that high amphetamine doses increase DA levels in the nigrostriatal system whereas the lower dose has more effect on mesolimbic DA levels; it is the mesolimbic DA system which is likely, along with mesocortical DA projections to mediate selective attention. In contrast, administration of high doses of antipsychotic and neuroleptic drugs (e.g. haloperidol), which act as blockers on D_2 receptors, normalised latent inhibition in both animal and (healthy) human studies (Kumari, et al., 1999; Weiner et al., 2003). Therefore, the fact that LI is a DA-dependent phenomenon and it seems to be affected by positive schizotypy is consistent with the proposed DAergic nature of positive schizotypy.

Moreover, studies exploring latent inhibition (LI) have also supported the claim that ImpAss and positive schizotypy are inter-correlated scales as they independently mediated performance on the task (Pickering & Gray, 2001; Gray et al., 2002). For example, a study carried out by Gray and colleagues (2002) showed that high scores on ImpAss and positive schizotypy components (of the OLIFE questionnaire, Mason, Claridge & Jackson, 1995) were positively associated with reduced LI. In this study, positive schizotypy was measured by the OLIFE scale known as Unusual Experiences. The ImpAss measure constituted scores on the Impulsivity Nonconformity OLIFE scale that captures behaviours related to impulsivity and risk taking. Hence, the effect of ImpNon on LI may be caused by its close association with ImpAss measures (i.e. sensation seeking, disinhibition and psychoticism) which have been found to have a blunting effect on this phenomenon (Gibbons & Rammsayer, 1999). However, further analysis also indicated the contribution of measures of positive schizotypy (i.e. schizotypal traits) on LI was independent from the contribution of ImpAss measures; whereas the ImpAss contribution was not additive to that of schizotypal personality traits. It is possible that the relationship observed between LI and ImpAss may be spurious and, hence, purely due to the fact that ImpAss measures are significantly inter-correlated with schizotypal personality traits.

Measures of trait Anxiety were also included in this thesis because, as indicated by the revised version of RST, the personality traits underlying the reward and punishment systems, although orthogonal, interact when determining behaviour (Corr, 2004b; Smillie et al., 2006a). Experimental and psychometric data have also indicated that the personality traits reflecting the level of activation of the reward and punishment systems are correlated (Zinbarg & Revelle, 1989; Carver & White, 1994). Similarly, formal and neural modelling has indicated the BAS output is not only mediated by the BAS activation but also by the activation of the punishment systems (Pickering, 2008). Therefore, these data indicate the importance of including anxiety measures even in studies that manipulate reward and aim to explore the BAS activation and reactivity. For example, Neuroticism scales were administered to participants as part of the questionnaire battery since N scores have been found to load onto the BIS/FFFS component (Smillie & Jackson, 2006; Franken & Muris, 2006).

4.2.1 Aim

The PCA analysis was run in order to summarise the various personality measures collected in the several studies and extract the four main personality components that have been discussed above. These measures are: extraversion and ImpAss, which may be related to the BAS activation, positive schizotypy, as a dopaminergic trait dimension, and neuroticism which may be related to the BIS/FFFS activation.

4.3 Method

4.3.1 Personality measures

Six personality questionnaires were collected during the experiments reported in the following sections. The questionnaires used are briefly reported below.

Big Five Inventory (BFI; John, Donohue & Kentle, 1991)

Table 4.1. BFI subscales

Big Five Inventory (BFI)	
Extraversion	**b5e**
Neuroticism	**b5n**
Conscientiousness	**b5c**
Agreeableness	**b5a**
Opennes	**b5o**

The Big Five Inventory (BFI) consists of 44 items scored on 5-point Likert scales and it comprises five subscales that correspond to five personality traits (table 4.1).

The extraversion subscale is characterised by traits such as activity and energy, dominance, sociability and positive emotions. The agreeableness scale represents a pro-social attitude towards those with a spirit of antagonism. This subscale includes traits such as altruism, trust and modesty. Conscientiousness describes the impulse control that facilitates goal-directed behaviour and it is characterised by traits such as organisation and good planning skills. Neuroticism is the opposite of emotional stability and it is characterised by traits such as anxiety, sadness and irritability. Finally, openness represents the breadth and depth of people's mental life. Thus, this dimension includes having wide interests, and being imaginative and insightful.

Behavioural inhibition/activation scales (BIS/BAS; Carver & White, 1994).

Table 4.2. BIS/BAS subscales

Behavioural inhibition/activation scales (BIS/BAS)	
BAS - Reward Responsiveness	BAS-RR
BAS – Fun Seeking	BAS-FS
BAS - Drive	BAS-D
BAS – subscales total score	BAS - Sum
BIS	BIS

The BIS/BAS Scales consist of 20 self-administered questions scored on 5-point Likert scales. The BIS scale includes seven items assessing anxiety about negative external events (e.g., ''I feel worried when I think I have done poorly at something''). The BAS scale includes 13 items, which are subdivided into Drive (4 items, e.g., ''I go out of my way to get things I want''), Reward Responsiveness (5 items, e.g., ''When I get something I want, I feel excited and energized''), and Fun Seeking (4 items, e.g., ''I crave excitement and new sensations''; table 4.2). A total score can also be calculated by adding the individual scores from the three subscales.

Eysenck Personality Questionnaire – Revised (EQP-R; Eysenck, Eysenck & Barrett, 1985)

Table 4.3. EPQ subscales

Eysenck Personality Questionnaire - Revised (EQP-R)	
Extraversion	EPQ-E
Neuroticism	EPQ-N
Psychoticism	EPQ-P

The EPQ-R is a 100 item scale containing three personality subscales: Neuroticism (EPQN), Introversion-Extraversion (EPQE), and Psychoticism (EPQP; table 4.3). The EPQ-R also contains a validity scale, the Lie Scale (EPQL). Extraversion (24 items) is characterized by being outgoing, talkative and in need of external stimulation. Psychoticism (32 items) was initially proposed as a measure of the liability to the psychosis spectrum but contains items resembling certain personality disorders (antisocial, borderline and schizoid). It is characterised by traits such as aggression, non-conformity, inconsideration, anger and impulsiveness. Neuroticism (24 items) is a measure of negative trait emotionality, which is characterised by high levels of negative affect such as depression and anxiety. Moreover, it is supposed to be associated with the activation of the autonomic nervous system (Phillips, Gossop & Bradley, 1986). Finally, the Lie Scale (20 items) is considered a measure of social conformity. Participants are invited to express their opinion by responding either yes or no to the questionnaires item whether they believe the item describes their behaviour or not.

Sensation seeking scale V (SSS; Zuckerman, 1979)

Table 4.4. SSS subscales

Sensation seeking scale (SSS)	
Thrill and Adventure Seeking	TAS
Experience Seeking	ES
Disinhibition	D
Boredom Susceptibility	BS
SSS - subscales total score	SSS - tot

The Sensation Seeking Scale measure four subscales: Boredom Susceptibility (BS); the Thrill and Adventure Seeking (TAS); Experience Seeking (ES); and Disinhibition (Dis) scales. The scale contains 40 items, 10 for each subscale. Participants are invited for each item to select one of the two options presented that they consider applies to their way of thinking. Zuckerman (1994) defines sensation seeking as a trait describing the tendency to seek novel and intense sensations and experiences and the willingness to take risks for the sake of such experience. A total score can also be calculated by summing the scores from the four individual subscales (table 4.4).

Oxford-Liverpool Inventory of Feelings and Experiences Scale (OLIFE; Mason, Claridge & Jackson, 1995)

Table 4.5. OLIFE subscales

Oxford-Liverpool Inventory of Feelings and Experiences Scale (OLIFE)	
Unusual Experiences	UnEx
Cognitive Disorganisation	CogDis
Introvertive Anhedonia	Intan
Impulsive Nonconformity	ImpNon

The OLIFE is a broad measure of schizotypal personality. It contains 159 items which require participants to express whether they agree or disagree with each items by responding either yes or no. It contains four personality factors that closely correspond to the symptoms reported by patients affected by schizophrenia (table 4.5). Specifically, these factors relate to positive reality distortion symptoms, negative psychomotor poverty symptoms and positive thought disorder symptoms. The scales yielded by the OLIFE are: unusual experiences (positive schizotypy), cognitive disorganisation (disorganised schizotypy) and introverted anhedonia (negative schizotypy). In addition the OLIFE yields a fourth factor that consists of impulsive and antisocial/nonconformist character traits, which closely resembles Eysenck's psychoticism. The scale is referred to as impulsive nonconformity. As already noted, the content of this scale is related to particular personality disorders rather than schizophrenia *per se*, leading some to argue that this scale is not a measure of schizotypal personality but rather a factor of the ImpAss component (e.g. Pickering, 2004). The questionnaire also includes the social desirability (i.e. LIE) and Extraversion scales of the Eysenck Personality Questionnaire as well as the Schizotypal Personality Scale (STA) developed by Clardige and Broks (1984; Mason et al., 1995; Avons, Nunn, Chan & Armstrong, 2003).

Schizotypal Personality Questionnaire (SPQ; Raine, 1991)

Table 4.6. SPQ subscales

Schizotypal Personality Questionnaire (SPQ)		
Subscales	**Factors**	
Ideas of Reference	**Cognitive-Perceptual (CogPer)**	**IoR**
Odd beliefs/Magical thinking		**OddBel**
Suspiciousness		**Susp**
Unusual Perceptual Experiences		**UPE**
Excessive Social Anxiety	**Interpersonal**	**ESA**
No Close Friends		**NCF**
Constricted Affect		**CA**
Suspiciousness		**Susp**
Odd Behaviour	**Disorganised**	**OddBeh**
Odd Speech		**OddSp**
Total score		**SPQ**

The SPQ is a self-report scale modelled on DSM-III-R criteria for schizotypal personality disorder and containing subscales for all nine schizotypal traits. The SPQ consists of 74 items. Participants are requested to express whether they agree with the items by responding yes or no. There are nine separate subscales, each reflecting a DSM-III-R criterion: ideas of reference, excessive social anxiety, odd beliefs or magical thinking, unusual perceptual experiences, odd or eccentric behaviour, no close friends, odd speech, constricted affect, and suspiciousness (table 4.6). Additionally, these nine scales can be further classified into three factors, which are: Cognitive-Perceptual, Interpersonal and Disorganised. As can be seen in table 4.6, suspiciousness loads onto both the intepersonal and the Cognitive-Perceptual factor. These three factors make broad reference to three of the OLIFE scales: Unusual Experiences, Introvertive Anhedonia and Cognitive Disorganisation. Finally, it is also possible to obtain an overall score of schizotypy (SPQ tot).

4.3.2 Data analysis

A principal component analysis (PCA) with Varimax rotation was carried out as an exploratory tool in order to obtain a summary of the personality data collected over the several studies. The use of the PCA aimed to reduce the observed personality variables into a limited number of factors which are more easily interpreted. Reducing the number of personality factors also decreases the number of potential follow-up comparisons and, therefore, reduces the possibility of incurring type-I errors.

The choice of adopting a Varimax solution was driven by the intention of simplifying the solution and the structure of the extracted components and their discernability. In fact, following a Varimax rotation each scale tends to be associated with one of the extracted components. Overall, the presence of orthogonal factors renders interpretation of results easier in the analyses which are carried out. The regression analysis maximum power is obtained by having orthogonal predictors.

Several exploratory PCA were run in order to extract the more meaningful and complete trait components. In particular, these exploratory analyses included the three BAS subscales as well as subscales of the SSS. However, results showed that none of these subscales loaded strongly on any of the extracted components. Following these observations, it was decided to enter their total scores (i.e. BAS sum and SSS total), which gave a better solution (i.e. with fewer complex variables). The Chronbach's alphas for the total scales were 0.63 and 0.61 for the BAS and SSS scales, respectively. These alpha values suggest reasonable internal consistency of the scales and, therefore, reasonable communality among the subscales.

Fifteen personality scales were entered in the PCA which were: BFI extraversion (b5e) and neuroticism (b5n) subscales, EPQ extraversion (E), psychoticism (P) and neuroticism (N) subscales, SPQ cognitive/perceptual (CogPer) and disorganised factors, the OLIFE unusual experiences (unex), cognitive disorganisation (CogDis), impulsive/non conformity (ImpNon), introvertive anhedonia (intan) and extraversion (ext) subscales, the BIS scale, a total score of the sensation seeking (SSS_tot) and the BAS scales (BASsum). The Chonbach's alpha reliability values for the scales entered in the PCA are reported in table 4.7.

Table 4. 7. Chronbach's alpha reliability values for the fifteen scales entered in the PCA

Scale	Alpha score
EPQ- Extraversion	0.86
EPQ- Neuroticism	0.88
EPQ- Psychoticism	0.72
Big five –Extraversion	0.86
Big five - Neuroticism	0.85
BIS	0.79
BAS total	0.63
SPQ – Cognitive perceptual	0.80
SPQ – Disorganised	0.85
OLIFE – Unusual Experiences	0.91
OLIFE – Cognitive Disorganisation	0.87
Impulsive Nonconformity	0.72
OLIFE – Introvertive Anhedonia	0.77
OLIFE Extraversion	0.86
SSS total	0.61

4.4 Results

Data were collected from a total of 232 participants. However, five participants took part in more than one of the studies. Data from these individuals were identified and their scores on the two sessions were replaced by the average score across the two sessions. The PCA analysis was run on a data set that consisted of 227 cases.

The distribution of these scales was not very far from a normal distribution and even a small variation has very little effect on the component solution offered by the PCA. Additionally, normality is not a requirement when PCA is used to summarise a large set of variables into a smaller number of factors (as in the present case; Tabachnick & Fidell, 2001).

A few outliers were identified on a few of the scales: BIS, BASsum, EPQ-E, SPQ: cognitive perceptual, OLIFE:unex, OLIFE:intan, OLIFE:ext and SSS_tot. However, when comparing the original mean values to the 5% trimmed mean values they did not seem to be significantly different, which suggests that the more extreme values did not have a great impact on the mean, so they were left in the analysis (Pallant, 2001). Additionally, personality traits extend over a continuum and individuals are expected to score along the whole spectrum even though only a few individuals are expected to score near the extremes.

Prior to performing the principal component analysis, the suitability of the dataset for the analysis was assessed (Tabachnick & Fidell, 2001). Inspection of the correlation matrix showed that there were several correlation coefficients higher than .3 (see appendix 1). In fact 35% of the total correlations showed a correlation coefficient of .3 and above. Moreover the KMO test value was .82 (at a cutoff of 0.6) and the Barlett's test of Sphericity was significant ($p < .001$) which indicated that the data was suitable for factorising.

PCA identified four factors with an eigenvalue higher than one (4.46, 3.78, 1.78 and 1.06 respectively). The scree plot also showed a clear break after the second and the fourth components (figure 4.1).

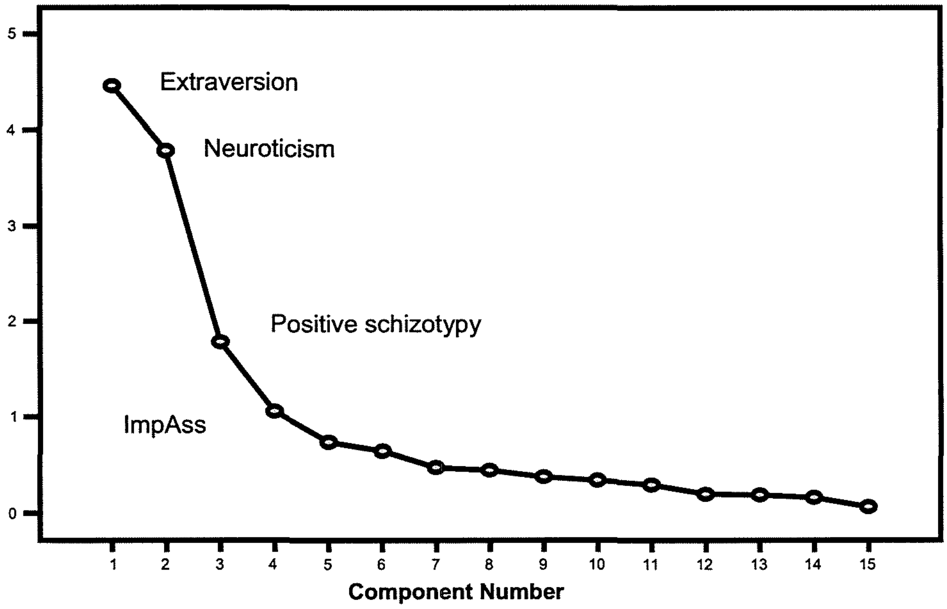


Figure 4.1. Scree plot

The Varimax rotation (table 4.7) revealed the presence of a simple structure. In fact, all components showed a number of strong loadings and all variables loaded largely on only one component. The four factors explained 73.8% of the total variance, with the first component (i.e. extraversion) contributing 29.7%, the second component (i.e. neuroticism) 25.2%, the third component (positive schizotypy) 11.9% and the fourth component (i.e. ImpAss) 7.0% to the total variance. The extracted components resemble the expected factors.

By considering the variables that loaded into each component, the four extracted components can be identified as measures of the four main personality components discussed in the introduction, which are: extraversion, neuroticism (anxiety), positive schizotypy (PS) and impulsive antisocial sensation seeking traits (ImpASS), respectively (table 4.8).

Table 4.8. Loading of the 15 scales on the 4 components following a Varimax rotation

Rotated Component Matrix

	Component			
	E	N	PS	ImpAss
b5 - extraversion	.829			
epq_e	.903	-.242		
olife-intan	-.749		.293	
olife-ext	.926			
bassum	.588	.229		
b5 - neuroticism		.838		
epq_n		.824	.285	
bis		.855		
olife-cogdis	-.217	.684	.310	
spq-cognitive/perceptual			.900	
spq-disorganised		.293	.687	.361
olife-unex			.907	
epq_p			.255	.775
olife-impnon	.290		.223	.571
sss_tot	.287			.755

* Loadings below 0.2 are omitted

4.5 Summary

A principal component analysis was run on relevant personality scales in order to extract four personality components to use in the analysis of the data obtained in the empirical studies that are reported in the following chapters. The PCA was implemented in order to reduce the number of personality measures collected during the experimental sessions. This was done to reduce multiple testing and therefore reduce the chance of incurring type-I errors. Therefore, the extracted four components are used in the following chapters to explore the relationship between performance and personality individual differences.

The PCA aimed to extract personality traits that affect BAS activation and, therefore, may represent the system underlying traits. For this reason, both extraversion and ImpAss components were extracted since the debate is still on-going on which one of the two is the BAS-related trait. Additionally, a positive schizotypy component was also extracted because evidence shows that there is some overlap between ImpAss and positive schizotypy measures. Finally, a component representing the anxiety personality component was also extracted due to evidence that BAS output may also be mediated by the activation of punishment systems. The anxiety component is labelled neuroticism since N measures loaded strongly onto this component.

The PCA results show that four scales load onto more than one component, which are: the BASsum, the Olife-cogdis, the OLIFE-impnon and the SPQ-disorganised scales. Another PCA was run after excluding these scales and its results showed a cleaner components' solution (table 4.9).

Table 4.9. Loading of the 11 scales following exclusion of those scales with multiple loadings

Rotated Component Matrix

	Component			
	1	2	3	4
b5 - extraversion	.846			
epq_e	.924			
olife-intan	-.775		.283	
olife-ext	.943			
b5 - neuroticism	-.219	.871		
epq_n		.819	.264	
bis		.856		
spq-cognitive/perceptual			.919	
olife-unex			.917	
epq_p			.273	.795
sss_tot	.306			.797

* Loadings below 0.2 are omitted

A correlation was run between the original components and those components extracted by the PCA with reduced scales in order to establish their relationship. Results showed that the components were strongly correlated (Rs => 0.96; table 4.9). These results indicate that the two PCAs extract roughly identical personality components.

Table 4.10. Correlations between the components extracted in the original PCA (e, n, p and i) and the 'reduced' PCA (e_reduced, n_reduced, p_reduced and i_reduced)

Correlations

		e_reduced	n_reduced	p_reduced	i_reduced
e	Pearson Correlation	.983	.041	.026	-.003
	Sig. (2-tailed)	.000	.535	.695	.966
	N	227	227	227	227
n	Pearson Correlation	-.064	.979	-.023	-.022
	Sig. (2-tailed)	.336	.000	.731	.743
	N	227	227	227	227
p	Pearson Correlation	-.022	.006	.983	.000
	Sig. (2-tailed)	.738	.929	.000	.995
	N	227	227	227	227
i	Pearson Correlation	-.006	-.004	-.037	.967
	Sig. (2-tailed)	.925	.953	.581	.000
	N	227	227	227	227

Following these observations, it was then decided to retain the four components extracted in the original PCA since its solution offers a more exhaustive summary of the personality traits that are of interest in the thesis. Moreover, the multi-loading scales loaded much more strongly onto one component (i.e. the predicted components), as most of the secondary loadings had loading coefficients lower than 0.3.

The component solution obtained from the PCA shows that EPQ-P loads onto the ImpAss component rather than on the positive schizotypy component. These results are counterintuitive to Eysenck's main assumption that P is a measure of the psychoticism spectrum (Eysenck, Eysenck & Barrett, 1985). Nonetheless, the present results are in line with previous factor-analytical studies that observed that P failed to load onto psychotic-like components (Bentall, Claridge & Slade, 1989; Mason et al., 1995; Claridge et al., 1996). Bentall and colleagues (1989) carried out one PCA on the four EPQ scales and ten schizotypy scales. The PCA extracted three components, one component captured the positive symptoms typical of schizotypy, the second component captured the negative symptoms whereas the third factor that indexed a mixture of social anxiety and cognitive disorganisation. The EPQ-P scale was found to load on the first component (i.e. positive schizotypy). Subsequently, these researchers carried out a second PCA on the same schizotypy trait scales and four schizophrenic symptoms scales. The PCA extracted a four-factor solution. The EPQ-P scale together with Chapman's hypomanic personality scale loaded on the forth factor which resembles disinhibited and antisocial-like behaviours. In a follow-up study, Claridge and colleagues (1996) extracted the same four-factor solution despite the fact that the four symptom scales were not included in the factor analysis. Vollema and van den Bosch (1995) reviewed nine factor analytical studies and they observed that, overall, these studies obtained the same 4-factor solution described above. In particular, the EPQ-P scale was found to load highly on a component that measures impulsivity and asocial ideas, which the authors labelled 'nonconformity'. This 'nonconformity' component captures behaviours that are more typical of patients with personality disorders (e.g. antisocial personality disorder) than those of schizophrenics (Pickering, 2004). Overall, these results and considerations are in line with the present results, since they suggest that the EPQ-P scale is a better measure of antisocial and disinhibited behaviours (as indexed by the ImpAss component) rather than positive schizotypy.

Chapter 5

Impulsivity and delay discounting

5.1 Abstract

There is evidence that indicates that reward sensitivity (indexed by extraversion, psychoticism and other possible BAS measures) leads to disinhibited responding during performance on passive avoidance learning tasks (Patterson and Newman, 1993; Avila, 2001). Hence, the present study aimed to explore individual differences during a decision-making task that assessed both reward delay discounting and disinhibition. The task is a variation of the task implemented by Newman and colleagues (1992) where participants had to choose to respond to one of two stimuli in order to win money. The choice is between a stimulus that offers frequent reward (80% of responses, high reward stimuli) but is presented following a 10-second delay and a stimulus that offers more infrequent reward (40%, low reward stimuli) but allows immediate responding. Results indicated that there were two opposing processes in action and they were mediated by different aspects of impulsivity. In particular, one form of impulsivity was found to moderate disinhibited responding in response to low reward (LR) stimuli (i.e. proportion of LR responses) whereas reward-related impulsivity affected responding in relation to high reward (HR) responding (i.e. RT to HR stimuli). Results are discussed in light with this observation. The findings support the view that impulsivity is a multi-dimensional factor characterised by at least two major subtypes; these bear some resemblance between reward-related and rash impulsivity (Dawe, Gullo & Loxton, 2004), which differently affected performance. Results are discussed in light of this observation.

5.2 Introduction

Behavioural (motor) impulsivity is characterised by the individual's inability to inhibit responding even though responding may not be advantageous. Inhibition has been studied by using behavioural tasks with a go/no-go or reversal learning paradigm that requires passive avoidance learning (Arce & Santisteban, 2006). During a passive avoidance learning task, the individual needs to learn to withhold a response that had previously been associated with reward but it is subsequently associated with punishment, extinction or contingency reversal (Patterson, Kosson & Newman, 1987; Patterson & Newman, 1993). In other words, during performance on a passive avoidance task, the individual needs to interrupt ongoing (dominant and/or rewarded) behaviour in order to avoid punishment and learn a new response strategy; s/he can do so by paying attention to feedback signals. Failure to do so results in perseverative responding and, hence, experiencing more punishment or non-reward signals.

Patterson and Newman (1993) developed a four-stage model to explain disinhibited behaviour. In the first stage, performance starts and the individual has to establish a stimulus-response association (i.e. dominant response). However, during the second stage contingencies change and so hinder ongoing responding; this, therefore, requires a behavioural change to cope with the new task contingencies (i.e. third phase; during the fourth phase, participants implement, or not, the new strategy). The authors suggested that coping with the task's new requirement is dependent on the individual's inhibition levels. Indeed, disinhibited individuals are less likely to change their responding pattern than their inhibited counterparts, especially if the reward-related cues remain present. Moreover, they suggested that disinhibition is associated with high reward and low punishment sensitivity. In contrast, inhibited individuals are more likely to interrupt dominant responding to modify their behaviour according to the new contingencies' requirements (i.e. they show response modulation).

Several studies explored the impact of "disinhibited" personality traits on inhibition; they have indicated that extraverts and psychopaths have lower levels of inhibition and show impaired performance during reversal learning tasks (Patterson & Newman, 1993; Avila, 2001). For example, Newman and colleagues (1985) investigated passive avoidance during a go/no-go task. During the task participants were presented with six go (S_+) and six no-go (S_-) cues, and they had to learn to make or withhold a response depending on the predictive cue that appeared on the screen. Participants earned money for correct responses and lost money for incorrect ones whereas no money was earned or lost when no response was made. They found that extraverts showed a deficit in inhibition and, as a consequence, they committed more errors of commission than introverts[3]. In other words, extraverts found it hard to withhold a response that was strongly associated with reward even in the presence of no-go cues. The same paradigm was used to compare performance between individuals who scored high and low on psychoticism measures. Results indicated that high scorers on the psychoticism measures made more passive avoidance errors (PA; i.e. errors of commission) than their low counterparts (Newman, Widon & Nathan, 1985).

These findings were further replicated in a study conducted by Patterson and colleagues (1987). They had participants perform on a variation of the go/no-go task used in the study by Newman and colleagues (1985). In a first experiment, these researchers also found that extraverts committed more PA errors than introverts. In a second experiment, they had participants perform on the go/no-go task in one of two conditions that varied the length of the feedback interval. In one condition, the feedback was presented for five seconds (fixed-feedback interval). In the other condition, participants had control over the feedback presentation time (flexible-feedback interval). Hence, in the latter condition, they could move onto the next trial before the five seconds had expired. Results showed that the number of PA errors committed by individuals scoring high and low on the extraversion scale differed in the flexible feedback condition only. Even though there was a general tendency to initiate the next trial before the five seconds had expired, extraverts waited less than introverts to move onto the next trial. Moreover, this difference was also moderated by scores on the neuroticism scale (i.e. there was a significant E x N interaction). Indeed, neurotic extraverts made significantly more PA errors than neurotic introverts. No difference was observed between stable introverts and stable extraverts. These results suggest

---

[3] Pickering (2004) noted that it is a shame that analyses have focused on errors of commission to index inhibition deficits. He suggested that application of signal detection theory (SDT) models to the behaviour would be preferable and more informative. In fact, the pattern of results can change dramatically when analyses are run using SDT models rather than the error scores (see Pickering, 2004 for an example)

that extraverts and, in particular, neurotic extraverts spend less time processing feedback. This may account for the greater response perseveration observed in decision-making tasks.

Moreover, a study by Newman and colleagues (1992) looked at inhibition in psychopathic and non-psychopathic offenders performing on a decision-making task where participants had to learn to inhibit immediate responding. During the task, participants were presented with two boxes on the screen and they had to decide which one they wanted to respond to in order to win money. Responding to one box led to more frequent winning than the other (high vs. low frequency reward). Participants could respond to the low frequency reward box straight away whereas they had to wait ten seconds to be able to respond to the high frequency box. There were two feedback conditions. There was a reward only condition, where participants won money in the rewarded trials, and a mixed-incentive condition, where participants earned money in the rewarded trials but lost money on the other trials. There was also a control condition where participants had to wait a 10-second delay before they could respond to either box. The experimental condition is a measure of both response inhibition (i.e. passive avoidance) and delay gratification (i.e. delay discounting). Indeed, in the experimental condition participants have to learn to withhold an immediate response in favour of a delayed response which is associated with long-run winning (Newman, Kosson & Patterson, 1992). In a typical delay discounting task, participants are presented with the choice between an immediate but small reward and a delayed but great reward (Monterosso & Ainslie, 1999; Arce & Santisteban, 2006).

Overall, the results showed that, compared to performance on the control condition, psychopaths, performing under the mixed-incentive condition, were less likely to inhibit immediate responding. This led to infrequent reward relative to the delayed response that offered more frequent reward. This effect was significantly moderated by levels of anxiety. In fact, it was low anxious psychopaths that were found to show a deficit in inhibition compared to low anxious controls. However, there was no significant difference in performance between these two groups during performance on the reward-only condition.

It has been observed that there seems to be some fairly close similarities between the inhibition model suggested by Patterson and Newman and the Reinforcement Sensitivity Theory (RST; Patterson & Newman, 1993; Avila, 2001). Avila (2001) suggested that inhibition can be explained in terms of the RST applied to the inhibition model. In particular, he suggested that the type of response produced in the third stage of the model may be mediated by inter-individual variation on the reward and punishment systems. Avila refers to the BIS when discussing the punishment system; however according to the revised version of RST, the punishment system corresponds to the fight-flight-freeze system (FFFS) while the BIS only becomes active in conflict situations (e.g. mixed incentive conditions; McNaughton & Corr, 2004; Smillie et al., 2006a). Avila (2001) further suggested that disinhibition, as observed in the experiments reported above, would be expected in individuals with a hyperactive BAS and a weak BIS (weak FFFS according to the revised version of the reinforcement sensitivity theory). Indeed, high BAS individuals are more prone to pursue pleasure, and experience positive affect, in situations when they are presented with reward stimuli. Therefore, according to the disinhibition model they should be less likely to show passive avoidance even after contingencies have changed. Variations in the reactivity of the BAS determine the individual's sensitivity to reward and propensity to approach behaviour; a highly reactive BAS leads to disinhibited behaviour.

Indeed, high BAS individuals were found to show impaired passive avoidance during procedural learning tasks (Avila, 2001). High BAS individuals showed response perseveration during passive avoidance tasks despite the fact that the stimulus-reward association had changed. Indeed, in a study conducted by Avila and Parcet (2000) participants performed on a procedural learning task where they had to learn to press one of two buttons in order to earn points. Responses to one button led to continuous winnings while responses to the other led to partial winnings. In a second phase, the dominant response underwent extinction. BAS measures were obtained by administering the reward sensitivity scale of the sensitivity to punishment and sensitivity to reward questionnaire (SPSRQ; Torrubia et al., 1995). Results showed that high BAS individuals were more likely to respond to the continuous reward button in the first phase of the task. Moreover, they were less likely to inhibit a dominant response that had been previously associated with reward, in spite of extinction.

Similar results were obtained in other studies conducted by Avila which looked at the relationship between BAS scores and response rate during a continuous odd-even discrimination task (Avila, 2001). Participants earned points for correct responses and lost points for incorrect ones. The amount of points lost was constant whereas the amount of money earned was proportional to the RT. Results indicated that individuals with high BAS scores showed faster responding than low BAS individuals. In a second phase of the discrimination task (punishment phase), participants were informed that any response emitted in the presence of a punishment cue (i.e. a red circle presented together with the test stimuli) would result in losing 50% of the point total. High BAS individuals failed to inhibit their responding when the punishment cue appeared on the screen.

Data collected on Parkinson Disease patients (PD) showed that administration of L-Dopa medication increased impulsive responding during a decision making task, similar to the one employed by Newman and colleagues (1992), to explore disinhibition in decision making (Cools, Barker & Sahakian, 2003). They compared PD patients on and off medication with healthy controls. Participants were presented with ten red or blue boxes (ratio of red to blue boxes varied over trials) and they had to decide whether a yellow token was hidden under the blue or red boxes. Participants had to accompany each decision with a bet. Depending on the outcome of their bet they would either win or lose the amount of points they had gambled. The task had two betting conditions, ascending and descending. In the ascending condition, the amount of points that participants could bet increased in a stepwise fashion (5, 25, 50, 75 and 95% of the total points) every five seconds. Participants had to press the 'bet' response to indicate that they were prepared to bet that amount of points. The opposite pattern was true in the descending condition. In other words, participants had to withhold from making the 'bet' response, if they intended to bet a large amount of points in the ascending bet condition, and also if they wanted to bet a small amount of points in the descending bet condition. Results showed that patients on medication showed more disinhibited responding. Indeed, they made smaller bets in the ascending condition and greater ones in the descending condition compared to patients off medication and controls, the last two groups showed similar patterns. However, results also indicated that patients on medication had accuracy levels equivalent to the ones obtained by the other two groups.

PD is characterised by decreased levels of DA neurotransmission. In particular, in the early stages of the disease, DA activity is strongly depleted in the nigrostriatal system whereas DA depletion in the mesolimbic and mesocortical systems is more typical of later stages of the disease (Pinel, 2002, Cools, 2006). Hence, the researchers suggested that administration of the DA precursor (L-Dopa) would have a detrimental effect on performance of the ascending/descending decision making task. This was due to the fact that L-Dopa led to DA over-activity in the intact mesolimbic system and, subsequently, impulsive/disinhibited-like responding. Indeed, the mesolimbic system is associated with reward-related learning as well as motor performance, which are both necessary aspects of the behavioural approach (BAS) system in the presence of reward (Patterson & Newman, 1993; Schultz, 1998; Avila, 2000). According to the four-stage inhibition model (Patterson & Newman, 1993), higher reward sensitivity (i.e. high BAS reactivity) should lead to greater disinhibited behaviour (i.e. more responding).

The studies reported above, which indicated that extraverts or psychopaths show disinhibited behaviour during decision-making tasks, can also be explained according to the RST model. Patterson and Newman (1993) suggested that increased levels of disinhibition are greater in individuals who are sensitive to reward. Indeed, there is evidence indicating that extraversion is the underlying trait of the behavioural activation system (BAS); and thus extraversion is sensitive to stimuli associated with reward (Depue & Collins, 1999; Pickering, 2004; Cohen, et al., 2005; Smillie et al., 2007). According to these considerations it is not surprising that extraversion scores are positively associated with performance impairments during reward-related decision making tasks that require disinhibition (e.g. reversal learning). As previously observed, the impact of extraversion on inhibition is moderated by anxiety/neuroticism scores. Again, this is not surprising since many studies used mixed incentive payoff matrices which activate both the reward and punishment system, and the conflict resolution (BIS) system as well. Hence, the final BAS output would be a by-product of the reactivity of these systems rather than the BAS alone (Corr, 2002b; Smillie et al., 2006a; Pickering, 2008).

Similarly, the study by Newman and colleagues (1992) indicated that psychoticism moderates inhibition in the same fashion as extraversion. This is also predicted by RST since psychoticism is captured by measures such as the EPQ-P scale which is one of the main scales that loads onto the ImpAss component (Zuckerman, 1993; Pickering, 2004). ImpAss traits have been identified as a possible candidate for the BAS trait. Gray originally identified impulsivity as the underlying personality trait of the BAS (Gray, 1970; Pickering, 2004; Corr, 2006) and there is empirical evidence that supports this claim (e.g. Stuettgen et al., 2005; Franken & Muris, 2006; Pardo, Aguilar, Molinuevo & Torrubia, 2007; see chapter 3 for a review).

Overall, these considerations and Avila's studies (Avila & Parcet, 2001; Avila, 2001) suggest that inhibition tasks can be a useful tool to explore the BAS and identify its underlying traits. However, the results reported above in this chapter are still not conclusive in identifying the BAS-related personality trait(s). This is due to the fact that the inhibition paradigm has not been used directly to assess the RST assumptions and, subsequently, the methodology implemented by some of the studies reviewed in this chapter may have not been ideal. For example, most studies used mixed-incentive matrices and they did not fully measure and analyse personality measures in relation to the BAS.

In the present study, it was decided to explore BAS reactivity (indexed using several possible BAS-related trait measures) by using the same decision-making task adopted in the study by Newman and colleagues (1992). Hence, it allows one to investigate the relationship between individual differences and reward sensitivity. The task adopted in the study by Newman and colleagues (1992) not only looked at inhibition but also at delay discounting, which is a measure of how quickly individuals discount a reward as a function of its delay in time

(Monterosso and Ainslie, 1999; Bickel, et al., 1999). The delay discounting task is generally considered to be a behavioural measure of impulsivity and greater discounting is observed in individuals who score high on impulsivity (ImpAss measures; Mitchell, 2004; Field, Santarcangelo, Sumnall, Goudie & Cole, 2006). In fact, impulsive individuals are more likely to prefer a small (less frequent) but immediate reward over a larger (more frequent) but delayed one (Monterosso & Ainslie, 1999; Kalenscher & Pennartz, 2008).

Smokers have been found to show low levels of inhibition during response inhibition, gambling and delay discounting tasks (Grant, Contoreggi & London, 2000; Powell, Sawkins & Davis, 2002). There is wide evidence that indicates that smokers as well as alcohol and drug users show greater discounting for monetary rewards compared to matched controls (Vuchinic & Simpson, 1998; Mitchell, 1999; Kirby, Petry & Bickel, 1999; Field et al., 2006). Abstinent smokers have also been found to show steeper discounting of delayed cigarettes than monetary rewards (Field et al., 2006; 2007). Bickel and colleagues (1999) also found that current smokers showed steeper delay discounting than ex-smokers and never smokers. Additionally, both heavy and light smokers showed greater delay discounting compared to never-smokers (Johnson, Bickel & Baker, 2007). These results indicate that smokers show disinhibited behaviour typical of impulsive individuals (i.e. high "rash" impulsiveness and reward sensitivity; Mitchell, 1999; Dawe & Loxton, 2004). Similar results have been observed during performance on the IOWA gambling task when substance users were found to prefer stimuli that led to immediate, large rewards but determined long-run losses compared to stimuli that led to long-run winnings in spite of more modest immediate rewards (Dawe et al., 2004). Similarly, heavy drinking adolescents were found to show steeper delay discounting of both monetary and alcohol-related rewards compared to light drinking adolescents (Field et al., 2007). Grant and colleagues (2000) suggested that this could be related to the user's inability to inhibit drug use (immediate reward) in favour of long-term well-being (reward). Hence, smoking status was also recorded in the present study since it seems to mediate disinhibition and, possibly, BAS activity.

5.2.1 Aims of the present study

According to the literature reviewed above, it was decided to use an inhibition paradigm to explore BAS activation and identify the BAS-related trait. It was decided to use the same task adopted by Newman and colleagues (1992) since the task explores both inhibition and delay discounting, which may be both mediated by BAS outputs. Only the control and the reward-only conditions, used in the original study, were included in the present study. This choice was determined by the main aim of the study to identify the BAS trait. Indeed, the mixed-matrix condition may activate both the reward and punishment systems and, therefore, it would make it harder to assess the BAS output and discern the pure BAS trait (Corr, 2004b)[4]. According to previous findings, it was predicted to find: i) more inhibited responding during the experimental condition than the control one; ii) high BAS individuals to show greater sensitivity to reward and, according to the four-stage inhibition model, to show more disinhibited responding; iii) BAS-mediated behaviour to be associated with scores on the extraversion or the ImpAss personality component; and iv) smoking status was expected to modulate disinhibited responding, in particular smokers were expected to show more disinhibited responding than non-smokers during the experimental condition.

Additionally, several measures of impulsivity were taken, in order to assess the relationship between the BAS-related personality trait(s) and measures of typical impulsivity. In fact, on top of the typical battery of questionnaires administered in every testing session (see chapter 4), participants completed the Dickman Impulsivity Inventory (DII) and also performed on a standard delay discounting (DD) task. The latter, as previously mentioned, is a behavioural measure of impulsivity and it closely resembles the paradigm used in the decision-making task. Hence, the study aimed to assess how performance on the DD task related to performance on the decision-making task and, consequently, how the two tasks related to the various personality measures.

---

[4] Moreover, the thesis aims to explore the impact of reward on learning and motivation and, subsequently, the neural model developed and implemented in the later studies is a reward-only model

The DII is a questionnaire which measures two sub-types of impulsivity, reckless behaviour (dysfunctional impulsivity) and reward-driven behaviour (functional impulsivity; Smillie & Jackson, 2006). Empirical evidence has indicated that functional impulsivity affects behavioural responses made by participants during performance on tasks that offer the chance to earn rewards (e.g. IOWA gambling task, the BART and go/no-go tasks: Franken & Muris, 2005; Smillie & Jackson, 2006; Vigil-Colet, 2007). In contrast, dysfunctional impulsivity (DI) has been found to be associated with reduced inhibition in withholding a disadvantageous dominant response (e.g. reversal learning) which is a behaviour typically observed in individuals with high scores on impulsivity measures such as the I_7 questionnaire (Franken, van Strien, Nijs & Muris, 2008). For example, a study by Smillie and Jackson (2006), participants had to learn to respond to the 'good numbers' (go cues) and withhold responses to 'bad numbers' (no-go cues) during a go-no/go task. Individuals with high scores on the DI scale showed disinhibited responding to the no-go cues regardless of payoff manipulation (reward-only vs. punishment-only). In contrast, individuals with high scores on the FI scale learnt to correctly respond to go cues in the reward-only condition but not in the punishment-only condition (Smillie & Jackson, 2006). These results suggest that there may be impulsivity measures that are related to reward sensitivity (i.e. the impulsivity of RST, a BAS-trait), and impulsivity measures that capture a more typical facet of impulsivity, characterised by a lack of forethought and reckless behaviour ("rash impulsivity"; Dawe & Loxton, 2004; Dawe et al., 2004; Franken & Muris, 2005, 2006).

Hence, a further aim of the present study was to explore the relationship of these two scales with the personality components included in the study and the BAS outcome as well as measures.

5.3 Method

5.3.1 Participants

A total sample of 92 participants took part in this study. The age of the sample ranged from 18 to 39 years (mean = 23.82, s.d. = 5.13). There were 57 females and 35 males in the sample. The sample was recruited over two separate waves of testing, a summer and an autumn session. During the first wave of testing (Summer session), 32 participants (15 females and 17 males) were recruited and tested. They were recruited from the student population at Goldsmiths (University of London) and among non-students who lived in the local area. They were predominantly non-psychology students and received payment for their participation. They earned £10 for their participation (approximately 2 hours).

During the second wave of testing (Autumn session) 60 (42 females and 18 males) participants were tested. The sample was constituted mainly by first year undergraduate Psychology students at Goldsmiths who took part to earn course credits. However, 38.3% of the second wave consisted of non-psychology students who received payment for their participation (£10 for approximately 2 hours). Additionally to payment or credits, participants could also earn up to £2 extra depending on their performance on this task.

Participants in the Autumn study were slightly but significantly younger than those in the Summer session [22.87 (4.9) cf. 25.59 (5.12); $t(90) = 2.50$; $p = .014$].

All participants spoke fluent English that enabled them to complete the personality questionnaires. However, participants were invited to ask for clarification on any of the questionnaire items.

5.3.2 Design

The present study assessed the impact of personality and reward frequency (payoff structure manipulation) on a decision-making task. In particular, it aimed to assess the impact of high and low reward frequency on participants' ability to inhibit responding and delay gratification as a function of their scores on BAS measures.

The primary independent variable was reward frequency and it was run as a within-subjects factor with two levels, i.e. high and low reward frequency. In order to control for order effects, the two tasks were intertwined so that they were performed at the same time. Thus, any carry-over effects between the tasks should not occur as a result of order of presentation. The study assessed how different frequencies of reward affected the pattern of responding to stimuli associated with low and high reward frequency.

The task consisted of an experimental and a control condition. The latter represented a baseline measure to assess whether participants were actually able to detect the different reward rates and, hence, maximise long-run winnings. The task simply aimed to assess reward-maximising strategies in relation to individual differences.

5.3.3 Task and apparatus

The task was run on a Mesh PC and a Mitsubishi 21" monitor with 1024 x 768 pixel resolution in an artificially lit room. Each stimulus was computer generated by using Matlab routines from Brainard's (1997) Psychophysics Toolbox. Stimuli were flashed on a black background that filled the entire screen. Responses were made using the four (appropriately labelled) arrow keys on a standard keyboard.

The stimuli were 2 squares that measured 4.5 cm * 4.5 cm and were displayed side by side with an edge-to-edge distance of 3.5 cm. The squares were presented at the centre of the screen 9 cm from the bottom and top of the screen and also 9 cm from each side of the screen.

The stimulus labels, button 1/A or button 2/B, appeared below each corresponding stimulus box, depending on the task condition participants were performing and in accordance to the timing parameter of each condition. Figure 5.1 illustrates the format of the task for those participants allocated to the first counterbalancing (CB) order (see below for CB details).

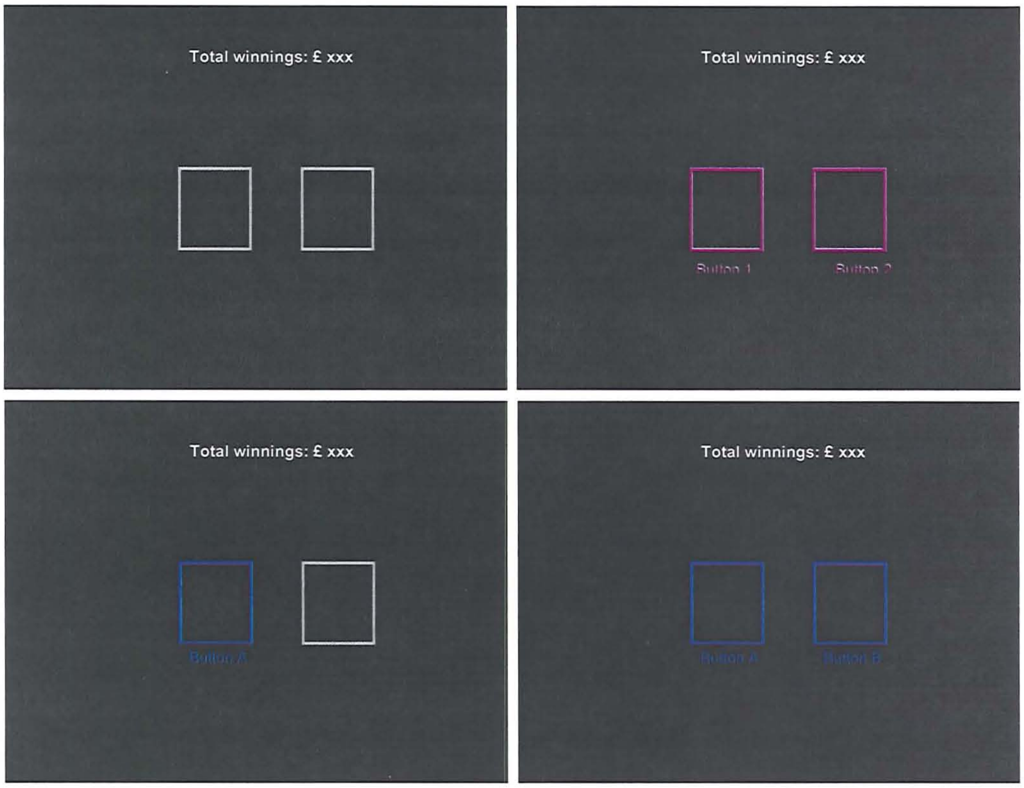


Figure 5.1. Screen shots for the two tasks for those participants performing under CB 1. The two panels at the top represent the control condition (i.e. pink boxes) before and after the 10-second delay had expired on the left and the right, respectively. The two bottom panels represent the experimental condition. The colour-task combination was reversed for those participants allocated to CB 2.

The expression 'Total winnings = £xx.xx' appeared above the two boxes. This expression kept participants continuously updated on the total amount of money they had earned. Every time participants responded to either box they also received a feedback message that was displayed for 2 seconds and then was immediately followed by the next trial. The feedback message informed participants on whether they had earned 3 pence or not. During reward trials, participants received a message in green stating 'You win 3 pence' whereas during no reward trials, participants received the following message '_ _ _ _ _ _ _ _', drawn in white (i.e. 8 hyphens, as in the original study).

Participants performed the task under two conditions, the experimental and the control condition. In order to control for possible order effects, the two tasks were intertwined. During the experimental condition, participants could respond either to the box on the left-hand side of the screen or to the one on the right-hand side of the screen. Responses to the box on the left were rewarded 40% of the trials (i.e. low reward response, LR) whereas responses to the box on the right were rewarded 80% of the total trials (i.e. high reward response, HR). The box on the left-hand side of the screen lit up as soon as the trial began, so that participants could respond straight away. In contrast, participants could respond to the high reward box only after a 10-second delay, when the box lit up. The control condition offered the same reward frequency structure as the experimental one. However, in the control condition in order to respond to either box participants had to wait until a 10-second delay had expired, whenever both boxes lit up simultaneously. If they responded to the high frequency feedback box on every single trial, they earned £1.92 (which was then rounded up to £2).

Owing to the fact that the two tasks were essentially identical and intertwined, they were visually distinguished by the colour used to draw the boxes presented in the two conditions. This was done in order to render the two conditions more easily discernible. In fact, one of the tasks consisted of pink boxes and the other task consisted of two blue boxes. During the pink boxes task, in order to respond to the box appearing on the left-hand side of the screen participants had to press the up arrow key which was labelled as '1' and for the alternative choice they had to press the down arrow key that was labelled as '2'. During the blue task, to respond to the box on the left-hand side of the screen participants had to use the left arrow key labelled as 'A' or press the right arrow key, labelled as 'B', to respond to the box on the right-hand side of the screen. Depending on the counterbalancing (CB) order participants experienced the pink boxes task as the experimental or control condition (see counterbalancing section below).

The whole real task was made up of 80 trials, 40 trials per each task. The task was programmed so that the order in which trials from the two tasks were presented was on a quasi-random sequence, identical for all participants. Participants could respond only using the appropriate task keys, and only once the boxes had lit up and the corresponding labels had appeared. Attempts to make premature key presses to unlit boxes were ignored and had no effect on the task processing or data storage. In both conditions, the feedback frequency was also controlled as the rewards were presented in a fixed quasi-random order[5].

Prior to starting the real task, participants underwent a guided practice trial where they had to follow the instructions appearing on the screen. The practice trials were intended to get participants acquainted to the key presses required to respond to the boxes during the two tasks as well as to let them become familiar to the consequences of each press type (and, in particular, experience the different probabilities of feedback). Participants made 20 guided responses to the control condition boxes and 20 presses to the experimental condition boxes. They were instructed to respond ten times to the box appearing to the left-hand side of the screen (i.e. low reward frequency) and ten times to the box appearing on the right hand-side of the screen (i.e. high reward frequency) using the corresponding keys. They followed the same procedure for the control and the experimental condition. The probabilities of winning and the time delays were exactly the same as the ones adopted for the real trials. However, during the practice trials participants had no decisional power but simply had to follow the instruction on the screen. Participants were told to use the index finger of their dominant hand to press the key corresponding to the selected box to which they needed to respond.

---

[5] The mean proportion of reward received following HR and LR choices are close to the expected values (i.e. 40% and 80%) in both conditions as represented in the table below which reports the number of LR and HR responses, the % of rewarded response and the number of participants. Participants who made fewer than 8 LR responses, regardless of task condition, were excluded from this table since they have too few trials to meaningfully estimate proportions.

	LR-exp	HR-exp	LR-con	HR-con
# made	27.3(8.4)	15.0 (10.7)	17.2 (5.8)	27.6 (7.8)
% reward	36%	82%	37%	81%
N	78	82	55	89

5.3.4 Counterbalancing

Counterbalancing was necessary in order to control for possible artefacts related to a possible emotional valence related to the colours of the squares, as well as a possible key bias. Participants were allocated to one of the CB orders depending on whether they had an even or odd ID code. Participants with an odd code were allocated to the first CB condition (CB 1), where the blue-box task was the experimental condition and the pink-box task was the control task. Participants with an even code were allocated to the second CB condition (CB 2), where the pink-box task was the experimental condition and the blue-box task represented the control condition task.

5.3.5 Personality measures

Participants completed several personality measures, which were: the Eysenck Personality questionnaire revised (EPQ-R), the Oxford-Liverpool inventory of feelings and experiences scale (OLIFE), the sensation seeking scale (SSS), the schizotypal personality questionnaire (SPQ), the big five inventory (BFI) and the BIS/BAS questionnaire. Four personality components were extracted after running a PCA with Varimax rotation on these scales (see chapter 4 for details). The four components extracted are: Extraversion (E), Neuroticism (N), Positive Schizotypy (PS) and impulsivity-antisocial (ImpAss). The PCA was run on 232 participants, who constitute the overall number of participants tested. The components scores used in the present study are the corresponding scores extracted in the overall PCA involving 232 participants.

5.3.6 Additional impulsivity measures

Gray suggested that the underlying personality trait of the BAS was impulsivity. As already noticed, there is reason to believe that the impulsivity described by RST is one aspect of the multi-dimensional factor identified as impulsivity. Indeed, RST describes the BAS as reactive to reward and, hence, the BAS-related impulsivity is assumed to reflect reward sensitivity while rash impulsivity describes disinhibited and thoughtless behaviour (Smillie and Jackson, 2006; Franken and Muris, 2006). In the present study, it was decided to include further measures of impulsivity that were not collected in the other studies and, therefore, were not included in the PCA analysis. These measures were included in order to see how they related to performance on the decision-making task and how they were associated with the other personality components which, as described in the PCA chapter, may be related to the BAS functioning. It was expected that the Dysfunctional Impulsivity scale of the Dickman Impulsivity Inventory would be positively correlated with the ImpAss component extracted in the PCA. The relationships between the additional impulsivity measures and the extracted personality components, especially ImpAss, are reported in table 5.2.

Dickman Impulsivity Inventory (DII; Dickman, 1990)

The inventory consists of 23 items, twelve of the items measure functional impulsivity and 11 of the items measure dysfunctional impulsivity. Participants have to express how much they agree with the statement of each item using a 6-point Likert scale. The items that measure functional impulsivity capture impulsive behaviour directed to personal gain (e.g. item 18: 'I am good at taking advantage of unexpected opportunities, where you have to do something immediately or lose you chance'); whereas items measuring dysfunctional impulsivity capture reckless behaviour (e.g. item 12: 'I often say and do things without considering the consequences').

A study by Smillie and Jackson (2006) indicated that dysfunctional impulsivity (DI) was mainly associated with more typical measures of impulsivity and psychoticism (e.g. EPQ-P and sensation seeking), while functional impulsivity was mainly associated with measures such as extraversion, BAS (sum total score), but negatively with neuroticism and BIS scores. Empirical evidence, reviewed in the introduction, observed that the FI scale is positively associated with reward-driven responses whereas the DI scale is associated with disinhibited responding (Franken and Muris, 2005; Vigil Colet, 2006; Smillie and Jackson, 2006). Hence, it supports the claim that there may be different types of impulsivity and, in particular, one that refers to sensitivity to reward (BAS) and one that underlies reckless behaviour (Smillie and Jackson, 2006; Franken and Muris, 2006; Vigil-Colet, 2007). In the present study, the Chronbach's alpha for the FI scale was 0.79 and the alpha value for the DI scale was 0.86, which indicates that the two scales have a reasonable internal consistency.

Delay discounting task

The delay discounting task required participants to choose between a large but delayed reward and a small but immediate one. In the case of the immediate reward, the magnitude is varied until the individual perceives the delayed reward to be as valued as the immediate one and, therefore, s/he would not mind waiting. This is the indifference point for a particular delay interval. Once the indifference point is reached for one time delay, the time interval is further increased and the same procedure is followed to calculate the indifference point of the new time delay.

The delay intervals and reward magnitude were printed on separate 28 cm x 13 cm cards. There were 27 cards for the reward magnitude and the values used were: 990, 960, 920, 850, 800, 750, 700, 650, 600, 550, 500, 450, 400, 350, 300, 250, 200, 150, 100, 80, 60, 40, 20, 10, 5 and 1. These values have been used in other studies although in the current version they expressed money in GB pounds rather than US dollars (Bickel, et al., 1999). There were also two cards indicating £1000. There were 7 cards that represented the delay intervals, which were: 1 week, 2 weeks, 1 month, 6 months, 1 year, 5 years and 25 years. There was also one 'immediate' card. The cards were placed in front of the participants. The 'immediate' card with the variable reward magnitude cards were placed on the participant's left hand side whereas the '£1000' card and the variable delay interval cards were placed on the participant's right hand side. The first scenario presented to participants required them to choose between receiving £1000 pounds immediately and waiting a week for the £1000. The researcher turned through the 27 magnitude cards to alter the value of the immediate reward until the participants indicated that they preferred the delay reward. At this point, the delay interval was increased by turning the cards to the next time delay. This same procedure was employed through the seven time delays. The opposite procedure was followed to collect discounting measures in the ascending order.

The last immediate reward selected over the delayed reward in the descending sequence together with the first immediate reward selected over the delayed reward in the ascending sequence were recorded by the experimenter. The average of these values was taken as the indifference point for a specific delay interval. The 'temporal discount function' can be calculated using the obtained indifference points (Monterosso and Ainslie, 1999).

The instructions given to the participants at the beginning of the task are as follows:

> "I am going to ask you to make some choices involving money. You will not receive any of the amounts shown, but I want you to make your choices as though you could really have the amount of money displayed. The options are printed on cards on the table in front of you. To your left is an amount of money you can have immediately and to your right is an amount of money you can have after waiting a period of time. There are no right or wrong answers. I do not expect you to select one choice or the other. The choices you make are completely up to you. Just select the consequence you want, not the consequence you think that I want".

The researchers gave the following instructions to guide choice through the actual task. These instructions were given during the descending condition:

> "For your first choice, you can choose between £1000 delivered immediately and £1000 given to you after a delay of 1 week. Point to the choice you would prefer".

When participants pointed to the immediate £1000, as most participants did, instructions followed:

> "Now I will change the amount that you are being offered today. Keep choosing the option that you want".

In the few instances where participants preferred the delayed reward, instructions followed:

> "Now I will change the amount of time you would have to wait to obtain the £1000. Between £1000 delivered immediately and £1000 given to you after a delay of 2 weeks, which one would you prefer."

5.3.7 Procedure

The study method and procedure were approved by the Psychology Department Ethics Committee (DEC) at Goldsmiths, University of London. Testing was conducted in a testing room in the department. Participants sat at a desk in front of a computer screen, a keyboard and stereo speakers. On entering the testing room, participants were reminded that all the collected data and information would remain anonymous and be treated with confidentiality. They were also informed that they had the right to withdraw from the study at any time and for any reason. In order to guarantee anonymity, each participant was allocated a unique ID code that became the only means of identifying the data.

Participants were automatically allocated by the task programme to one of the two counterbalancing order depending on whether they had an even or odd ID code. Depending on the CB order, the colour of the two boxes for the experimental and control condition varied (i.e. either blue or pink; see counterbalancing section above).

Prior to the practice trials, participants were presented with some information on the two tasks they were going to carry on during both the practice and the real tasks. They were told which key presses were required and the response-timing issues related to the two conditions. Additionally, during the practice trials, participants had to follow the instructions on the screen, telling them which key press was required. Participants were invited to keep track of the winnings in relation to each key press.

Once the practice trials were over, participants were invited to express any doubt that may have risen during the practice and to ask for clarification. If no doubt was expressed, participants started the real task straight away. An information screen was presented prior to the actual task to remind participants of the two tasks they were about to perform and the key presses required. The instructions presented to participants were as follows:

> "You will carry out two tasks at the same time. They are the same tasks that you have seen during the practice trials. One task consists of two pink boxes whereas the other task consists of two blue boxes.
>
> Responses to the blue boxes can be made using the keys labelled as A and B for the box on the right and the box on the left, respectively.
>
> Responses to the pink boxes can be made using the keys labelled as 1 and 2 for the box on the right and the box on the left, respectively.
>
> ALL RESPONSES SHOULD BE MADE BY USING THE INDEX FINGER OF YOUR DOMINANT HAND.
>
> Despite the fact that the two tasks are presented at the same time, they are independent and should be considered separately. Your strategy should not rely on the tasks' inter-relationship.
>
> In this part of the experiment, after you press a key, the computer will give you monetary feedback. In fact, you will earn 3 pence on a randomly selected percentage of your responses.
>
> You will be able to monitor the number of money you earned for that trial and your running total. **You should try to earn as much money as possible."**

Since they were quite long, instructions were presented over two separate screens. Participants were also verbally reminded that their strategy should not be affected by the order of presentation of the two tasks. The researcher also stressed the fact that in this part of the experiment participants had to decide which box they wanted to respond to and that they could earn up to £2 depending on their performance. Once again they were urged to earn as much as possible.

Participants were invited to sit in front of the computer in a comfortable way, in order to move as little as possible during the trials. They were encouraged to express any doubts about the task and to start only once they felt completely confident about the task instructions. At this point, the experimenter left the testing room to let participants carry out the task on their own.

The delay discounting task was always performed as the last task of the session. The descending condition was run prior to the ascending one. The researcher manually noted the responses given by the participant in order to subsequently calculate the indifference points.

5.3.8 Data analysis

A data processing programme was developed to create an overall summary file that could be exported from Matlab into SPSS and, therefore, allow statistical analysis on the data. The programme extracted three main measures for each of the two conditions. They were: the number of low reward and high reward responses and their corresponding reaction times. Analyses were carried out on the overall 80 trials.

Additionally, the data obtained from each participant on the delay discounting task was individually fitted by both the exponential and the hyperbolic decay fitting model using non-linear regression in SPSS.

The hyperbolic function assumes that discounting decreases in proportion to the time delay, in particular discounting is greater with short time delay and smaller as the time delay increases (equation 10.3):

$$\mathbf{V_d = V_i/(1+k*D)} \qquad \textbf{Eq.10. 1}$$

In contrast, the exponential function assumes that the value of the reward decreases by a fixed amount constantly over time (equation 10.4):

$$\mathbf{V_d = V_i *exp^{-k*D}} \qquad \textbf{Eq.10. 2}$$

In both equations $\mathbf{V_d}$ is the value of the delayed reward, $\mathbf{V_i}$ represents the value of the immediate reward (i.e. 1000), **D** represents the delay and **k** is the discounting rate constant. Previous studies indicated that the hyperbolic function is more effective at capturing human data (Kalenscher et al., 2008). Prior to fitting the data via non-linear regression all delay intervals were transformed into months [e.g. 1-week = (1/30)*7 = 0.23].

5.4 Results

The main dependent measure was the proportion of responses to the box which offered less frequent but immediate rewards both in the experimental and in the control conditions. These responses are referred to as the low reward (LR) responses in the rest of the chapter. Results showed that a few participants made LR response after the 10-second delay had expired. In fact, a few participants reported having initially withheld LR responses for ten seconds because they had noticed that waiting seemed convenient regardless of task condition. However, most of them made only a few of such responses. Seventeen participants out of the total sample (N = 92) made delayed LR responses between 1 and 5 times, however 2 participants made 8 delayed LR responses and one participant committed delayed LR responses 22 times out of the total 40 trials. It was decided to exclude participants who had committed more than 5 of those responses because they cannot be considered as impulsive reward-driven choices.

Hence, the final sample included in the following analyses comprised 89 participants. All three participants excluded had performed under the second counterbalancing (CB 2; experimental task had pink boxes). Additionally, it was decided to explore the impact of individual differences on RTs in response to the LR stimulus in the experimental condition and in response to the high reward (HR) stimulus (i.e. following the 10-second delay).

5.4.1 Task performance – responses to the low reward (LR) stimuli

5.4.1.1 Behavioural data

A mixed-design ANOVA was run to explore the pattern of responding between the control and the experimental condition across all the trials. The primary dependent measure was the proportion of LR responses made by participants in the two task conditions. Each task consisted of forty trials. Task condition was entered as a repeated-measure factor with two levels (control vs. experimental). The counterbalancing condition was entered as a between-subject variable with two levels (e.g. experimental task used blue boxes vs. pink boxes). This was done in order to assess whether the task features affected performance.

Results showed that there was an extremely significant main effect for the amount of LR response made across the two task conditions [F(1,87) = 123.94, p <.001]. Participants responded significantly more frequently to the low reward box during the experimental condition (mean = 24.6, s.d. = 10.7) than during the control one (mean = 12.4, s.d. = 7.8). Moreover, the main effect of CB was significant [F(1,87) = 10.02, p = .002]. As visually represented in figure 5.2, participants in CB 1 tended to commit more LR responses than those in CB 2 regardless of the task condition. However, the 2-way interaction between task condition and CB condition was non-significant [F(1, 87) = 1.14, ns].

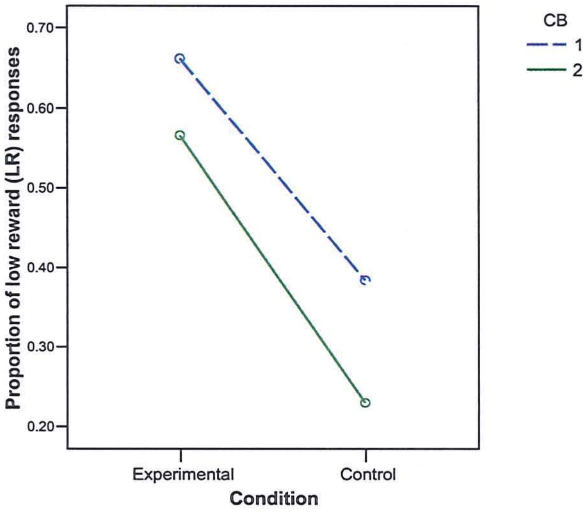


Figure 5.2. Proportion of low reward responses in the control and experimental condition as a function of CB group

5.4.1.2 Personality data

Prior to running any analysis to assess the impact of personality during performance on the decision-making task, scores on the four personality components obtained in the PCA were compared across the two CB conditions. The independent t-tests showed that there were no significant sampling error effects and that scores across the two CB conditions did not differ (p values greater than .26). Following these results and the previous results, which showed that performance on the tasks was not critically influenced by CB condition, CB was not included in the following analyses. Table 5.1 summarises the correlations between the four personality components extracted in the PCA and the proportion of LR responses (out of 40 trials) in the two task conditions. No correlation was found to be significant.

Table 5.1. Correlations between the personality components and the proportion of LR responses in the control (con) and the experimental (exp) conditions

		Extraversion	Neuroticism	Positive schizotypy	ImpAss
LR_con	Pearson Correlation	-.017	-.102	.149	-.104
	Sig. (2-tailed)	.876	.340	.163	.334
	N	89	89	89	89
LR_exp	Pearson Correlation	-.021	-.100	.140	-.033
	Sig. (2-tailed)	.845	.350	.190	.761
	N	89	89	89	89

An ANCOVA was run to explore the effect of personality on performance. As in the analysis reported in the previous section, the main DV was the number of LR responses and task condition was entered as a repeated measure. The four personality components were entered as continuous covariates. None of the personality components was found to mediate performance on the control or experimental task ($ps > 0.5$).

5.4.1.3 Performance and other impulsivity measures

As previously mentioned in the method section, other measures of impulsivity were collected in the present study. A behavioural measure of impulsivity was collected using the delay discounting task and further psychometric measures of impulsivity were also collected by administering the Dickman's Impulsivity Inventory (Dickman, 1990). Two participants from the first wave of testing did not perform on the delay discounting task so that the sample size on that test is equal to 90. The obtained hyperbolic k scores were positively skewed so a log transformation was applied to the data to obtain a more normal distribution. The recoded scores are used in the analyses reported below. Finally, in the second wave of testing ($N = 60$), smoking status was also recorded.

Correlations were run to assess the relationship between the four personality components extracted in the PCA and the extra impulsivity measures taken during the testing session (table 5.2). After correcting for multiple tests, the correlations showed that functional impulsivity was still positively correlated with extraversion ($r = .50$, $p < .001$) and negatively with neuroticism ($r = - 40$, p

<.001); dysfunctional impulsivity was also positively correlated with extraversion (r = .30, p = .005).

Table 5.2. Correlations between the impulsivity measures and the four personality components (FI = functional impulsivity; DI =dysfunctional impulsivity; recoded_hyp_k = delay discounting)

		Extraversion	Neuroticism	Positive schizotypy	ImpAss	Recoded_hyp_k
FI	Pearson Correlation	.500	-.402	-.179	.221	.051
	Sig. (2-tailed)	.000	.000	.094	.038	.635
	N	89	89	89	89	88
DI	Pearson Correlation	.294	.015	.186	.239	.217
	Sig. (2-tailed)	.005	.891	.081	.024	.042
	N	89	89	89	89	88
Recoded_hyp_k	Pearson Correlation	-.091	.019	.062	-.146	1.000
	Sig. (2-tailed)	.397	.861	.563	.174	.000
	N	88	88	88	88	88

As can be noticed in table 5.2, scores on the dysfunctional scale are positively correlated with scores on the delay discounting task (r = .22, p = .042) and the ImpAss component (r = .24, p = .024).

The scores on the functional and dysfunctional scales and the delayed discounting measure are generally considered to be impulsivity measures (Monterosso and Ainslie, 1999). Hence, it was decided to establish their relationship with the number of disinhibited (impulsive) responses made by participants during the decision-making task. Impulsive behaviour in the decision-making task is characterised by LR responses in the experimental condition. Hence, as a first step, correlations were conducted as an exploratory tool to observe the relationship between impulsivity measures and LR responses in the experimental task (labelled as pc_1e in table 5.3). Results showed that scores on the dysfunctional scale and the delay discounting task were significantly, although weakly, correlated with disinhibited responding in the experimental condition. In contrast, the relationship was absent in the control condition (the percentage of low reward responses is denoted pc1_c).

Table 5.3. Correlations between the impulsivity measures and the proportion of LR responses in the experimental (pc_1e) and the control (pc_1c) condition

		pc_1e	pc_1c
FI	Pearson Correlation	.074	.136
	Sig. (2-tailed)	.489	.204
	N	89	89
DI	Pearson Correlation	.223	.044
	Sig. (2-tailed)	.035	.679
	N	89	89
Recoded_hyp_k	Pearson Correlation	.227	.053
	Sig. (2-tailed)	.033	.621
	N	88	88

A multiple regression was run with dysfunctional impulsivity and discounting scores as predictors of the number of LR responses in the experimental condition. The model explained a significant 6.3% of the DV variance [F(2,85) = 3.91, p = .024]. The delay discounting scores contributed a proportion of unique variance equal to 17% while the dysfunctional scores contributed a unique variance equal to 18%. However, both contributions only showed a trend [dysfunctional: t(85) = 1.77, p = .085; discounting: t(85) = 1.79 p = .083]. As predicted by the correlation results, dysfunctional impulsivity and the discounting scores combined were found to be non-significant predictors of the number of LR responses made in the control condition [F(2,82) =.17, ns].

Further analyses were carried out by running two separate ANCOVA to explore the relationship between these two impulsivity measures (i.e. dysfunctional impulsivity and delay discounting) and task condition. The ANCOVAs' results showed that there was a trend for the 2-way interaction between condition and the delay discounting measures [F(1,86) = 3.21, p = .077] while the condition*dysfunctional impulsivity (DI) interaction also just failed formal significance [F(1,87) = 3.51, p = .064]. Post-hoc analyses were carried out to explore these interactions after transforming both the delay discounting and the DI measures into binary-valued factors. In spite of the loss of power caused by applying a median-split technique on the continuous data, results showed that the delay-discounting*condition interaction was still significant [F(1,86) = 7.57, p = .007]. Figure 5.3 indicates that individuals who showed a greater tendency to discount delayed rewards (high impulsive participants) made significantly more responses to the LR stimuli than their low counterparts. Both groups made similar numbers of LR choices in the control condition.

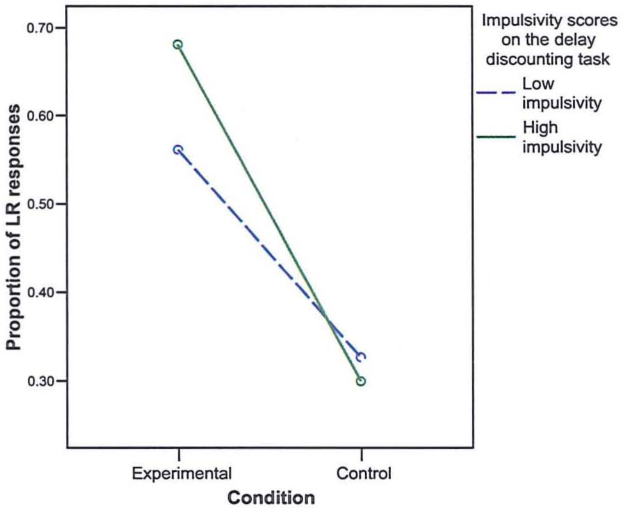


Figure 5.3. Relationship between the proportion of LR responses and delay discounting (i.e. impulsivity) scores

Results also showed that the DI*condition interaction was just significant [F(1,87) = 3.94, p = .050] and individuals with high scores on the scale responded to the LR stimuli more frequently than their low counterparts. Moreover, both groups made similar numbers of LR choices in the control condition (figure 5.4).

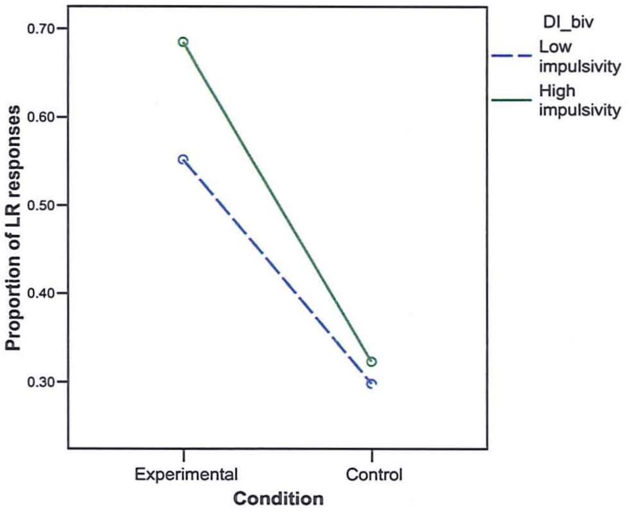


Figure 5.4. Relationship between dysfunctional impulsivity (DI) scores and the proportion of LR responses

5.4.1.4 Smoking status

Fifteen of the 60 participants tested in the second wave of the experiment identified themselves as smokers. On average they reported having smoked for 8.9 years (s.d. = 3.5) and smoking 10.3 (s.d. = 3.7) cigarettes a day. There were 9 female and 6 male smokers in the total sample. Their age ranged from 18 to 36 years, with a mean equal to 23.9 years (s.d. = 5.69).

One of the non-smokers made 8 delayed LR responses and, therefore, it was excluded from the analyses reported below (non-smokers: N = 44). It was decided to compare smokers and non-smokers scores on the impulsivity measures (delay discounting, functional and dysfunctional impulsivity) and the four personality components. Levene's test showed that the variances of these variables were not statistically different between smokers and non-smokers.

Independent sample t-tests indicated that there was a significant difference between smokers and non-smokers on the ImpAss component scores [t(57) = -3.44, p = .001] and there was a non-significant trend on the extraversion component scores [t(57) = 1.58, p = .12]. In particular, smokers scored higher than non-smokers on the ImpAss component but non-smokers scored higher than the smokers on the extraversion component (table 5.4).

Table 5.4. Comparison between smokers and non-smokers mean (s.d.) scores on the extraversion and the ImpAss component (scores are standardised scores from the PCA)

		Mean	S.d.
Extraversion	**Smokers** (N= 15)	-.29	1.31
	Non-smokers (N = 44)	.20	.95
ImpAss	**Smokers** (N = 15)	.72	.77
	Non-smokers (N= 44)	-.24	.98

5.4.1.5 Smoking status and performance

Further analyses were run in order to assess whether the relationship between individual differences and performance varied as a function of smoking status. Smoking status was treated as a between-Ss factor and each personality component and the impulsivity measures were entered individually as a covariate in several mixed-design ANCOVAs. The analyses indicated that only ImpAss and the delay discounting scores interacted with smoking status in their effects on performance on the task as described below.

ImpAss component

As predicted, results showed that when ImpAss was entered into the ANCOVA there was a significant main effect of task [F(1,55) = 11.10, p = .002]. The two-way ImpAss*task interaction was significant [F(1,55) = 4.5, p = .038] and the 3-way task*smoking status*ImpAss interaction was also significant [F(1,55) = 6.23, p = .016]. Additionally, there was a non-significant trend for the 2-way interaction between smoking status and task [F(1,55) = 3.70, p = .06].

Post-hoc tests were conducted to explore the interactions observed. A median split technique was implemented to transform the ImpAss component into a binary-valued factor. A mixed-design ANOVA was run with smoking status and the ImpAss bivariate variables as independent factors and task condition as the repeated-measures factor. In spite of the reduced power caused by recoding the ImpAss variable into a binary factor, results showed that there was still a significant trend for the 2-way-interaction between ImpAss and task [F(1,55) = 3.13, p = .08] and the 3-way interaction between ImpAss, task and smoking status [F(1,55) = 3.32, p = .07].

Figure 5.5 visually exemplifies the interaction between task and ImpAss scores. High ImpAss participants were those who showed the largest difference between the control and the experimental conditions with only a small difference being evident for the low ImpAss participants. These results resemble the previous findings for dysfunctional impulsivity but only when smoking status is taken into account.

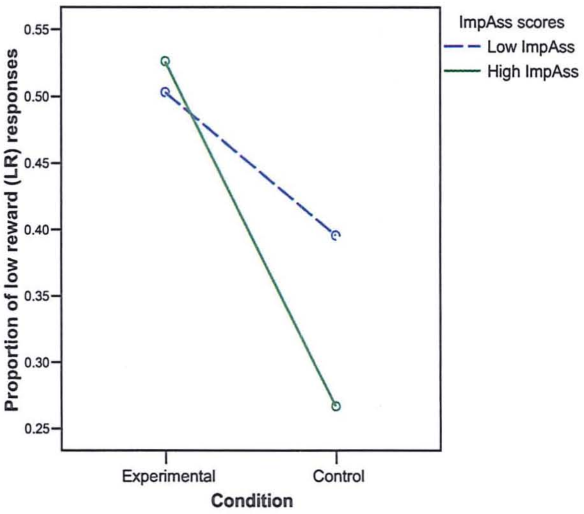


Figure 5.5. Relationship between ImpAss scores (median split) and the proportion of low reward responses (LR) in the two conditions

The task*smoking interaction was not significant [F(1,55) = 2.12, ns], although there was a trend in the earlier analysis. Figure 5.6 shows that smokers tended to respond to the LR stimuli less frequently than non-smokers especially in the experimental condition.

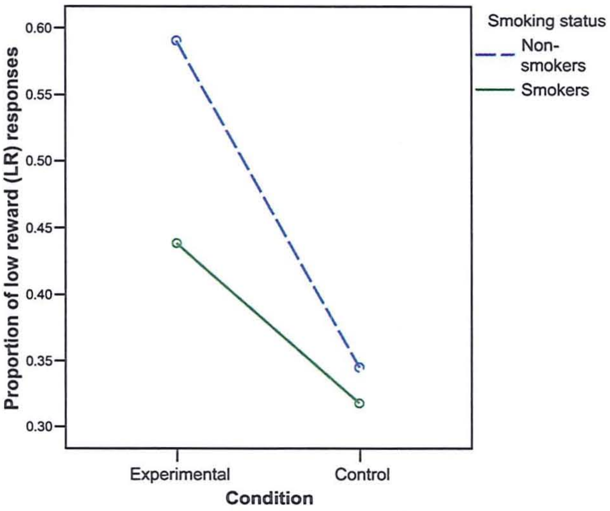


Figure 5.6. Relationship between smoking status and the proportion of low reward responses (LR) in the two conditions

In order to explore the 3-way interaction, the two smoking sub-samples were analysed separately after applying a split file filter on smoking status. A mixed-design ANCOVA with the ImpAss component entered as a covariate and task condition as a within-Ss variable.

The results indicated that the task*ImpAss interaction was significant in the smokers' sub-sample [F(1,13) 5.32, p = .038] but not in the non-smokers group [F(1,42) = .22, ns]. In the 'smoker' sub-sample, high ImpAss individuals tended to respond more frequently to the LR box in the experimental than in the control condition while low ImpAss individuals responded to the LR box with approximately equal frequency across both tasks (figure 5.7).

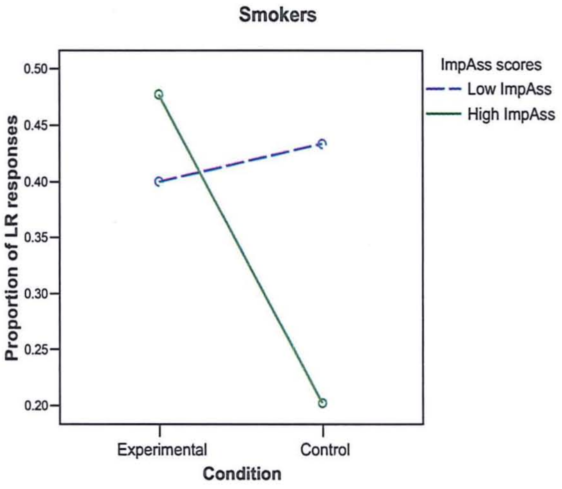


Figure 5.7. Relationship between ImpAss scores and the proportion of LR responses in the smokers' sub-sample

In contrast, among non-smokers participants with high and low scores on the ImpAss component showed equivalent performance in the two task conditions (figure 5.8). Indeed, high and low ImpAss individuals made significantly more LR responses in the experimental than the control condition.

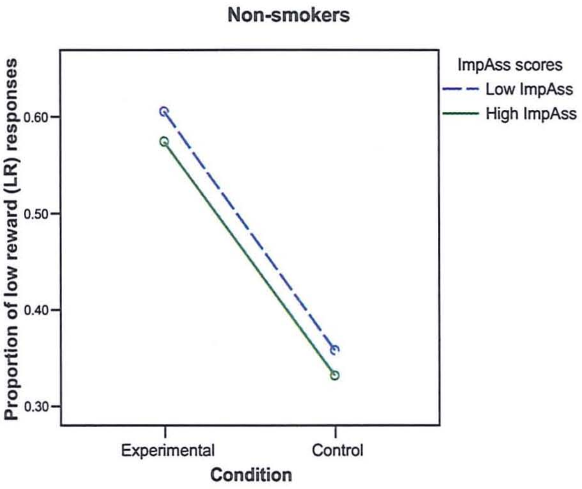


Figure 5.8. Relationship between ImpAss scores and the proportion of LR responses in the non-smokers' sub-sample

Hyperbolic K (discounting measure)

The ANCOVA results indicated that there was a significant main effect of task [F(1,55) = 32.35, p < .001]. Task condition was found to interact with the discounting measure [F(1,55) = 7.35, p = .009] and the 3-way interaction between smoking-status, task type and discounting measure was also significant [F(1,55) = 5.27, p = .025]. The 2-way task*smoking status interaction only showed a trend [F(1,25) = 2.76, p = .102].

Post-hoc analyses were run to explore these interactions. The delay discounting measure was transformed into a binary-valued variable by applying a median split on the factor. The new binary hyperbolic k factor was then entered into the ANOVA as a between-Ss variable. The 2-way interaction between task and discounting measure was still highly significant [F(1,55) = 9.28, p = .004]. However, due to the loss of power caused by applying a median split technique on the discounting measure, the 3-way interaction was no longer significant [F(1,55) = 2.08, p = .16]. Figure 5.9 illustrates the two-way interaction and it indicates that the increase in LR responses from the control to the experimental condition was significantly different for low k and high k (impulsive) participants. The increase was significant for both the high k [t(29) = 6.86, p < .001] and the low k [t(28) = 4.08, p < .001]. However, those participants who obtained a high impulsivity score made significantly more LR responses than those with low impulsivity scores in the experimental task [t(57) = -2.07, p = .043]. In the control condition, the groups did not differ [t(57) = .58, ns].

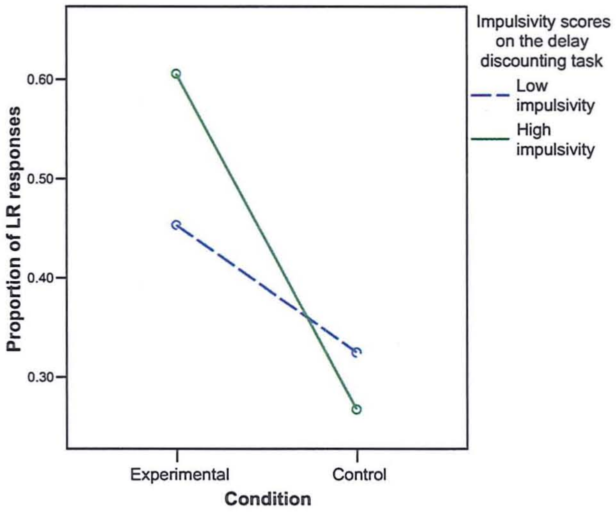


Figure 5.9. Relationship between the proportion of LR responses and delay discounting (impulsivity) scores

After applying a split-file filter on smoking, one further mixed-design ANOVA was carried out to explore the 3-way interaction. Results showed that the interaction between the discounting task scores and task was significant for the smokers [F(1,13) = 7.14, p = .019] but not the non-smokers [F(1,42) = .20, ns]. Figure 5.10 indicates that high impulsive smokers made more LR responses in the experimental condition than in the control condition. In contrast, those who scored low on the discounting measures showed a very similar pattern of responding on the two tasks.

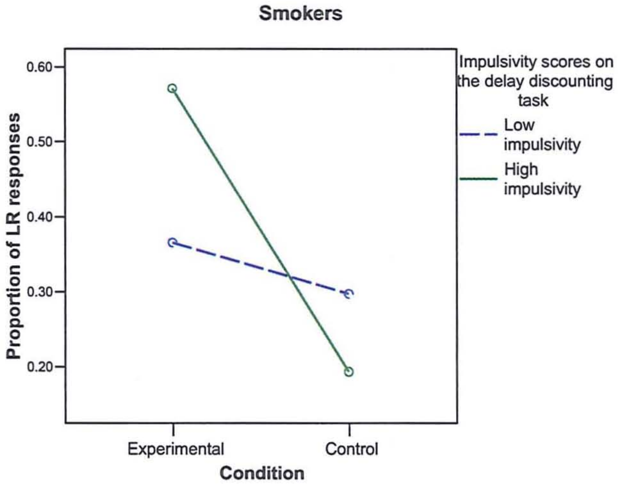


Figure 5.10. Relationship between delay discounting (impulsivity) scores and the proportion of LR responses in the smokers' sub-sample

Figure 5.11 illustrates that non-smokers showed a similar pattern of responding in the task regardless of their impulsivity scores on the delay discounting task condition.

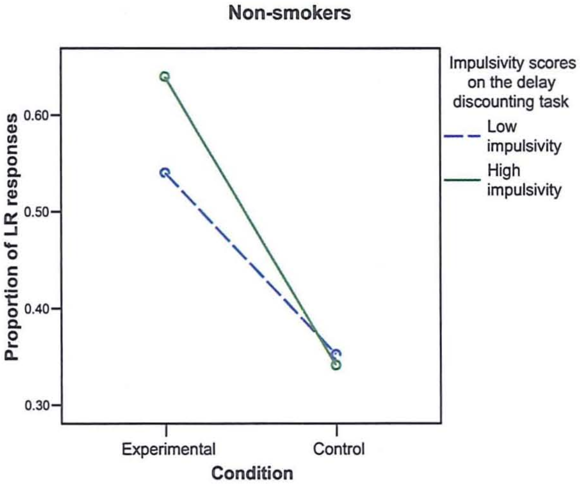


Figure 5.11. Relationship between delay discounting (impulsivity) scores and the proportion of LR responses in the non-smokers' sub-sample

Even though the impact of the ImpAss component and the impulsivity discounting measures seem to affect smokers' performance on the task in a very similar fashion, these two factors do not correlate with each other among smokers (N = 15; r = .20, p = .47). So their impact on performance is likely to be independent.

Dysfunctional impulsivity (DI)

A further ANCOVA was conducted to assess whether the mediating effect of dysfunctional impulsivity was moderated by smoking status. Results showed that, contrary to the ImpAss component and the impulsivity measure indexed by the delay discounting task, any effect of dysfunctional impulsivity upon performance on the task was not affected by smoking status. Indeed, the 2-way interaction between DI and condition showed only a non-significant trend [F(1,55) = 2.45, p =.12] whereas the 3-way interaction between DI, condition and smoking status was non-significant [F(1,55) = 1.01, ns].

Overall, the results show that impulsivity scores affected performance on the task. However, the impact of the ImpAss component on the task seems to be detectable only in smokers, possibly due to the fact that smokers had higher ImpAss levels than non-smokers. As a result the effect of ImpASS component was detectable only when smoking status was included in the analyses. By contrast, the impact of delay discounting impulsivity was found to be significant without including smoking status. Nonetheless, for the latter the effect of impulsivity on the task seems to be significantly present in the smokers, rather than the non-smokers. The effect of dysfunctional impulsivity showed yet another pattern: a trend was found for DI which mediated performance independently of smoking status.

Since the present task represents an approach-approach conflict, it may activate the BIS system and, therefore, it is possible that BIS-related traits (e.g. N) predict disinhibited responding (i.e. LR choices). Indeed, Newman and colleagues (1992) found that the effect of psychopathy on LR choices was dependent on anxiety scores. In order to test whether N interacted with impulsivity scores on LR choices, 3 further ANCOVAs were run with each of the impulsivity measures (separately) and N as a covariate. Results showed that N scores did not significantly interact with impulsivity scores to predict disinhibited responding (ps > .65).

5.4.2 Task performance - reaction time (RT) and personality

5.4.2.1 RTs to high reward (HR) responses

It was decided to run a few correlations to assess the relationship between RTs for the HR responses and the various personality components. This decision was driven by the consideration that reaction time scores related to HR responses are a good measure of reward reactivity since they show how quickly participants respond to stimuli associated with reward. In a few instances, participants made no response to the high reward stimuli in the experimental condition. Hence, the corresponding RTs (equal to 0) were labelled as missing data and excluded from the analysis. After applying this exclusion criterion, the sample used in the next analyses was equal to 82. The HR RTs in both conditions are measured from the actual onset of the stimuli (i.e. excluding the 10-second delay).

A repeated-measure t-test showed that the RTs in the control condition were significantly longer (mean = 1.69, s.d. = 3.59) than in the experimental one (mean = 1.39, s.d. = 3.91; t(81) = -4.06, p <.001). Table 5.5 indicates that there was a positive correlation between RTs and neuroticism (r = .22, p = .043) and a trend for a negative correlation between functional impulsivity and RTs (r = -.21, p = .061). Similar correlations were obtained between RTs to the HR in the control condition and functional impulsivity (r = -.22, p = .037) and neuroticism (r = .21, p = -.047). These results show consistency and suggest that both functional impulsivity and neuroticism may affect reward reactivity during performance of decision making tasks of the type used in the present study.

Table 5.5. Correlations between personality and impulsivity measures with reaction times to the high reward stimuli in the experimental (RT_2e) and the control (RT_2c) conditions (FI = functional impulsivity and DI = dysfunctional impulsivity; DD = delay discounting)

		Extraversion	Neuroticism	Positive schizotypy	ImpAss	FI	DI	DD
RT_2e	Pearson Correlation	.154	.224	.126	-.030	-.208	.022	-.084
	Sig. (2-tailed)	.167	.043	.258	.788	.061	.842	.456
	N	82	82	82	82	82	82	81
RT_2c	Pearson Correlation	.132	.211	.137	-.001	-.222	-.004	-.189
	Sig. (2-tailed)	.218	.047	.199	.992	.037	.967	.077
	N	89	89	89	89	89	89	88

Two separate multiple regressions were run with functional impulsivity and neuroticism (N) scores as predictors of RTs to the HR stimuli in both conditions. The model explained a significant 6% of the DV variance in the control condition [F(2,79) = 3.58, p = .033]. The functional impulsivity scores contributed a proportion of unique variance equal to 18% while the neuroticism scores contributed a unique variance equal to 17%. However, both contributions were non-significant [FI: t(79) = 1.56, p = .12; discounting: t(79) = 1.50 p = .14]. The model explained 5% of the DV variance in the experimental condition and showed a non-significant trend [F(2,79) = 2.93, p = .06]. Both predictors made small contributions to the dependent variable [standardised beta coefficients: FI = 15% and neuroticism = 17%] and were non-significant [FI: t(79) = 1.27, p = .21; discounting: t(79) = 1.48 p = .14]

Overall, these results seem to show that there may be two processes active simultaneously during the task. Indeed, on the one hand, it seems that high scores on 'rash' impulsivity (as indexed by the DI scale, scores on the delay discounting task and ImpAss scores) may lead participants to choose the low reward stimuli in the experimental condition. On the other hand high scores on functional ('reward-sensitive') impulsivity do not lead to responding to the immediate LR stimuli (in the experimental condition) rather than the delayed high reward stimuli. Additionally, reward-sensitivity (assuming this is what is indexed by high functional impulsivity and low neuroticism scores) speeds responding once the HR stimuli are present and responding is possible (in both conditions).

5.4.2.2 RTs to low reward (LR) responses

In a few instances, participants made no response to the low reward stimuli in the control condition and they were excluded from the analysis (since they represent missing data). After applying this exclusion criterion, the sample used in the next analyses was equal to 88.

It was decided to conduct further correlations to explore the relationship between the personality measures collected in the study and reaction times (RTs) to the low reward stimuli during the experimental condition. Indeed, LR responses during the experimental condition are a measure of rash impulsiveness as discussed earlier. However, results did not show any significant relationship between LR responses in the experimental task and any of the personality traits (table 5.6)

Table 5.6. Correlations between LR responses in the experimental task and the personality components collected in the study (FI = functional impulsivity and DI = dysfunctional impulsivity; DD = delay discounting)

		Extraversion	Neuroticism	Positive schizotypy	ImpAss	FI	DI	DD
RT_1e	Pearson Correlation	.038	.031	.029	.142	.071	.031	-.102
	Sig. (2-tailed)	.727	.774	.789	.187	.511	.776	.345
	N	88	88	88	88	88	88	87

5.4.2.3 RTs and smoking status

Further analyses were run in order to assess whether the relationship between individual differences and reaction times varied as a function of smoking status. Smoking status was treated as a between-Ss factor and each personality component and the impulsivity measures were entered individually as a covariate in several mixed-design ANCOVAs. Smoking status was not found to moderate the relationships between RTs in response to HR stimuli and the personality measures (ps > .2). In addition, the inclusion of smoking status in the analysis did not uncover any relationship between the personality factors and RTs in response to the low reward (LR) stimuli in the experimental condition (ps > 2).

5.5 Discussion

The present results replicated the original findings by Newman and colleagues (1992) and showed that that overall participants tended to respond to the low reward (LR) stimuli more frequently in the experimental condition compared to the control condition (i.e. disinhibited behaviour). This pattern of responding allowed participants to achieve two independent goals, which were on the one hand the attempt to speed up performance by responding straight away (LR in the experimental task) and on the other hand to increase their winnings (HR response in the control task).

Additionally, results indicated that disinhibited responding observed in the experimental condition was mediated by individual differences. In particular, impulsivity was found to enhance responding to LR stimuli, which is a measure of behavioural disinhibition. There were three impulsivity measures that mediated performance during the experimental task, which were: dysfunctional impulsivity, the hyperbolic k scores obtained from the delay discounting task and, once smoking status was taken into account, the ImpAss component. In fact, smokers who scored high on the ImpAss component were more likely to respond to LR stimuli than smokers with low scores on the ImpAss component. Not only the ImpAss effect on performance was dependent on smoking status but also the effect of both ImpAss and delay discounting impulsivity was found to be stronger in smokers. Finally, the results indicated that the speed at which HR responses were made (i.e. a reward sensitivity measure) was mediated by individual differences and, in particular, by scores on the functional impulsivity scale and the neuroticism component. Indeed, participants who scored high on the functional impulsivity (FI) scale or low on the neuroticism (N) scale tended to respond faster to the HR stimuli (in either condition) than those participants who scored low on FI or high on N (under both conditions).

The present results suggest that two opposite processes were active during performance on the decision-making task and each process was mediated by a different subtype of impulsivity. Hence, the results are in line with evidence that indicates that impulsivity is a multi-dimensional factor and it suggests that there are several subtypes of impulsivity, which are: rash impulsiveness and reward sensitivity (Dawe & Loxton, 2004; Franken & Muris, 2005 & 2006; Smillie et al., 2007; Vigil-Colet, 2007; Franken et al., 2008). In particular, the present results suggest that these two subtypes of impulsivity mediate the two opposing processes involved during performance on the task, disinhibited responding to LR stimuli in the experimental task and approach behaviour to reward in response to HR stimuli in the control task. Rash impulsiveness measures were found to capture disinhibited responding as indexed by the proportion of low reward (LR) responses in the experimental condition whereas reward sensitivity measures captured reward approach behaviour as indexed by the reaction times in response to the high reward (HR) stimuli.

Therefore, the findings are in line with the suggestion made by Dawe and Loxton (2004) that reward sensitivity (a BAS-related process) may account for the tendency to respond in the presence of reward cues whereas rash impulsiveness may be responsible for perseverative, disinhibited responding in spite of punishment or non-reward (i.e. thoughtless behaviour). However, these researchers conclude by saying that rash impulsiveness is a better predictor of disinhibition than reward sensitivity whereas the latter is necessary only for the establishment of the dominant response (Dawe et al., 2004).

The task adopted in the present study is the type of task most commonly employed to explore 'typical' impulsivity (Monterosso & Ainslie, 1999; Evenden, 1999; Dawe et al., 2004; Arce & Santisteban, 2006). Subsequently, the three impulsivity measures found to moderate LR responding in the experimental condition seem to be a good measure of rash impulsiveness rather than reward sensitivity (Dawe & Loxton, 2004; Dawe et al, 2004). Indeed, the dysfunctional impulsivity scale and scores on the delay discounting task are well-established measures of impulsivity and disinhibition (Monterosso & Ainslie, 1999; Eveden, 1999; Smillie & Jackson, 2006; Franken & Muris, 2006). It is not surprising that the impulsivity measure indexed by performance on the delay discounting task turned out to be the best predictor of performance on the experimental condition, since it is the closest measure in process terms. In fact, both tasks look at the impact of response disinhibition in response to immediate but smaller (less frequent) rewards and delayed but larger (more frequent) rewards. The ImpAss component was found to be a good predictor of disinhibited behaviour only once smoking status was taken into account. This may be due to the fact that disinhibition can only be observed with very high scores of impulsivity. Indeed, in the second wave of testing the ImpAss scores were significantly lower among non-smokers (mean = -.24, s.d. = .98) than smokers [mean = .72, s.d. = .77; t(57) = -3.44, p = .001]. The three impulsivity measures, which predict disinhibited responding, have been previously found to correlate with personality traits that capture behaviours in the healthy population that resemble mild versions of behaviours typical of psychopaths. For example, the dysfunctional impulsivity (DI) scale has been found to positively correlate with personality traits such as EPQ-P and Sensation Seeking (Smillie & Jackson, 2006). Additionally, both EPQ-P and SSS scores were positively loaded onto the ImpAss component extracted in the principal component analysis (PCA; chapter 4). Thus, the results seem to be consistent with the original findings obtained by Newman and colleagues (1992) who found that (low-anxious) psychopaths made more disinhibited responding as high impulsivity scores (DI scale, delay discounting scores and ImpAss component) were found to mediate disinhibited responding, in the present study.

Overall, the results indicate that disinhibition is moderated by smoking status since impulsive smokers showed the greatest levels of disinhibition (as indexed by an increased choice of LR responses in the experimental condition). These results are in line with previous research that found that smokers and other drug users show high levels of disinhibition during response inhibition, gambling and delay discounting tasks (e.g. Grant et al., 2000; Powell et al., 2002). Disinhibited behaviour in smokers is generally associated with high impulsivity levels (Bickel et al., 1999; Field et al., 2006). The positive relationship observed between performance on the decision-making task (i.e. proportion of LR response choice) and impulsivity index on the delay discounting task is very interesting and reassuring. Indeed, it indicates that even though the scenarios in the delay discounting task are based on hypothetical decisions between immediate rewards and rewards delayed over long time periods (e.g. weeks, months and years), the impulsivity score obtained from the task is a good predictor of thoughtless, disinhibited behaviours and choices (Field et al., 2006). Moreover, delay discounting scores were found to be correlated with impatience for real rewards over a 10-second interval. Hence, the present results confirm the utility of the delay discounting task as a measure of impulsivity despite the fact that it only presents imagined rewards and reward-scenarios (Kalenscher & Pennartz, 2008). Additionally, among smokers these scores are related to relapse, following abstinence (Bickel et al., 1999; Mitchell, 1999, 2004; Field, et al., 2006). Nonetheless, exposure to smoking-related cues was not found to increase impulsive responding during performance on a delay discounting task even though it increased self-reported craving scores (Field et al., 2007).

As previously mentioned, reward sensitivity may have been captured in the present study by reaction times (RTs) to high reward (HR) stimuli. These findings are in line with results obtained by Nichols and Newman (1986) who observed that when performing a pattern-matching task with a reward only feedback, extraverts showed faster reaction times than introverts. Hence, in the study extraversion (as an index of reward sensitivity) was found to mediate approach behaviour. These results seems to confirm the assumption made by the reinforcement sensitivity theory (RST; Gray, 1987; Nichols & Newman, 1986) that individuals who are sensitive to reward (high BAS) show enhanced approach behaviour to reward as indexed by faster reaction times to reward-related stimuli. The study conducted by Nichols and Newman (1986) supports the claim that extraversion (rather than impulsivity) is the BAS underlying personality trait (Pickering, 2004; Smillie et al., 2006a). In the present study, it was found that individual differences related to reward sensitivity were prominent when participants responded to high reward stimuli, regardless of condition. Indeed, results indicated that there was a negative relationship between RTs to the HR stimuli and functional impulsivity. In other words, individuals who scored high on the functional impulsivity scale tended to respond faster to the box that offered more frequent reward than their low scoring counterparts. Functional impulsivity (FI) was the personality trait that mediated approach behaviour and, therefore, the results suggest that FI may be a measure of the BAS. Smillie and Jackson (2006) have found that functional impulsivity correlates with several BAS-measures (i.e. Carver & White's BAS measures and the Torrubia and colleagues' Sensitivity to Reward scale). Moreover, they observed that high scores on the FI scale were also found to be positively associated with reward approach behaviour during performance on a go/no-go task. Therefore, these data are consistent with the idea that FI is a good BAS measure. This effect is seen once the 'choice' issues were resolved, either by waiting in the experimental condition or by being in the control condition. In the control condition, the response can be planned during the 10-second delay and there is a simple choice to pick the one that pays more; whereas, in the experimental condition, the choice is more complex (i.e. amount vs. delay) and it seems to allow disinhibited responding.

It may be argued that the experimental condition represents an approach-approach conflict (i.e. participants need to decided whether to respond immediately for a low-probability reward or whether to withhold responding for a high-probability reward). According to the revised RST, an approach-approach conflict should activate the BIS (Corr, 2006; 2008). Disinhibited participants, who respond to the LR stimuli more frequently than the HR stimuli in the experimental condition, should have a weak BIS (together with a strong BAS). A strong BIS, by contrast, would inhibit immediate responding and gather additional information to inform their choices. Therefore, individuals with high rash impulsivity and low anxiety (i.e. low BIS) scores are expected to show disinhibited responding. This is in line with the findings obtained by Newman and colleagues. Nonetheless, the neuroticism mediating effect, predicted by the above argument, was missing in the present study.

It is also possible that the presence of no-reward trials may have been perceived by some (especially high N) participants as non-rewarding (punishing), despite the fact that they had been informed that on some pre-defined trials they would receive no feedback (reward) and reassured that these instances were independent from their performance. Indeed, during the debriefing some participants referred to no-reward trials as 'wrong' since they had won no money. It may be possible in future studies to replicate the findings and assess a pure BAS activation effect by manipulating payoff magnitude rather than frequency in a reward only payoff structure. In this way, participants would earn money on every trial and this would remove the 'punishing' effect of no-reward trials.

The present results fail to replicate the studies that indicate that the BAS mediates disinhibition as indicated by studies carried out by Avila (2001). However, Avila administered a purpose-built BAS measure to assess individual differences in reward-reactivity (i.e. the sensitivity to reward (SR) scale from the SPSRQ, Torrubia et al., 1995) rather than standard personality components whose relationship with the BAS is still not well-established (e.g. extraversion and ImpAss; see chapter 3). The goal of the present research is interested in understanding how reward sensitivity relates to broad measures of personality and less interested in purpose-built questionnaires (e.g. the SR scale). Purpose-built scales include self-report items on how much one responds to rewards and, therefore, it is of much less interest to investigate how such scales relate to experimental measures, which index objective responses to rewards. Such studies are inherently circular. However, Franken and Muris (2006) have also observed that the SR scale was found to load onto both reward-related impulsivity and rash impulsiveness. In the present study scores on the dysfunctional impulsivity (DI) scale, which has been identified as a measure of rash impulsiveness (Franken and Muris, 2006; Smillie and Jackson, 2006), was found to be a predictor of rash impulsive responding, characterised by a tendency to respond more frequently to the low reward (rather than the high reward) stimuli in the experimental condition.

Moreover, the decision-making task employed in the present study may have been a measure closer to response disinhibition as indexed in the delay discounting task rather than passive avoidance tasks (Patterson and Newman, 1993; Avila, 2001). In fact, the passive avoidance (PA) tasks may be able to capture the processing related to reward-sensitivity and the personality traits underlying reward sensitivity (e.g. SR scale and E). These tasks may reflect the four-step disinhibition model, developed by Patterson and Newman (1993). Indeed, PA tasks require participants to withhold a dominant response that has been acquired in the first stage of the task in order to establish an alternative strategy that can allow efficient learning once the task contingencies have changed. The four-stage model suggests that individual differences related to reward sensitivity determines how efficient participants are in their response modulation. The task implemented in the present study does not follow the pattern typical of PA tasks and, therefore, the processes involved may be different.

In conclusion, the results indicate that there were two opposing processes that were active during performance on the decision-making task and that each process was related to one of the two subtypes of impulsivity (i.e. rash impulsivity and reward sensitivity; Dawe et al., 2004). More specifically, high scores on measures of rash impulsiveness (i.e. ImpAss, dysfunctional impulsivity and delay discounting scores) led to disinhibited responding whereas reward sensitivity measures mediated approach behaviour in response to high reward stimuli once any response conflict had been resolved (BAS-related process; McNaughton & Gray, 2004). The two processes are opposite in direction and, therefore, it is possible that rash impulsiveness and approach behaviour compete for control, in individuals depending on their scores on one of the two impulsivity scores.

The results are in line with the literature which suggests that disinhibited behaviour (e.g. delay discounting task) is typical of impulsivity and, in particular, rash impulsiveness. Indeed, performance on the decision-making task was dependent on scales which are well-established measures of rash impulsiveness (dysfunctional impulsivity and delay discounting scores). Moreover, the disinhibition effect is greatest in smokers and in the case of the ImpAss component it is dependent on smoking status. The study further replicates the findings obtained in the original study (Newman et al., 1992) where overall participants tended to respond to the HR stimuli less frequently in the experimental than the control condition. The original study also showed that low-anxious psychopaths made more LR responses than low-anxious controls. Nonetheless, the original study did not analyse potential inter-group differences in RTs to the HR stimuli, which may have captured BAS-related processes. The present findings on RT measures are exploratory and require replication.

Finally, the results, therefore, support the view of impulsivity as a multi-faceted factor and indicate the importance of using the appropriate tools (e.g. behaviour, cognitive and psychometric measures) in order to assess the aspect of impulsivity that is being explored in the study (Franken et al., 2008).

Chapter 6

The impact of partial feedback during a rule-based and an information-integration category learning task

6.1 Abstract

Previous studies have indicated that individuals who score high on ImpAss-related measures can sometimes perform better on simple rule-based (RB) CL tasks (Zuckerman and Ball, 1990; Pickering, 2004; Tharp, 2007 [phd]). These results were interpreted as suggesting that high scores on ImpAss measures are associated with higher levels of a relevant cognitive ability (e.g. working memory or selective attention). The present study aimed to replicate these findings on an RB task and it also assessed the impact of individual differences during performance on an information-integration (II) task. Additionally, it also assessed the impact of partial feedback during performance on the RB and II task. According to COVIS (Ashby et al., 1998; Ashby et al., 1999) performance on the II task should be impaired by partial feedback whereas RB performance should not be affected by this manipulation. Results on the RB task do not fully replicate the original studies on individual differences but they do support the assumptions made by COVIS. However, results indicated that the II task was too difficult for participants to learn and, therefore, it is not possible to draw final conclusions. These findings are discussed in terms of task and stimuli complexity.

6.2 Introduction

Performance on CL tasks often requires the participants to classify new and continuously distributed stimuli into separate categories. Performance on CL is initially based on guessing but over time categorisation decisions become more accurate as learning occurs. One of the main multiple systems models of CL is the COmpetition between Verbal and Implicit System (COVIS) model (Ashby et al., 1998). As discussed in chapter 1, COVIS suggests the existence of two independent systems that compete during learning: an explicit, rule-based system and an implicit system. The explicit system relies on logical reasoning and working memory whereas the implicit system is closely related to motor activity and procedural learning (Ashby et al., 1998; Maddox & Ashby, 2004; De Caro, Thomas & Beilock, 2008). Thus, one of the main assumptions made by the COVIS is that feedback facilitates the operations of both systems although it operates in different ways for the two systems (Maddox & Ashby, 2005). In particular, it has been suggested that the explicit system uses feedback to test whether the selected rule is correct or rule-switching is required. In contrast, in the implicit system positive feedback works as a reward signal that automatically reinforces the adequate response (i.e. DA-driven procedural learning).

One main assumption made by the COVIS model is that category learning tasks are differently affected by reward depending on the system that is involved. For example, it has been found that people can learn complex, non-verbal rules (i.e. procedural learning), when feedback about accuracy is offered on a trial-by-trial basis. However, in the absence of feedback, people adopt simple verbal rules for both rule-based and information-integration tasks (Ashby et al., 1998). Indeed, Ashby and colleagues (1999) carried out a study where participants had to learn either a rule-based or information-integration task without any form of accuracy feedback. Results showed that during unsupervised learning participants were able to efficiently learn the rule-based task whereas they were not able to master the more demanding information-integration task. (Participants were able to learn the information-integration task in the feedback condition.) Therefore, these findings indicate that procedural learning requires reward. Additionally, they also indicate that in the absence of a trial-by-trial feedback, participants adopt verbalisable, uni-dimensional rules that are not effective in solving information-integration-like tasks (Waldron & Ashby, 2001).

Similarly, Ashby and colleagues (2003) replicated the findings obtained in the previous study. In fact, they found that observational training impaired learning on an information-integration task compared to feedback training. However, feedback did not improve performance on a rule-based category task more than observational training. Further support to the assumption that the implicit system is more highly dependent on reward comes from a study carried out by Maddox and colleagues (2003). They compared the impact of delayed feedback on both a rule-based and an information-integration task. Results showed that during performance on an information-integration task, delayed feedback caused a decrease in response accuracy and an increase in the use of rule-based strategies. On the contrary, the feedback manipulation did not affect performance on the rule-based task.

Moreover, CL tasks have been identified as a useful tool to observe how different personality traits mediate learning and also to identify the BAS-related traits that mediate the activation of the system (Pickering, 2004).

Ball and Zuckerman (1990) offered the first study which used a CL task to assess individual differences related to BAS activation. The CL task used in the study presented participants with a pair of stimuli on each trial, one being the target category and the other the distractor. Participants had to learn to discriminate the target stimulus from the distractor. The visual stimuli varied on 8 bivariate dimensions, such as letter type (X or T), letter size (large or small), letter colour (black or white), border shape (circle or square) and so on. Two of the eight dimensions (letter type and border shape) were relevant to determine whether the stimulus was a target or not; in fact the target stimulus was always a T letter with a square border. The other 6 dimensions were irrelevant (i.e. uncorrelated with being a target). Learning was reinforced by feedback (verbal vs. monetary) which was delivered either in a reward-only or a punishment-only fashion. Once participants had reached the appropriate performance criterion (i.e. five correct responses in a row) the relevant dimensions became irrelevant and a new dimension became predictive of target status.

Results showed that individuals who scored high (in the top decile of a large sample) on the sensation seeking scale (SS; a measure of ImpAss) required fewer trials to reach criterion compared to those who scored low (bottom decile) on the same scale. However, this difference was not affected by the feedback manipulation (reward-only vs. punishment-only). Results also showed that participants who scored high on Neuroticism (N) learnt faster in the pre-shift phase but the opposite was found in the post-shift phase (stable individuals were better). The researchers suggested that the difference observed between high and low SS individuals may be due to higher cognitive abilities of high SS-scoring individuals compared to low SS scorers. In particular, the authors offered two explanations to account for the results. It was suggested that high SS individuals either adopted a more beneficial risk-taking strategy during the first few trials of the task that allowed them to discover the correct rule specifying the target more quickly; or had superior executive attention abilities, which allowed them to focus on the relevant and ignore the irrelevant dimensions (Ball & Zuckerman, 1990). According to COVIS, these results indicate that compared to individuals with low ImpAss scores, those with high scores may possess higher levels of cognitive abilities which are advantageous during RB tasks.

Pickering (2004) replicated Ball and Zuckerman's findings in two studies assessing the impact of ImpAss measures on performance on CL tasks of the RB kind. The first task presented visual stimuli that varied on two dimensions, which were: the height of a rectangle (relevant dimension) and the position of an internal line segment (irrelevant dimension). The personality measures included in the study were Novelty Seeking (NS) as a measure of the ImpAss construct and Harm Avoidance (HA) as a (control) measure of Anxiety. Both measures are part of the Tridimensional Personality Questionnaire (TPQ) developed by Cloninger (1989). HA was not found to correlate with overall performance on the task, whereas participants scoring high on NS were found to perform better than those scoring low on NS. A follow-up study was carried out using the same stimuli. The second study consisted of two phases, a learning phase and a rule-switch phase. In the learning phase, the position of the internal line segment was the relevant dimension while the height of the rectangle was the irrelevant one. In the rule-switch phase, an extra dimensional switch occurred so that the height of the rectangle became the relevant dimension. Participants were not informed about the rule-switch. The personality measures included in the second study were the EPQ-P scale (EPQ, Eysenck, Eysenck & Barrett, 1985), a measure of ImpAss, and the Unusual Experiences scale (UnEx from the OLIFE, Mason et al., 1995), a measure of schizotypal personality. Results showed that EPQ-P scores were positively correlated with performance during the leaning phase but not the rule-switch phase of the task, while UnEx was negatively correlated with performance during the rule-switch phase of the task but not the learning one.

These studies support the assumption that high ImpAss measures (e.g. SS, NS and EPQ-P) are associated with greater executive attention ability that leads to optimal performance on a CL task whose solution requires a simple uni-dimensional rule to be selected from several possibilities. However, Tharp (2007) has found that high ImpAss individuals were impaired during performance on CL tasks that required the use of a more complex conjunctive rule. Indeed, high ImpAss participants in Tharp's study showed a preference for simple, uni-dimensional rules. Thus, Tharp speculated that high ImpAss participants may be relatively cognitively inflexible. It is possible that a relative lack of flexibility may be beneficial in very simple uni-dimensional tasks (like the ones used by Ball and Zuckerman and Pickering) since too much flexibility may over-complicate the search for possible solutions. However, the lack of flexibility is 'exposed' and counterproductive in more complex tasks that require a multidimensional solution.

6.2.1 Aims of the present study

The present study aims to replicate Pickering's findings (2004) that showed that high ImpAss scores are associated with greater cognitive performance and better performance on simple RB tasks. The study aimed to extend the analysis to explore the impact of individual differences during performance on an information-integration task. Following the results obtained by Tharp, it is expected that high ImpAss scores may be associated with cognitive inflexibility, which was found to impair performance during a CL task that required a conjunctive-rule strategy. Therefore, high ImpAss scores might be associated with impaired performance on II tasks. However, in Tharp's study, participants could perform reasonably well (around 80% accuracy) using a simple uni-dimensional rule-based strategy. In the present study, high ImpAss-scoring participants may persevere with a simple RB solution only if it achieves a reasonable, although suboptimal, level of performance. Therefore, the prediction about ImpAss impairing performance is dependent upon how well the suboptimal strategy is at solving the categorisation problem.

Additionally, the study aims to explore the impact of feedback manipulations during performance on an RB and an II CL task. In particular, the study manipulated feedback frequency during CL. According to previous research, feedback manipulation should not affect performance on the RB task although it is expected to impair performance on the II task and, possibly, induce participants to resort to an RB strategy.

6.3 Rule-based task

The present study attempted to observe the effects of probabilistic feedback on category learning during a rule-based learning task. According to the COVIS model's assumptions, performance on rule-based task should not be affected by feedback manipulation (Ashby et al., 1998; Maddox & Ashby, 2005). Hence, participants should still learn the optimal rule even when receiving feedback on a low frequency basis. In fact, they should be able to compensate for the limited amount of information about their performance accuracy by maintaining cognitive control over the task by employing working memory and attention. Participants also performed on a working memory (WM) task in order to assess the relationship between performance on the RB task and WM scores. De Caro and colleagues (2008) found that performance on a rule-based (RB) task was positively associated with WM scores. These results are in line with the literature that suggests that the explicit system is dependent on working memory for hypothesis generation and testing (Ashby et al., 1998; De Caro et al., 2008).

Finally, at odds with the above predictions, there are possible mechanisms through which the feedback manipulation in the current study might affect RB task performance. DAergic activity has been found to modulate neural processes in the head of the caudate and this structure of the explicit system is argued to be responsible for rule selection and switching. Indeed, low levels of DA in the caudate have been found to be associated with a greater level of preservative errors during RB tasks (Ashby et al., 1998; Ashby & Spiering, 2004; Ashby & Ennis, 2006). Therefore, it may be possible that probabilistic feedback may have an impact in DAergic levels and, therefore, impair rule-based learning.

6.4 Method

6.4.1 Participants

An opportunity sample was drawn from the student population at Goldsmiths (University of London). Participants were recruited through advertisement around the college campus. The sample was made up of 64 participants, 32 males and 32 females. All but one were right-handed. All participants were aged between 18 and 39 (mean age = 26.7; s.d. = 4.41). Participants were guaranteed confidentiality. They were tested in one sitting that lasted approximately 1 hour and 15 minutes. They all received £7.50 for their participation.

6.4.2 Design

The present study was interested in assessing the impact of partial feedback on learning during a rule-based category learning task. The independent variable of interest was, therefore, represented by feedback frequency. This IV was a between-subjects factor with two levels. In fact, in one condition participants received feedback on 37.5% of their responses (i.e. low feedback frequency condition). In the high feedback frequency condition participants received feedback on 75% of their responses.

The study was interested in assessing the impact of the IV on learning (i.e. the dependent variable). Learning was assessed by accuracy levels that were recorded by the computer as correct responses.

6.4.3 Personality measures

Participants completed several personality measures, which were: the Eysenck Personality questionnaire revised (EPQ-R), the Oxford-Liverpool inventory of feelings and experiences scale (OLIFE), the sensation seeking scale (SSS), the schizotypal personality questionnaire (SPQ), the big five inventory (BFI) and the BIS/BAS questionnaire. Four personality components were extracted after running a PCA with Varimax rotation on these scales (see chapter 4 for details). The four components extracted are: Extraversion (E), Neuroticism (N), Positive Schizotypy (PS) and impulsivity-antisocial (ImpAss). The PCA was run on 232 participants, who constitute the overall number of participants tested. The components scores used in the present study are the corresponding scores extracted in the overall PCA involving 232 participants.

6.4.4 Working memory measure (WM task)

It was decided to include a measure of WM ability in the battery of measurements, since explicit system is dependent on working memory for hypothesis generation and testing (Ashby et al., 1998; De Caro et al., 2008). Thus, high scores on the WM task were expected to be associated with greater performance on the task. In particular, the task used is a measure of memory scanning ability (Stenberg, 1966). The task is composed of 14 trials. During each trial the participant is presented with a set of letters to memorise. The first four trials consist of sets of 4 letters and they function as practice trials. The 10 experimental trials present sets of 6 letters.

Following the presentation phase (2.5 seconds), participants were presented with single letters on the screen and had to decided whether they were members of the preceding set or not by responding ‘yes’ or ‘no’ using the corresponding keys that were the y and the n key on the keyboard, respectively. Participants scored one point per correct identification of targets and distracters. During the experimental trials they were presented with 12 single letters (6 targets and 6 distractors), hence they could score a maximum of 12 points per trial. There were 10 experimental trials so that the maximum total score was equal to 120.

6.4.5 Task and apparatus

The task was run on a Mesh PC and a Mitsubishi 21" monitor with 1024 x 768 pixel resolution in an artificially lit room. Each stimulus was computer generated by using Matlab routines from Brainard's (1997) Psychophysics Toolbox. Stimuli were flashed on a black background that filled the entire screen. The stimuli were constituted by sine wave gratings (Gabor patches) that was enclosed in a 200 x 200 pixel frame. Each stimulus varied in spatial orientation and spatial frequency. The optimal rule that maximised performance was unidimensional and was determined by spatial frequency. In fact, the rule was as follow: 'when spatial frequency is high, the stimulus belongs to B; whereas when spatial frequency is low it belongs to category A'. The category discriminabilities (d') were equal to 3. The category distribution parameters are summarised in table 6.1.

Table 6.1. Mean, standard deviation and covariate of category A and category B

		Mean	SD	Covariance
Category A	**Angle**	90	37	.0001
	Frequency	0.05	0.01	
Category B	**Angle**	90	37	-.0178
	Frequency	0.08	0.01	

The task consisted of 140 trials including the CL (real) and the filler task trials and it was so designed to have 70 stimuli that belonged to category A and 70 that belonged to category B. The two category stimuli were created by sampling from two normal distributions, which were randomly generated and described variation in spatial frequency. The mean score of the normal distribution used to generate category A stimuli was equal to 0.05, whereas for the category B distribution was 0.08. Both distributions had a standard deviation equal to 0.01. The two normal distributions overlapped to some extent. However, the spatial frequency values (x_1) were rounded to two decimal points before computation of the spatial frequency, f, using the formula below. The result was the following distribution of values of x_1 (table 6. 2). The choice of using numbers with only two decimal points for the spatial frequency dimension was made to render the task perceptually easier and, therefore, create a task that was not extremely difficult to ensure learning across participants. This was done to compensate for the lack of feedback and, especially, the added filler task (see below), which might disrupt performance.

Table 6.2. Spatial frequency of the Gabor patches

	Spatial frequency							
	.03	**.04**	**.05**	**.06**	**.07**	**.08**	**.09**	
Cat A	5	13	25	24	3	0	0	N=70
Cat B	0	0	0	1	24	28	17	N=70

Spatial orientation, o, (irrelevant dimension) for category A and category B stimuli was sampled from one normal distribution that had a mean equal to 90 and a standard deviation equal to 40. Each random sample value described above (x1, x2) was converted to a stimulus by deriving the frequency, $f = x1 * 2 * \pi$, and orientation, $o = x2 * \pi /180$.

Stimuli were presented until participants pressed either the A key or the B key. In order to facilitate learning, participants were given visual, auditory and monetary feedback (in feedback trials). In fact, following each response, participants received a feedback message that informed theme whether they were 'correct' or 'wrong' together with a high-pitched or a low-pitched noise, respectively. For each correct response, participants were rewarded with the gain of 5 pence. All these monetary gains would cumulate across each correct trial. All different types of feedback were simultaneously presented for 600 msec and followed by 1,000-msec ITI.

Participants in the two feedback conditions received the same absolute amount of feedback responses (i.e. feedback messages). However, those allocated to the low feedback received half the frequency of feedback trials as those in the high condition. In fact, participants in the low feedback condition received feedback on 37.5% of their responses whereas those in the high condition received feedback on 75% of their responses.

However, if feedback occurred on half the amount of trials in the high feedback condition compared the low feedback one and, therefore, the feedback trials were further apart in the low feedback condition. This artefact might, thus, place a higher demand on working memory during the low feedback condition to retain the rule across trials.

In order to counteract this possible higher working memory workload in the low feedback condition, extra filler trials were added. These extra trials asked participants to perform a 'task' which required them to press either the category A or the category B key following the instructions that appeared on the screen (e.g. 'Press the category A response key'). No feedback was given for these trials and each response was simply followed by 1,000-msec ITI. The number of filler task trials in the high condition was the equivalent of the extra no-feedback trials in the low feedback condition. Table 6.3 represents the proportion of feedback, no-feedback and filler task trials out of a sample of 10 trials.

Table 6.3. Ratio of feedback (fbk), no-feedback (no fbk) and filler task trials out of a sample of 10 trials

	Fbk	**No fbk**	**Filler**
High	3	1	6
Low	3	5	2

Hence, over the whole task (140 trials), participants in both conditions received the same total amount of feedback trials (42). However, those participants in the high condition performed half (56) the amount of CL trials compared to those in the low feedback condition (112). Nonetheless, this difference was counteracted by the 56 extra filler-task trials (see table 6.4). After introducing the filler task trials in the task, there were still approximately equivalent number of stimuli from both categories in both feedback conditions (high: 26 A vs. 30 B; low: 58 A vs. 54 B).

Table 6. 4. Trial structure in the two feedback conditions (FBK = feedback; CL = category learning)

	CL trials	**FBK trials for CL task**	**FBK %**	**Filler trials**	**Total trials**
High FBK	56	42	75%	84	140
Low FBK	112	42	37.5%	28	140

The 56 trials in the high feedback presented the same stimuli that were used in the corresponding 56 trials in the low feedback condition version of the task (matched trials). Additionally, the two conditions were matched to have 28 filer task trials in common; this was done to ensure that the two conditions were matched as far as possible.

Trials in both conditions were presented according to a fixed quasi-random order, the same for all participants, in order to avoid any possible order effects.

6.4.6 Procedure

The study method and procedure were approved by the Psychology Department Ethics Committee (DEC) at Goldsmiths, University of London. Testing was conducted in a testing room in the department and it took place in one session that lasted 1 hour and 15 minutes. Participants sat at a desk in front of a computer screen, a keyboard and stereo speakers.

On entering the testing room, participants were reminded that all the collected data and information would remain anonymous and be treated with confidentiality. They were also informed that they had the right to withdraw from the study at any time, for any reason. In order to guarantee anonymity, each participant was allocated a unique ID code that became the only means of identifying the data.

Participants were randomly assigned to one of the two conditions. They were presented with written instructions on the computer screen. The instructions informed participants that a series of stimuli would be individually presented on the screen. Their task was to learn, by trial and error, to classify these stimuli into category A and category B. They had to do so by pressing the 'z' key or the'/?' key for category A and category B respectively. They were also informed that in order to facilitate their performance, they would receive visual feedback on whether their response was either correct or incorrect. In fact, the computer would flash the word 'correct' when the participant's response was correct or 'wrong' when their response was wrong. These two messages were accompanied by noise feedback. Additionally, they also received monetary feedback. In fact, for each correct response participants were rewarded with the gain of 5 pence and did not lose any money for incorrect responses (i.e. it was a reward-only condition). Each time they made a correct response, they were presented with the total amount of their winnings. Hence, a promotion focus was used in the task (Maddox, Baldwin & Markman, 2006; Markman, et al., 2005).

They were also informed that feedback was not always available and, therefore, in some of the trials they would not receive any feedback following their response. The researcher verbally stressed the fact that the lack of feedback was completely independent of the participants' performance and was simply devised to make the task more complicated. The instructions also introduced the filler task, so participants were informed that from time to time the computer would have presented them with written instructions that would have requested them to press either the category A or the category B key. They were invited to simply follow the instructions. They were also informed that no feedback would have been given during these trials. However, in case of an incorrect key press (i.e. any key other than the two being used; ?/ and 'z') they would receive a 'wrong key' message.

Participants were invited to sit in front of the computer in a comfortable way, in order to move as little as possible during the trials. Participants were encouraged to express any doubts about their understanding of the task and to start only once they felt completely confident about the task instructions. At this point, the experimenter left the testing room to let participants carry out the task on their own.

6.4.7 Data analysis

A data extraction programme was developed in Matlab to create an overall summary file that could be exported from Matlab into SPSS and, therefore, allow statistical analysis on the data. The main extracted measure was accuracy indexed by the proportion of correct responses (pc) calculated both across the overall task and on the matched 56 trials. Additionally, formal modelling was also applied to these trials to explore participants' performance and establish the type of strategy that participants used compared to the strategy they were expected to use.

These 56 matched trials consisted of all the real task trials experienced by the participants in the high feedback condition and only half of the trials for those in the low condition. Because of the way the task was programmed, the stimuli on these trials were identical across the two conditions.

One other aspect of the design is worth noting. In the high feedback condition participants received the total amount of feedback trials (i.e. N = 42) but in the matched trials in the low feedback condition they received only 36 feedback messages. This difference was due to the fact that the low feedback condition was programmed first and then the same feedback sequence was applied to the high condition. However, since some of the real trials in the low condition were actually filler trials in the high feedback the sequence had to be slightly modified for the high feedback condition. Nonetheless, across the whole task both groups received 42 feedback signals. The key feature is that both the high and low feedback conditions offer the same amount of feedback signals (42) over the whole task, which contained the same number of trails (140) and, hence, lasted the same time in the two conditions owing to the filler trials. The critical difference between the two conditions is the feedback frequency. Indeed, in the low feedback condition, the feedback trials were interspersed with larger number of non-feedback trials compared to the high feedback condition.

6.5 Results

6.5.1 Behavioural data

A t-test was carried out to assess whether the accuracy scores, across the overall task, varied between the high and low feedback frequency conditions. The t-test showed that there was no statistical difference between the two feedback conditions [t (62) = .482, ns].

However, as overall levels of performance were quite low (proportion of correct responses equal to 59% and 57% for the low and high feedback condition, respectively), it was decided to re-run the analysis after removing the 'non-learners'. Participants were identified as 'learners' when they showed a proportion of correct responses significantly greater than or equal to chance scores. This was done using the binomial distribution which allows one to calculate for a single participant the minimal levels of performance that should be regarded as significantly ($p < 0.05$) above chance (0.5) across the whole task. Participants in the low feedback condition were identified as 'learners' when they showed a proportion of correct responses equal or greater than .58, whereas those in the high feedback condition were identified as learners if they achieved accuracy scores equal or greater than .607. The filter indicated that 27 participants from the whole sample size (N = 64) performed above chance levels, i.e. 41% of the sample performed above chance overall. These results indicate that the task was too hard for participants to perform optimally. Thus, they suggest that the task was not ideal to explore the impact of partial feedback on learning since the learning rate is at or close to chance level. Equivalent numbers of participants from the two feedback conditions were classified as learners: 14 from the low and 13 from the high feedback conditions. All the analyses reported below are based on 'learners' performance.

Once 'non-learners' were removed, results indicated a small difference between the two feedback frequency conditions that did not reach statistical significance [t(25) = 1.55, p = .134, 2-tailed]. Participants in the low feedback condition seem to perform numerically better than those in the high feedback condition as they made a higher proportion of correct responses (pc; table 6.5). This difference could be an indirect result of the fact that participants in the low feedback condition had twice as many trials to learn the strategy and improve their performance. Hence, extensive practice on an RB task could have enhanced performance.

Table 6.5. Proportion of correct responses (pc) after excluding non-learners

	Condition	**Mean**	**SD**
PC_overall	**Low (N=14)**	.737	.056
	High (N=13)	.701	.065

It was subsequently decided to test whether this difference would be significant (for participants who had received equivalent amount of feedback) when participants were responding to the same stimuli. In order to test this hypothesis, it was necessary to compare participants' performance on the 56 trials that were equivalent for participants in the high and in the low feedback condition (i.e. matched trials).

As summarised in table 6.6, the performance of learners did not differ on the matched trials across the two conditions [t (25) = .410, p = .68]. In other words, performance does not seem to be significantly enhanced by higher feedback frequency.

6.6. Proportion of correct (pc) responses comparing performance on the matched trials and after excluding non-learners

	Condition	Mean	SD
PC_matched	**Low (N=14)**	.711	.06
	High (N=13)	.701	.06

Post-hoc power analyses were carried out using the G*power program (Buchner et al., 1997). It was necessary to enter three types of information, the alpha value (.05 for a one-tailed test), the sample size of the two groups (n_1 = 14 and n_2 = 13) and the effect size (d). According to Cohen's effect size conventions, the d value was set equal to 0.3 since a 'small' to 'medium' effect was expected. Results showed that the power was very low [power(1- beta) = 0.1876] which suggests that the test may have not been powerful enough to detect any difference between the two groups. The loss of power is due to the small sample size retained once non-learners were removed.

6.5.2 Association between performance and covariate (CV) measures

Correlations were run to assess any possible relationship between CV measures (i.e. age and gender) and performance on the task, indexed by the correct proportion of responses on the overall task (pc_overall) and on the matched 56 trials (pc_matched). Given the similarity of performance across high and low feedback condition, the two conditions were pooled.

Table 6.7, shows that there is a negative correlation between age and the overall proportion of correct responses for the overall task (r = -.41, 0 = .035) and the matched trials (r = -0.46, p = .017). The results indicate that older participants made a higher proportion of correct responses than younger participants.

Table 6.7. Correlations between CV measures, WM scores and accuracy levels (i.e. proportion of correct responses, pc) for all trials and the matched trials

Correlations

		wmtestcorr	pc_matched	pc_overall
gender	Pearson Correlation	.044	-.312	-.302
	Sig. (2-tailed)	.828	.114	.125
	N	27	27	27
age	Pearson Correlation	.117	-.457	-.408
	Sig. (2-tailed)	.562	.017	.035
	N	27	27	27

Contrary to COVIS which states that performance on an RB task is dependent on executive attention and working memory (WM), the scores on the WM task were not found to be significantly correlated with the accuracy scores across the overall task (r = .13, ns) or across the matched trials (r = .03, ns).

6.5.3 Performance and personality

Four between-subject ANOVAs were carried out to explore the impact of the four personality components and the feedback manipulations on the degree of learning, measured as accuracy levels by the proportion of correct responses (pc). Because of the observations made above, it was decided to run the ANOVA using the pc on the matched trials as the DV. Results showed that the 2-way interaction between Positive Schizotypy (PS) and feedback condition just failed to be significant [F(1,23)= 3.86, p = .062]. None of the other personality components was found to have any affect on accuracy [Fs < 1.4, p > .24].

In order to explore this interaction, a multiple regression was run with PS as a predictor of pc scores on the matched trials after applying a split file on the feedback condition variable. Results indicated that the model was not a good predictor of the DV variance in the high feedback condition [F(1,11) = 1.035, ns] but it explained a significant 24% of the DV variance [F(1,12) = -2.245, p = .044] in the low feedback condition. Positive schizotypy scores contributed 54% of unique variance in the low feedback model. Results indicated that individuals with low scores on the PS component showed greater accuracy (higher pc scores) than their high counterparts (beta = -.54).

6.5.4 Model fitting

As described in the method section, the stimuli presented in the task varied on two dimensions, i.e. Gabor frequency and angle. The relevant dimension was the frequency of the Gabor sine-waves hence, participants were expected to use this dimension to perform effectively on the task. Nonetheless, it is possible that participants may have used the angle dimension or a combination of the two dimensions, or worse case scenario the may have been guessing throughout the task.

Hence, the data had to be fitted by 5 different models, these were:

1 – a guessing model, which has 1 free parameter
2a – a unidimensional rule model, which uses the frequency dimension (2 parameters)
2b – a unidimensional rule model, which uses the angle dimension (2 parameters)
3 – a two-dimensional rule model, which uses a conjunctive rule (4 parameters)
4 – a two-dimensional rule model, which uses an information integration (II) rule (3 parameters)

All five models were applied to each participant's data using maximum likelihood methods. Each of the models 2a-4 estimated two types of parameters, which were: decision criterion in the relevant dimension and 'noise'. The decision criterion is a boundary, in the perceptual space occupied by the stimuli, which may be being used by a participant to separate stimuli in category A from those in category B. In a single dimensional model it is the value on that particular dimension which is used to separate categories. In a conjunction rule model two such boundaries are required, one for each dimension, so that category A might for example be described as being above value x on dimension 1 and also above value y on dimension 2. For a two-dimensional II model the decision boundary is a line (specified by 2 parameters: a particular slope and intercept) in the plane representing the two dimensions upon which the stimuli vary. The noise parameters are the standard deviation of normal distributions. Noise is assumed to exist in perceptual processes (i.e. a participant does not have perfectly accurate perception of where a particular stimulus lies in perceptual space) and in decision processes (a participant may place their decision boundary with some degree of unintended variation from trial to trial). A single noise parameter represents the combination of these independent sources of noise variance. Each model is initially compared to a saturated model which has no free parameters and describes the data perfectly.

The goodness of each model was compared to the goodness of fit of the other models by comparing the Akaike Information Criterion[6] (AIC; Dayton, 2003). The AIC score was calculated on the basis of the free parameters and it is an estimate of the goodness of fit. It penalises the model with extra free parameters so that the lower the AIC score, the better the fitting (i.e. closer to the saturated model; Maddox, Ashby and Bohil, 2003).

The models were fitted to 56 trials in total. The trials take into considerations were the overall 56 of the high feedback condition and the corresponding 56 matched trials in the low feedback condition. This trial matching for analysis purposes was possible since the stimuli presented on the matched 56 trials were equivalent (showed exactly the same Gabor patches) across the two conditions. The decision to apply the data fitting only to the matched trials was driven by the possibility that model fitting discrepancies would be more detectable if you used the greater number of task trials in total that were collected from participants in the low feedback condition. Using the matched trials allowed us to compare pc scores on equivalent trials that showed the same stimuli in both conditions. Moreover, compared to other solutions, the use of the matched trials allowed us to include all feedback trials and, therefore, be able to compare performance on the main IV, feedback frequency.

However, the analysis is slightly conservative since in the high feedback condition participants received the total amount of feedback (i.e. N = 42) but in

---

[6] AIC = 2r – 2lnL, where r is the number of free parameters and L is the log likelihood of the model (Maddox et al., 2003)

the matched trials in the low condition they only received 36 feedback messages. As already noted, this difference is due to the fact, that the low feedback condition was programmed first and then the same feedback sequence was applied to the high condition. However, since some of the real trials in the low condition are actually filler trials in the high feedback the sequence had to be slightly modified for the high condition. Fitting was applied to all 64 participants in the attempt to obtain greater insight into participants' strategies during the task.

6.5.4.1 Fitting results

1- Fitting using a guessing model

The guessing model describes the situation where participants are randomly assigning the stimuli to one of the two categories, regardless of the stimuli's dimensions. This model has only one free parameter (guessing probability for one of the category responses, from 0 to 1). Results showed that the guessing model was significantly poorer than the saturated model in all but 5 of the 64 total cases (7.8%). This confirms that most of the subjects were doing something other than guessing.

2 - Fitting through a uni-dimensional rule model

The uni-dimensional models reported below describe the situation in which participants categorise the stimuli into two categories in relation to a decision criterion set on one of the varying dimensions. After setting a criterion on one of the dimensions (e.g. frequency), participants assign the actual stimulus to one or the other category depending on whether the stimulus value on the relevant dimension exceed or is below the set criterion. Since the Gabor patches vary on two dimensions (spatial frequency and orientation), there are two possible uni-dimensional (UD) rules that could be implemented and, thus, two UD models were fitted. Each UD model has two free parameters, which are: decision criterion and noise (see below).

2a- Fitting through a unidimensional rule model based on the spatial frequency dimension

The present model describes the optimal situation where participants select spatial frequency as the relevant dimension and place a criterion on this dimension to categorise the Gabor patches. This would be the optimal rule since spatial frequency was set as the relevant dimension in the program.

Results indicated that the uni-dimensional model related to Gabor frequency was not significantly worse than the saturated one for 22 cases. This indicates that 34% of the whole sample used frequency as a relevant dimension to classify the Gabor stimuli.

2b- Fitting through a unidimensional rule model based on the spatial orientation (angle) dimension

The present model describes the situation where participants select spatial orientation as the relevant dimension and place a criterion on this dimension to categories the Gabor patches even though spatial orientation was not the relevant dimension in the actual task. Results of this fitting indicated that this UD model was not significantly worse than the saturated one for 19 cases. This indicates that 30% of the whole sample used angle as the relevant dimension to classify the Gabor stimuli.

Comparison of the two uni-dimensional models

Overall, it seems that a uni-dimensional model fits 40 out of the 64 participants (i.e. 62.5% of the overall sample). Table 6.8 indicates that data from one of the 40 participants who used a uni-dimensional rule was well fitted by both the angle and frequency models.

Table 6.8. Preliminary results obtained following the fitting of both uni-dimensional models (frequency vs. angle)

	Angle dimension	Frequency dimension	Both dimensions
N (tot= 64)	18	21	1

It was decided to compare the fitting of the two models to assess which one offered the best fitting for this participant. In order to do so the AIC values of the two models were compared. The model with the lowest AIC value was selected and identified as the best fitting uni-dimensional model.

Following AIC comparisons, 21 participants were found to use frequency as the relevant dimension as opposed to 19 participants who were using the irrelevant dimension (i.e. angle) as the relevant dimension to base their decisions. Finally, 24 participants were not fitted by a uni-dimensional model (table 6.9).

Table 6.9. The proportion of participants using either angle or frequency as the relevant dimension to categorise the Gabor stimuli

	Frequency	Percent	Valid Percent	Cumulative Percent
No rule	24	37.5	37.5	37.5
Angle	19	29.7	29.7	67.2
Frequency	21	32.8	32.8	100.0
Total	64	100.0	100.0	

Five of the 24 participants not fitted by any of uni-dimensional models were guessing. However, performance of 19 participants was not fitted by a uni-dimensional rule, which means that for 32% of the whole sample was not using any of the above fitted models. This suggests that they might have been using alternative strategies which are not based on the two dimensions (i.e. non-dimensional rules; e.g. sequence of key presses on the keyboard). Indeed, participants verbally reported having used alternative strategies following debriefing. None of the fitting models developed here accounts for non-dimensional rules since learning was defined as using the relevant dimension to classify the Gabor patches. It is also possible that those 19 participants could have used a more complex strategy (e.g. a conjunctive or an II rule). This possibility is explored by models 3 and 4 which test whether participants may have implemented a conjunctive or an II rule, respectively, to categorise the stimuli. The results are reported below.

3 – Fitting through a 2-dimensional model, using a conjunctive rule

A conjunctive rule model captures the situation where participants categories the stimuli using information from the two dimensions. Individuals adopting a conjunctive rule set a decision criterion on both dimensions, for example 'if the stimulus value is above value x_1 on dimension 1 (criterion 1) and below value x_2 on dimension 2 (criterion 2), then the stimulus is a member of category A; otherwise it is a member of category B')'. Hence, the information from the two dimensions is combined after the criteria are separately set on each dimension (post-decisional). The conjunctive model fitted below aims to assess whether some participants used information from both spatial frequency and orientation to categorise the Gabor stimuli.

Results indicated that the conjunctive model fitted data from 27 participants non-significantly worse than the saturated model. However, the 27 data files fitted by the conjunctive rule were all also fitted by one of the uni-dimensional models. It was decided to compare the conjunctive model to the best-fitting uni-dimensional rule. Comparison of the conjunctive and the uni-dimensional models was carried out by comparing their corresponding AIC values. The model with the lowest AIC value was identified as the best fitting uni-dimensional model. These comparisons showed that the conjunctive model was better than the uni-dimensional model for four of the 27 data sets compared. Table 6.10 below indicates that after fitting the conjunctive rule model, 17 participants were using the angle dimension, 19 used the frequency (relevant) dimension and 4 used both dimensions (in a conjunctive fashion) to formulate their decision rule.

Table 6.10. Frequency of participants implementing one of the two uni-dimensional rules or the conjunctive rule

	Frequency	Percent	Valid Percent	Cumulative Percent
No rule	24	37.5	37.5	37.5
Angle	17	26.6	26.6	64.1
Frequency	19	29.7	29.7	93.8
Conjunctive	4	6.3	6.3	100.0
Total	64	100.0	100.0	

4 – Fitting through a 2-dimensional model, using an II rule

The II model describes the situation where participants apply an information-integration (II) strategy to determine category membership of the different stimuli. As in the case of a conjunctive rule model, when using an II strategy participants combine information from the two dimensions but they do so at a pre-decisional level (i.e. implicit learning).

Results indicated that the II two-dimensional model was worse than the saturated model in all but two cases. The p values corresponding to these two cases were equal to .07 and .13. However, these two cases were also fitted by the angle uni-dimensional model. It was decided to compare the goodness of fit of the relevant uni-dimensional model with the goodness of fit of the two-dimensional one.

Comparisons were made by contrasting the AIC values for these two models. Results showed that the uni-dimensional model fitted the data better than the II model, hence table 6.10 above represents a good summary of the fitting results from these analyses. 40 participants (62.5% of the sample) used a uni-dimensional rule to learn the task whereas 24 (37.5%) participants failed to use an effective dimensional rule.

The results from the model fitting were calculated in relation to feedback condition and this indicated that the strategy used was partially affected by the feedback condition participants had been allocated to ($\chi^2_{(3)} = 13.2$ p = .004). Indeed, participants in the high feedback condition were less likely to apply a dimensional rule to solve the categorisation task but when they did, they were more likely (9/14) than participants in the low feedback condition (10/26) to use frequency as the relevant dimension (i.e. to use the correct rule). Moreover, they selected this dimension more frequently than the spatial orientation one, which was actually irrelevant. In contrast, almost equivalent number of participants in the low feedback condition used the relevant or the irrelevant dimension to formulate their rule. Tables 6.11 and 6.12 below summarise the strategy used by participants in the two feedback conditions.

Table 6.11. Summary of the strategy used in the low feedback condition

	Frequency	Percent	Valid Percent	Cumulative Percent
No rule	6	18.8	18.8	18.8
Angle	14	43.8	43.8	62.5
Frequency	10	31.3	31.3	93.8
Conjunctive	2	6.3	6.3	100.0
Total	32	100.0	100.0	

Table 6.12. Summary of the strategy used in the high feedback condition

	Frequency	Percent	Valid Percent	Cumulative Percent
No rule	18	56.3	56.3	56.3
Angle	3	9.4	9.4	65.6
Frequency	9	28.1	28.1	93.8
Conjunctive	2	6.3	6.3	100.0
Total	32	100.0	100.0	

As observed by comparing table 6.11 and table 6.12, participants in the low feedback condition were significantly more likely to use a dimensional strategy than those under high frequency feedback [$\chi^2(1) = 9.6$, $p = 0.02$]. However, participants receiving low frequency feedback were also more likely to implement the irrelevant angle dimension to formulate the optimal rule; however, this difference just failed to be significant [$\chi^2(1) = 3.6$, $p = 0.06$].

Participants who employed the correct dimension to formulate the uni-dimensional rule were significantly more likely to achieve accuracy scores above chance level than those who employed any of the other two strategies [i.e.

incorrect dimension or conjunctive strategy; $\chi^2_{(3)} = 44.70$, $p < .001$], as summarised in table 6.13.

Table 6.13. Proportion of participants using one of the three dimensional strategy across learners and non-learners

Count

		Accuracy scores above chance		
		.00	1.00	Total
rule	No rule	19	5	24
	Angle	17	0	17
	Frequency	0	19	19
	Conjunctive	1	3	4
Total		37	27	64

When considering the raw data, learners were identified as those participants who performed above chance during the task. According to this method, twenty-seven participants had been identified as learners whereas the model-fitting analysis indicated that only 22 participants used a dimensional strategy to classify the stimuli. This suggests that 5 of the 'learners' had managed to score above chance either using an alternative, non-dimensional strategy. Moreover, only 19 of the 22 individuals who scored above chance and used a dimensional rule actually used the appropriate uni-dimensional rule based on the frequency dimension. Three of them used a conjunctive rule which also allowed them to have higher accuracy levels possibly because they set the criterion correctly on the frequency dimension and only added some noise by combining this criterion with the criterion set on the irrelevant dimension. In other words, 19 of the 27 individuals who performed above chance used the correct rule, three used a conjunctive rule and the remaining 5 used a non-dimensional strategy.

Seventeen of the 37 participants who performed below chance had used the incorrect uni-dimensional rule based on angle and one used a conjunctive rule. Hence, 18 non-learners used an incorrect dimensional strategy whereas 19 of them must have used a non-dimensional strategy, which was sub-optimal.

The relationship between the type of strategy used and the proportion of correct responses (performance above chance) can be explained by considering the accuracy performance levels that could have been obtained using one of the three rule strategy throughout the task (i.e. the matched 56 trials). Indeed, if participants had implemented the optimal rule which relied on frequency as the relevant dimension, they could have obtained an accuracy score equal to 95%. The reason why 100% accuracy could not be reached is due to the fact that some of the trials (i.e. 3 trials of the total 56) were wrongly assigned to categories and did not follow the optimal rule (see table 6.2). Participants using the conjunctive rule (i.e. 'if frequency > .065 and angle > 90, then respond B') could have obtained an accuracy level of 68%. Therefore, accuracy levels obtained using either of these two strategies would have allowed participants to perform above chance (i.e. above the 61% cutoff employed in the high feedback condition). However, a uni-dimensional rule using angle as the relevant dimension could only allow one to reach an accuracy level equal to 52%, which is below the individual subject above-chance cut-off level.

6.5.5 Individual differences and strategy used

Further analyses were carried out to assess whether there was a relationship between personality and the strategy used, as indexed by the best fitting models.

As a preliminary analysis it was decided to assess whether the use of a dimensional strategy was associated with any of the personality components extracted in the PCA, scores on the WM task and accuracy scores as indexed by the proportion of correct responses on the matched trials (pc_matched). The variable that coded whether participants had used a dimensional rule or not was labelled strategy used and it coded the use of a dimensional rule as 1 and no dimensional rule as 0. In order to explore any possible relationship, a few point-biserial correlations were run to explore the data.

Results showed that there was a positive correlation between the number of correct responses in the WM and the strategy used ($r = .26$, $p = .035$). These results suggest that participants who perform better on the WM task were more likely to use a dimensional rule than those who scored low. These results are in line with the COVIS model that suggests that performance on an RB task requires logical reasoning and hypothesis testing which relies on the WM task. However, WM scores were not significantly associated with the UD implemented (angle vs. frequency coded as 1 and 2, respectively; r = .21, ns).

Results also indicated that the interaction between the strategy used (yes = 1 or not = 0) and neuroticism just failed to be significant (r = -.24, p = .060). There was also a trend for a negative correlation between positive schizotypy factor and the strategy used (r = -.21, p = .09). Finally, there was a positive correlation between the pc scores on the matched trials and the strategy used (r = .37, p = .003), which indicated that the use of a dimensional strategy as opposed to guessing and/or using non-dimensional strategies was associated to higher accuracy levels.

Following, the correlational results, a logistic regression was run with N, PS and WM scores entered as predictors and strategy used as the DV. Results suggest that the model entered is a good predictor of the strategy used ($\chi^2_{(3)} = 11.71$, p = .008) and that the N and WM covariates were good predictors in the model (N: Exp(B) = .54 $\chi^2_{(1)} = 4.80$, p = .028; WM: Exp(B) = 1.1 $\chi^2_{(1)} = 4.09$, p = .043). Positive schizotypy was found to be a weaker predictor in the model (Exp(B) = .63 $\chi^2_{(1)} = 2.41$, p = .12, 2-tailed). An independent t-test was run to explore the relationship between N scores (DV) and the strategy used (IV). Results showed that there was a non-significant trend for the effect of strategy on neuroticism [t(62) = 1.92, p = .060]. Indeed, participants who implemented a dimensional strategy scored lower on the N component (-0.32) than those participants who used a non-dimensional strategy (0.22).

Subsequently it was decided to assess whether personality and WM scores may have affected which one of the dimensional rule participants used as a relevant one during performance of the task.

As indicated in table 6.14, high Neurotics (cf. low N) were more likely to use the relevant dimension (frequency) to formulate their categorisation strategy whereas they did not differ in their choice of conjunctive rule. Hence, it was decided to run an independent-sample t-test to compare performance of high and low neurotics in relation to the two uni-dimensional models (angle vs. frequency). The t-test showed that neuroticism scores significantly differed among participants who used the relevant dimension and those who failed to do so [t(34) = 2.16, p = .038]. Indeed, participants who used the relevant dimension (frequency) scored higher on neuroticism (-.02) than those who used the irrelevant dimension (i.e. angle; N: -.77).

Table 6.14. Frequency of participants using one of the three dimensional rules (angle, frequency or conjunctive) across the two feedback conditions

Count

		Neuroticism		
		Low	High	Total
	Angle	12	5	17
	Frequency	9	10	19
	Conjunctive	1	3	4
Total		22	18	40

Finally, it was decided to compare performance on the task for those 19 participants who adopted the correct uni-dimensional strategy. Results showed that performance, measured by the pc on the matched trials, did not vary across the two feedback conditions [t(17) = .27, p = .79]. The proportion of correct responses on the matched trials did not correlate with any of the personality components taken into consideration in the study (p >.224).

6.6 Discussion

Preliminary analysis showed that participants in the low feedback condition did not perform statistically more poorly than those in the high feedback condition. These results are in line with COVIS that states that performance on an RB task should not be affected by feedback manipulations (Ashby et al., 1998; Ashby et al., 1999). However, it is possible that the study was not able to detect any difference between the two groups as an effect of feedback manipulation because of the low levels of power.

WM ability was positively associated with the correct use of a dimensional strategy, which is in line with COVIS' assumption that performance on an RB task relies on WM and executive attention abilities.

Fitting behavioural data with the formal model offered a greater insight into the strategy used by participants to categorise the stimuli during the task. In particular, it showed that a few of those participants who scored above chance may have not necessarily used the correct strategy. These results indicate that using a specific CL task does not necessarily predict the strategy that participants will actually use. Hence, it further supports evidence which highlights the importance of using formal models to obtain clearer and more reliable information about the task (Gluck, et al., 2002; Tharp, 2007)

Overall high neurotics were more likely to use the relevant dimension and so implement the correct rule. This may be due to the fact that the low feedback frequency condition may be perceived as punishing. Therefore, the regulatory fit between a BIS trait (N) and (punishing) partial feedback manipulation could enhance performance. Indeed, Maddox and colleagues (2006) have observed that regulatory fit enhances cognitive flexibility which, subsequently, facilitates performance on simple CL tasks which require a simple solution (e.g. uni-dimensional rule).

Results also found a weak trend for participants who scored low on the PS trait component to be more likely to use a dimensional rule than their high counterparts. Among the general population PS scores reflect behaviours and thoughts with resemblance to the positive symptoms typical of SZ. One of the main characteristics of SZ patients is impaired executive attention abilities which often result in impaired filtering of irrelevant information (e.g. dimensions and/or stimuli). One source of evidence for this claim is provided by studies that explored SZ patients and healthy controls performance on LI tasks. Results indicated that LI was reduced in SZ patients as opposed to controls and also in high PS participants (cf. low PS participants; see Pickering and Gray, 2001 for a review). It is possible that high PS individuals find it harder to filter out irrelevant information of the task (e.g. trial sequence) or other features of the stimuli (irrelevant dimension) and consider them to be relevant for efficient categorisation. Participants with high scores on PS may consider irrelevant information presented in the filler task as potentially relevant to classification, for example the sequence of filler trials or the instructions presented on the screen during this task (e.g. 'press the category A key').

Some of the PS measures are correlated with ImpAss measures (Pickering, 2004). In the present study positive schizotypy was found to be negatively correlated with the strategy used. These results suggest that individuals with high scores on this component were less likely to use a dimensional strategy to solve the task. Individuals who score high on ImpAss traits have been found to show superior performance than their low counterparts when performing on simple uni-dimensional tasks of the type administered in this study (Pickering, 2004; Tharp, 2007). However, it is possible that the Gabor patches presented in this study may be more complex than the stimuli used in the previous ones. This is suggested by the fact that participants performed around chance level.

In fact, even though the stimuli presented in the previous studies also varied on two dimensions (with only one of them being relevant) the dimensions that were used were perhaps more salient and easier to identify. For example, in Tharp study the stimuli varied on 4 binary valued dimensions and the relevant dimension was background colour. In Pickering's studies, the relevant dimensions were height of the rectangle or the location of an inner line. Although, the angle variation of the Gabor is quite salient the frequency of the waves is more complex to perceive. Some participants reported perceiving the Gabor patches as tri-dimensional (not just 2D) and they used the shade between the waves as the relevant dimension which is a slightly more complex perception of the stimulus variations than the simple line frequency. The shade is a by-product of the distance (frequency) of the lines as well as grey variation within the stimulus. Hence, the stimuli used may have been perceived to be complex and requiring a more complex rule. High ImpAss perceiving the task as complex could have shown impaired performance due to cognitive inflexibility that is typical of high ImpAss on more complex tasks (Tharp, 2007).

Participants may have also perceived the task to be complex because of its structure and in particular because of the inclusion of the filler task trials. During debriefing, it was reported by some participants that they had perceived the presence of the filler task to be relevant for the main categorisation task. Verbal reports by participants stated that they had developed rules related to features of the filler task. This occurred despite explicit instructions that the filler task and its features were irrelevant to task performance.

The fact that participants may have experienced the task as complex explains why four participants employed a more complex conjunctive rule and it also suggests that they may have used a different, non-dimensional strategy. In fact, they may have resorted to a more complex rule and eventually to procedural learning. This would not be surprising since previous research has found that an II rule may be employed to solve complex RB tasks (Maddox, Filoteo, Hejl & Ing, 2004; Ashby & Maddox, 2005). It is possible that a more complex strategy would be employed if the individual assumed that a complex rule is more appropriate.

6.7 Information-integration study

6.7.1 Aim of the study

The second study of this chapter aimed to explore individual differences when performing on an alternative version of the task. Because of limited cognitive flexibility, and/or a preference for simple uni-dimensional task solutions, high ImpAss participants are expected to be more impaired than their lower-scoring counterparts according to results obtained in previous studies (Tharp, 2007). Simultaneously, the study aimed to explore whether induced regulatory fit between personality trait and feedback manipulation can enhance performance during an II task.

Additionally, the task compared performance under high and low feedback frequency manipulation to investigate how feedback manipulation affects learning on an II task. According to the results obtained by Ashby et al. (1999), it was predicted that partial feedback should impair learning in the II task, especially under the low feedback condition. The study aimed to assess whether participants will automatically interpret no-feedback trials as 'non-reward', despite the fact that they were explicitly told that the no-feedback trials were unrelated to their performance level. The reason why no feedback was expected to be interpreted as non-reward relies upon one of COVIS main assumptions that feedback is automatically processes by the implicit learning system of the model (Ashby et al., 1998; Maddox & Ashby, 2005).

Finally, the study explored the relationship between performance on the II task and WM scores. COVIS suggests that performance on the II task is dependent on the implicit system which relies on procedural learning rather than more explicit hypothesis testing processes (Ashby et al., 1998; Ashby, Queller & Berrety, 1999; Maddox & Ashby, 2004). However, there is recent evidence that indicates that WM scores mediate performance on II tasks. De Caro and colleagues (2008) showed that scores on a WM task were negatively associated with performance on an II task. Indeed, participants who scored high on the WM task required more trials to reach the learning criterion on the II task. De Caro and colleagues argued that individuals with high WM capabilities might be engaging the explicit system rule system to test complex but ineffective rules and so perform less well than low WM participants who were relying more upon procedural learning better suited to performing the II task well. In contrast, Tharp and Pickering (under review) found that high WM scores were associated with greater accuracy scores and the use of a multi-dimensional (MD) strategy during the identical II task to that used by De Caro and colleagues. They concluded that De Caro and colleagues' results might be a product of procedural inadequacies concerning the learning criterion adopted in their study.

6.8 Method

6.8.1 Participants

An opportunity sample was drawn from the student population at Goldsmiths (University of London). Participants were recruited through the research participation scheme run by the psychology department and they all received 3 course credits for their participation. The sample was constituted of 76 psychology students, 66 females and 10 males. All students were aged between 18 and 48 (mean age 21.04 years; s.d. = 5.53). Participants were guaranteed confidentiality. They were tested in one sitting that lasted approximately 1 hour and 15 minutes.

6.8.2 Design

As in the study with the rule-based task, the main independent variable was feedback frequency which was treated as a between-subjects factor with two levels: high and low. In the low feedback frequency condition, participants received feedback on 37.5% of their responses, whereas in the high feedback frequency condition participants received feedback on 75% of their responses. The study was interested in assessing the impact of the IV on performance during the II task, which was assessed by considering the proportion of correct responses during the category learning task. In addition to course credits, participants could also earn money depending on their performance. They won 5 pence for each correct response and lost no money for incorrect ones. This was done to maintain a parallel with the RB version of the task and to have participants perform under a reward-only condition and be induced with a promotion focus. Participants were randomly allocated to the high or low feedback condition.

6.8.3 Stimuli and materials

The stimuli for the information-integration task were generated by rotating the original stimuli by 45° and then shifting the spatial frequency and spatial orientation by an amount that resulted in a d' equal to 3, as for the rule-based task (see table 6.15).

Table 6.15. Mean, standard deviation and covariate of category A and category B

		Mean	SD	Covariance
Category A	**Angle**	108	30	0.251
	Frequency	0.05	0.015	
Category B	**Angle**	72	29	0.272
	Frequency	0.08	0.014	

As in the RB task, the values used to draw both the spatial frequency and the orientation of the Gabor patches was rounded by two decimal points, which should render the task perceptually easier. This was done to compensate for the difficulty of the task, lack of 100% accuracy feedback and the inclusion of the filler task. The stimuli distribution for the II and the RB task are presented below in figure 6.1 below.

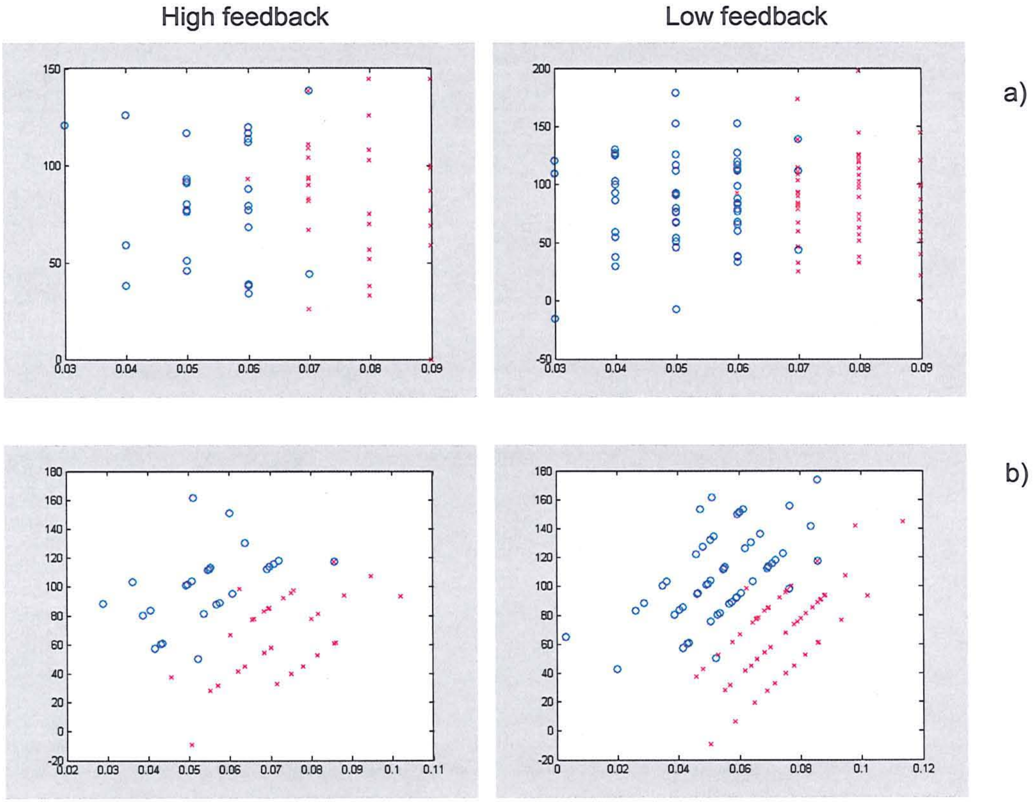


Figure 6. 1. The stimuli presented in the a panel (top) represent the stimuli created for the RB task while those in the b panel represent the stimuli created for the II task. Category A stimuli are represented by the blue circles whereas category B stimuli by the red crosses. The left-hand side panels represent the stimuli drawn for the high feedback condition and the right-hand side panels represents those stimuli drawn for the low feedback condition.

The Gabor stimuli were produced and presented in the same fashion as in the rule-based task by using Matlab routines from Brainard's (1997) Psychophysics Toolbox.

Owing to the fact that II tasks are harder than RB tasks, the number of total trials was doubled so that the whole task consisted of 280 trials. The ratio of feedback, no-feedback and filler task trials was identical to the ratio in the rule-based task (table 6.16).

Table 6.16. Ratio of feedback, no-feedback and filler task trials out of a sample of 10 trials

	Feedback	**No feedback**	**Filler task**
High	3	1	6
Low	3	5	2

The low feedback condition consisted of 224 trials presenting CL stimuli while the high condition presented half as many trials (112). The 112 trials presented the same stimuli as the matching 112 trials in the low feedback condition.

As in the RB version of the task, the original 280 trials were created to have equal numbers of stimuli sampled from the two category distributions. Once the filler task trials were introduced there were still roughly equivalent numbers of the two category stimuli in both conditions (high: 52 A vs. 60 B; low: 116 A vs. 108 B).

6.8.4 Procedure

The procedure followed in the information-integration task was identical to the one adopted in the rule-based task. Participants were presented with exactly the same instructions used in the rule-based task.

6.8.5 Personality measures

Participants completed several personality measures, which were: the Eysenck Personality questionnaire revised (EPQ-R), the Oxford-Liverpool inventory of feelings and experiences scale (OLIFE), the sensation seeking scale (SSS), the schizotypal personality questionnaire (SPQ), the big five inventory (BFI) and the BIS/BAS questionnaire. Four personality components were extracted after running a PCA with Varimax rotation on these scales (see chapter 4 for details). The four components extracted are: Extraversion (E), Neuroticism (N), positive schizotypy (PS) and impulsivity-antisocial (ImpAss). The PCA was run on 232 participants, who constitute the overall number of participants tested. The components scores used in the present study are the corresponding scores extracted in the overall PCA involving 232 participants.

6.8.6 Data analysis

A data extraction programme was developed in Matlab to create an overall summary file that could be exported from Matlab into SPSS and, therefore, allow statistical analysis on the data. The main measure was accuracy, indexed by the proportion of correct (pc) responses. The study was mainly interested in the pc measures both on the overall task and on the matched 112 trials. Additionally, formal modelling was used to explore participants' performance and establish the type of strategy that participants used compared to the strategy they were expected to use.

These 112 trials consisted of all the real task trials experienced by the participants in the high feedback condition and only half of the trials for those in the low condition (the matched trials). Because of the way the task was programmed, the stimuli on these trials were identical across the two conditions. According to the observations made for the RB version of the task, it was decided to run the majority of the analyses on these matched trials.

6.9 Results

6.9.1 Behavioural data

A t-test was carried out to assess whether there was a significant difference in the overall proportion of correct responses between the high and low feedback frequency conditions. The t-test showed that there was no statistical difference between the two feedback conditions [t (74) = 0.37, ns].

However, it was decided to re-run the analysis after removing 'non-learners'. Participants in the high feedback condition were identified as 'learners' when they showed a proportion of correct responses significantly higher than chance scores.

Participants in the low feedback condition were identified as 'learners' when they showed a proportion of correct responses equal or greater than .55, whereas those in the high feedback condition were identified as learners if they achieved accuracy scores equal or greater than .58, across the whole task. Fifty-four participants (71% of the overall sample) were identified as 'learners' and were included into the analysis whereas 'non-learners' were excluded. There were approximately equal numbers of learners in the two feedback conditions, there were 25 in the high and 29 in the low condition.

Even once 'non-learners' were removed, results still indicated that there was no significant difference between the category learning success of participants in different feedback frequency conditions [t (52) = .162, ns]. Indeed, table 6.17 shows that the proportion of correct responses is virtually identical in the two feedback conditions.

Table 6.17. Proportion of correct (pc) responses on the overall task after excluding non-learners

	Condition	Mean	SD
PC_overall	**Low (N=29)**	.646	.052
	High (N=25)	.643	.045

As for the analysis of the RB version of the Gabor task, it was decided to test whether this difference would still be non-significant after participants had received equivalent amount of feedback and were responding to the same stimuli. In order to test this hypothesis, it was necessary to compare participants' performance on the matched 112 trials. The matched trials consisted of all the real task trials experienced by the participants in the high feedback condition and only half of the trials for those in the low condition. Because of the way the task was programmed, the stimuli on these trials were identical across the two conditions.

However, the analysis is slightly conservative since in the high feedback condition participants received the total amount of feedback (i.e. N = 84) but in the matched trials in the low condition they received only 72 feedback messages. This difference is due to the fact, that the low feedback condition was programmed first and then the same feedback sequence was applied to the high condition. However, since some of the real trials in the low condition are actually filler trials in the high feedback the sequence had to be slightly modified for the high condition.

As summarised in table 6.18, learners in the two feedback conditions did not differ across the two conditions [t (52) = -1.10, ns]. In other words, performance was not enhanced by higher feedback frequency.

Table 6.18. Proportion of correct (pc) responses comparing performance on the matched trials and after excluding non-learners

	Condition	Mean	SD
PC_matched	**Low (N=29)**	.629	.055
	High (N=25)	.644	.045

Post-hoc power analyses were carried out using the G*power program (Buchner et al., 1997). It was necessary to enter three types of information, the alpha value (.05 for a one-tailed test), the sample size of the two groups (n_1 = 29 and n_2 = 24) and the effect size (d). According to Cohen's effect size conventions, the d value was set equal to 0.3 since a 'small' to 'medium' effect was expected. Results showed that the power was low [power = (1- beta) = 0.2878] which suggests that the test may have not been powerful enough to detect any small to medium difference between the two groups. The loss of power is due to the relatively small sample size obtained once non-learners were removed. Obviously, if the effect size were bigger than that assumed above, then this sample size would have much higher levels of power.

6.9.2 Performance and covariate (CV) measures

Correlations were run to assess any possible relationship between CV measures (i.e. age and gender) and performance on the task, measured by the correct proportion of responses on the overall task (pc_overall) and on the matched 112 trials (pc_matched).

Neither gender nor age correlated with the accuracy scores on the overall trials or on the matched trials (p >.5). Performance on the WM task positively correlated with the proportion of correct responses both on the overall trials (r = .34, p = .013) and on the matched trials (r = .26, p = .056). These data apparently go against the COVIS assumption that performance on II tasks relies on procedural learning and, therefore, does not rely on WM. Tharp and Pickering (under review) also showed that accuracy levels achieved during performance on an II CL task was positively correlated with WM scores.

6.9.3 Performance and personality

A between-subject ANCOVA was carried out to explore the impact of the four personality components and the feedback manipulations (between-Ss IV) on the degree of learning, indexed by the proportion of correct responses (pc). Because of the observations made above, it was decided to run the ANCOVAs using the pc on the matched trials as the DV. The four personality components were entered simultaneously as covariates. Analyses included learners only. Results showed that the main effect of feedback condition was non-significant in the ANCOVAs [Fs < .62 , ps > .44]. However, there were significant main effects of N [F (1,49)= 4.80, p = .033] and ImpAss [F(1,49) = 5.68, p = .021]. After applying a median split on the personality variables, two separate independent-sample t-tests were run in order to explore these main effects. Individuals with high scores on the N components were more accurate (mean = .65, s.d.= .05) than their low counterparts [mean = .62, s.d. = .046; t(52) = -2.27, p = .028]. Similarly individuals with high ImpAss scores also showed a trend to be better (mean = .65, s.d. = .05) at the task than those with low score [mean = .62, s.d. = .05; t(52) = -1.80, p = .078].

A multiple regression was run with N and ImpAss scores as predictors of pc scores on the matched trials. The model explained a significant 17% of the DV variance [F(2,51) = 6.24, p = .004]. N scores contributed a greater proportion of unique variance [32%; t(51) = 2.44, p = .018] while ImpAss made a smaller unique contribution of 23%, although this contribution showed only a trend [t(51) = 1.73, p = .090].

6.9.4 Model fitting

Like the RB version, the stimuli presented in the task varied on two dimensions, i.e. Gabor frequency and angle. The optimal rule required participants to integrate information from both dimensions. Nonetheless, it is possible that participants may have used the angle or the frequency dimension individually to formulate a uni-dimensional rule which would lead to poor performance.

Hence the data had to be fitted by the same five models used to fit the RB task although the parameters values were modified to account for the fact that the stimuli used in the II version of the task were different from the ones used in the RB task as they were produced by rotating the original ones. The models were:

1 – a guessing model, which has 1 free parameter
2a – a uni-dimensional rule model, which uses the frequency dimension (2 parameters)
2b – a uni-dimensional rule model, which uses the angle dimension (2 parameters)
3 – a two-dimensional rule model, which uses a conjunctive rule (4 parameters)
4 – a two-dimensional rule model, which uses an information integration (II) rule (3 parameters)

The Gabor task used in this study is a complex information-integration task and, therefore, participants are expected to use an II strategy that requires them to combine information from both dimensions at a pre-decisional level. However, all five models were fitted because it is possible that the task was too hard for participants to derive the correct strategy and participants may have implemented a simpler rule (uni-dimensional or conjunctive; Gluck et al., 2002)

The models were fitted to the matched 112 trials in total, for the same reasons elucidated in the RB section. Basing the data analysis on the matched trials allowed us to compare pc scores on equivalent trials that showed the same stimuli in both conditions. Moreover, compared to other solutions, the use of the matched trials allowed us to include all feedback trials and, therefore, to be able to compare performance on the main IV, feedback frequency. Fitting was applied to all 76 participants in the attempt to obtain greater insight into participants' strategies during the task.

6.9.4.1 Modelling results

1- Fitting through a guessing model

The guessing model was always significantly worse than the saturated model; it did not fit any of the data sets well.

2a- Fitting through a uni-dimensional (UD) rule model based on the spatial frequency dimension

Results indicated that the UD model using spatial frequency as the relevant dimension did not lead to a significantly worse fit than the saturated model for 9 cases. This indicates that 11.8% of the whole sample used frequency as a relevant dimension to classify the Gabor stimuli.

2b- Fitting through a uni-dimensional rule model based on the spatial orientation (angle) dimension

Results of this fitting indicated that the UD model using the spatial orientation (i.e. angle) as the relevant dimension did not lead to a significantly worse fit than the saturated model for 14 cases. This indicates that 18.4% of the whole sample used frequency as a relevant dimension to classify the Gabor stimuli.

Results on the uni-dimensional model indicate that both angle and frequency were used as relevant dimensions, although angle was used more frequently. Indeed, 23 participants out of the whole sample (N = 76) used a uni-dimensional rule to establish the membership of the stimuli.

3- Fitting through a 2-dimensional model based on a conjunctive rule

Results showed that the conjunctive rule on both dimensions model was not significantly worse than the saturated model for 16 cases. However, 14 of these 16 cases fitted by the conjunctive rule were also fitted by one of the UD models. Hence, the goodness of fit of the best-fitting uni-dimensional model was compared to the goodness of fit of the conjunctive model. Comparisons were made using the AIC values of the two models. The results obtained from the comparisons showed that of the 14 overlapping cases, 10 were better fitted by the conjunctive model than the UD models whereas 4 were best fitted by the UD models (see table 6.19 for a summary).

Table 6.19. Data files fitted by the uni-dimensional and the conjunctive rule model prior and following comparisons

		Prior to comparison	Following comparisons
Rule	1-dimension	9	13
	Conjunctive	2	12
	Both	14	0

4 – Fitting through a 2-dimensional model based on an information integration (II) rule

Results indicated that the II two-dimensional model was worse than the saturated model in all cases, which suggests that the actual task may have been too difficult for participants to switch to a linear rule that contained both dimensions. Even though the fitting results showed that the II model was significantly worse than the saturated model, further comparisons were carried out between the II model and the 1-dimensional and conjunctive models. In fact, comparisons of the AIC values indicated that the II model seemed to fit the data better than the UD models for seven participants. Nonetheless, these 7 participants were best fit by the conjunctive model.

Overall, the modelling results show that 25 of the 76 participants tested were using a dimensional rule to learn to classify the Gabor task. Strategy users were evenly sampled from both feedback conditions (high = 12 & low = 13). Nonetheless, the dimensional rule was suboptimal since those 25 participants used a UD or a conjunctive rule rather than the optimal II rule.

Table 6.20 below shows that participants who scored below chance levels did not use any dimensional rule. However, these results do not explain how the remaining 29 participants managed to score above chance without using any dimensional strategy.

Table 6.20. Cross-tabulation of learners (pc above chance) with the strategy used variable

Count

		Accuracy above chance		
		No	Yes	Total
Dimensional strategy	No	22	29	51
	Yes	0	25	25
Total		22	54	76

6.9.5 Individual differences and strategy used

Further analyses were carried out to assess whether there was a relationship between personality and the strategy used (dimensional vs. non-dimensional), as indexed by the best fitting models.

As a preliminary analysis it was decided to assess whether the use of a dimensional strategy was associated with any of the personality components extracted in the PCA and/or scores on the WM task – as observed when analysing the raw data on the proportion of correct responses. In order to do so a few point-biserial correlations were run to explore the data. The only personality trait that was found to be significantly correlated with strategy used (dimensional = 1 vs. non-dimensional = 0) was N ($r = .305$, $p = .007$). The positive sign of the correlation means that high neurotics were significantly more likely to use a dimensional rule to perform on the task than their low counterparts. The use of a dimensional rule was also very positively correlated with the proportion of correct responses made ($r = .68$, $p < .001$)

However, there was only a weak trend for a positive correlation between WM and the pc scores on the matched trials ($r = .26$, $p = .056$).

6.10 Discussion

Analysis of the behavioural data indicated that even though it was a difficult task, fifty-four of the 76 participants (71% of the sample) performed above chance ('learners') and that superior performance was associated with higher scores on the N and ImpAss personality components. Contrary to COVIS predictions, higher performance on the II CL task was associated with higher performance on the WM task. Moreover, the feedback manipulation was not found to alter performance in the two groups. These findings also contradict the COVIS model which states that performance on II tasks is highly dependent on feedback. As in the case of the RB version of the task, it is possible that no effect was detected due to low power.

However, once the data had been fitted with the formal models it was apparent that participants who had performed above chance were not using the optimal rule (an II rule). In fact, 23 of those 54 participants performing above chance were found to use a uni-dimensional rule that treated frequency or angle as the relevant dimension. The modelling data also showed that two participants were using a conjunctive rule that combined information from both the spatial frequency and the angle dimension. The remaining 31 'learners' were not using a dimensional rule, which suggests they may have being basing their categorisation on aspects of the task or the stimuli which were not directly related to the two stimulus dimensions. As for the RB version of the task, they could have considered the filler task relevant to categorisation. Some participants reported that they thought there was logic in the alternating of filler and real trials (e.g. 'after two filler trials in a row, the Gabor stimulus belonged to Category A'). Other participants considered relevant the instructions given in the filler trials (e.g. 'If I were asked to press the category A key, the following (Gabor) stimulus was a member of category B'; or vice-versa). This occurred despite specific instructions that such rules would not apply.

Participants who adopted this type of alternative strategy reported having generally adopted (sub-optimal) uni-dimensional rules at the beginning of the task but they gave them up as trials went on as they did not seem to enhance performance. Since results indicated that uni-dimensional rules were associated with above chance performance, the preliminary use of a uni-dimensional rule and the subsequent use of an alternative strategy may account for those 29 participants performing above chance. However, it is possible that due to the changes of strategy as the task progressed, none of the formal models could offer a good fit of the data across the whole task.

It is hence possible that not only different strategies could have been applied to solve the task but also that the alternative dimensional rules were sufficient to perform above chance, as observed in the weather prediction task (Gluck et al., 2002). In fact, results showed that the conjunctive rule appeared to be an effective strategy as, indeed, the participants using a conjunctive rule during the RB task could reach accuracy levels above chance (68%; 69% in the II task). The use of the uni-dimensional rule that used angle as the relevant dimension, allowed participants to reach accuracy levels that were only at chance (i.e. 52%; 70% in the II task). Nonetheless, the accuracy level obtained using the irrelevant dimension may have been sufficiently high for participants to have the impression that their strategy was effective.

Participants may have stuck to the uni-dimensional strategy, because there were not enough resources (feedback signals) for the implicit system to take over and procedural learning to occur. It is possible that even the feedback frequency used in the high feedback condition (75%) was not sufficient for implicit learning to occur. The lack of trial-by-trial feedback may have rendered the task harder and, similarly, the introduction of the irrelevant filler task trials may have had a similar effect. As noted, the presence of filler task trials led some participants to use alternative strategies (e.g. task trial sequence). Finally, pc scores indicate that the use of an RB strategy worked quite well which may have been another reason why participants continued using it throughout the task.

After making these considerations, it is not surprising to find WM ability to be associated with the proportion of correct responses scored by learners, which is in line with the results showing that learners used explicit rules that require WM and executive attention ability. Additionally, it is possible that WM resources may have been also employed when using alternative strategies, e.g. a rule based on the filler-real trials sequence. Hence, the results indicating that WM was positively associated with performance on the II task do not necessarily contradict COVIS assumption that performance on II tasks requires the activation of the implicit system that does not rely on working memory or logical reasoning. The results, instead, are consistent with the fitting findings that indicate that participants solved the II task using a uni-dimensional rule, a strategy which is more typical of the explicit system, which relies on WM and attention abilities. These results are in line with the results obtained by Tharp and Pickering (under review) who found that accuracy scores on an II CL task were positively related with WM scores. Moreover, they found that participants who scored high on the WM task were more likely to develop and implement an MD strategy to solve the task. These authors suggested that high WM scorers are more likely to abandon an incorrect strategy and, therefore, implement more complex ones. Thus, these authors found contrasting results to those obtained by De Caro and colleagues (2008), who found that participants with high WM scores required more trials-to-criterion to learn the II task. Tharp and Pickering (under review) suggest that De Caro and colleagues' results might be a by-product of the task procedures they adopted. Tharp and Pickering suggest that high WM scorers may be more likely to abandon an incorrect strategy and, therefore, implement more complex ones. By contrast, perseveration with a simple strategy, by low WM scorers may very well have allowed them to pass the easy criterion adopted by De Caro et al. This could well have occurred without the low WM participants having learned the task effectively and certainly without them engaging implicit system procedural learning (for more details see Tharp and Pickering, under review). However, it is not possible to draw such conclusions in the present study since, due to the complexity of the task and the small trial size, participants failed to implement the optimal II strategy.

Individuals with high scores on the N or ImpAss trait components were more likely to score above chance although high N, but not high ImpAss, participants were more likely to use a dimensional rule than their counterparts who, as previously mentioned, may have resorted to alternative strategies. However, ImpAss was generally a weaker predictor of participants' performance (i.e. accuracy; according to the results obtained from the multiple regression), which may explain why ImpAss scores did not predict whether participants would use a dimensional strategy or not. The analysis loses power because of the limited sample size once 'non-fitted' participants were excluded. Hence it may have not been able to capture the weaker impact of ImpAss.

Superior performance by high Neurotic participants may be a by-product of the feedback manipulation. Some participants may have experienced the no-feedback trials as punishing, especially if they were fairly confident of their responses. According to RST, N is the underlying trait of the behavioural inhibition system (BIS) and the fight-flight-freeze system (FFFS) that is sensitive to conflict or punishment, respectively. It is possible that the design of the study produced a regulatory fit between N scores and the feedback manipulation. This regulatory fit could increase cognitive flexibility in high neurotic individuals and, therefore, increase their performance.

6.11 Overall discussion

The results obtained in the present study highlight the importance of using formal models to further explore the data and assess what strategy participants are using. In this way, it is possible to identify learners from non-learners. In fact, the behavioural data analysed in the two studies identified several learners when considering individuals who scored above chance. However, after fitting the formal model to the data it was possible to see that some individuals, who had been identified as learners, were using highly suboptimal, non-dimensional strategies. This was particularly true in the II version of the task where 54 participants were found to perform above chance but the formal model indicated that none of them was using the optimal II rule and, approximately, half of the 'learners' was using a sub-optimal uni-dimensional rule. These observations suggest that accuracy scores are much less meaningful if they are not presented together with an analysis of the strategy used. These observations are in line with previous studies that indicated that model fitting offers a great insight into behaviour and, therefore, a better understanding of individual differences (Gluck, et al., 2002; Tharp, 2007).

Contrary to expectations ImpAss was not significantly associated with the type of strategy adopted by participants, but neuroticism was found to be associated with the strategy used. In fact, high N scorers were found to be more likely to implement the optimal strategy (cf. low scorers) in the RB task whereas high N scorers were more likely to use a sub-optimal (uni-dimensional) rule in the II. Nonetheless, the sub-optimal strategy was associated with levels of performance above chance. Thus, the use of the sub-optimal strategy was able to achieve satisfactory accuracy levels in both the RB and the II task, as mentioned earlier. Overall, these considerations indicate that neurotic participants showed superior performance (cf. stable participants). These results may be a by-product of the feedback manipulations. Indeed, no-feedback trials could have been perceived as punishing. Hence, even though a reward-only condition was adopted it is possible that it was perceived as a mixed-incentive condition. This may have produced a regulatory fit between the feedback (especially its punishing or conflict inducing aspect) and high FFFS/BIS individuals (high N). As mentioned before, the regulatory fit would enhance cognitive flexibility which, subsequently, improved performance. It is possible that, in order to induce a pure BAS activation, it is necessary to have purely rewarding manipulations rather than mixed-matrices, which elicit simultaneous FFFS or BIS activation (Pickering, 2004). The use of a gain-only matrix, thus, may be more appropriate to explore the BAS.

During non-feedback trials, individuals may have responded having the belief of being correct. Hence, they would have expected a 'correct' message and to earn 5 pence but received no feedback, which would have been perceived as punishing. According to the reward prediction error learning model, this mismatch between predicted and actual reward would have produced LTD on the stimuli-response association and, therefore, un-learning of the correct response. This mechanism could account for the limited learning.

Results also indicated that the feedback manipulation was not effective in either task. In fact, as predicted by COVIS, the feedback manipulation did not affect the way participants learnt to perform the RB task. However, the fact that only a few participants used the optimal type of rule (19 in the RB task and none in the II task) indicates that feedback frequency manipulation and, possibly, the introduction of a concurrent task impaired learning. Impaired learning was indexed by the fact that 17 participants used an incorrect dimensional strategy to set their rule and that 24 used no dimensional rule at all in the RB task. Moreover, contrary to COVIS assumptions, the feedback manipulation did not affect learning on the II task when comparing the two feedback conditions. However, power analyses showed that the statistical power was low in both studies and therefore, it may have not allowed detection of a significant effect of small to medium effect size. Low power may have been generally caused by the small sample size in the two studies. Indeed, power levels increase as a monotonic function of sample size, all other things being equal.

However, it is very difficult to draw exhaustive conclusions from these results since results suggest that many participants did not learn through the task: most participants did not base their strategy on a dimensional rule and the few who used a dimensional rule developed a sub-optimal (uni-dimensional) strategy. These results also suggest that the feedback frequency may have been too low for implicit learning to occur and it is also possible that the filler task interfered with performance on the II task.

These observations suggest that this type of feedback manipulation may be too drastic and, thus, not adequate to observe potential individual differences during performance CL tasks (as indicated by the low proportion of participants implementing the appropriate strategy). It may be necessary to implement more subtle variations of reward manipulations that may be more effective in uncovering small variations in performance mediated by individual differences.

Decision-bound models suggest that during categorisation, individuals learn to assign different responses to different regions of perceptual space, obtained by placing a decision bound (or criterion; Bohil & Maddox, 2001; Maddox &

Dodd, 2003). The theory further states that base-rate and payoff manipulations produce a bias in the criterion location. Markman and colleagues (2005) observed that asymmetric payoff matrices produced a bias in the criterion location and that the extent of the bias was mediated by regulatory focus, which characterises a sensitivity to reward or a sensitivity to punishment (promotion and prevention regulatory focus, respectively). These definitions indicate that there is a broad overlap between RST and the regulatory focus theory (RFT; Beauducel et al., 2005). Hence, for future studies it was decided to take into consideration the RFT literature in order to explore RST. In particular, a promotion focus is employed in the remaining studies in order to encourage the activation of the BAS. Moreover, a gain-only matrix is also implemented following earlier considerations.

Chapter 7

Neural model

7.1 Chapter aims

The present chapter introduces a biologically-constrained neural model able to simulate and predict human individual differences effects in learning data as well as to test certain theoretical assumptions. This dopaminergic model was developed to represent some of the biological mechanisms that underlie procedural learning and that may occur within brain structures thought to be part of the BAS (Gray, 1987). The present chapter describes the structure of the model and tests the validity of the model to simulate human behaviour. Simulation results showed that the model was able to simulate the response bias shown by participants performing on a category learning (CL) task with asymmetric reward payoffs (Markman et al., 2005). The numerous simplifications of the model are also described.

7.2 Introduction

The study reported in the previous chapter failed to observe any significant effect of a feedback manipulation on performance during CL tasks. Overall, the results showed that the task was hard and that only a small proportion of participants implemented the appropriate rule. Therefore, due to the low power it may have not been possible to observe any significant effect. Moreover, it is possible that the task design was not optimal to explore reward-related learning as a function of personality traits and motivation, characterised by a biological mechanism (i.e. BAS). Indeed, it is possible that the impact of reward on the task was small as indicated by the fact that participants implemented a simple dimensional rule in both RB and II tasks. It has been suggested that participants tend to implement simpler rules when they ensure above-chance performance, regardless of the task design (Gluck, 2002; Shohamy et al., 2008). There is evidence that indicates that the explicit system, which is responsible for developing simple verbal rules, does not require trial-by-trial feedback (Ashby et al., 1998; Ashby et al., 1999; Maddox & Ashby, 2004). Thus, these considerations suggest that performance on the Gabor CL tasks used in the previous chapter may have been largely reliant on the explicit system rather than the implicit system. Hence, future studies should develop a task whose performance is dependent on the implicit system, which is thought to be much more sensitive to reward manipulations than the explicit system. In this way, it is possible to explore reward-related learning and the mediating effect of personality traits.

BAS-related traits are believed to be associated with dopaminergic activity and, therefore, goal-driven approach behaviour (Gray, 1987). Moreover, DA activity has been found to mediate reward sensitivity and approach behaviour and, therefore, it has been suggested as a key substrate of the implicit system (Schultz, 1998; Ashby et al., 1998). Thus, it was decided to develop a DAergic neural model capable of testing the relationship between personality and behaviour, determined by their common biological mechanism (i.e. midbrain DA projections; Pickering & Gray, 2001). Indeed, inter-individual variations in the system reactivity could be used to simulate differences in personality traits (e.g. BAS-like traits).

There is a growing literature that suggests that personality traits may represent the endophenotypes for psychiatric disorders (Benjamin, Ebstein & Belmaker, 2001; Gottesman & Gould, 2003; Rommelse et al., 2008). Indeed, personality traits are heritable and extreme scores are associated with psychopathology (e.g. high impulsivity is associated with ADHD, addiction and pathological gambling; Eisenberg et al., 2007). Therefore, personality traits may represent individual's characteristics closer to the genetic/biological mechanism underlying the BAS than the pathological conditions. However, there is evidence that indicates that this relationship is not so straightforward. Personality dimensions (i.e. complex higher-order trait measures) represent the phenotype for the present research and their relationship with underlying biology and genotypes is pretty modest. Hence, there is little evidence that they are stronger than the relationships with psychopathology.

In the thesis, personality dimensions (indexed by complex higher-order trait measures) are referred to as the phenotype and their relationship with the underlying biology and genotypes is pretty modest. In contrast, endophenotypes are measures of interest, associated with the actual phenotype, that are expected to be simpler and, therefore, more strongly correlated with genes and biological mechanisms than phenotypes (Gottesman & Gould, 2003; Rommelse et al., 2008). They are, thus, easier to measure reliably and are more amenable to scientific study. Thus, the present research aims to identify the endophenotypes that are closely related to the biological substrate of interest (i.e. dopaminergic system, BAS).

There is evidence that shows that the association between behaviour (e.g. task performance; i.e. endophenotype) and personality is quite weak and difficult to detect in empirical studies, owing to the weak relationship between phenotypes and genotypes (cf. stronger relationship between endophenotypes and genotypes). Since the relationship between endophenotypes and phenotypes is driven by their common biological/genetic basis, the endophenotype-phenotype relationship is weakened by the weak phenotype-genotype relationship. This latter relationship is attenuated by the interplay between genetic and environmental factors (Munafo, Clark, Payne, Walton & Flint, 2003; Pardo, Aguilar, Molinuevo & Torrubia, 2007). In fact, it has been found that 30-60% of variance in personality traits is determined by genetic factors while the rest of the variance is influenced by environmental factors (Blum et al., 2000; Reuter, Schmitz, Corr & Hennig, 2006; Reuter, 2008).

For example, there is evidence that indicates that A1+ polymorphism on the dopamine D2 receptor (DRD2) gene is associated with high scores on personality traits such as extraversion and novelty seeking (Ozkaragoz & Noble, 2000). The presence of the A1 allele on the DRD2 gene is associated with a 30-40% reduction in DRD2 receptor sensitivity (Lee et al., 2007; Davis et al., 2008). A meta-analysis conducted by Munafo and colleagues (2003), however, indicates that these results are inconsistent. Indeed, there are several studies that have failed to capture the relationships between genotype and individual differences (e.g. Jonsson et al., 1997; Sullivan et al., 1998). Moreover, the meta-analysis highlighted the fact that genetic effects on complex traits are modest, especially when compared to the relationship between genotype and endophenotypes (i.e. behavioural measures collected in lab-based studies). Together with genetic variance, environmental factors, such as stress, have been found to mediate pathological behaviour (e.g. alcoholism; Bau & Salzano, 1995; Bau, Almeida & Hulz, 2000). For example, stress factors together with DA levels (indexed by DRD2 polymorphism) were found to be good predictors of extraversion scores, although they accounted for only 8% of the variation on the personality scores (Ozkaragoz & Noble, 2000).

Cohen and colleagues (2005) conducted an fMRI study to explore the relationship between extraversion and reactivity of the mesolimbic DA system as a function of the A1+ polymorphism during a go-no/go gambling task. Results showed that extraversion did not affect decision making, although extraverts showed greater brain activation in reward-processing areas (i.e. nucleus accumbens, amygdala and orbitofrontal cortex) during reward delivery (cf. reward anticipation). Indeed, extraversion scores explained 33% of the inter-individual differences observed in brain activation. In contrast, the presence of the A1 DRD2 allele was associated with lower brain reactivity to reward delivery. The results illustrate that there is a significant relationship between the genotype and brain activation (as indexed by the fMRI measures). Additionally, fMRI measures are also associated with measures of extraversion (i.e. phenotypes). Overall, the results show that brain activation is a possible endophenotype of extraversion. There was a weak trend correlation between the presence of the A1 allele and low scores on extraversion; however, power was extremely low as the sample included fewer than 20 participants. The weak relationship might also be due to the inherently weak relationship between genotype and phenotype. The authors also suggested that this relationship may have been weakened by environmental factors (i.e. stress).

Another study by Reuter and colleagues (2006) identified the relationship between scores on Carver and White's BAS scales and DA activity (indexed by prolactin levels). These researchers found that individuals who scored high on the BAS scales (especially, BAS-drive, BAS-FS and BAS-total) had greater DA levels than their low scoring counterparts. Nonetheless, this effect was quite weak as indicated by an eta squared value equal to 0.083 (i.e. 8% of the variance is explained by personality). Similarly, the DRD4 polymorphism was found, in the original studies, to explain 4% of the variance of the Novelty Seeking scale (Benjamin et al., 2001). Overall, these studies illustrate that the relationship between genotypes and phenotypes is weak and that even biological markers, which are closely related to the genotype (e.g. prolactin), show only a moderate relationship with personality traits (i.e. phenotype).

In contrast, genetic polymorphisms have been found to be better predictors of specific behaviours (e.g. alcohol use) than they are predictors of broad individual differences. Indeed, Blum and colleagues (2000) have suggested that alcoholism is associated with the presence of the A1+ allele on the DRD2 gene. Comings and colleagues (1991) found that among alcoholics 42.3% carried the A1+ allele whereas 14.5% of the non-alcoholics carried the A1+ allele. Similar results had been reported by Bau and colleagues (2000) and Ponce and colleagues (2003). The latter researchers found that individuals who carried the A1+ allele were three times more likely to become severe alcoholics compared to those carrying the A1- allele. A meta-analysis by Noble (1998) also reported that severe alcoholics were 1.5 to 3-times more likely to carry the A1+ allele than non-alcoholics.

There is also evidence of a weak relationship between personality traits and behaviours related to the same genetic variance. This can be observed, for example, when considering reward-related personality traits (BAS-traits) and alcohol/drug use in relation with low DRD2 receptor sensitivity (Blum et al., 2000). BAS-like traits have been found to be positively and significantly correlated with alcohol use and frequency among healthy participants. For example, Franken and Muris (2006) found that Carver and White's BAS-fun seeking scale was significantly correlated with the number of illegal drugs used ($r = .35$, $p < .01$), alcohol use ($r = .24$, $p < .01$) and binge-drinking ($r = .25$, $p < .01$). BAS-drive was also found to be correlated with the number of illegal drugs used (r = .18, p <.01). Similar results were obtained in a study by Pardo and colleagues (2007) that showed that Corr's BAS scale [BAS = (Ex2)+N+P] was not only significantly correlated with alcohol use frequency and quantity, but it was also a reasonable predictor of these variables. Indeed, the BAS factor accounted for 13% and 12% of the variance of frequency and use, respectively. These results indicate that alcohol use is related to both DRD2 genotypes and to personality scores; although the relationship between personality and genotype is usually weak. The reason why the relationship between personality and behaviour (e.g. alcohol use) is quite weak and difficult to detect may be due to two main factors: the weak relationship between the measured broad trait and genetic variance and the added noise typical of lab measures further weakens the relationship (Cohen et al., 2005).

Following these observations, the use of a neural model represents a useful tool to explore the efficacy of laboratory tasks at capturing the relationship between the biological mechanism and associated behaviours. Given the weak relationships described above, it is necessary to ensure that the task-biology relationship is strong. Thus, a biologically-rooted model may therefore provide a useful test to of the extent to which the task design really is sensitive to the phenomenon of interest. Moreover, the neural model is a useful tool to explore a theory's main assumptions (Pickering, 2008). Hence, a neural model was developed that captures the impact of DAergic function during performance on reward-related learning tasks. Since, midbrain DA has been proposed as the BAS substrate, individual differences in DAergic function are good candidates to represent BAS variations and, therefore, mirror individual differences on reward sensitivity and approach (Pickering, 2004). In other words, BAS-related personality traits are expected to represent the phenotype of the biological, DA-based mechanism(s). Therefore, a strong relationship between the model's simulated biological mechanism and its simulated behavioural responding on the task is likely to be necessary for the task to be useful in testing biological theories of personality. With typical sample sizes, it is only with a strong relationship of this kind that it will be possible to detect a significant relationship between real lab behaviours and the participants' measured personality traits. The latter real relationship is weakened by the fact that the simulated biological parameter in the model is only weakly related to personality traits (as the studies reviewed above demonstrate). It is further weakened by added noise in task performance due to empirical testing (Lee et al., 2007) and the gene-environment interaction (described earlier). Hence, only a strong simulated relationship will survive through to the laboratory. The present chapter presents the rationale and the structure of the DAergic neural model and, additionally, it tests its ability at predicting performance on a CL task with asymmetric payoffs.

7.3 Model description

The computational model developed in the thesis is an actor-critic model which resembles the architecture of the basal ganglia and simulates reinforcement learning (Suri & Schultz, 1999; Joel, Niv & Ruppin, 2002). Thus, the model consists of two sub-networks, the critic sub-network which is responsible for predicting future rewards in relation to sensory stimuli and the actor sub-unit which is responsible for choosing how to respond to the presented stimuli (Sutton & Barto, 1988; Suri & Schultz, 1998; Joel et al., 2002). The reward prediction signal, formulated by the critic sub-network, is constantly updated thanks to comparison of the predicted reward with the actual reward (i.e., to give a reward prediction error, RPE).

The model is a neurally-based model that suggests that learning during performance on the CL task is mediated by reward prediction errors (RPEs). As described in chapter 2, RPEs are determined by the difference between the expected and the actual reward (equation 7.1).

RPE = reward which occurred – reward which was expected **Eq. 7.1**

where a positive RPE indicates that the reward occurrence is not fully predicted (Schultz, et al., 1993). A negative RPE, by contrast, indicates that the actual reward is smaller than predicted. Existing evidence has indicated that DA cells change their firing rate in response to reward unpredictability, rather than reward *per se*, and that DA-driven learning occurs only when rewards are unpredictable (Schultz, 1998; 2002). Moreover, changes in DA firing, in response to unpredicted reward, represents an RPE signal. The present model calculates the RPE signal by subtracting the reward prediction signal from the actual reward received following a categorisation response. Hence, following a response that leads to a reward greater than the expected one (positive RPE), the DA cells increase their firing and the rewarded stimulus-response association is strengthened; by contrast, this association is weakened when the actual reward is smaller than the predicted reward (i.e. negative RPE, and associated DA cell firing decrease). Learning ceases when the actual reward is equal to the expected one (i.e. when RPE = 0).

The present model is a DAergic model and it is in line with evidence that indicates that DA/RPE signals play a key-role in synaptic plasticity, as indicated by studies investigating long-term potentiation and depression (LTP & LTD; Carlsson, 2002; Corr, 2006; Calabresi, Picconi, Tozzi & DiFilippo, 2007; Surmeier, Ding, Day, Wang & Shen, 2007; Schotanus & Chergui, 2008). The DA signal acts as a reinforcement signal that is responsible for adapting the strength of synapses (i.e. weights); this mirrors the reinforcement signal employed in computational models using the temporal-difference (TD) algorithm (Sutton & Barto, 1988). Learning, in the corticostriatal synapses represented in the model, occurs according to the so-called 3-factor learning rule (Pickering & Gray, 2001; Schultz, 2002; Ashby et al., 2007). Indeed, the DA signal has been identified as one of three factors that mediate learning within the 3-factor model. It can account for both LTP/LTD and procedural learning in the striatum (Joel et al., 2002; Ashby & Ennis, 2006).

The three factors operating in this model are: 1) the pre-synaptic activation (glutamate signal), 2) post-synaptic activation (NMDA receptor depolarisation) and 3) DA signal (for more details refer to chapter 2). In the present DAergic model, the three factors are: 1) cortical input to the striatal units, 2) activation/depolarisation of the striatal units and 3) DA signal (i.e. RPE). Therefore, in line with COVIS the current model is an implicit model, since it relies on gradual procedural learning (Ashby et al., 1998; Pickering & Gray, 2001; Ashby & Ennis, 2006).

As noted, the present model has an actor-critic architecture (Joel et al., 2002; Seger, 2008). The critic sub-network is loosely based on the model developed by Brown and colleagues (1999), which aimed to simulate classical conditioning. Thus, the present model is an extension of this model since it aimed to simulate procedural learning and, therefore, a response unit (actor) is also included (Suri & Schultz, 1998; Joel et al., 2002). The actor sub-network is based on the so-called visual corticostriatal loop, which represents one of the main corticostriatal loops involved in category learning (Seger, 2008). Figure 7.1 illustrates the basic structure of the model.

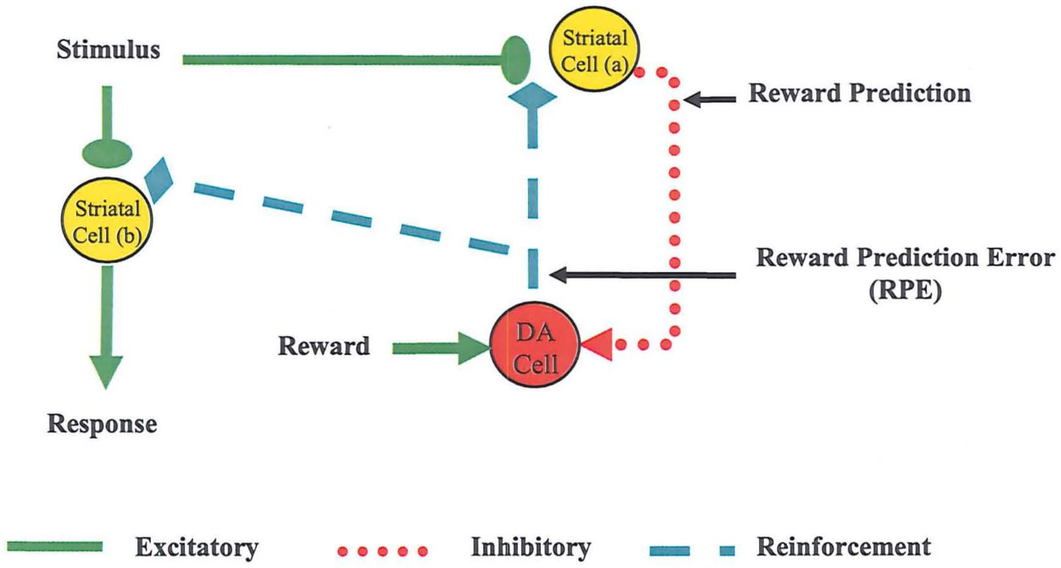


Figure 7. 1. The basic structure of the neural model

The model encompasses two striatal cells and a midbrain DA cell (red cell). The dopamine cell is either in the ventral tegmental area (VTA) or in the substantia nigra pars compacta (SNc). The model includes a ventral striatal cell and a dorsal striatal cell (labelled 'a' and 'b', respectively), following evidence that indicates that both types of striatal cell communicate with DA neurons to mediate implicit learning (O'Reilly and Munakata, 2000; Joel et al., 2002; Seger, 2008). The ventral cell is a striatal cell within the so-called motivational (limbic) corticostriatal loop (Seger, 2008). It has been suggested that this cell lies within the nucleus accumbens (Nac; Schotanus & Chergui, 2008). The dorsal striatal cell lies within body and tail of the caudate, within the so-called visual (associative) corticostriatal loop (Seger, 2008). The dorsal striatum is considered to be responsible for motor functions and habit (implicit) learning whereas the ventral striatum serves motivation and reward processing (Schotanus & Chergui, 2008; Seger, 2008). The two striatal cells receive coded stimulus input from the visual cortex (Brown et al., 1999; Seger, 2008). There is evidence that indicates that each striatal cell receives up to 10000 inputs from the cortex and that the compression ratio between cortical neurons and striatal neurons is 10:1 (Ashby & Ennis, 2006; Ashby et al., 2007; Seger, 2008).

The ventral striatal cell, which represents the critic network, receives input from the visual cortex (Brown et al., 1999). The cortical input is glutamatergic in nature and leads to depolarisation of the striatal cells (Brown et al., 1999). The critic network is responsible for formulating the reward prediction (RP) signal which is a GABAergic inhibitory signal (projecting to the VTA/SNc cell; Brown et al., 1999; Joel, et al., 2002; Seger, 2008). The dorsal striatal cell represents the actor sub-network implicated in a neural pathway supporting stimulus-response (SR) associations (Ashby & Ennis, 2006; Seger, 2008). The actor also receives cortical input from the visual cortex and it, subsequently, associates an "abstract" response selection with the stimulus information (Ashby et al., 2007; Seger, 2008). The abstract response selection is projected, via the thalamus and the globus pallidus (GPi and GPe; i.e. internal and external), to the premotor cortex where it forms the input into the so-called 'motor' corticostriatal loop (especially, Brodmann Area 8) which produces an actual motor response (Ashby et al., 2007; Seger, 2008; figure 7.2). Figure 7.2 illustrates that the striatum sends inhibitory GABAergic signals to the globus pallidus and this inhibitory signal releases the tonic inhibition on the thalamus which, subsequently, sends an excitatory signal to the (premotor) cortex.

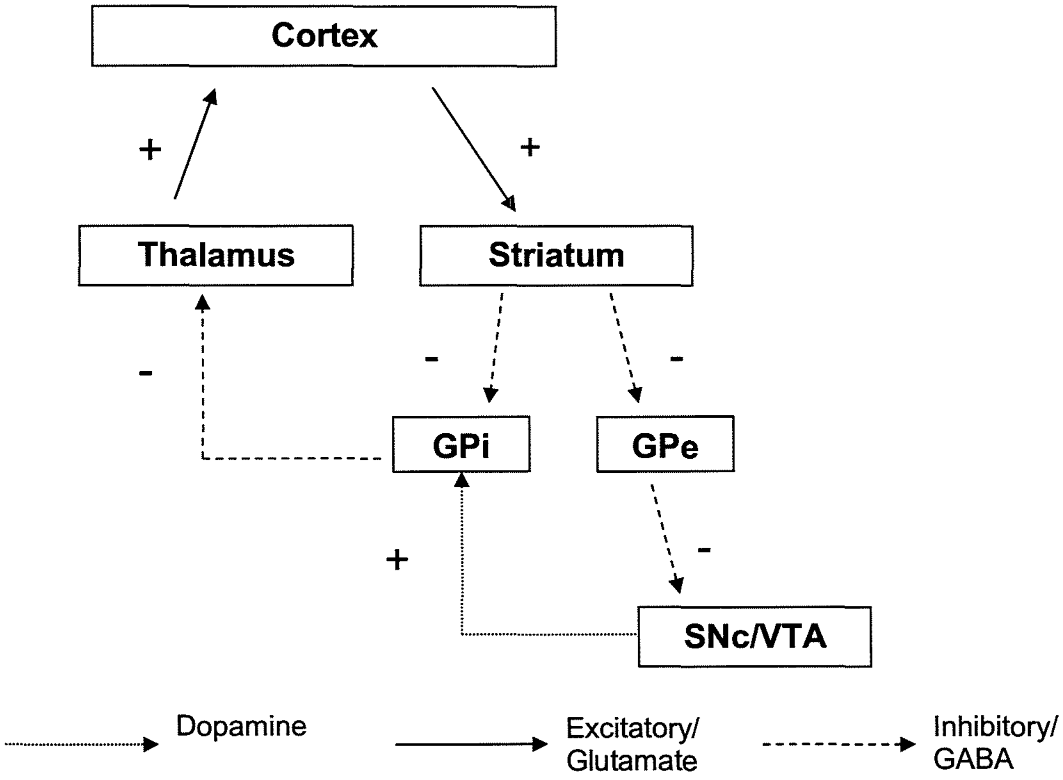


Figure 7. 2. Structure of the corticostriatal visual loop

The actor in the present model represents a simplification of the visual loop presented by Seger (2008) since it only includes the excitatory glutamatergic cortical projections into the striatum (i.e. body and tail of caudate). The present model makes a sweeping simplification that once a response choice (i.e. to respond with a category A response) is made within the dorsal striatal (caudate) cell then that choice is accurately transmitted through the output side of the visual loop to the input of the motor loop and, hence, out to the appropriate effector systems (in this example, those responsible for pressing the category A key). The loop structure was grossly simplified since the model is not interested in reproducing the complex processes involved in response-making.

According to Brown and colleagues (1999), midbrain DA cells receive both fast indirect excitatory input from the pedunculopontine tegmental nucleus (PPTN) and delayed transient inhibitory signals from striosomal neurons in the ventral striatum. The PPTN receives excitatory input from the hypothalamus, which responds to primary rewards, and matrisomal neurons of the ventral striatum, which respond to both primary and secondary rewards (Brown et al., 1999;

Schotanus & Chergui, 2008). In the present model (see Figure 7.1), the DA cell processes the excitatory PPTN input (i.e. "actual" reward) and the inhibitory striosomal ventral striatal input (i.e. the output of the reward prediction unit) in order to compute the reward prediction error (RPE) signal. The RPE signal then projects to both kinds of striatal cells although it has a different function on each cell. On the corticostriatal synapses of the limbic striatal cell, the co-occurrence of the RPE signal at a D1 receptor and glutamatergic activation (mediated by timed calcium spikes in response to the stimulus input) are involved in updating the reward prediction signal for future trials and, therefore, learning (i.e. it is a so-called "adaptive" critic). In contrast, the RPE signals, acting on both the D1 and the D2 receptors of the caudate cell, act as a learning signal that strengthens the correct SR association (i.e. adapting another set of corticostriatal synapses; Joel et al., 2002; Seger, 2008; Schotanus & Chergui, 2008).

Learning on both striatal cells is expected to follow the 3-factor Hebbian rule (i.e. with both LTP and LTD; Schultz, 1998; Brown et al., 1999; O'Reilly & Munakata, 2000; Joel et al., 2002; Ashby et al., 2007). Long-term potentiation (LTP) and long-term depression (LTD) are both responsible for long-term plasticity of glutamatergic synapses in corticostriatal areas (Calabresi et al., 2007; Surmeier, et al., 2007; Schotanus & Chergui, 2008). Schotanus and Chergui (2008) have observed that dopamine D1 receptors play a critic role in LTP in the NAc (i.e. ventral striatum) and the dorsal striatum; moreover, they observed that a synergistic interaction between D1 and D2 receptors evoke LTD in the dorsal striatum. Thus, synaptic plasticity (and learning) in the dorsal striatum is dependent on both D1 and D2 receptors.

As described below, the debated mechanisms (of which various kinds have been discussed; Joel et al., 2002) which mediate reward timing (which is essential for implicit learning to occur; see chapter 2) are not relevant in the present model. Therefore, they are stripped out of the present model. The model assumes that, at the time when the DA cell produces an RPE on a particular trial, there are a dorsal striatal cell and a ventral striatal cell which are in the appropriate 'active' state so that a burst of DA cell firing, or a brief period of inhibition of DA cell firing, can produce synaptic weight changes. However, for simplicity in the present model both the actor and the critic were updated on a trial-by-trial basis rather than in real time during each learning trial. This was done since the model is not attempting to capture reaction times (RTs) within trials.

7.3.1 Model algorithms

The stimuli are coded as a vector of N units, $\mathbf{x}^{\mathbf{out}}$, each unit of which represents output from cortical neurons; the cortical units are coding perceptual representations. The stimuli in the task to be simulated were small dots which appeared either in one of the 650 horizontal locations (i.e. along an imaginary horizontal line) or in one of 650 vertical locations (i.e. along an imaginary vertical line). The dots were displayed within a 650*650 display box shifted somewhat from the centre towards the left-hand side of the screen. Thus, the stimuli were coded in a 650-element vector which represented the pixels lying along the imaginary horizontal (or vertical) line on the screen. The stimuli consisted of dots with a 7-pixel radius (parameter $\mathbf{K_r}$ = 7 to reflect this) centred at a pixel location on trial n denoted by $\mathbf{K_n}$ ($\mathbf{K_n}$ values lie between 8 and 643 inclusive)[7]. The small (plus or minus 7-pixel) extensions of the dots to the left/right of (or above/below) the horizontal (vertical) line, where the stimuli lay, were ignored by the stimulus coding adopted. Hence, $\mathbf{x}^{\mathbf{out}}$ was formulated as:

$$\mathbf{x_i^{out} = 1 \text{ for } i = K_n - K_r : K_n + K_r; \text{ otherwise } x_i^{out} = 0} \qquad \textbf{Eq. 7.2}$$

Ashby and colleagues (2007) have implemented a radial basis function (RBF) to produce an alternative coding pattern. The authors suggest that RBFs are an effective way to model sensory units during a categorisation task. The use of an RBF determines that the cortical units that project to the striatum are either off (activation = 0) or they are positively activated within the interval [0 1]. According to the RBF, the activation of a cortical unit is dependent on the distance between the stimulus location (i) and its prototypical centre ($\mathbf{K_n}$) at trial n:

---

[7] It was decided the use the "colon" notation which is typical of many programming systems (e.g. Matlab); for example 1:5 means 1 to 5 (1,2,3,4 &5); so K_n =8:643

$$\mathbf{RBF(n) = exp[-1.*dist / (2.* ά * ά)]} \qquad \textbf{Eq. 7.3}$$

where **ά** is a constant that represents the width of the stimulus and **dist** is equal to:

$$\mathbf{dist = (K_n - i) ^\wedge 2} \qquad \textbf{Eq. 7.4}$$

Simulations showed that the coding expressed in equation 7.2 and the RBF approach (equation 7.3) produced essentially similar results and, therefore, it was decided to adopt the simpler coding method (equation 7.2) throughout the thesis.

In the model, the cortical stimulus coding outputs all project to a single (limbic) striatal cell which represents the critic, reward prediction (RP) unit (cell a in figure 7.1). The input into the RP (**RP^{in}**) unit is equal to:

$$\mathbf{RP^{in} = rptonic + \Sigma\, x_i^{out} * w_{rp}} \qquad \textbf{Eq. 7.5}$$

where $\mathbf{w_{rp}}$ represent the weights from the cortical units into the ventral striatal "reward prediction" unit (i.e., the critic) and **rptonic** is a tonic input into the RP unit. For simplicity, the RP output was set equal to the input in the model. The $\mathbf{w_{rp}}$ values are initially set equal to 0 for each simulated participant. Moreover, they were updated following each response as a function of the 3-factor rule (i.e. error correction learning). The tonic reward prediction input (**rptonic**) is assumed to be computed early in training by frontal brain structures (e.g. orbitofrontal cortex, OFC; Cromwell & Schultz, 2003; O'Doherty, 2004). The processes involved in formulating this signal are not included in the model. However, the inclusion of this parameter is important since it ensures that some reward prediction signal is present from the first trials (if **rptonic** were not present or/and $\mathbf{w_{rp}}$ were initialised at 0, there would be no reward prediction error in the first trials). For this reason, the **rptonic** value was set at, or around, 0.5 since this value represents the average reward values on offer in the task (see below).

Dayan has used a parameter which closely resembles the rptonic parameter implemented in the present model and he labelled it as the reinforcement comparison signal (Dayan, 1990). Moreover, Williams (1992) refers to a similar parameter which he labelled reinforcement baseline term. These parameters are based on Sutton and Barto's work (1981). Overall, this evidence indicates that reward prediction signals have been consistently employed in neural models that attempt to simulate reward-mediated learning. Dayan has observed that it speeds up or slows down learning of a specific behaviour although it does not affect the chosen end-behaviour (personal communication).

The cortical units also project outputs to a set of dorsal striatal units ($\mathbf{y_j}$) which is responsible for selecting a response (i.e. categorization) when a visual stimulus is present. The input to $\mathbf{y_j}$ (where $\mathbf{j}$ is either equal to 1 or 2, in a 2-category task) is given by the equation below:

$$\mathbf{y_j^{in} = \Sigma x_i^{out} * w_{ij}} \qquad \textbf{Eq. 7.6}$$

where $\mathbf{w_{ij}}$ represent the SR weights which are initialized using a normal distribution with a mean equal to 0.3 and a standard deviation equal to 0.02. The SR weights are updated following the 3-factor learning rule throughout the task, on a trial-by-trial basis.

The response selection units are activated in a winner-take-all (WTA) fashion. It has been suggested that this could be included in the model by implementing lateral inhibition between striatal response units (Suri & Schultz, 1998; 1999). However, it has recently been suggested that lateral inhibition may not represent a biologically plausible process in the striatum (Bar-Gad et al., 2000; Joel et al., 2002). Brown and colleagues (1999) have suggested that lateral inhibition in the striatum may be too weak to affect decision making. Therefore, it was decided to simplify this step by implementing a simple function to reproduce a winner-take-all rule. This was obtained by setting the activation of the unit with the largest net input (as specified in equation 7.6) equal to 1 [cf. the other(s) which are set to 0]; the winning unit determines the to-be-executed response for a particular trial (e.g. $\mathbf{y_j^{out} = 1}$ and $\mathbf{y_i^{out} = 0}$, for $\mathbf{i \neq j}$).

As already noted, the model does not include a response effector module responsible for expressing the selected response by pressing the button that corresponds with the response selected by the response module ($\mathbf{y_j}$). Indeed, it was assumed that the transfer from the response selection unit to the response effector unit was error-free. Hence, the selected response (i.e. **j**) corresponds to the active response selection unit ($\mathbf{y_j}$), which has an output equal to 1 (cf. $\mathbf{y_i} = \mathbf{0}$ for the other units).

One of the main assumptions is that following a response, participants generally receive a reinforcement from the environment as a direct consequence of the action. The goal of the individual is to produce actions that maximise reward-earning (O'Reilly & Munakata, 2000). During performance on a CL task, the reinforcement is usually provided by the task itself (i.e., as computer-generated feedback); participants generally receive such reinforcement following each response they produce. Hence, the computational model needs to include a neural reinforcement signal (denoted as **rf**) that codes the actual reinforcement (see simulation section for more details).

The **rf** values in the model represent the size of the excitatory reward input from the brain's reward system (e.g. PPTN) into the dopaminergic midbrain cells (SNc/VTA). The literature shows that these DA cells compute a reward prediction error (RPE) signal and, therefore, respond to novel and unpredicted stimuli (Schultz, 1998; Pickering & Gray, 2001; Schultz, 2002). The RPE signal represents the difference between the actual reward received following each response (i.e. **rf**) and the predicted reward computed by the critic (i.e. $\mathbf{rp^{out}}$). Hence, in the computational model, the raw RPE is denoted as **RPE*** and equal to:

$$\mathbf{RPE^* = (rf - rp^{out})} \qquad \textbf{Eq. 7.7}$$

The raw **RPE*** was, subsequently, thresholded so that small positive or negative signals (i.e. those with an absolute value smaller than 0.05) were set equal to 0 and, therefore, produced no effect. The thresholded **RPE*** signals (**RPE**) are calculated using equation 7.8:

Eq. 7.8

$$\text{If } \mathbf{RPE^* \geq 0, RPE = max\,(RPE^* - \theta_r, 0)}; \text{ else if } \mathbf{RPE^* < 0, RPE = min(RPE^* + \theta_r, 0)}$$

where $\boldsymbol{\theta_r}$ is the threshold with a value equal to 0.05. Therefore, for positive RPE* the recoded RPE is recoded as 0 if the raw signal is between 0 and $\boldsymbol{\theta_r}$, otherwise it is set equal to (**RPE*** – $\boldsymbol{\theta_r}$). For negative raw RPE values that are below -$\boldsymbol{\theta r}$ the RPE value is set equal to (**RPE*** + $\boldsymbol{\theta_r}$); otherwise 0.

According to the 3-factor learning rule, the RPE signal represents the teaching signal that mediates learning and the updating of the learning weights (i.e. $\mathbf{w_{ij}}$ and $\mathbf{rp^{wt}}$), which are adjusted over trials. Models that implement the 3-factor learning rule stress the importance of the timing of the reward prediction error to ensure strengthening of the appropriate synapse (i.e. weights; Schultz, 1998; Suri & Schultz, 1998; Brown et al., 1999; Pickering & Gray, 2001; Ashby & Ennis, 2006; Ashby et al., 2007). For this reason, computational models generally aim to generate a reward prediction signal able to predict both the size and the timing of the reward (Sutton & Barto, 1988; Suri & Schultz, 1998; O'Reilly & Munakata, 2000). Therefore, these models propose that RPE signals update the weights that code both size and timing. Nonetheless, the timing considerations were not critical to the present model and, therefore, were excluded as part of the model simplification. The RP unit coded only the reinforcement value (i.e. magnitude) in the present model. The model assumed that both the limbic and the caudate striatal units are activated at an appropriate time (in relation to the reward timing) so that their weights could be appropriately updated.

The change in weights from the cortical stimulus coding unit into the reward prediction unit at trial n is given by equation 7.9:

$$\text{if } \mathbf{RPE \geq 0, \Delta rp^{wt}(n) = b_1 {*} RPE {*} g(rp^{out}) {*} x_i^{out}};$$

Eq. 7.9

$$\text{else if } \mathbf{RPE < 0, \Delta rp^{wt}(n) = b_2 {*} RPE {*} g(rp^{out}) {*} x_i^{out}}$$

where $\mathbf{b_1}$ is a learning parameter that sets the learning rate of the RP unit with positive RPE signals; whereas $\mathbf{b_2}$ is an extinction parameter of the RP unit with negative RPE signals. The $\mathbf{g(rp^{out})}$ function represents the activation of the reward prediction unit and it is equal to:

$$\mathbf{g(rp^{out}) = [max(rp^{out},0) \; \& \; 1]} \qquad \textbf{Eq. 7.10}$$

so that any output (greater than 0) from the RP unit is coded as 1 and the unit is considered active and ready to have its weights adjusted by an RPE (reinforcement) signal. Once the weight change ($\mathbf{\Delta rp^{wt}}$) had been calculated for trial n, the weight value was updated at the end of each simulated trial:

$$\mathbf{rp^{wt}(n) = rp^{wt}(n\text{-}1) + \Delta rp^{wt}(n)} \qquad \textbf{Eq. 7.11}$$

The change in weights to the response selection unit was computed in a similar fashion (equation 7.12):

$$\begin{aligned} &\text{if } \mathbf{RPE \geq 0, \Delta w_{ij}(n) = b_3 * RPE * y_j^{out} * x_i^{out}};\\ &\text{else if } \mathbf{RPE < 0, \Delta w_{ij}(n) = b_4 * RPE * y_j^{out} * x_i^{out}}; \end{aligned} \qquad \textbf{Eq. 7.12}$$

where $\mathbf{b_3}$ and $\mathbf{b_4}$ represent the learning and extinction rates of the actor (SR) unit in response to positive and negative RPE signals, respectively. Only one of the actor (response selection) units is active on any trial (with an activation and output value of 1), as a result of the WTA algorithm described earlier.

Finally, the weights are updated at the end of each trial according to equation 7.13:

$$\mathbf{w_{ij}(n) = w_{ij}(n\text{-}1) + \Delta w_{ij}(n)} \qquad \textbf{Eq. 7.13}$$

7.3.2 Main model parameters

The model description indicates that the model contains several parameters and their values affect learning during reward-related learning. Indeed, several parameters indirectly affect the magnitude of the RPE' signal and, therefore, modulate the learning rate. The main parameters and the values adopted in the various simulations are reported in appendix 1.

7.4 Preliminary simulations

The present RPE-model was used to simulate the data obtained in a study by Markman, Baldwin and Maddox (2005), in order to assess the efficacy of the model at capturing human performance during performance on a 'notionally' rule-based CL task with asymmetric payoffs. The category learning stimuli consisted of single small dots whose location on the screen varied along one of two imaginary lines (vertical vs. horizontal) on the screen (from the left or right of the screen or from the bottom to the top of the screen). Participants were instructed to learn to classify the stimuli via trial and error through the use of feedback. The simple rule required participants to classify stimuli on one side of the screen as A and those on the other half of the screen as B. In the task, the positions of each category A stimuli on the screen was sampled from a random distribution centred around 275 pixels from the left-hand side of the screen and a s.d. of 100; whereas category B were sampled from a random distribution with mean 375 and the same s.d. of y. The mean and s.d. value were chosen to ensure an overlap between the two distributions. Thus the task is probabilistic and it is impossible to achieve 100% accuracy as there is overlap in the distributions from which the two stimuli were created. However, participants can learn to place a simple decision criterion to optimise task performance. In order to maximise accuracy, participants should place their decision bound midway between the means of the distributions from which category A and category B stimuli were sampled (i.e. at 325 pixels). However, the best criterion was difficult to identify through learning as there were many errors even with the optimal accuracy criterion because of the probabilistic nature of the task.

Participants performed under three biased payoff matrix conditions: a mixed, a loss and a gain payoff matrix. Under all matrices category A stimuli offered a higher payoff. In the mixed matrix, participants earned points for correct responses and lost points (-100) for incorrect ones. However, they earned more points for making correct responses to category A (200) than category B (0) stimuli. In the loss payoff matrix condition, points were lost for all responses. Nonetheless, participants lost fewer points for being correct than incorrect (-411) but the points lost for correct category A responses (-111 points) were smaller than those lost for correct category B responses (-311 points). In the gain payoff matrix condition, participants won points for all responses. They gained more points for being correct than incorrect (100) but earned more points for making correct category A responses (400) than correct category B ones (200). Taking into consideration RST, the gains condition seems likely to be the purest BAS-activating condition since it delivers only rewards.

According to the decision bound theory, in each of the three asymmetric payoff conditions, the criterion bound should be shifted away from the optimal accuracy criterion towards the optimal reward criterion (Maddox, Bohil & Dodd, 2003). In the study, the optimal criterion for maximising reward (or minimising losses) should be placed further away from the centre of the high payoff category from the optimal bounds for maximising accuracy (see figure 7.3). The bias towards the optimal reward criterion enabled participants to miss very few high payoff category A stimuli at the expense of making more errors on the low payoff category B stimuli. Therefore, participants gained more points than if they had maximised accuracy.

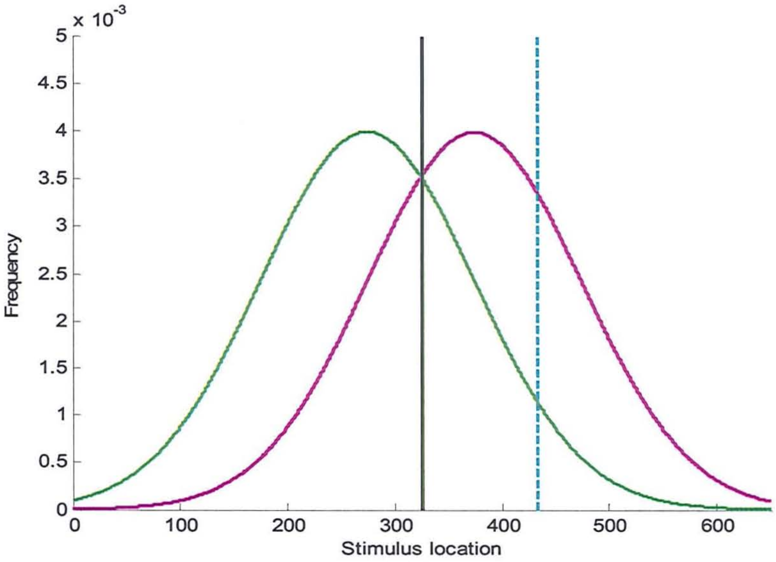


Figure 7. 3. Category distributions and the optimal accuracy criterion (solid black) and the optimal reward criterion (dashed blue)

Markman and colleagues (2005) also manipulated regulatory focus. In fact, they had half of their participants perform under a so-called 'promotion' regulatory focus and the other half under a 'prevention' regulatory focus. The promotion focus was induced by telling participants that they could win tickets for a $50 draw if their points total exceeded a specific score at the end of the task. The prevention focus was induced by giving the draw tickets at the start and telling them that they could keep the tickets only if their score at the end was above the criterion. Hence, the framing of the task urged participants to maximise their winnings (and so sacrifice accuracy).

Results showed that participants in the mixed matrix did not show any response bias regardless of the regulatory focus into which they were induced. Nonetheless, under the gain payoff matrix, participants with a promotion focus placed their decision bound closer to the optimal reward criterion (criterion at approximately 430 pixels on average) compared to participants with a prevention focus. In fact participants with a prevention focus placed their criterion bound closer to the optimal accuracy criterion (approximately 325 pixels). The complementary pattern of results was observed under the loss payoff matrix.

7.4.1 Simulation aims

The present simulations aim to capture the response bias in favour of the high-payoff category under the gain-only matrix observed among those participants who had been induced with a promotion focus. These simulations aimed to assess whether this effect can be captured by a realistic biological, DA-based mechanism. The DA mechanisms within the model are those that are thought to mediate reward sensitivity and are part of a system which controls approach behaviour (i.e. the BAS according to RST; see chapter 3).

7.4.2 Methods

In the task implemented by Markman and colleagues (2005), participants received a reinforcement on every trial following their response. The feedback in the task was constituted by the amount of points earned on each trial. In order to simulate performance on the task, each response was followed by a reinforcement in the model, too. Thus, the actual reinforcement values (i.e. points) were recoded into a neural reinforcement signal (i.e. parameter **rf** described above). In other words, several **rf** values were mapped onto the different points (i.e. rewards) during the simulations. The simplest mapping assumes a direct, linear relationship between the reward and the **rf** values. Table 7.1 offers an example of the reward-**rf** mapping during a 2-category task with an asymmetric payoff matrix.

Table 7. 1. Linear transformation of the reward values (i.e. points) into rf values for an asymmetric payoff matrix (✔ = correct vs. ✖= incorrect)

	High payoff category		**Low payoff category**	
	✔	✖	✔	✖
Points	400	100	200	100
rf values	1	0.25	0.5	0.25

An alternative, linear **rf**-reward mapping is presented below for a 2-category task with a symmetric payoff matrix where the two categories offer the same amount of points for correct (300 points) and incorrect (100 points) responses (table 7.2).

Table 7. 2. Linear transformation of the reward values (i.e. points) into rf values for a symmetric payoff matrix (✓ = correct vs. ✗= incorrect)

	High payoff category		**Low payoff category**	
	✓	✗	✓	✗
Points	300	100	300	100
rf values	0.75	0.25	0.75	0.25

The two payoff matrices produce the same total reward with equal numbers of responses to each category.

7.4.2.1 Model parameters

The **rptonic** parameter was set to be equal to 0.5. In order to capture the response bias, the **rf** values (tables 7.1 and 7.2) were multiplied by a constant value (=0.5) so that the actual rf values were 0.5, 0.375, 0.25 and 0.125 (for 400, 300, 200 and 100 points, respectively). Moreover, the RP learning rates were set equal to 0.10 (**b1** and **b2** = .05 * **scaler1**, where **scaler1** = 2) and the SR learning weights were set equal to 0.10 (**b3** and **b4** = .05 * **scaler2**, where **scaler2** = 2). The parameters and their values are summarised in table 7.3.

Table 7. 3. Parameters' values implemented in the simulations

Symbol	**Description**	**Value**
rptonic	**Reward prediction tonic value**	0.5
m	**Scales the rf values**	0.5
b_1 and b_2	**Learning rates of the RP unit**	0.05*scaler1
b_3 and b_4	**Learning rates of the SR unit**	0.05*scaler2
Scaler1	**It scales the learning rate of RP unit**	2
Scaler2	**It scales the learning rate of RP unit**	2

7.4.5 Simulation results

Three-hundred simulated subjects were created to capture the human behaviour. The actual criterion location was extracted from the simulated data by fitting each data set via a uni-dimensional formal model. The model assumes that participants based their categorisation on the relevant dimension (i.e. location on the screen). Data from each simulated 'participant' was fitted independently and, using maximum likelihood methods, the model estimated two parameters (i.e. pixel position of the decision bound and 'noise'; see chapter 6 for previous examples of formal model fitting of decision boundaries in uni-dimensional models). The goodness of the model solution is estimated against an imaginary saturated model with no free parameters (the saturated model captures the data perfectly). The middle 100 trials were included in the analysis (i.e. trials 26 to 125 out of the total of 150) as in the original study.

Results showed that the mean (s.d.) simulated deciosn bound (aka "criterion") was equal to 423.23 (19.2) while the noise level was equal to 284.55 (25.9) under the asymmetric payoff. This is a good replication of the criterion shown by participants under a gain-only asymmetric payoff matrix in the study by Markman and colleagues (2005) participants in the asymmetric gains condition with a promotion focus. Figure 7.4 shows the typical learning pattern displayed by simulated participants.

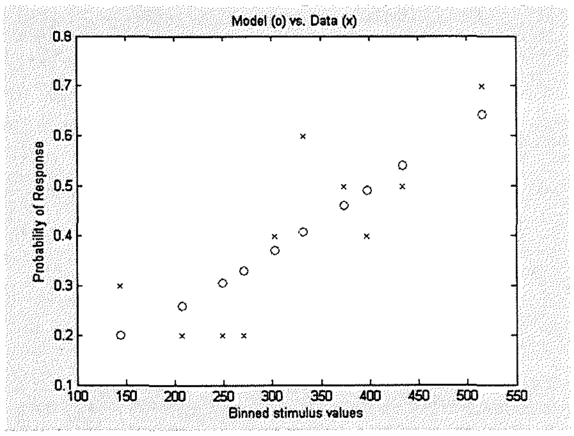


Figure 7. 4. Typical learning pattern observed in the simulated data

In contrast, under a symmetric payoff matrix, the criterion was equal to 326.75 (s.d. = 14.7) while the noise level was 292.66 (s.d. = 18.5). The latter condition was not included in the study by Markman and colleagues; however, it was interesting to simulate this condition to confirm that it does not produce any response bias and, therefore, confirm the biasing 'force' specifically produced by the asymmetric payoffs.

7.5 Discussion

The present chapter presents a DAergic model based on actor-critic models of reinforcement learning that resembles the biological nature of both the BAS and the implicit system postulated by the COVIS model. The model was able to simulate the mean behavioural data obtained in a study by Markman and colleagues (2005) where participants showed a response bias during a CL task with asymmetric payoffs. Nonetheless, the formal model showed that the simulated data does not look very "categorical" as indicated by the high noise levels and the observed learning pattern (figure 7.4). The formal model that was fitted to the simulated data is one in which participants set a criterion (pixel position) and decide if the stimulus is one side of that criterion (category A) or the other side (category B). The formal model should therefore be applied only when the underlying data produce a somewhat step-like function in response choices at the criterion. The step-like function is an intrinsic property of a boundary-based formal model. However, the formal model includes perceptual noise in perceiving the true position of the stimulus and also noise in the positioning the criterion on a particular trial. This noise acts to "soften" the step function into a sigmoidal shape. However, even data which lack any step or sigmoidal shape can be captured by the formal model, as long as the noise parameter is allowed to grow unrealistically large.

It is clear from the simulated data shown above that there is no step-like behaviour in these simulated data and the estimated noise parameters are so high because the best fitting decision curve is almost linear. Under these conditions the formal model does not meaningfully apply to the simulated data even though it produces estimates for the decision criterion and the noise. It is unlikely that real human data on this task will look like this [although Markman and colleagues (2005) do not comment on the steepness of the category boundaries in the data from individual participants].

However, despite the above problems with the nature of the "decision curve" in the simulated data, the neural model seems to have the potential to simulate the reward-related effects of payoff matrices. Therefore, the neural model may be able to capture a more categorical ('steep') learning curve following the addition of a categorical, perceptual module to the basic model described in this chapter. Indeed, the RPE neural model presented in this chapter is an implicit learning model of the kind proposed in COVIS. The categorical behaviour lacking in the simulations seems likely to be a property more naturally occurring within the other, rule-based explicit system described by COVIS. It is, thus, likely that the current single system implicit model will therefore need the addition of a second (rule-based) system in order to develop a better simulation of real human data. This will be discussed in the next chapter in relation to the data collected in the next empirical study.

The model is used in the remaining empirical chapters to simulate the relationship between inter-individual variations on the biological mechanism and performance on RB CL tasks. These simulations allow one to assess a priori the effectiveness of the task in capturing the personality-behaviour relationship, and also offer further insights into the data in relation to RST.

The model is biologically realistic although it has a very simplified structure; for example it is constituted by individual cells (cf. nuclei) and some processes are not represented (e.g. thalamic projections). Additionally, the model simplifies the response selection processes (and especially those involved in the to-be-added explicit system; see chapter 8 for details). Nonetheless, it is potentially effective at capturing human behaviour and, therefore, it represents a good starting point for developing more biologically-valid models. Indeed, simplification is an essential element of neural modelling (Pickering, 2008).

Chapter 8

The interaction between personality and payoff structure in category learning

8.1 Abstract

A study carried out by Markman and colleagues (2005) assessed the impact of asymmetric payoff matrices during learning on a rule-based category learning task. The task required participants to learn to classify probabilistic stimuli into two categories via trial and error through the use of feedback. In order to solve the categorization problem, participants had to learn to place a simple decision criterion to sort the stimuli into two categories and so optimise the points gained during the task. Participants performed under a gain payoff matrix where both correct and incorrect responses earned points, although one category offered higher payoffs for correct responses than the other. Consistent with signal detection theory (SDT), participants were found to place their decision criterion away from the optimal accuracy bound and closer to the optimal reward bound (i.e. they showed a response bias). The present study replicated these findings. Indeed, participants were found to show a response bias under asymmetric payoffs that, although less marked, resembled the bias observed by Markman and colleagues. However, no response bias was observed under symmetric payoffs. Additionally, the study explored the impact of individual differences on criterion placement in relation to the reinforcement sensitivity theory. Positive schizotypy was found to be positively associated with the criterion shift. Results were interpreted in the light of RST using the neural model described in the previous chapter. However, due to the design of the study, the results were not conclusive.

8.2 Introduction

As stated in the chapter 1, in order to fully understand the processes underlying individual differences in learning and decision-making, it is necessary to understand the interplay between motivation and cognition (Higgins, 1997; Maddox, Markman & Baldwin, 2007). Signal detection theory suggests that during categorisation, individuals learn to assign different responses to different regions of perceptual space, obtained by placing a number of decision bounds (or criteria; Maddox, Bohil & Dodd, 2003); for n categories, one needs to employ (n-1) decision bound(s). Evidence further indicates that base-rate and payoff manipulations produce a bias in the criterion location (Maddox & Bohil, 1998; Bohil & Maddox, 2001; Maddox & Bohil, 2005; Maddox, Baldwin & Markman, 2006). Hence, according to these considerations, and following the conclusions drawn from the CL study reported in chapter 6, reward payoff and/or base-rate manipulations may be an effective way to explore individual differences in motivational and cognitive factors.

The study by Markman et al. (2005) explored the relationship between motivation and cognition by assessing the impact of motivation during performance on a CL learning with asymmetric payoff matrices (see chapter 7 for details). In particular, the study looked at the influence of regulatory focus on classification learning. Beauducel and colleagues (2005) have highlighted the parallel between regulatory focus theory (RFT; Higgins, 1987) and reinforcement sensitivity theory (RST; Gray & McNaughton, 2000). Indeed, both theories suggest that there are different types of goals depending on their valence (positive or negative). Goals with a positive valence are referred to as rewards and as approach goals by RST and RFT, respectively, and both theories suggest that they lead to approach behaviour towards the desired goal. RST suggest that the behavioural activation system (BAS) is responsible for processing reward-related stimuli and mediating approach behaviour. Hence, the results obtained in the study by Markman and colleagues (2005) raise some interesting questions for RST, particularly in relation to the BAS.

Of particular interest are the results obtained under a gain-only matrix, since inter-individual differences in BAS reactivity should affect performance in this condition. The results showed that under the gain payoff matrix, participants placed their decision criterion away from the optimal accuracy bound and towards the optimal reward bound, but only if they had been induced with a promotion focus. In other words, they tended to over-classify probabilistic stimuli as members of the high payoff category (i.e. they showed a response bias). Participants obtain higher winnings when placing the criterion close to the optimal reward, compared to the optimal accuracy, criterion. Those participants who had been induced with a prevention focus, by contrast, failed to show this response bias and they placed their criterion bound close to the optimal accuracy criterion. The complementary pattern was observed under the loss payoff matrix. Therefore, the study found that having participants performing under an asymmetric payoff matrix would make their decision bound deviate from the optimal accuracy criterion towards the optimal reward one, but only when there was a regulatory fit between the induced regulatory focus (promotion or prevention) and the payoff manipulation used (i.e. rewards or losses).

Maddox and colleagues (2006; 2007) suggested that regulatory fit enhances cognitive flexibility, which allows participants to consider alternative strategies to maximise their performance depending to the task specifications. Due to the study frame and instructions, the participants' goal is to maximize gains and, consequently, to sacrifice accuracy. In this case, cognitive flexibility should increase the shift from the optimal accuracy criterion to the optimal reward criterion. Maddox and colleagues (2007) have speculated further on these results and suggested that 'regulatory fit' between the situational focus and the payoff structure may enhance cognitive flexibility by increasing the release of DA from the ventral tegmental area (VTA) to brain areas involved in working memory (e.g. anterior cingulate and prefrontal cortex) and episodic memory consolidation (hippocampus). Dopamine release from the VTA also projects widely to the limbic system which is involved in emotions and emotion-related memory formation (and so it may support positive affect generally; Carlsson, 2002). Additionally, as noted earlier, midbrain DA neurons are involved in the processing of rewarding stimuli (Schultz et al., 1997).

Even though the study by Markman and colleagues did not test RST, the results indicate that this paradigm may be an effective tool to explore this model and, in particular, the BAS. Indeed, as already noticed, these authors have suggested that regulatory fit leads to greater phasic DA activity in the mesolimbic system (Maddox et al., 2006), which is considered to be a major part of the neural basis of the BAS (Gray, 1987; Pickering, 2004; Corr, 2006). Therefore, under these conditions, the BAS should become strongly activated and, subsequently, should take control of responding. In particular, RST predicts that high BAS trait individuals should show a greater shift towards the optimal reward criterion than low BAS trait individuals (Smillie & Jackson, 2006; Smillie et al., 2007). Based on these considerations, the task appeared to represent a good tool to explore the BAS system and, in particular, to assess how different levels of BAS activation are captured by differences in BAS-related personality traits.

The neural model, described in chapter 7, was able to capture the response bias, observed in the original study. Moreover it offered an entirely different account of the data than the flexibility account presented by Markman and Maddox. The simulations indicated that the bias is driven by a DAergic mechanism that closely resembles components of the implicit system described by the COVIS model and the BAS motivational system described by RST (Ashby et al., 1998; Gray, 1987; Pickering & Gray, 2001). The response bias should, therefore, be correlated with inter-individual differences in reactivity of the DAergic system and, therefore, the personality traits stemming from such reactivity (e.g. BAS-like traits, according to RST; Pickering and Gray, 2001).

Simulations, run using the neural model described in the previous chapter, showed that the task was theoretically able to reveal a strong association between the neurobiology (i.e. DA activation) and the response bias on the task implemented by Markman and colleagues ($r = -.71$, $p <.001$; see later in this chapter for more details). Since certain personality traits may represent the phenotype of the BAS/DA system, a strong correlation between the response bias measure and the variation in the reactivity of the biological mechanism suggests that the response bias might be a good endophenotype of BAS/DA system variability. Thus, this increases confidence that it might be possible to detect the relationship between the endophenotype (response bias) and the phenotype (personality trait scores) in real empirical data, as both are affected (to differing extents) by the same underlying biological variations. A strong biology-endophenotype relationship is needed as phenotype*gene and phenotype biology associations are usually quite weak in personality research.

This occurs because both genetic and environmental factors account for personality variance (Blum et al., 2000; Ozkaragoz & Noble, 2000; Benjamin et al., 2001; Munafo et al., 2003; Pardo et al., 2007). This relationship is further weakened by empirical noise in the measurement of the phenotype (e.g. self-report bias, Lee et al., 2007; Eisenberg et al., 2007).

8.2.1 Aims of the study

The present study aims to replicate and extend the findings obtained by Markman and colleagues who found that participants shifted their criterion towards the optimal reward bound to maximise their winnings. These results are somewhat counterintuitive to COVIS assumptions, which suggest that reward manipulations should not affect performance during an RB CL task. As discussed in chapter 1, there is evidence that indicates that participants are able to implement the optimal verbal rule when performing on an RB task without accuracy feedback (e.g. Queller et al., 1999). It is possible that partial feedback manipulations do not affect learning on an RB task since efficient learning can occur even without any form of accuracy feedback (Ashby et al, 1999; Ashby et al., 2002). However, more subtle manipulations of feedback may actually affect learning in the way observed in the study by Markman and colleagues (2005). As noted above, these authors have interpreted the findings by arguing for dopaminergically-mediated changes in cognitive flexibility triggered by the feedback signals in their general motivational context. Thus, they suggested that reward feedback can affect performance on RB CL tasks, but not directly through its reinforcing properties. Their account rather suggests that reward feedback affects the executive attention and working memory processes that are critical to performance on RB CL tasks.

The present study has two main aims. The first aim is to explore how the BAS mediates performance on the task in relation to the feedback manipulation and the second aim is to identify the underlying BAS-related traits. In order to ensure a pure activation of the BAS system, it was decided to have participants perform under a gain-only payoff matrix and participants were all induced with a promotion focus.

In order to assess how BAS activation mediated the response bias from the optimal accuracy criterion towards the optimal reward criterion, a baseline measure was also collected. In order to do this, performance on the asymmetric condition was compared to performance on a symmetric payoff condition. According to RST, high BAS activation, which is associated to increased levels of phasic DA activity, leads to (reward) approach behaviour. Therefore, according to RST high BAS individuals should show the greatest shift in bias across the two task conditions. Nonetheless, evidence from the field of addiction shows that rather than increased reward sensitivity, it is reduced reward sensitivity that leads to approach behaviour (i.e. drug use; Honkanen, Hytia, Korpi & Ahtee, 1999; Volkow et al., 1999; Blum et al., 2000; Davis & Woodside, 2002). This model of addiction is known as reward deficiency syndrome (RDS; Blum et al., 2000). According to RDS, reduced reward sensitivity is determined by hypodopaminergic function[8]. Additionally the model predicts that low BAS individuals should be the ones who show the greatest shift in approach behaviour as indexed by the response bias measure. Simulation of the data via the neural model will allow one to investigate these two models in relation to the empirical data.

The empirical and theoretical evidence is currently inconclusive on which personality trait underlies the BAS system. Indeed, there is an ongoing debate on whether extraversion or impulsivity is the true BAS-trait. Thus, the present study aims to shed some light on this debate. According to the original theory (Gray, 1970; 1987; Corr, 2006), impulsivity-like personality traits should mediate performance on the task used by Markman and colleagues (2005) and determine the shift from the optimal accuracy criterion towards the optimal reward criterion in the asymmetric payoff condition. Nonetheless, there is

---

[8] Genetic and imaging studies have shown that DA hypofunction is caused by lower density of D2 receptors in the mesolimbic system (Volkow et al., 1997; Noble, 1998; Pohjalainen et al., 1998). In particular, the presence of the A1 allele on the D2 receptor gene (i.e. DRD2) is associated with a reduction (30-40%) in D2 receptors density (Pohjalainen et al., 1998; Blum et al., 1999; Reuter, et al., 2006).

growing evidence that extraversion is the personality trait underlying the BAS (Depue and Collins, 1999; Smillie, Jackson & Dalgleish, 2006) and, therefore, one may predict that extraversion might mediate the shift in bias across the two task conditions.

8.3 Method

8.3.1 Participants

An opportunity sample was recruited around Goldsmiths campus. It comprised both students and non-students over the summer holiday.

The sample was made up of 64 participants, 32 males and 32 females. All participants but one were right-handed. All participants were aged between 18 and 39 (mean age = 26.7; s.d. = 4.41).

Participants were guaranteed confidentiality. They were tested in one sitting that lasted approximately 1 hour and 15 minutes and they received £7.50 for their participation.

8.3.2 Design

The present study assessed the impact of payoff matrices on performance during a category learning task. In particular, the study sought to compare the impact of a symmetric and an asymmetric payoff matrix. Therefore, the primary independent variable was payoff matrix and it was run within-subjects with two levels, which were: symmetric and asymmetric. The order in which the two payoff matrices were encountered was counterbalanced across participants and represented the other IV which was run as a between-Ss factor. In order to obtain optimal winnings, the asymmetric payoff matrix was designed to require adoption of a signal detection decision index (β) equal to 3, whereas the symmetric matrix required a neutral response bias signified by a β equal to 1. The structure of the two payoff matrices is summarised in table 8.1.

Table 8.1. Structure of the payoff matrices and performance criteria adopted for the two task conditions (✓ = correct vs. ✗= incorrect; Asym = asymmeritc vs. Symm. = symmetric)

	High payoff category		**Low payoff category**		**Performance criterion[9]**
Matrix	✓	✗	✓	✗	
Asym.	400	100	200	100	33700
Symm.	300	100	300	100	31600

The study assessed the impact of different types of payoff matrices (IV) on decision bound location that constitutes the main dependent variable.

8.3.3 Task and apparatus

The task was run on a Mesh PC with a Mitsubishi 21" monitor with 1024 x 768 pixel resolution in an artificially lit room.

Each stimulus was computer generated by using Matlab routines from Brainard's (1997) Psychophysics Toolbox. Stimuli were flashed on a black background that filled the entire screen. Each stimulus was a small circular dot (14 pixels in diameter) that appeared on the computer screen. The stimuli appeared within a 650*650 pixels box that delineated the attention space participants had to concentrate on. The dots would appear at varying positions along an imaginary line 650-pixel in length running either horizontally or vertically along the box's vertical or horizontal midlines. Participants performed two perceptual classification tasks one after the other and, in each task, one spatial dimension (vertical or horizontal) varied whereas the other was kept constant. The order of the spatial orientation was counterbalanced over the two tasks, across participants (see counterbalancing section below).

The stimuli for category A and category B were generated by sampling from two independent but overlapping normal distributions, thus making the task

---

[9] The performance criterion for each condition was set at 80% of the points that the optimal (i.e. reward maximising) classifier obtains over 150 trials, relative to 0% accuracy.

probabilistic. The distribution used to generate one category of stimuli had a mean equal to 275 whereas the one used to generate the other category of stimuli had a mean equal to 375. Both distributions had a standard deviation equal to 100. These numbers were used to determine the position (in pixels) of the stimulus dots, as described below. As in the original study, the mean and standard deviation values were chosen so that the signal detection discriminability of the two categories (d'; pronounced d prime) was equal to 1, which rendered the task very difficult. This decision was guided by evidence showing that when the categorisation task is difficult, motivational factors have a greater impact on the way participants perform (Markman et al., 2005).

The dots were drawn, centred at these random, varying positions, measured from the left-hand end of the imaginary horizontal box mid-line (in the horizontal task) or from the top of the vertical box mid-line (in the vertical task). For half of the participants, in the first task category A stimuli had a mean position of 275 whereas category B stimuli had a mean position of 375. This assignment was reversed for the other half of the participants. The stimulus location was reversed in the second task. Similarly, for half the participants, in the first task category A stimuli were high payoff stimuli. This assignment was reversed for the other half of the participants. For each participant, this assignment was reversed in the second task they completed. In the symmetric task condition both categories had an equal payoff, therefore assigning one category as high payoff had no effect on the payoffs earned, but it was done for symmetry. This is summarised in the counterbalancing section below.

The stimuli in both tasks were presented in one of two fixed quasi-random orders to all participants. Each stimulus appeared on the screen until the participant responded.

8.3.4 Counterbalancing

It was decided to implement counterbalancing across participants in order to avoid possible order and handedness effects. There were four crossed counterbalancing factors: order of task condition (vertical or horizontal given first); order of payoff matrices (asymmetric vs. symmetric given first); order of mean location for the categories (i.e. category A stimulus on left/top of screen [i.e. mean category location 275 pixels) in first or second task]; allocation of high payoff to the two categories (i.e. category A = high payoff vs. category B = high payoff). Therefore, there were 16 possible combinations (cells) and each cell had 4 participants per cell, allocated at random.

8.3.5 Personality measures

Participants completed several personality measures, which were: the Eysenck Personality questionnaire revised (EPQ-R), the Oxford-Liverpool inventory of feelings and experiences scale (OLIFE), the sensation seeking scale (SSS), the schizotypal personality questionnaire (SPQ), the big five inventory (BFI) and the BIS/BAS questionnaire. Four personality components were extracted after running a PCA with Varimax rotation on these scales (see chapter 4 for details). The four components extracted are: Extraversion (E), Neuroticism (N), Positive Schizotypy (PS) and impulsivity-antisocial (ImpAss). The PCA was run on 232 participants, who constitute the overall number of participants tested. The components scores used in the present study are the corresponding scores extracted in the overall PCA involving 232 participants.

8.3.5 Working memory measure (WM task)

A measure of WM ability was included in the battery of measures. In particular, the task used is a measure of memory scanning ability (Stenberg, 1966). The task is composed of 14 trials. During each trial, the participant was presented with a set of letters to memorise. The first four trials consisted of sets of 4 letters and they functioned as practice trials. The experimental trials presented sets of 6 letters.

Following the presentation phase (2.5 seconds), participants were presented with single letters on the screen and had to decided whether they were members of the preceding set or not by responding ‘yes’ or ‘no’ using the corresponding keys (i.e. the 'y' and the 'n' key on the keyboard). Participants scored one point per correct identification of targets and distracters. During the experimental trials they were presented with 12 single letters (6 targets and 6 distractors), hence they could score a maximum of 12 points per trial. Since there were 10 experimental trials, the total possible score was equal to 120.

8.3.6 Procedure

The study method and procedure were approved by the Psychology Department Ethics Committee (DEC) at Goldsmiths, University of London. Testing was conducted in a testing room in the department (University of London).

Participants sat at a desk in front of a computer screen, a keyboard and stereo speakers. On entering the testing room, participants were reminded that all the collected data and information would remain anonymous and be treated with confidentiality. They were also informed that they had the right to withdraw from the study at any time and for any reason. In order to guarantee anonymity, each participant was allocated a unique ID code that became the only means of identifying the data. Participants were randomly allocated to one of the 16 cells, depending on their ID number.

All participants were given a so-called promotion focus. Indeed, they were told that the task consisted of two tasks and that, for each task, they could win an entry into a £25 draw. In order to win the entry, they had to exceed a certain amount of points (i.e. performance criterion; table 8.1). One draw was held for all 64 participants.

The participants were presented with written instructions on the computer screen. Instructions informed participants that a series of stimuli would be individually presented on the screen. Their task was to learn, by trial and error, to classify these stimuli into category A and category B. They had to do so by pressing the 'z' key or the'/?' key for category A and category B respectively. The two keys were covered by a sticky label marked as A or B, in order to make it easier for participants to remember which key corresponded to which category.

For each response they made, participants received auditory, textual and visual feedback. The auditory feedback was represented by the sound of a ringing cash register for correct responses, whereas for incorrect responses it was a mildly unpleasant buzzer sound. The textual feedback presented participants with information on their performance for each response. Participants were not explicitly told whether their response was right or wrong (i.e. accuracy feedback). Instead, they were told how many points they had won and the maximum they could have won on that trial. Feedback was presented for 300 msec. and followed by a 100-msec inter-trial interval (ITI).

The visual feedback displayed all the points that participants had won which were accumulated on a vertically oriented "point meter" displayed on the right-hand side of the computer screen. The meter consisted of a 765 pixel tall x 50 pixels wide rectangle and it was set to zero at the beginning of each task. The performance criterion was presented as a horizontal line across the meter and was labelled 'Bonus'. The performance criteria for each matrix are shown in Table 8.1. The region above the criterion line was labelled 'Yes' whereas the one below it was labelled 'No', which indicated whether the participant had won the ticket or not (figure 8.1). The screen shot is taken on the first trial of the task where no points have yet been won. The point meter was updated following each response at a rate of 1 pixel per 50 points. The part of the meter that changed flashed three times to stress the idea that the points were increasing.

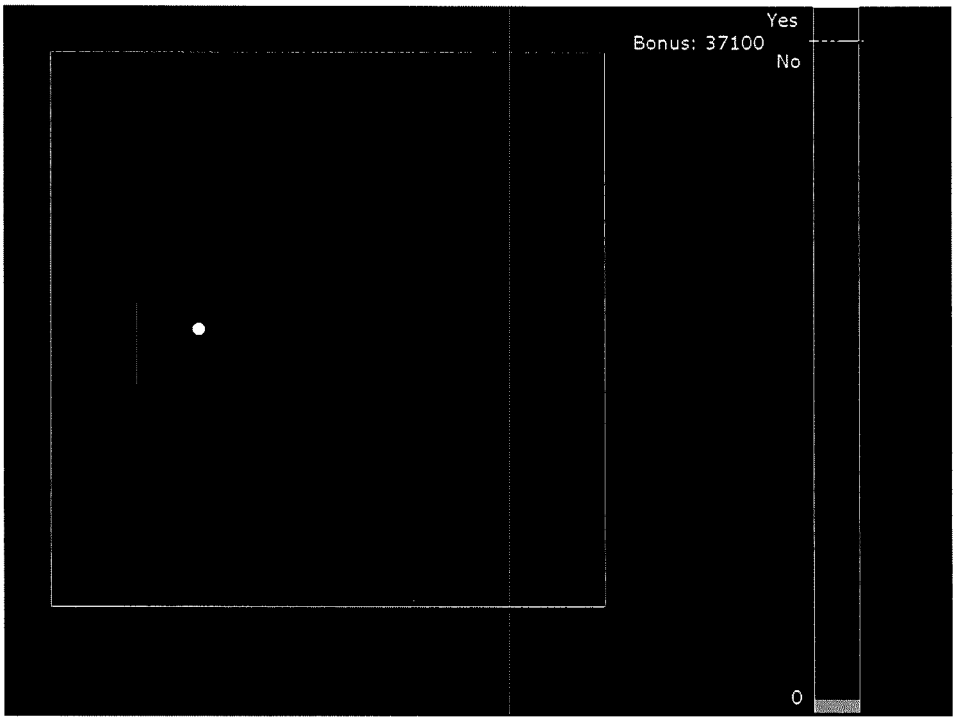


Figure 8.1. Screenshot of the task layout

Each task consisted of 3 blocks of 50 trials each. At the end of each of the first two blocks, participants received an inter-block message that informed them of their performance. The message represented a lottery ticket that was struck through by a red cross and had a message saying:

> 'If that had been the end of this section of the experiment, you would not have earned an entry into the lottery'.

However, it was not possible to have won the lottery ticket at this stage of the task. At the end of each task, participants were presented with a message that represented the same ticket as the one in the inter-block message. The ticket was either crossed out or not depending on whether participants had reached the performance criterion or not. The window dialog also reported a message saying:

> 'That is the end of this section of the experiment and you earned enough points to get an entry into the £25 lottery'.

if participants had reached the performance criterion. Otherwise, the message would read:

> 'That is the end of this section of the experiment and you did not
> earn enough points to get an entry into the £25 lottery'.

The subsequent window offered participants a summary on their performance, the message said:

> 'That is the end of this section of the experiment.
> [Sorry, but] you earned (number of) points, which means
> that you did [did not] earn an entry into the lottery for £25'.

At the end of the first task participants also received a message that informed them:

> 'You can take a short break now if you would like. Press any key when
> you are ready to continue on to the next section. You need to earn xxxxx
> points in order to win an entry.'

At the end of the two tasks, participants were debriefed by a message which said:

> 'This is the end of this experiment. Thanks for your participation!
> You earned x entries into the £25 lottery,'
> Please tell the experimenter you have finished, and
> that you earned x entries'.

Depending on their performance over the two tasks, x could be 0, 1 or 2.

Participants were invited to sit in front of the computer in a comfortable way, in order to move as little as possible during the trials. Participants were encouraged to express any doubts about the task and to start only once they felt completely confident about the task instructions. At this point, the experimenter left the testing room to let participants carry out the task on their own.

8.3.8 Data analysis

A data processing programme was developed to create an overall summary file that could be exported from Matlab into SPSS and, therefore, allow statistical analysis on the data. The programme extracted five main measures for each of the asymmetric and symmetric conditions, which were: proportion of correct responses (PC), mean reaction time (RT), the location parameter (C), beta and dprime. These last three values were calculated using formulae from signal detection theory (SDT; Stanislaw and Todorov, 1999; Maddox & Dodd, 2001; Bohil and Maddox, 2001). The location parameter (C) reflects where subjects place their decision boundary. The C parameter is such that a score equal to zero indicated an optimal accuracy criterion placement. The value was recoded (in relation to the counterbalancing condition) so that a positive value reflected a bias away from the centre of the high payoff category towards the maximal reward criterion.

However, calculated SDT measures do not offer a rigorous measure of how payoff manipulations affect the decision criterion location. They can be calculated on a set of data even when the participant did not perform in accordance with the assumptions of SDT. Hence, a formal SDT model was fitted to the data. The use of a formal model allows one to establish whether each participant was actually solving the categorisation problem by implementing the appropriate uni-dimensional rule. The appropriate rule is based on the relevant dimension, which was the location of the stimulus on the screen (i.e. left or right in the horizontal condition vs. up or down in the vertical condition). The formal modelling was able to identify individuals who did not use a criterion to solve the task (as implied by the SDT model) and they were excluded from the analysis as they are a source of noise. Simply calculating SDT scores does not assess whether a particular participant has used an SDT-like strategy to guide their performance. Including the SDT scores from individuals who are not using an SDT-like model is comparable to measuring the diameter of (orange) tennis balls in a study that aims to assess the impact of a fertiliser on the size of oranges. The measure itself would be easily computed (i.e. a diameter) but it would be meaningless.

Following Markman and colleagues' data analysis, only the middle 100 (out of the total 150) trials were analysed. The decision was made since the first 25 trials constitute practice and familiarisation trials where the appropriate rule was being established by participants, whereas the last 25 may constitute noisy responses since participants had often started reaching criterion by then.

Finally, the results obtained in the empirical study were further explored by simulating the data through the neural model described in the previous chapter. The neural model allows one to obtain insight into the neuro-biological processes that may underlie the behaviour choices as well as the relationship between individual differences and behaviour.

8.4 Results

8.4.1 Human data analysis

Owing to equipment failure, data from one participant (out of the total N = 64) was not recorded during performance on the asymmetric task, when the symmetric task had been performed first (CB matrix order 2).

8.4.2 Preliminary analysis

8.4.2.1 Proportion of correct responses

Preliminary analyses were conducted on the proportion of correct responses in the two task conditions only taking into consideration the mid-100 trials. The analysis observed that the proportion of correct (pc) responses was close to chance levels in both task conditions. Indeed, it was equal to 51% in the symmetric condition and 47% in the asymmetric condition. Even after excluding those participants who did not use the adequate uni-dimensional rule, pc levels were equal to 52% and 47% in the symmetric and the asymmetric condition, respectively. The proportion of correct responses did not differ as a function of matrix condition [t(62) = 1.16, ns]. These pc scores were not much higher when looking at all 150 trials or the last 100 (table 8.2).

Table 8.2. Proportion of correct (pc) responses in the two conditions tasks when analysing all 150 trials (All 150T) or the last 100 trials (last 100T)

	Asymmetric		Symmetric	
	Last 100 T	All 150 T	Last 100 T	All 150 T
PC mean (s.d.)	.48 (.15)	.48 (.13)	.52 (.14)	.51 (.14)

8.4.2.2 Reaction time scores

Table 8.3 summarises the reaction time (RT) scores in the two task conditions and, as expected by looking at the table, a paired-sample t-test indicated that RT scores were not significantly different across the two matrix conditions [t(62) = -1.65, ns].

Table 8.3. RT scores across the two task conditions in seconds

	Asymmetric	Symmetric
RT mean (s.d.)	1.20 (.79)	1.07 (.46)

8.4.2.3 Points earned

Participants gained points throughout the task and aimed to reach the performance criterion to be able to earn entries into the £25 lottery. The performance criterion for the asymmetric condition was equal to 33700 while the criterion for the symmetric condition was equal to 31600. Table 8.4 shows the amount of points earned across the two conditions by the whole sample and by those participants who had implemented the appropriate uni-dimensional strategy[10].

---

[10] This is assessed by the formal model fitting described below

Table 8.4. Points (mean and standard deviation) earned by the overall sample (N = 63) and by those participants who implemented the optimal uni-dimensional rule (UD; N = 35).

	Asymmetric	Symmetric
Overall sample (N=63)	33523.8 (2683.0)	33231.7 (2298.4)
UD-users (N=35)	35045.7 (1210.1)	34497.1 (1451.5)

Those participants who implemented the appropriate rule gained more points than the overall sample (i.e. which included participants who used alternative strategies). Moreover, results showed that, amongst those participants who used the appropriate uni-dimensional rule (i.e. UD-users), the difference of gained points across conditions was almost significant [t(34) = 2.03, p = .051, 2-tailed]. Participants gained more points in the asymmetric rather than in the symmetric condition (table 8.4).

8.4.2.4 Model fitting

The stimuli presented in the two task conditions consisted of dots that varied on one dimension, which was their location on the screen. Location was the relevant dimension and participants needed to use it to develop the appropriate strategy to solve the categorisation problem. Hence, a uni-dimensional model which used location as the relevant dimension was fitted to the data as well as a guessing model. The two formal models were run for each participant's data individually and separately on the data collected in the asymmetric and the symmetric task conditions. Moreover, the goodness of each model was compared to the goodness of fit of the saturated model with no free parameters. The saturated model described the data perfectly. All models' parameters were fit using maximum likelihood estimation.

1 – Fitting by a guessing model

The guessing model describes the situation where participants are randomly assigning the stimuli to one of the two categories, regardless of the location of the stimuli on the screen. This model has only one free parameter (guessing probability for category A, from 0 to 1). Results showed that the guessing model's fit to the data was worse than that of the saturated model in all cases.

2 – Fitting by a uni-dimensional rule model based on spatial location

The present model describes the situation where participants select spatial location as the relevant dimension and place a criterion on this dimension to categorise the dots.

Results indicated that the one-dimensional (UD) model was not significantly worse than the saturated model for 35 cases. These 35 participants belonged equally to the two payoff matrix conditions, in fact 18 performed on the asymmetric task first and 17 performed on the symmetric task first. This indicates that 55% of the whole sample clearly used spatial position as a relevant dimension to classify the stimuli. Figure 8.2 illustrates the typical distribution of data that was well-fitted by the uni-dimensional model (figure 8.2a) and the distribution typical of data not fitted by the uni-dimensional model (figure 8.2b).

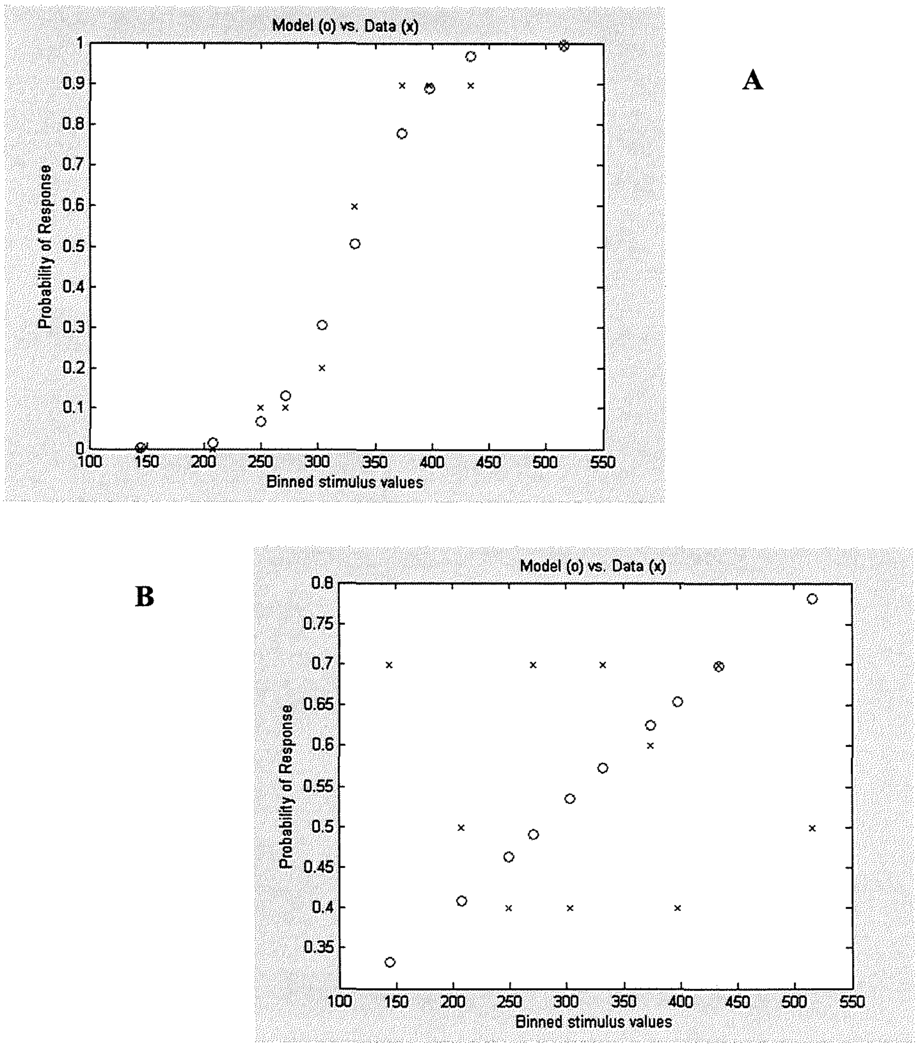


Figure 8.2. It illustrates the distribution typical of data that was well-fitted by the uni-dimensional model (figure a) and the distribution of data that was not fitted by the data (figure b). The red stars represent the actual data whereas the blue circles represent the best-fitting uni-dimensional model.

The remaining 45% of the sample did not use the appropriate strategy to perform on the task but it was not found to be guessing either. These results suggest that those participants who did not use the correct dimensional rule may have been using alternative strategies. During debriefing some participants verbally reported having treated the distance between the last stimulus on the screen and the latest stimulus as the relevant dimension to formulate their decisions. Hence the rule they developed could have been: 'if the distance of the latest dot from the last one is great the category belongs to category A, otherwise it belongs to category B'. These participants are occasionally referred to as non-learners in the chapter to indicate the fact that they failed to implement the appropriate UD strategy.

The UD formal model identified the actual (fitted) decision criterion placed by each individual who employed the appropriate rule. The decision criterion was also calculated using signal detection theory (C[11]). Correlations were run to establish the relationship between the calculated and the fitted criteria and they were found to be highly positively correlated (table 8.5). Since the fitted criterion is used to eliminate poor-fitting participants, the correlations were conducted only including those 35 participants who were well-fitted by he UD model.

Table 8.5. Correlations between the fitted and the calculated criteria only for those participants well-fitted by the uni-dimensional model

		calculated_a	calculated_s
fitted_s	Pearson Correlation	-.209	.967
	Sig. (2-tailed)	.227	.000
	N	35	35
fitted_a	Pearson Correlation	.962	-.230
	Sig. (2-tailed)	.000	.184
	N	35	35

These results suggest that the calculated criteria are close to the actual fitted criteria. Nonetheless, the latter scores are a more reliable index of performance (pixel vs. standardised measure) and they identify those participants who are well-described by the optimal rule and, therefore, exclude those who were using an alternative strategy (i.e. source of noise). Thus, the dependent variables, used in the following analyses, primarily use the fitted, rather than the calculated, decision criterion.

---

[11] C = -0.5.*(norminv(phit)+norminv(pfa)), where phit represents the proportion of correct responses made to category B stimuli whereas pfa represents is the proportion of false alarm responses made to category A items

8.4.2.5 Behavioural data

Some preliminary analyses were carried out to explore the behavioural data and observe where participants placed their decision criteria under the two payoff matrix conditions. In a first stage, analyses were carried out with criteria obtained from the formal model (i.e. fitted criteria) since, as previously discussed, the fitted criteria are a more informative measure of performance.

A mixed-subjects ANOVA was run with two main IVs, which were: payoff matrix and payoff matrix order. Payoff matrix had two levels (symmetric vs. asymmetric) and it was a within-subject factor. The payoff matrix order had two levels (symmetric first vs. asymmetric first) and it was a between-subject factor. The analysis included only those participants who used the appropriate strategy (N = 35). Results indicated that there was a significant 2-way interaction between payoff*matrix order [F(1, 33) = 6.09, p = .019]. The main effect of payoff was non-significant [F(1,33) = 1.73, ns] and neither was the main effect of matrix order [F(1,61) = .23, ns]. Figure 8.3 indicates that when participants performed on the asymmetric task condition first, they showed a positive response bias in the asymmetric task and they showed no bias in the symmetric task. This is the pattern expected if participants shift their criterion to maximise winnings in the asymmetric task condition. However, participants who experienced the symmetric task first showed a bizarre pattern of responding. In fact, they showed a positive bias in the symmetric condition and no bias in the asymmetric condition.

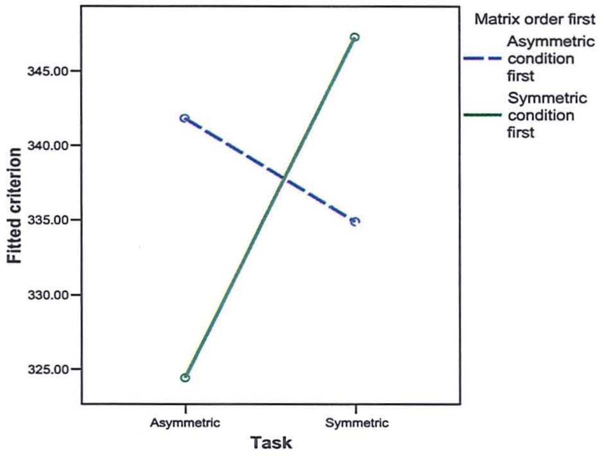


Figure 8.3. Relationship between the fitted decision criterion across the two task conditions and matrix order (decision criterion equal to 325 pixels indicates that the criterion is placed at the optimal accuracy criterion)

Post-hoc t-tests were carried out to assess whether the difference between criterion location in the asymmetric and the symmetric conditions was statistically different depending on the order of the payoff conditions. A repeated t-test showed that, for those participants who experienced the symmetric task first, the difference between the asymmetric and the symmetric criteria was significant [t(16) = -3.32, p = .004]. In contrast, this difference was not significant for those individuals who experienced the asymmetric condition first [t(17) = .71, ns].

Owing to the observation that performance of those who experienced the symmetric task first was bizarre and counterintuitive (i.e. higher shift in bias in the symmetric payoff matrix condition), data from these participants were excluded from further analysis.

In the asymmetric condition, these 18 participants placed their criterion at a mean of 341.84 pixels (s.d. = 31.1) and showed a mean noise level equal to 98.28 (s.d. = 47.2). These results indicate that the response bias in the present study is much lower than the bias observed in the original study by Markman and colleagues (approximately 430). In the symmetric condition the criterion was set at 334.89 pixels (s.d. = 22.3) with a noise level 73.49 (s.d. = 41.9). The asymmetric criterion was statistically above the optimal accuracy criterion [325; t(17) = 2.30, p = .035 2-tailed] and the symmetric criterion was non-significantly above the optimal accuracy criterion [t(17) = 1.89, p = .077]. As already noted, there was no significant difference in criterion placement between the two payoff conditions.

8.4.2.6 Correlations between criterion location measures and the amount of points earned during the task

A set of correlations was carried out to assess the relationship between the fitted criterion measures, the proportion of correct responses and the number of points earned during the two task conditions. This set of correlations indicated that there was a significant trend for a positive correlation between the amount of points earned and the criterion used during the asymmetric payoff matrix condition (r = .42, p = .08). Hence, consistent with expectations, the greater the response bias (i.e. towards the optimal reward criterion), the greater the amount of points earned. Results also showed a negative correlation between the number of points earned and the decision criterion set during the symmetric payoff matrix condition (r = -.55, p = .02). Those participants who were closer to the optimal accuracy criterion in the symmetric task (mean = 325; i.e. they showed lower criterion biases) were more accurate and, as a direct consequence, gained more points.

Overall these results indicate that behavioural performance on the two conditions required different strategies in order to maximise winnings. Indeed, participants had to place their criterion closer to the optimal accuracy criterion in the symmetric condition but shift their criterion away from it, closer to the optimal reward criterion, in the asymmetric condition in order to maximise their winnings.

8.4.2.7 Correlations between measures of task performance and scores on the working memory task

Correlations were run between the scores on the working memory task (i.e. the proportion of correct responses) and scores on the category learning task, indexed by the decision criteria and the number of points earned in the two conditions. Results showed that there was no significant correlation between any of these measures (p > .23; table 8.6).

Table 8.6. Correlations between scores on the working memory (WM) task and task performance measures (fitted = criterion calculated by the formal model; score = points earned; a, asym = asymmetric; s, symm = symmetric)

		fitted_a	fitted_s	asym_score	symm_score
WM scores	Pearson Correlation	-.013	-.300	.238	.293
	Sig. (2-tailed)	.960	.226	.342	.239
	N	18	18	18	18

8.4.2.8 Performance and personality

The analyses reported below aimed to establish the relationship between the four personality components extracted through the PCA (chapter 4) and performance during the two task conditions, in particular the fitted criterion. These analyses were also only conducted for the data of those participants who experienced the asymmetric task first and who developed the appropriate strategy (N = 18).

8.4.2.9 Correlations between personality and the criteria set during the symmetric and the asymmetric task

A series of correlations assessed the relationship between the four extracted personality components and the criterion location scores. Results showed that there was a significant positive correlation between scores on the positive schizotypy (PS) component and the fitted criterion in the asymmetric task (r = .56, p = .016; table 8.7). Hence, participants with high scores on the PS component tended to show a greater criterion bias than their low counterparts during performance on the asymmetric task. None of the personality components was significantly correlated with the decision criterion shown during the symmetric conditions.

Table 8.7. Correlations between the four personality components and decision criteria in the two task conditions (fitted_s = criterion set in the symmetric condition; fitted_a = criterion set in the asymmetric condition)

		fitted_a	fitted_s
Extraversion	Pearson Correlation	-.354	.006
	Sig. (2-tailed)	.150	.980
	N	18	18
Neuroticism	Pearson Correlation	.092	.328
	Sig. (2-tailed)	.718	.184
	N	18	18
Positive schizotypy	Pearson Correlation	.557	-.115
	Sig. (2-tailed)	.016	.649
	N	18	18
Impass	Pearson Correlation	-.150	-.124
	Sig. (2-tailed)	.552	.625
	N	18	18

8.4.2.10 Mixed between-within measures ANCOVAs

Several ANCOVAs were run to assess the impact of personality factors on the decision criterion during the two tasks. Since the personality components were continuous factors, it was necessary to treat them as covariates. Owing to the small sample size included in the analyses once participants who did not apply the appropriate positional decision rule were excluded (non-learners), each personality component was entered separately.

Results supported the observations made in the exploratory correlational analyses and they showed that positive schizotypy affected the shift in bias across the two task conditions. In fact, there was a significant 2-way interaction between positive schizotypy scores and payoff [F(1,16) = 4.76, p = .044]. As observed in the behavioural analyses, the main effect of payoff was not significant [F(1,16) = .43, ns].

These results support the correlation results and, therefore, it was decided to further explore the impact of positive schizotypy on performance.

In order to gain a better understanding of the significant interaction between positive schizotypy and payoff, it was decided to calculate the criterion shift across conditions (i.e. asymmetric criterion - symmetric criterion). The correlation identified a significant correlation between the two factors (r = .48, p = .044), which is expected since this analysis is mathematically identical to the interaction between PS and payoff, observed in the ANCOVA analysis. The Scatterplot visually exemplifies the present results (figure 8.4).

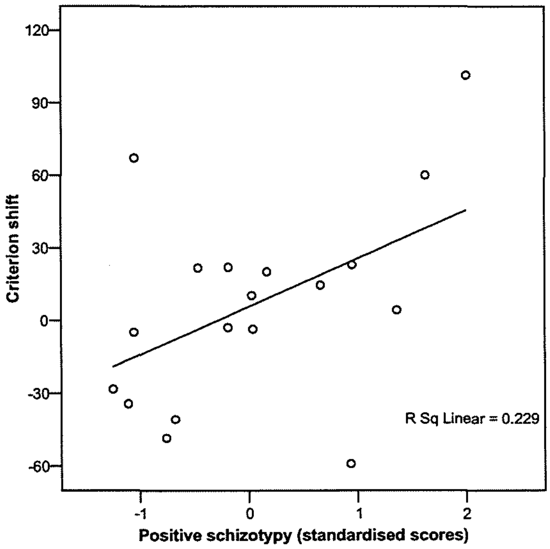


Figure 8.4. Scatterplot summarising the relationship between positive schizotypy and the criterion shift across condition (shift = asymmetric – symmetric)

A further graphical aid to understand the present results was used by applying the median split technique on the PS component, in order to transform it into a binary valued factor, which was coded as 1 for low scores and as 2 for high scores. However, the median split technique was used purely as a visual aid (figure 8.5) rather than as an analysis tool since, owing to the small sample size, power was very low.

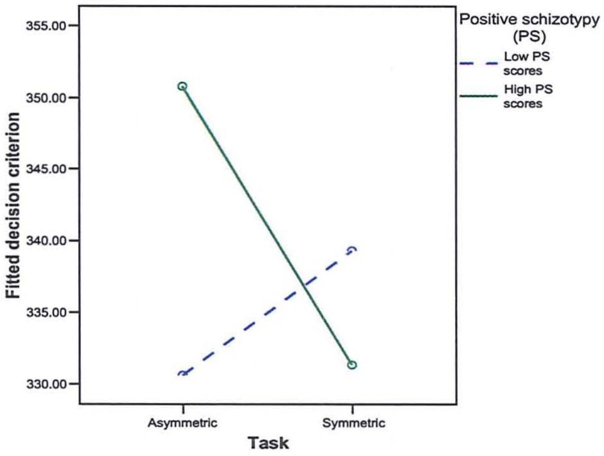


Figure 8.5. Relationship between the fitted decision criterion across the two task conditions and scores on the positive schizotypy component

8.4.2.11 Relationship between covariance measures, personality components and strategy used

Point-biserial correlations were run in order to establish whether the type of strategy used (appropriate uni-dimensional rule coded as 1 vs. alternative strategies coded as 0) in the two conditions was associated with scores on the WM task and/or any of the personality components. Results showed that scores on the WM task were not associated with the type of strategy implemented during performance (r = -.01, ns). Moreover, a trend for a significant negative correlation was found between scores on the extraversion component and the type of strategy used in the symmetric condition (r = -.33, p = .06). Hence, according to these results it seems that introverts were more likely to use the appropriate uni-dimensional rule in the symmetric task. No personality component was found to be associated with the type of strategy implemented in the asymmetric task (ps > 0.18).

As expected, the use of the appropriate unidimensional strategy was positively correlated with the amount of points earned in both the asymmetric condition (r = .84, p < .001) and the symmetric (r = .66, p < .001) condition.

8.4.2.12 Non-learners

Results showed that among non-learners who performed the asymmetric task first, the accuracy levels were non-significantly below chance levels in both task conditions {asymmetric: 48%, [t(13) = -.96, ns]; symmetric: 47%, [t(13) = -.66, ns]}. Hence, in the present study the use of alternative strategy was associated with lower accuracy and greater response bias in the asymmetric task; similarly, they showed a negative bias in the symmetric condition (table 8.8). The noise levels were significantly higher among those participants who did not implement a uni-dimensional strategy than among those who used the appropriate dimensional strategy in both the asymmetric [t(30) = 8.0, p <.001] and the symmetric condition [t(30) = 5.6, p <.001].

Table 8.8. Criteria and noise levels shown by those participants who implemented a uni-dimensional strategy (UD-users; N = 18) and those who implemented an alternative strategy (non-UD users; N = 14).

	Crit_a	Noise_a	Crit_s	Noise_s
UD users	341.84 (31.1)	98.28 (47.2)	334.89 (22.3)	73.49 (41.9)
Non-UD users	344.70 (35.5)	260.52 (68.2)	304.95 (53.3)	214.43 (97.0)

Overall, the results are inconclusive due to the aberrant behaviour observed in participants who experienced the symmetric task condition first and due to the limited number of participants who learnt to implement the optimal rule.

8.4.3 Preliminary simulations

8.4.3.1 Preliminary simulations with the RPE-based neural model

Preliminary simulations were conducted in order to ensure that the RPE-model could capture human behaviour. Indeed, in the previous chapter, it was observed that the model could simulate the response bias in the asymmetric task but could not capture category learning, as indicated by the fact that the simulated data had very high noise levels and did not reproduce the step-like learning function observed in the human data (figure 8.7). The simulated data were fitted using the same formal model used to extract the fitted criterion in the human data and chapter 7.

The RPE-based model was able to reproduce the shift in bias across the two task conditions observed by Markman and colleagues (2005) and in the present study. The values used for the SR and RP learning rate parameters were the same as the ones used in chapter 7, to simulate the data from the Markman and colleagues' study. The reinforcement (**rf**) values are the same used in the simulations reported in chapter 7 (table 8.9)

Table 8.9. Reinforcement (rf) values used for the two payoff conditions

	High payoff category		**Low payoff category**	
	✓	✗	✓	✗
Asymmetric	1	0.25	0.5	0.25
Symmetric	0.75	0.25	0.75	0.25

These **rf** values were multiplied by a constant value (i.e. **m**) equal to 1 (cf. **m** was set equal to 0.5 in chapter 7) and the **rptonic** value had to be set equal to 0.3 (cf. **rptonic** = 0.5 in chapter 7) to simulate the human data. All other parameter values, however, were maintained from the simulations in the previous chapter data (table 8.10).

Table 8.10. Parameters' values implemented in the simulations

Symbol	**Description**	**Value**
rptonic	**Reward prediction tonic value**	0.3
m	**Scales the rf values**	1
$\mathbf{b_1/b_2}$	**Learning rates of the RP unit**	0.05*scaler1
$\mathbf{b_3/b_4}$	**Learning rates of the SR unit**	0.05*scaler2
Scaler1	**It scales the learning rate of RP unit**	2
Scaler2	**It scales the learning rate of RP unit**	2

As predicted, in the asymmetric condition, simulated 'participants' showed a response bias in favour of the high payoff category as captured by a criterion equal to 340.30 (s.d. 53.3). In contrast, they showed no response bias in the symmetric task as indicated by a criterion location mean equal to 325.43 (s.d. 16.0).

However, the simulated 'participants' showed noise mean levels much higher than those shown by those human participants, who were well-fitted by the uni-dimensional model (table 8.11).

Table 8.11.Simulated and actual (human) criteria and noise levels obtained in the asymmetric and the symmetric condition (asy = asymmetric, sym = symmetric)

	Criterion asy	Noise asy	Criterion sym	Noise sym
Simulated (N =300)	340.30 (53.3)	248.38 (52.1)	325.43 (16.0)	292.66 (18.5)
Human (N=18)	341.84 (31.1)	98.28 (47.2)	334.89 (22.3)	73.49 (41.9)

Owing to the high noise levels, the model showed quite flat learning functions rather than the step-like learning functions displayed by human participants (figure 8.6).

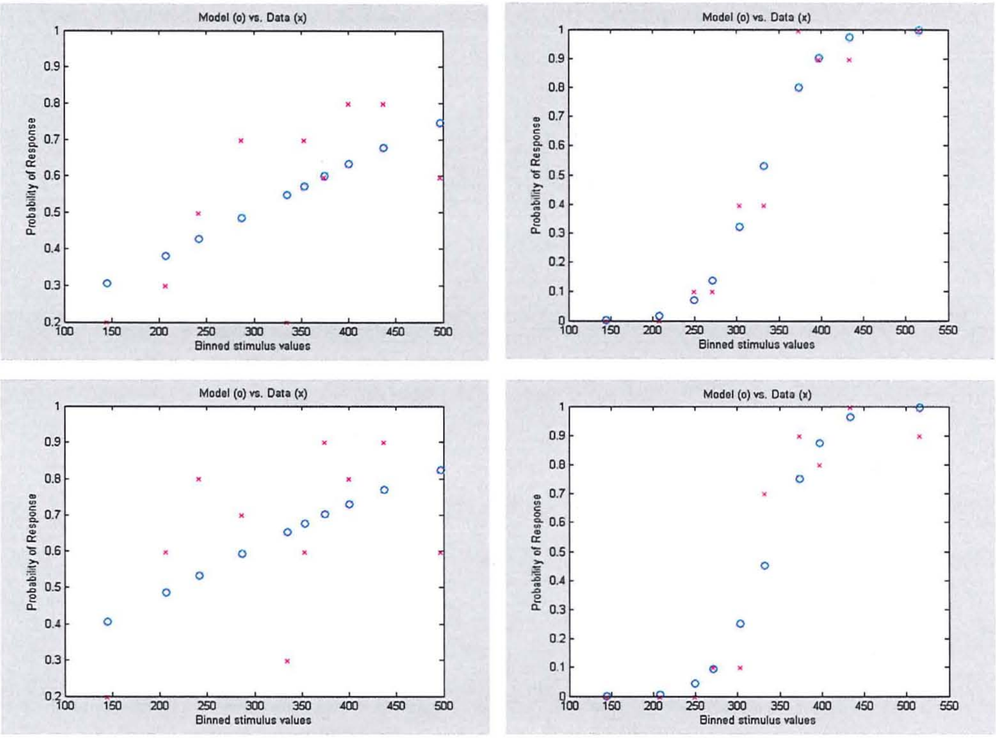


Figure 8.6. The figures represent the learning functions observed across typical participants from the simulated (left) and the human (right) data in the asymmetric (top) and the symmetric (bottom) task conditions. The red stars represent the actual data whereas the blue circles the model fitting data.

These observations indicate that even though the model is able to capture the response bias observed in the empirical data, it was not able to capture the learning process shown by a significant number of the human participants. Nonetheless, the simulated flat learning curves resemble the learning patterns observed among those participants who were not fitted by the uni-dimensional model (i.e. participants using an alternative strategy; figure 8.7).

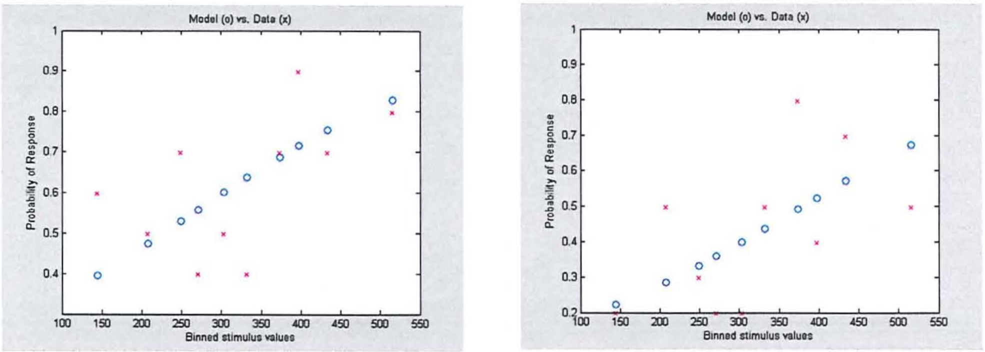


Figure 8.7. Learning functions observed among those participants who were not fitted by the uni-dimensional formal model (i.e. were guessing or using an alternative strategy). The red stars represent the actual data whereas the blue circles the model fitting data.

Additionally, both the simulated 'participants' and the human participants, who failed to implement the appropriate dimensional strategy, showed similar high levels of noise (i.e. noise > 200; table 8.12).

Table 8.12. Noise levels shown by simulated 'participants' and those participants who did not use a dimensional strategy (non-UD; asy = asymmetric vs. sym = symmetric)

	Noise asy	Noise sym
Simulated (N =300)	248.38 (52.1)	292.66 (18.5)
Non-UD (N=14)	260.52 (68.2)	214.43 (97.0)

Overall, these results suggest that two processes may be simultaneously active and mediate different aspects of learning. Indeed, the data shows that the RPE-based process may be involved in mediating reward-related learning which establishes the relationship between stimuli and appropriate responses. According to these results, reward-related learning is mediated by DA activity but it is not able to capture the actual categorical, step-like decision function. In other words, the DA-based, implicit system is involved in gradually fine-tuning the solution to the categorisation problem (i.e. which stimuli belong to which category), but does not develop a uni-dimensional strategy. A rule-based process may be active during performance on this task (which is formally at least a RB CL task) and this may allow participants to develop a dimensional strategy more rapidly, and one which has sharper category boundaries.

Thus, a revised version of the model was developed to capture the learning process observed in the behavioural data. The revised model adds an explicit rule-based module to the more implicit, RPE-mediated, process described in this section.

8.4.3.2 Rule-based component of the Model

The revised version of the model includes both the implicit process from the original model and an explicit, rule-based process that is able to capture the step-like categorisation function observed in much of the human behavioural data. As described in the previous section, the implicit learning is captured by the RPE model and is, thus, mediated by the BAS.

The revised model relies on participants applying an explicit, rule-based strategy to learn to categorise the different stimuli into the relevant categories. For simplicity, the explicit module is not modelled in a neural fashion but it is implemented mathematically. This simplification of the model is justified by the fact that the functioning of the explicit model is not at the heart of the simulations. The explicit module solves the categorisation problem by implementing a simple and verbal rule (e.g. stimuli on the left of the screen belong to category A; otherwise B). This rule is based on signal detection theory (SDT) and it accounts for the patterns of generalisation observed during the task.

According to signal detection theory (SDT), the participants are expected to split the perceptual space into two regions associated with category A and category B by identifying the 'optimal decision bound' (Bohil & Maddox, 2001; Maddox & Dodd, 2001). This perceptual decision criterion facilitates the task of classifying stimuli as members of either category by applying a uni-dimensional strategy. In the present neural model, the optimal rule requires participants to split the screen into two equal spatial zones and use this as a perceptual decision bound.

Hence, equation 7.6 is modified in order to include an explicit rule unit that ensures that the one-dimensional strategy has been implemented with a sharp categorical boundary to solve the categorisation problem (equation 8.1):

$$\mathbf{y_j^{in} = \Sigma x_i^{out} * w_{ij} + rulebias{*}c_j} \qquad \textbf{Eq. 8. 1}$$

where c_j (which is either 0 or 1) is defined below in equation 8.2. The **rulebias** parameter is a biasing output which represents the output from a "rule" module and it determines how strongly participants apply the explicit rule to solve the categorisation problem. This parameter determines the relative influence of the rule module during performance on the task with the actor units receiving the biasing input. The size of **rulebias** parameter (which can take any positive value), relative to the inputs from the "raw" cortical representation of the stimulus, determines the degree of rule influence. Hence, in the revised neural model, the inputs into the actor units are a combination of inputs from the cortical stimulus coding units (gated by SR weights $\mathbf{w_{ij}}$) and a perceptual rule bias signal applied to one of the response selection units. The same winner-take-all rule for the response selection that was used in the RPE-only model is implemented in the revised version of the model. Once again, the model includes a simplification that is probably at odds with the real underlying neural architecture. In the model, convergence of the rule and SR procedural information occurs at the striatum where response selection occurs between competing responses. The anatomy of corticostriatal projections (Seger, 2008) suggests that the likely convergence of different kinds of cortical inputs will take place on the output side of (i.e. downstream of) the striatum. The behaviour of the model is unlikely to be greatly affected by this simplification.

A perceptual rule is used to determine which actor unit receives the rulebias input. The rule is quite simple and it relies on the assumption that category membership is determined by comparing the perceived position of the stimulus on trial n (**K'[n]**) to the decision bound (threshold, **θ**; equation 8.2):

$$\textbf{if } \mathbf{K'[n] <= \theta} \textbf{ then } \mathbf{c_1 = 1} \textbf{ and } \mathbf{c_2 = 0}\textbf{; otherwise, } \mathbf{c_1 = 0} \textbf{ and } \mathbf{c_2 = 1} \qquad \textbf{Eq. 8. 2}$$

This decision process is in line with the signal detection theory (SDT) model, the simulated participants are expected to split the perceptual space into two regions associated with category A and category B by identifying the 'optimal decision bound' (Bohil & Maddox, 2001; Maddox & Dodd, 2001). However, the present model does not assume that the participants are striving for a perceptual division of the screen which would optimise their decision making accuracy. Rather, it assumes that the perceptual division is based on principles of simplicity, i.e. dividing the screen into equal-sized zones. In many tasks, as in the current one, this perceptual division coincidentally does correspond to the optimal accuracy decision bound.

For a task with x categories, the rule model requires x – 1 threshold(s) to solve the categorisation problem. This perceptual decision criterion facilitates the task of classifying stimuli as members of each category. In the present neural model, the simulated 'participants' are expected to split the screen into x spatial zones after identifying a decision bound that divides up the stimulus space into equal-sized zones. According to equation 8.2, during performance on a 2-category task, stimuli with pixel location values lower than the threshold are associated with the response selection unit denoted by the value **j**=1. For simplicity it is assumed that this is the response selection unit that most often leads to the correct response for that stimulus. In doing this the counterbalancing used in the real experiment is ignored. A further assumption is that the correct mapping (e.g., 'left' screen stimuli should receive a button A response), for a particular subject, is acquired rapidly during the first few trials of the task and this process is not modelled.

There are simple ways to capture the rule module in a somewhat more neural fashion, in future refinements of the model. For example, one might employ a cortical module which recodes the raw pixel value of the stimulus into a simple high-level representation of its global screen position (e.g. as a pair of binary inputs [10] if the stimulus is on the left; and [01] if the stimulus is on the right). This could then serve as another source of input to the striatal actor and critic units, with modifiable synapses trained in the same fashion as the synapses from the pixel-position, topographic representation which is currently employed.

Maddox and Bohil (1998), amongst many others, have indicated that human signal detection performance is characterised by trial-by-trial error which is caused by both perceptual and criterial noise. Perceptual noise is variability in the way participants perceive identical stimuli over trials whereas criterial noise indicates the variability in the way participants place their criterion over trials (i.e. the spatial zones discussed above). Over trials, participants can show different responses to similar stimuli due to both types of noise. Therefore, the present neural model includes noise in order to simulate more human-like performance data. The effects of the noise parameter were added to the actual stimulus location on trial n (**K[n]**) in order to capture its effect on performance (equation 8.3):

$$\mathbf{K'[n] = K[n] + E[n]} \qquad \textbf{Eq. 8. 3}$$

where **E[n]** represents the error element, determined by combined perceptual and criterial noise, on trial n. **E]n]** is a random variable drawn from a normal distribution with mean equal to 0 and standard deviation equal to **pnoise**. This is one of the main parameters in the model whose value varies across the simulations. Having set the **pnoise** value equal to 50 implies that on 95% of trials participants will perceive the centre of the stimulus as being within +/- 100 pixels of the true value. Thus, over trials perceptual and criterial noise (which are combined here into a single error term here) vary stimulus perception and so, indirectly, affect the criterion location process.

8.4.3.3 Preliminary simulations with the rule-based model

Preliminary simulations were conducted in order to capture the behavioural data displayed by human participants (i.e. decision bound and noise). The simulations were conducted using a learning parameter (**scaler**) equal to 2, **rulebias** was set equal to 0.6, **rptonic** was set equal to 0.5 and **rf** values were set equal to 0.8, 0.6, 0.4 and 0.2 for 400, 300, 200 and 100 points, respectively (since **m**, the **rf** scaler, was set equal to 0.8). This set of parameter values (set A) was chosen due to its ability to capture the behavioural data of those individuals who were well-fitted by the formal model and, following previous considerations, those who had experienced the asymmetric task first[12]. Indeed, the simulated criterion and noise level values were close to those values observed among human participants (table 8.13). Thus, the stimulus set implemented in the neural model produced a good simulation of the data from the human participants who were well-fit by the uni-dimensional formal model.

Table 8.13. Simulated and actual (human) criteria and noise levels obtained in the asymmetric and the symmetric condition (asy = asymmetric, sym = symmetric)

	Criterion asy	Noise asy	Criterion sym	Noise sym
Simulated (N =300)	341.67 (8.9)	69.82 (8.8)	327.15 (9.0)	59.79 (10.0)
Human (N=18)	341.84 (31.1)	98.28 (47.1)	334.89 (22.3)	73.49 (41.9)

A possible reason why the human participants showed a slight bias in the symmetric condition may be as a result of a carry-over effect from having experienced the asymmetric task first.

8.4.3.4 Procedural module parameters

The model was also found to be able to simulate various magnitudes of response bias as a function of the SR learning weight parameter that modulates the learning rate (Δw_{ij}) by scaling the RPE signal (equation 8.4).

Eq. 8. 4

$$\text{if } \mathbf{RPE} \geq \mathbf{0},\ \Delta \mathbf{w_{ij}(n)} = \mathbf{b_3} * \mathbf{RPE} * \mathbf{y_j^{out}} * \mathbf{x_i^{out}}\ ;\ \text{else if } \mathbf{RPE} < \mathbf{0},\ \Delta \mathbf{w_{ij}(n)} = \mathbf{b_4} * \mathbf{RPE} * \mathbf{y_j^{out}} * \mathbf{x_i^{out}}$$

where $\mathbf{b_3}$ and $\mathbf{b_4}$ represent the learning and extinction rates of the actor (SR) unit with positive and negative RPE signals, respectively. The $\mathbf{b_3}$ and $\mathbf{b_4}$ parameters are equal to:

---

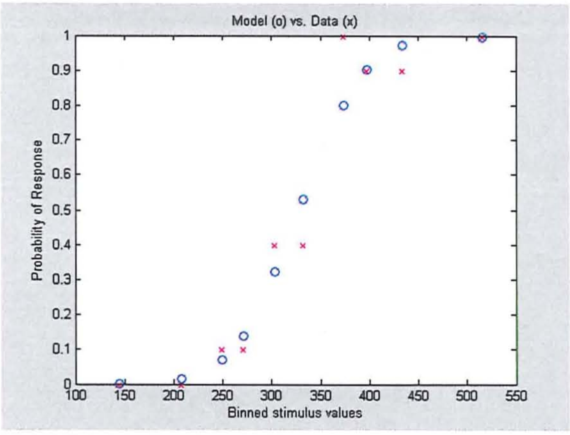
[12] Simulations were run to capture performance of those participants who had experienced the asymmetric task first; in fact the model could not capture the aberrant behaviour displayed by those participants who experienced the symmetric task first. Moreover, the model was insensitive to the order in which the two conditions were tested

$$\mathbf{b_3 = b_4 = 0.05*scaler1} \qquad \textbf{Eq. 8. 5}$$

where the scaler1 parameter scales the learning rate of the SR association. In these simulations reported above **scaler1** was set equal to 2, so that the $\mathbf{b_3}$ and $\mathbf{b_4}$ parameter were equal to 0.1. Decreasing the scaler1 value from 2 to 1 and, therefore, $\mathbf{b_3}$ and $\mathbf{b_4}$ to 0.05 significantly reduced the response bias to 326.45 (9.90) and the noise level was also reduced (mean = 49.09, s.d. = 9.2); in contrast increasing $\mathbf{b_3}$ and $\mathbf{b_4}$ to 0.2 produced a response bias equal to 346.20 (10.96) and a noise level equal to 110.85(12.60).

Similarly, it is possible to modify the degree of response bias by varying the value of the tonic reward prediction input (i.e. **rptonic**). Indeed, as shown above an **rptonic** value equal to 0.5 and scaler equal to 2 were found to produce a response bias on average equal to 341.67 (s.d. = 8.90) pixels and a mean noise level equal to 69.82 (s.d. = 8.84). However, when leaving all parameters constant but increasing the reward prediction input to 0.6, both the mean response bias and noise level are increased [mean criterion = 355.63 (s.d. = 9.34); mean noise level = 105.06 (s.d. = 11.29)].

As implied by the dramatically reduced noise parameters in the simulations with the revised neural model, this was now also able to simulate the step-like learning function observed in the empirical data (figure 8.8).



Asymmetric

Symmetric

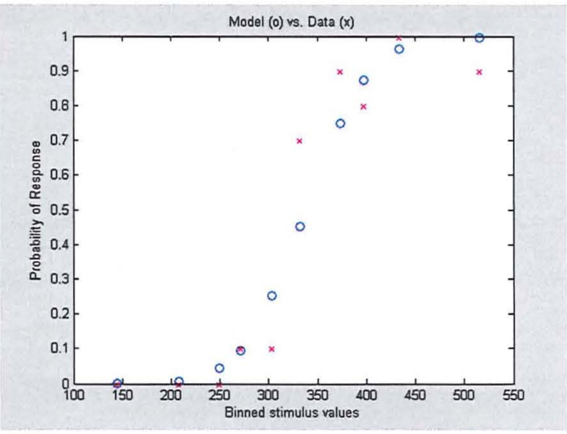


Figure 8.8. Simulated 'behavioural' data in the two task conditions. The red stars represent the actual data whereas the blue circles the model fitting data.

An alternative set of parameter variation was also tested (set B), which had a **scaler** value of 1.5, the **rf** values were 0.25, 0.5, 0.75 and 1 (for 100, 200, 300 and 400 points, respectively, since **m** was set equal to 1) while **rulebias** was set equal to 0.35. These values were chosen due to their ability to capture human behaviour. Simulations obtained using the alternative set of variables (set B) showed very similar results to the ones obtained by using set A[13]. This is not surprising since the two variable sets are not entirely independent (cf. inter-related) since both reward magnitude and the scaler parameter indirectly mediate learning by affecting the size of the RPE signal on each trial. Hence, simulations using the different parameter sets offered very similar simulation results.

---

[13] The simulated asymmetric criterion was equal to 342.69 (s.d. = 10.0) and the noise level was 58.75 (s.d. = 11.7), whereas the symmetric criterion was equal to 322.21 (s.d. = 7.8) and the noise level was 48.94 (s.d. = 9.5)

Explicit module parameters

Figure 8.9 visually represents the distribution of the decision bound for the asymmetric condition, following 300 simulations using parameter set A. It was noticed that there was a great variation across the decision bound set by the simulated 'participants' and this variation must be caused by individual variation across the simulations. This variation could, therefore, be due to the fact that the initial input weights into the actor unit are individually randomized, as this was one of the only features that was different between simulated 'participants'.

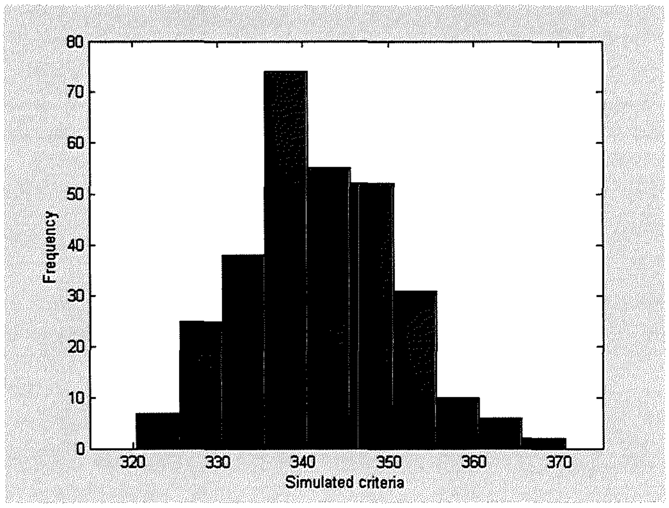


Figure 8.9. Histogram representing the tabulated frequencies of the criteria simulated with pnoise = 50

The parameter that set perceptual noise also added inter-individual differences on the decision bounds. Indeed, each simulated participant was due to experience a stimulus centred at 310 pixels on trial **n** of the task, but the perceived position was of the stimulus was cantered at 310 ± perceptual noise. Therefore, the same stimulus could be classified differently by different 'participants' depending (randomly) on the amount of noise experienced. Thus, it was decided to reduce the value of the perceptual noise (**pnoise** = 50). Indeed, the greater the noise, the greater the difference in allocation of category membership across 'participants'. By trial-and error, it was found that one could decrease the criterion variation by setting **pnoise** equal to 12.5.

However, simulation results obtained with **pnoise** equal to 12.5 were not very efficient at reproducing the human data. In fact, the simulated noise level was much lower (mean = 52.64 and s.d. = 2.3) than the one set by human participants (mean = 98.28 and s.d. = 47.1). Subsequently, the step-like learning function was steeper than the one observed in humans (figure 8.10).

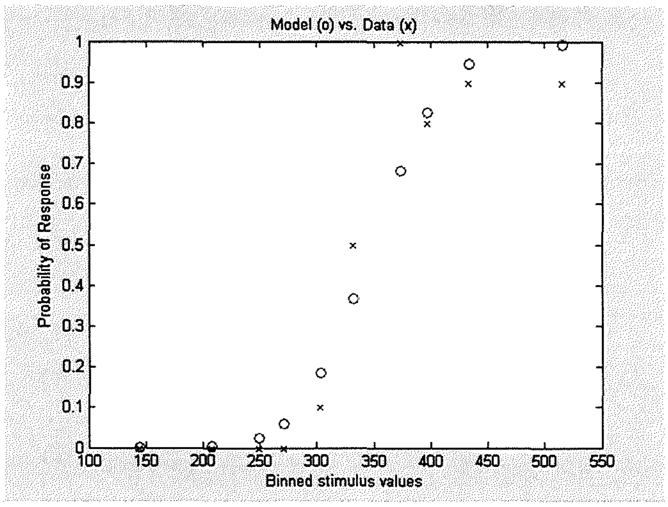


Figure 8.10. Simulated learning function with pnoise = 12.5

Following these considerations, it was decided to run the simulations using the original **pnoise** value (50) since it seems to capture the human data better than **pnoise** equal to 12.5.

The **rulebias** parameter determines how strongly participants apply the appropriate explicit rule, where the higher the value of **rulebias**, the more the performance is influenced by the rule module which sets the decision boundary symmetrically at 325 pixels in both conditions. Thus, if the **rulebias** value is decreased from 0.6 to 0.3, the response bias increases (357.99 and s.d. = 11.1) and so does the noise level (114.6, s.d. = 14.5). This reflects the relatively greater influence of the implicit RPE learning system on performance. Additionally, using a **rulebias** value of 0.3 also simulated much flatter learning curves (figure 8.11), explaining the increase in the noise parameters.

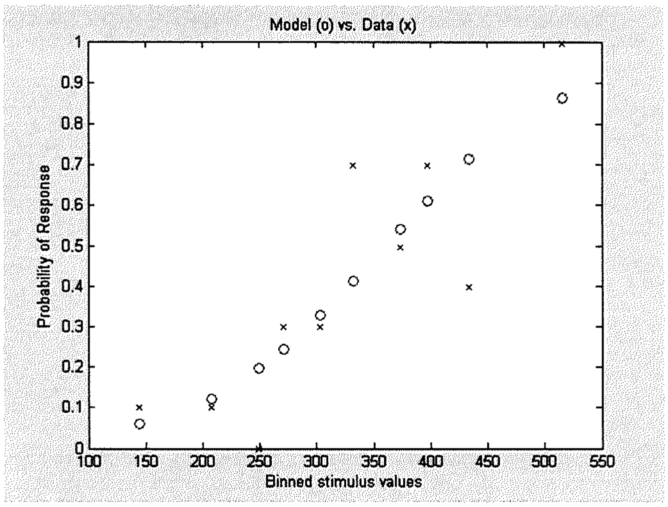


Figure 8.11. Simulated learning function with rulebias = 0.3

Thus, when the **rulebias** values is decrease, the impact of the explicit module is small and, simultaneously, the implicit module takes over and produces a greater response bias (as well as increased noise level).

In conclusion, the present results indicate that the revised version of the neural model, which combines an explicit and an implicit system, is capable of capturing both the response bias and the learning function displayed by participants in the two task conditions. Moreover, as shown above, the model is able to simulate variation in response bias as a function of the learning rate (mediated by **scaler**). After the parameter values that offered the best simulations had been identified, individual differences (BAS-variation) were then included in the model to explore whether it could also simulate the behaviour-personality relationship observed in the empirical data. It was decided to report the simulation results obtained using the parameter set A[14].

---

[14] Simulations run using parameter set B produced similar results and, therefore, led to the same conclusions

8.4.3.6 Simulations with individual differences

Since, the BAS has been identified as a motivational-emotional system that relies on the DA system (Gray, 1987; Pickering, 2004; Corr, 2006), the present model assumes that the magnitude of RPE signals may be modulated by inter-individual variation, which resembles BAS variation. Thus, the added variation is referred to as simulated 'BAS' variation since the BAS is a plausible candidate for individual differences in reward-related learning.

There are 4 sub-models that were tested, which were: BAS variation on the effectiveness of the reward signal projecting to the DA cell (model 1); BAS on both striatal cells and, subsequently (model 2); the separate impact of BAS variation on each striatal cell [i.e. actor (model 3) and critic sub-units (model 4)]. Adding BAS variation on the reward cell captures individual differences in sensitivity to actual rewards. This sub-model resembles the classical view in which RST describes the BAS system as sensitive to reward and reward-related stimuli (Corr, 2006). In particular, individuals with a highly responsive BAS are expected to perceive the reinforcement signal more strongly and, therefore, the coded reinforcement value (**rf**) has a larger impact on the DA cell (cf. low BAS individuals).

The other neural sub-models assume that 'BAS' variation acts directly on the effectiveness of the RPE signal (i.e. **RPE** = **RPE*BAS**) and, by doing this, it is able to add variance to the strengthening or weakening of active synapses under a positive or negative RPE. The sub-model with the BAS acting on the striatal cell denoted as 'a' in figure 7.1 (i.e. model 4; see chapter 7) captures the impact of individual differences on the RPE signal acting on the synapses of the reward prediction (RP) critic cell. In contrast, the sub-model with the BAS acting on the striatal cell denoted as 'b' (model 3; see figure 7.1) captures the impact of individual differences on the RPE signal acting on the actor synapses of the SR cell.

In each sub-model, the 'BAS variance' is added simply by multiplying the relevant parameter (i.e. **rf** or **RPE**) by a 'BAS value' for each simulated participant. Initial simulations are run with a random uniformly distributed BAS variation which ranged between 0 and 2.

The two striatal cells are innervated by RPE-DA signals (Brown et al., 1999; Joel et al., 2002; Seger, 2008). Thus, the weights of these corticostriatal synapses are mediated by DAergic firing according to the 3-factor learning rule (Ashby et al., 2007; Seger, 2008). Hence, the strengthening of these synapses (i.e. learning) is highly dependent on DAergic activity and, more specifically, on D1 and D2 receptors sensitivity (Calabresi et al., 2007; Seger, 2008; Schotanus and Chergui, 2008). Schotanus and Chergui (2008) have observed that D1 receptors seem to mediate learning in the areas where the critic is located; whereas both D1 and D2 receptors mediate learning in areas when the actor is located. BAS variation on the SR (actor) cell captures the inter-individual variation in sensitivity to the RPE signal with high BAS scores associated to greater responsiveness to RPE and, therefore, enhanced SR associative learning and unlearning. Finally, BAS variation on the RP (critic) cell captures individual differences in sensitivity to RPE and, thus, learning and unlearning the reward prediction for each stimulus. The parameter values used in the simulations reported below are summarised in table 8.14.

Table 8.14. Parameters' values implemented in the simulations

Symbol	**Description**	**Value**
rptonic	**Reward prediction tonic value**	0.5
rulebias	**Strength of explicit module**	0.6
m	**Scales the rf values**	0.8
Pnoise	**Perceptual and criterial noise**	50
b_1 and b_2	**Learning weights of the RP unit**	0.05*scaler1
b_3 and b_4	**Learning weights of the SR unit**	0.05*scaler2
Scaler1	**It scales the learning rate of RP unit**	2
Scaler2	**It scales the learning rate of SR unit**	2

Model 1: BAS on the reward cell (inter-individual differences in sensitivity to rewards)

Simulations using this sub-model were run first, since this sub-model represents the action of the BAS according to the classical assumptions made by RST (Gray, 1987; Corr, 2006). Simulations with a model with the BAS acting as a multiplier on the reward inputs to the DA cell showed a negative correlation between the BAS value and the fitted criterion ($r = -0.50$, $p < .001$) in the asymmetric task. The relationship between the BAS and the fitted criterion is visually represented in figure 8.12 and appears to be curvilinear. Indeed, the relationship between the BAS and performance is negative for those 'individuals' with BAS scores between 0.4 and 1.2, whereas it is flat above 1.2 and positive for those with BAS scores below 0.3.

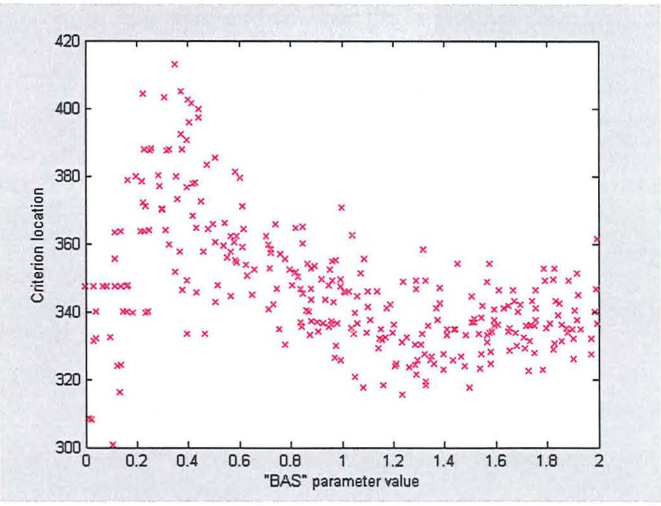


Figure 8.12. The scatterplot represents the relationship between the BAS variation and the fitted criterion in the asymmetric task for 300 simulated subjects with random uniform BAS parameter variation

If the relationship between 'reward sensitivity' and the behaviour in the asymmetric condition is strongly curvilinear as represented in figure 8.13, then it is likely that this relationship would be difficult to detect in real data using linear techniques (e.g. correlation and regression analyses). This difficulty would become more acute when one attempts (in real data) to observe the relationship between a personality trait and the task behaviour, assuming that the two would be related by virtue of their shared relationship with the underlying biology. As the personality trait relationships with underlying biological parameters are relatively weak (Blum et al., 2000; Reuter et al., 2006; Lee et al., 2007; see chapter 7), and the task and biology may relate non-linearly, it is possible that the trait*task relationship may be undetectable in an experimental study with a typical sample size.

During the symmetric condition, the weak correlation between the BAS variation and the fitted criterion was of borderline significance ($r = -0.11$, $p = 0.051$). Figure 8.13 suggests that this correlation is mainly an effect of the fact that 'individuals' with BAS scores lower than 0.5 on average showed a slight positive bias (i.e. criterion > 325) that is not observed amongst those who scored higher than 0.5.

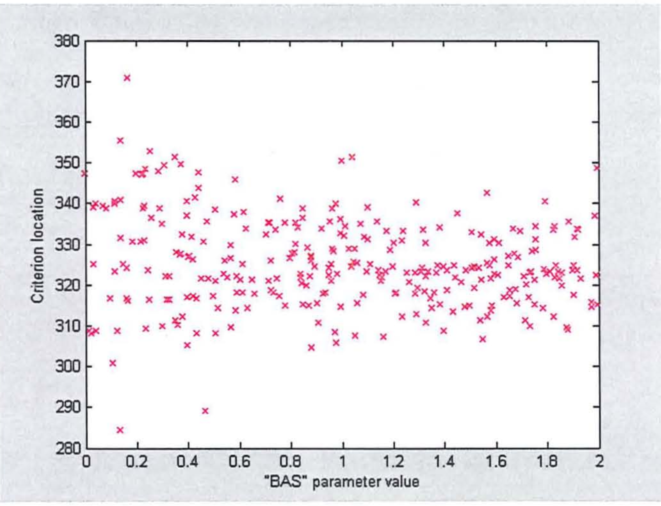


Figure 8.13. The scatterplot represents the relationship between the BAS variation and the fitted criterion in the symmetric task for 300 simulated subjects with random uniform BAS parameter variation

Model 2: BAS variation on both of the striatal cells

The next simulation conducted with the revised model was the model with the BAS acting as a multiplier on both striatal cells. Results showed that, for the asymmetric condition, the rule model could capture a positive relationship between the BAS variation and the fitted criterion ($r = 0.59$, $p <.001$). This relationship is presented in the scatterplot below (figure 8.14).

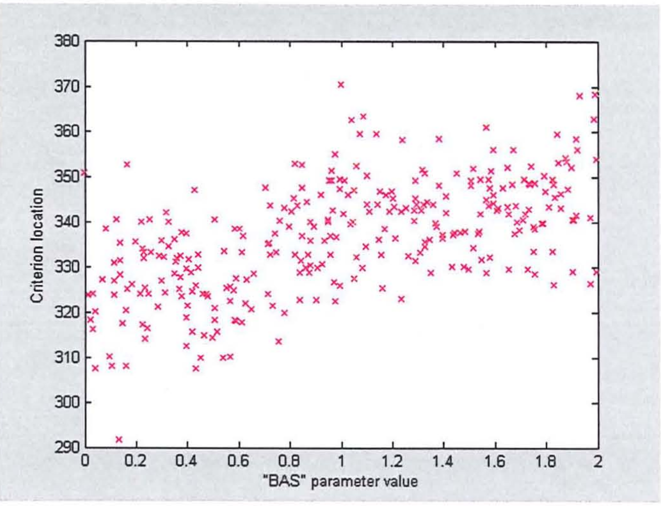


Figure 8.14. BAS*behaviour relationship in the asymmetric task for 300 simulated subjects with random uniform BAS parameter variation

The relationship between the BAS and the criterion was non-significant in the simulated symmetric condition (r = -.07, ns; figure 8.15).

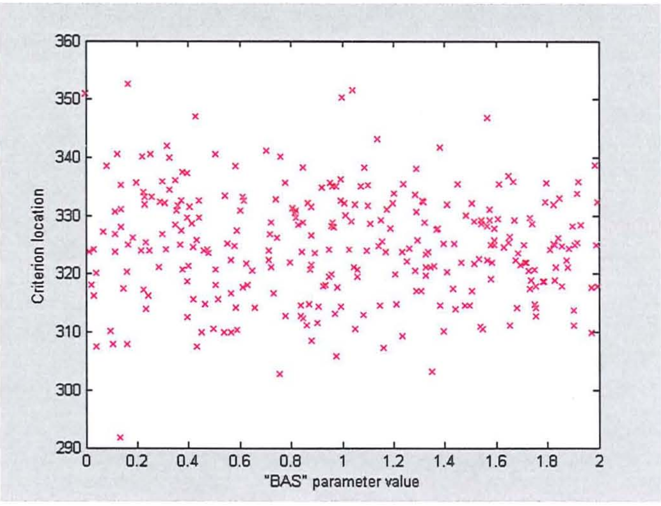


Figure 8.15. BAS*behaviour relationship in the symmetric task for 300 simulated subjects with random uniform BAS parameter variation

Model 4: BAS on the critic reward prediction cell (RP; striatal cell 'a'; see figure 7.1 in chapter 7).

The results showed that this model was also able to capture the step-like learning function but the correlation between the BAS and the fitted criterion was significant but weak (r = -0.14, p = .012). The relationship between the

BAS and the criterion was non- significant in the symmetrical condition (r = 0.08, ns).

Model 3: BAS on the actor stimulus-response cell (SR cell; striatal cell 'b'; see figure 7.1 in chapter 7)

The simulated data showed a positive correlation between the BAS and the asymmetric criterion (r = .76, p <.001), as visually represented in figure 8.16.

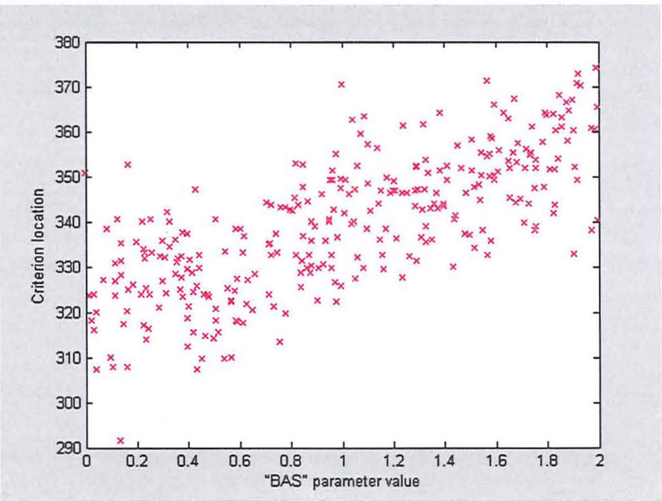


Figure 8.16. The scatterplot represents the relationship between the BAS variation and the fitted criterion in the asymmetric task for 300 simulated subjects with random uniform BAS parameter variation

As expected, the correlation between the BAS and the fitted criterion was non-significant (r = -0.09, ns) in the symmetric condition (figure 8.17).

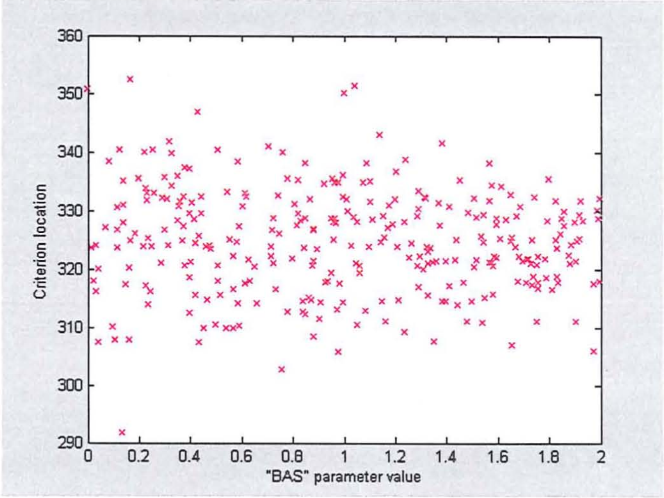


Figure 8.17. The scatterplot represents the relationship between the BAS variation and the fitted criterion in the symmetric task for 300 simulated subjects with random uniform BAS parameter variation

It is possible to conclude that the correlation observed from the model with BAS acting as a multiplier on both of the striatal cells was driven from the strong correlation observed in the model where BAS acts on the SR cell. Moreover, the results indicate that as a strong correlation exists in the model, it should be possible, using the same paradigm, to detect a similar relationship between performance on the task and the BAS-trait in an empirical study. However, as the correlation between personality trait scores and the biological equivalent of the BAS parameter is likely to be quite modest, the correlation between trait scores and behaviour on the task is likely to be even more modest. For example, imagine that the correlation between personality trait scores and the biological equivalent of the BAS is equal to 0.3. In addition, the model shows that the correlation between the simulated behaviour and the BAS parameter may be of the order of 0.75. Thus, the upper limit of the correlation between the personality trait scores and the actual behaviour is equal to 0.23 (= 0.3*0.75). This is a weak correlation, but it could be captured with 80% power in a sample size of 113 in an empirical study. Such sample sizes are plausible in personality research. However, due to various exclusions in the present experiment, the current study is underpowered and it was unlikely to detect such a small effect. These considerations indicate the importance of the model in determining the upper limit of the effect size (e.g. Cohen's d or Pearson's r) and, therefore, determining the sample size necessary to detect the trait*behaviour relationship in an experimental study. Nonetheless, the present results showed that the correlation between positive schizotypy and behaviour had a correlation coefficient greater than 0.5 and it was significant, in spite of the small sample (N = 18).

Overall, these simulations further support previous observations that indicate that the present CL task may be a useful tool to explore the relationship between the BAS and behaviour owing to the results that show a strong simulated relationship between the biology and behaviour. Moreover, it appears that the best-simulating models are the sub-model where the BAS acts as a multiplier on the reward cell (model 1) and the sub-model where the BAS acts as a multiplier on the SR cell (model 3).

8.5 Discussion

The results obtained in the study replicate the findings, obtained by Markman and colleagues (2005) that showed that reward manipulations affect performance on a nominally RB task. Indeed, participants in the present study showed a tendency to place their decision criterion away from the optimal accuracy criterion and closer to the optimal reward criterion under asymmetric payoffs. In other words, they identified probabilistic stimuli more frequently as members of the high payoff category than as members of the low feedback category. The response bias shown by participants in the present study was much more conservative than the bias in the original study. Obviously, this may result from sample differences across the two studies. Participants in the present study seemed to have been less willing to sacrifice accuracy as much as those in the original study. The current results are in line with evidence that shows that humans show smaller response biases in favour of winnings than the biases produced by the optimal classifier (Maddox & Bohil, 2005; Maddox, Bohil and Dodd, 2003).

Under the symmetric payoff matrix condition, by contrast, participants placed their criterion close to the optimal accuracy criterion (i.e. criterion shift across payoff conditions). However, this shift in bias across conditions was observed only for those participants who experienced the asymmetric condition prior to the symmetric one. Those participants who experienced the symmetric task prior to the asymmetric task showed an aberrant pattern of responding (i.e. positive bias in the symmetric task). Since these latter results were not simulated by the neural model and counterintuitive to the initial predictions, data from this CB order was excluded. Results further indicated that the shift in bias, observed for those participants who experienced the asymmetric task first, was mediated by individual differences. In particular, individuals who scored high on positive schizotypy were found to show a greater shift in bias than their low scoring counterparts. However, the fact that the order of task presentation had such a profound effect upon task behaviour limits the conclusiveness of these findings.

The behavioural findings are somewhat counterintuitive to the COVIS model which postulates that the reinforcing aspects of reward manipulations should not have affected performance on RB tasks of the type used in this study. Nonetheless, the present results showed that an asymmetric payoff matrix led to a response bias in favour of the high payoff category. Queller and colleagues (1999) found that reward frequency manipulations did not impair learning on a rule-based CL task. Similarly, several studies have indicated that learning on an RB task could occur in the absence of trial-by-trial feedback (Waldron & Ashby, 2001; Ashby et al., 2003). These results may be explained by the fact that due to their simple structure, participants can learn RB tasks regardless of feedback manipulations. Nonetheless, it is possible that more subtle reward manipulations (e.g. variations in reward magnitude) may affect performance on an RB task by activating the implicit system. Additionally, in the present study participants were openly asked to respond quickly. Ashby and colleagues (1998) have suggested that such an instruction may lead to the activation of the implicit system.

There are various ways to resolve this tension with COVIS. The first would be to argue that reward feedback from the asymmetric payoff matrix acts as a source of information as well as a source of reinforcement. The explicit system could, in principle, manipulate feedback information consciously and explicitly, in working memory, to adjust the decision criterion. However, this does not seem very plausible and it does not fit with the verbal reports from participants who did not mention any deliberate asymmetry in their criterion placement. A second approach, followed by Maddox and colleagues (2006) is to propose an indirect effect of the reward feedback, via regulatory fit, on cognitive flexibility in the explicit system (discussed earlier). Finally, the approach, encapsulated in the modelling, is to argue that the criterion placement is an implicit learning process driven by the reinforcing properties of reward and this is superimposed upon the workings of the explicit system. The latter is perceptually-guided and it solves the task in the way described by COVIS.

Overall, the modelling results indicate that the explicit and implicit learning systems mediate different aspects of performance on the CL task. The explicit system is involved in solving the categorisation problem whereas the implicit system mediates the response bias. Evidence to support the co-activity of the two systems during performance on the task also comes from participants' verbal reports. Indeed, several participants were able to verbally describe the dimensional strategy that they had implemented to categorise the stimuli. In spite of showing a response bias in the asymmetric task (which was detected by the formal model in the data), they reported not using feedback information in their decision-making process. Thus, these verbal reports also seem to indicate that feedback processing was implicit. In the present model, the bias is produced by the 'statistical' properties of the RPE-based learning rules.

The reason why the two systems are active at the same time may be due to the fact that the structure of the task never completely favoured activation of the explicit system over the implicit system. Indeed, it is possible that the explicit system becomes active during the first trials in order to solve the categorisation problem whereas the implicit system may become active at a later stage (i.e. once a perceptual criterion has been adopted). In this second stage of learning, the criterion is perhaps fine-tuned by the implicit learning process and so can be moved away from the optimal accuracy criterion and closer to the optimal reward boundary to maximise winnings. Indeed, this response bias is mediated by the implicit system according to simulations run using the neural model.

Scores on the WM task were found not to be associated with the decision bound shown by participants during the task. This is not surprising since the working memory ability is a process typically involved in the functioning of the explicit system while decision bound position was dependent on the functioning of the implicit system. Since both the implicit system and the BAS rely on the same midbrain DA projections, it was expected that BAS activation would mediate decision location.

Measures of positive schizotypy were included in the study due to the overlap with the impulsivity construct. Moreover, positive schizotypy is likely to have a partly dopaminergic basis (Kumari, et al., 1999; Weiner et al., 2003; Pickering, 2004; see chapter 4 for more details). Hence, association between PS scores and criterion location consistent with the idea that PS may represent a DAergic trait. Positive schizotypy scales capture the positive symptoms typical of schizophrenia (Pickering, 2004). Schizophrenia is characterised by a dysregulation of the DA system and, in particular, by higher tonic DA levels (Julien, 2003; Juckel et al., 2006a, 2006b). High tonic DA levels have been found to hinder reward sensitivity by blunting the phasic DA response to reward (prediction error; Juckel, et al., 2006b; Knutson et al., 2004). In a study by Juckel and colleagues (2006a) reduced phasic DA activity was found to be positively associated both with negative and, trend-wise, with positive symptoms of schizophrenia. In a study by Murray and colleagues (2007), patients with positive psychotic symptoms showed reduced RPE signals in the ventral striatum and a non-significant trend to respond to high frequency reward stimuli less frequently than healthy controls. According to these results, positive schizotypy could modulate reward sensitivity due to its relationship with DA activity (i.e. lower phasic DA firing; see chapter 2 for more details). Positive symptoms are associated to DA dysfunction in the striatum, which also processes reward (Carlsson, 2002; Murray et al., 2007). Hence, scores on the positive schizotypy component might be expected to relate to performance on the present task, which is mediated by the striatum (Heinz, 2002; Murray et al., 2007). Future studies investigating the BAS and its underlying trait should include measures of extraversion, impulsivity and positive schizotypy in order to assess which trait mediates performance on the task according to RST predictions.

However, hypodopaminergic-induced reward deficiency in schizophrenia is generally identified as one of the negative symptoms of schizophrenia (i.e. emotional and motivational deficits, such as 'apathy' and 'avolition'; Heinz, 2002 p.14; Julien, 2003; Juckel, et al., 2006b), which are not indexed by the positive schizotypy construct (cf. the introvertive anhedonia scale of the OLIFE questionnaire; Mason et al., 1995). Introvertive anhedonia is a measure of negative schizotypy and was not included in the present study. Negative symptoms are caused by a DAergic dysfunction in the prefrontal cortex (Carlsson, 2002; Heinz, 2002; Juckel et al., 2006a). Thus, negative schizotypy may be related to the BAS and, in particular, anhedonia may correspond to an underactive BAS. It would be more complex to simulate prefrontal hypodopaminergic activity and its effects using the current modelling. It is possible that reducing the **rptonic** signal may (indirectly) capture that effect.

As already indicated, the main limitation of the present study is that the sample size is quite small, especially once those participants who were not using the appropriate strategy or showed an aberrant response pattern (i.e. participants who experienced the symmetric task first) were excluded from the analyses. Nonetheless, the model simulations indicated that the CL task could be a useful tool to explore the impact of reward manipulation during learning and the relationship between behaviour and the trait. Indeed, the neural model identified a strong biology*behaviour relationship that suggests it should be possible to detect the weaker personality-behaviour relationship in the empirical data even with a moderate-size small sample.

Furthermore, the neural simulations indicated that the best sub-model candidates were the model with the BAS on the SR cell (i.e. actor unit; especially model 3) and the model with the BAS on the DA cell (model 1). 'Best' here indicates the model revealing individual differences that are more likely to be detectable in real data. Obviously, the real relationship between personality and the biological parameter might lie in the critic, within the ventral striatum. If this were true, the model would strongly suggest that the current task would not show measurable relationship with personality. Thus, the modelling and behaviour results combined might allow us to rule out some locations as places where the biology*personality relationship is located.

The results obtained in the simulations offer a deeper understanding of the behavioural data obtained in the study. The models make differing assumptions about the way reward affects performance during the task and, in particular, how the BAS responds to reward manipulations and mediates reward-related learning. Indeed, simulations obtained using model 1 showed that there was a complex curvilinear relationship between the BAS parameter (reaction to the excitatory reward signal) and the criterion set in the asymmetric condition. Over the greatest linear portion of this curve, the relationship was negative suggesting that low reward sensitivity might be associated with greater approach behaviour (indexed by the response bias). The human data, instead, captured a positive correlation between positive schizotypy and the criterion location in the asymmetric condition. Therefore, these observations could suggest that high scores on PS are associated with low reward sensitivity. The direction of these relationships are, therefore, consistent with evidence that suggests that high scores on positive schizotypy, as an index of the positive symptoms of schizophrenia, are associated with lower phasic DA activity (Mason et al., 1995; Juckel et al., 2006a,b). These results are also broadly in line with the reward deficiency syndrome (RDS) model which postulates that greater approach behaviour towards reward stimuli is determined by DAergic hypofunction (Blum et al., 2000). This argument follows because, all other things being equal, low sensitivity to excitatory reward input to DA cells would reduce DA activity.

In contrast, model 3 simulations captured a linear positive correlation between the BAS variation (on RPE signals) and the decision bound. BAS variation in this model determined the effectiveness of the reward prediction error (RPE) signals for SR learning. Thus, high scores on positive schizotypy (as a BAS-trait) would need to be associated with greater RPE signals and, subsequently, more efficient SR learning during the task, in order to produce the direction of the association with the response bias observed in the asymmetric condition.

According to this model, 'BAS variation' mediates the individual's sensitivity to RPE signals rather than sensitivity to reward *per se*. This would represent an alternative process to account for BAS-driven individual differences in reward-related learning and, especially, approach behaviour. The simulations offer plausible explanations to account for the empirical data although these results and these considerations are not conclusive and further testing and simulations are required to obtain a deeper insight into the data.

Formal modelling was a very useful exploratory tool to analyse the data since it allows one to identify those participants who are using the appropriate uni-dimensional strategy and those who are either guessing or using an alternative strategy. This is very important since, as pointed out by Gluck and colleagues (2002), task structure is not normally a good predictor of the strategy actually implemented by participants. Additionally, as mentioned in previous chapters, inclusion in the analysis of data from those participants not performing according to the task is a source of noise since the data do not capture the phenomenon tested in the study. Hence, in future studies formal modelling should be implemented where possible to ensure that the desired phenomenon has been captured and it is analysed.

In conclusion, the present results replicate the original findings obtained by Markman and colleagues (2005) as they show that participants place their decision criterion away from the optimal accuracy criterion towards the optimal reward criterion under asymmetric payoff compared to symmetric payoff. Moreover, the study identified the personality trait that mediates the response bias as positive schizotypy. Owing to the great overlap between PS and the impulsivity construct and the simulation results obtained with model 3, it is possible that the PS component in the present study may represent a surrogate measure of impulsivity. Due to the low statistical power, in a sample of 18 participants, it was not possible to test this hypothesis using regression techniques. Thus, PS may have appeared to be the BAS-trait purely due to its strong association with impulsivity. As discussed earlier these results are far from conclusive and replication is required to draw any conclusion. Nonetheless, the simulations add validity to the present results and the need for replication. The main aim of the next study is to replicate the present results with a bigger sample size and trying to remove any order effect. The next study should also aim to identify the BAS-related trait and test whether positive schizotypy is an alternative BAS-trait candidate or just a proxy measure of impulsivity.

Chapter 9

Motivation and personality during the learning and reversal phase of category learning task

9.1 Abstract

The present study aimed to replicate and extend the findings obtained in the previous chapter after addressing issues in the task design. Hence, the primary goal of the study was to explore the impact of asymmetric payoffs on decision bound and the mediating effect of individual differences on performance (i.e. on the response bias). Additionally, the task included a reversal learning phase where an inter-dimensional rule-switch was implemented in the symmetric condition whereas a payoff switch was implemented in the asymmetric condition. In particular, this aspect of the study explored the impact of the BAS-related personality dimension(s) on reversal learning since reward sensitivity has been identified as one of the factors that affects response moderation in passive avoidance tasks (Patterson and Newman, 1993; Avila, 2001). Due to the complexity of the task and the relatively small number of trials, only a few participants learnt to perform appropriately in the reversal phase. However, in the learning phase, participants were found to over-classify the probabilistic stimuli as members of the high payoff category and this bias was mediated by scores on the impulsivity component. In particular, low impulsivity scores were associated with greater response bias in the asymmetric task. Results were discussed in relation to RST and following the simulation results obtained with the neural model described in chapter 7.

9.2 Introduction

The present study aims to replicate and extend the findings reported in chapter 8. The study showed that during performance on a CL task with an asymmetric payoff matrix, participants showed a tendency to place their criterion bound away from the optimal accuracy criterion and closer to the optimal reward criterion (i.e. show a response bias). The study also observed that, when the asymmetric payoff matrix was experienced first, such a response bias was not present during performance under a symmetrical payoff matrix. These results are in line with the main assumptions of signal detection theory that states that everyday categorisation is influenced by the costs and benefits of correct and incorrect responses (payoffs; Maddox, Bohil & Dodd, 2003; Maddox and Dodd, 2003). Additionally, individual differences were found to mediate the shift in bias across the two task conditions (asymmetric vs. symmetric matrix). In particular, participants who scored high on the positive schizotypy (PS) component were found to show a greater shift in bias than their low counterparts under asymmetric payoffs. In the previous chapter, it was suggested that the modulating effect of PS on performance may be due to the fact that PS is a proxy measure for impulsivity which, together with extraversion, is a main BAS-trait candidate. However, the previous results are not conclusive due to the small sample size and the fact that these results were observed only in one counterbalancing group and, therefore, may be caused by order effects.

Thus, the main aim of the present study is to replicate the previous behavioural findings (i.e. shift in bias) and further explore the relationship between personality and behaviour after trying to eliminate any possible complicating order effects. Additionally, the study aimed to explore the impact of personality under changing task contingencies. For this purpose, the task is composed of two phases: a learning phase that resembles the original study and a reversal learning phase. The learning phase requires participants to learn to solve the categorisation problem whereas the reversal-learning phase requires participants to interrupt a dominant response to develop a new, optimal strategy. Personality traits that underlie the BAS are expected to mediate performance on the learning phase (as rationalised and observed in chapter 8) and perhaps in the reversal phase, too. Indeed, rule-switching has been found to be mediated by DArgic firing in the nucleus accumbens (NAc; Joel et al., 2002; Cools et al., 2006; 2007), which RST considers to be a major substrate of the BAS biological system (Pickering and Gray, 2001; Corr, 2006).

Moreover, it has been suggested that perseveration is caused by hypersensitivity to reward (Patterson et al., 1987; Avila, 2001). Therefore, personality traits that are related to reward-sensitivity (i.e. BAS-related traits) might be expected to be associated with a reduced ability to inhibit a response previously accompanied by reward (Patterson and Newman, 1993; Avila, 2001). In particular, extraverts have been found to show a deficit in their ability to inhibit a dominant response (i.e. show response perseveration; Pearce-McCall and Newman, 1986; Patterson and Newman, 1993). In a study by Patterson and colleagues (1987), participants performed on a go/no-go discrimination task. The task required participants to learn to respond when presented with one of the six go stimuli (S_+) and to withhold their response when presented with one of the six no-go stimuli (S_-). Extraverts were found to respond to the S_- stimuli more frequently than introverts (i.e. passive avoidance errors). These results indicate that extraverts are less likely than introverts to inhibit a response (a go button press) that is associated to reward, when responding to the S_- stimuli. There is also evidence that indicates that neuroticism modulates the effect of extraversion in response perseveration (Nichols and Newman, 1986).

In line with the assumption that response perseveration is caused by reward sensitivity and enhanced by approach motivation, high BAS individuals were found to show a reduced ability at inhibiting a dominant response once the contingencies of the task changed (Avila, 2001). In a series of tasks, participants had to perform on a continuous odd-even discrimination task, where they had to learn to withhold responding when an aversive cue (i.e. a red circle) appeared on the screen together with the target stimulus. As predicted, results showed that high BAS individuals who scored high on the Sensitivity to Reward (SR) scale were less likely to inhibit their responding when the red circle appeared on the screen. However, these results are not specific to reward. Indeed, the lack of a control condition does not allow us to assess whether disinhinbited responding would occur when no (or small) rewards are administered. Thus, it is not possible to conclude whether the disinhibited responding is due to reward sensitivity or purely to impulsivity. Moreover, there is evidence that indicates that impulsivity is associated with reduced error processing as indexed by event-related potentials; in particular high impulsive individuals have been found to show low amplitude error-related negativity (ERN; Ruchsow et al., 2005; Franken et al., 2007). ERN represents a negative RPE deflection usually observed following erroneous responding. Moreover, ERN has been found to reflect a negative RPE and to mediate response switching (Asako et al., 2004; Atsushi and Asako, 2004).

Pickering (2004) conducted a study where participants were presented with visual stimuli that varied on two dimensions, i.e. the height of a rectangle and the position of an internal line segment. The CL task had two phases, which were a learning and rule-switch phase. In the learning phase, the position of the internal line segment was the relevant dimension while the height of the rectangle was the irrelevant one. In the rule-switch phase, an extra dimensional switch occurred so that the height of the rectangle became the relevant dimension. Results showed that EPQ-P scores (i.e. ImpAss measure) were significantly and positively correlated with performance during the leaning phase but not the rule-switch phase of the task, while scores on the Unusual Experience scale were significantly, negatively correlated with the rule-switch phase of the task but not the learning phase. However, these differences between the phases were not significant. Thus, Pickering's results tentatively suggest that ImpAss enhances performance on a simple RB task but not during rule-switch. Similarly, Tharp (2007) found that ImpAss scores were negatively associated with performance on a more complex, conjunctive RB task and also during reversal learning of a simple, uni-dimensional RB task. Indeed, he found that extraversion was positively associated with performance during the learning phase. In contrast, during the rule-switch phase impulsivity was found to be negatively associated with performance while scores on the schizotypal personality trait (Unusual Experiences; Mason et al., 1995) were positively associated with performance. Tharp suggested that impaired performance observed across the high impulsive individuals during both initial learning and reversal learning was due to low cognitive flexibility.

The neural model, implemented in the previous study, was found to be an effective tool to simulate the human data and offers an explanation of the personality-behaviour relationships observed in the data (see chapter 8). The neural model suggests that performance on the CL task employed by Markman and colleagues (2005) was mediated by both the explicit and the implicit systems. The explicit system seems to be involved in learning the gross perceptual basis of category-membership whereas the implicit system seems to process the feedback signals and, subsequently, mediate the process involved in setting the decision bound to fine-tune category decisions in order to increase winnings. However, these conclusions seem to be at odds with the COVIS model that suggests that performance on an RB task, such as this one, should be guided by the explicit system alone and not affected by feedback manipulations. Therefore, it was decided to include several practice trials prior to the experimental task in order to get participants acquainted to the task and develop their knowledge of the appropriate uni-dimensional rules for use in the experimental task. During the practice trials, which otherwise followed the same procedure used in the experimental task, participants received accuracy feedback following each response. Accuracy feedback was used in order to aid learning and activate the explicit system. Thus, the task aims to assess whether under an asymmetric payoff matrix, participants show a response bias during the experimental task as a result of the activation of the implicit system that takes control of performance as learning occurs over trials.

9.2.1 Study aims

Following from the previous findings and the simulation results, it was expected that in the learning phase participants would show a shift in bias across the two task conditions and that the shift would be mediated by individual differences. Hence, the task should allow us to identify the underlying personality trait(s) mediating the response bias. According to RST, the personality trait(s) should reflect BAS functioning (i.e. impulsivity or extraversion) or, if the previous findings are reliable, positive schizotypy.

The second aim of the study was to observe how individual differences mediated performance during the reversal learning phase, which required response inhibition of a previously reinforced response in favour of a new strategy. This new strategy should maximise performance once the task's contingencies have changed. According to the existing literature, failure to exhibit response modulation (i.e. to show perseveration) may be due to reward over-sensitivity and/or low cognitive flexibility (Patterson and Newman, 1993; Pickering, 2004; Tharp, 2007). Both accounts suggest that response perseveration is mediated by midbrain DA activity and, therefore, suggest that the activation of the BAS should mediate response modulation.

The study makes a further distinction in the type of reversal learning required by participants. In fact, during the reversal phase of the symmetric task participants experience an intra-dimensional shift (i.e. the stimulus-response assignments are swapped) whereas in the reversal phase of the asymmetric task, they experience a reward switch. Indeed, the payoff matrix is reversed so that the category that offered low payoffs in the learning phase offers high payoffs in the reversal phase. This dissociation may identify different processes and personality traits activated during reversal learning.

9.3 Method

9.3.1 Participants

Participants were drawn from an opportunity sample recruited over the summer holiday around Goldsmiths and it consisted of both students and non-students.

The sample was made up of 32 participants, 16 males and 16 females. All but three participants in the sample were right-handed. All participants were aged between 18 and 38 (mean age = 25.2; s.d. = 4.7).

Participants were guaranteed confidentiality. They were tested in one sitting that lasted approximately 1 hour and they each received £8.50 for their participation.

9.3.2 Design

The present study assessed the impact of payoff matrices on performance during a category learning task. In particular it was interested in comparing the impact of a symmetric and an asymmetric payoff matrix on response bias. Therefore, the independent variable was payoff matrix and it was run within-subjects with two levels, i.e. symmetric and asymmetric. The study assessed the impact of different types of payoff matrices (IV) on decision bound location (i.e. the dependent variable). The two payoff matrices were identical to the ones used in the previous study (table 9.1).

Table 9.1. Structure of the payoff matrices and performance criteria adopted for the two task conditions (✓ = correct vs. ✗ = incorrect)

	High payoff category		**Low payoff category**		**Performance criterion**
Matrix	✓	✗	✓	✗	
Asym.	400	100	200	100	35900
Symm.	300	100	300	100	33700

9.3.3 Task and apparatus

The task was run on a Mesh PC and a Mitsubishi 21" monitor with 1024 x 768 pixel resolution in an artificially lit room. Each stimulus was computer generated by using Matlab routines from Brainard's (1997) Psychophysics Toolbox. Stimuli were flashed on a black background that filled the entire screen.

Participants performed the symmetric and the asymmetric task in an intertwined fashion, in order to avoid any possible order effects caused by the order of presentation of the two task conditions (as was observed in chapter 8). The stimuli in both tasks were presented in a fixed quasi-random order to all participants. Each stimulus appeared on the screen until participants responded.

In order for participants to be aware of the fact that they were performing on two different tasks, each task presented different stimuli although they were equivalent in presentation. It was expected that in this way, participants would be encouraged to use different approaches to deal with the two conditions.

In one of the tasks participants had to learn to categorise graphical 'bars' that appeared on the screen; in the other task participants learnt to classify small circular dots (as in the previous chapter). In both tasks, the stimuli appeared within a 650*650 pixel box that delineated the stimulus space which participants had to concentrate on. The bars always appeared along an imaginary line 650-pixel in length running horizontally. The bars all started at pixel number 1 on the left-hand side of the stimulus box and ended at a horizontal position which corresponded to the centre of the circular dots presented in task implemented in chapter 8. Hence, the relevant dimension that determined which category the stimuli belonged to was bar-length, whereas the bars' width was constant and equal to 40 pixels (approx. 1.5 cm). In the other task, as in the task presented in chapter 8, the stimuli were small circular dots (14 pixels in diameter, approximately .525 cm). The dots would appear at varying positions along an imaginary line 650-pixel in length running vertically. The relevant dimension for this task was location of the dot along the imaginary vertical line.

The stimuli for the two tasks were created in the same way as the stimuli implemented in the previous study to ensure that the two tasks were equivalent. In the bar task, participants had to classify stimuli as members of category 1 or category 2. The stimuli for category 1 and category 2 were generated by sampling from two independent but overlapping normal distributions, thus making the task probabilistic. The distribution used to generate category 1 stimuli had a mean score equal to 275 whereas the one used to generate category 2 stimuli had a mean equal to 375. Both distributions had a standard deviation equal to 100.

In a similar fashion, the stimuli in the dot task had to be classified into either category A or category B. The stimuli used in the dot task were generated in exactly the same way as the ones used in the bar task. Category A stimuli had a mean equal to 275 pixels and category B stimuli a mean of 375, with both distributions having a standard deviation of 100. The dots were drawn, centred at these random, varying positions, measured from the top of the imaginary vertical line.

In the bar task, they had to press the 'up arrow' key or the 'down arrow' key to classify the stimuli into category 1 or category 2, respectively. In the dot task, they had to press the 'left arrow' key or the 'right arrow' key to classify the stimuli into category A or category B, respectively. The arrow keys were selected as response keys due to their proximity on the keyboard. The proximity, in fact, allows participants to press each one of them by just using one finger and, therefore, avoid any possible handedness effects. Participants were clearly instructed to use 'the index finger of your dominant hand'. The corresponding keys were covered by a sticky label marked as 1, 2, A and B as necessary. This was done in order to make it easier for participants to remember which key corresponded to which category. Participants were also informed that category membership was independent from the sequence/order in which trials of the two different tasks were presented. Hence, a strategy based on trial sequence (i.e. if two dots trials appear in a row, the subsequent stimuli) would not be effective.

The task consisted of 320 trials, the first 160 trials constituted the initial learning phase whereas the last 160 made up the reversal learning phase. Each phase consisted of an equivalent number of trials from the two task conditions. Prior to performing on the experimental task, participants carried out a practice session which consisted of 160 trials. During the practice trials, participants were exposed to 80 trials of the bar task and 80 of the dot task. The task presentation in the practice was identical to the one in the experimental task so that participants had the opportunity to learn the appropriate category memberships that ensured optimal performance. The only difference was in the feedback signals received by participants. Indeed, during practice, participants received only accuracy messages (i.e. correct/incorrect) following each response.

9.3.3.1 Practice trials

During the practice session, participants received visual and auditory accuracy feedback following each response. The visual feedback informed them on whether their response was 'correct' or 'wrong'. The word correct was written in green whereas the wrong word appeared in red to underline the nature of the feedback. The auditory feedback consisted of a high-pitched tone for correct responses whereas incorrect responses were accompanied by a low-pitched tone. Once the practice trials were over, participants were invited to express any questions and when any doubts were solved, the experimental session started.

The practice trials were included in order to ensure participants would develop a uni-dimensional, rule-based strategy and, therefore, ensure the activation of the explicit system prior to the experimental task. In this way, it can be assumed that any response bias observed in the experimental task is due to the activation of the implicit system under an asymmetric payoff matrix. Indeed, over trials the implicit system is expected to take control of performance and over-ride the explicit system, as indexed by the response shift.

9.3.3.2 Learning phase

The procedure implemented in the learning phase of the task was equivalent to the one used in the previous study and it aimed to replicate those previous findings. In fact, during the two tasks, participants had to learn by trial and error to classify the stimuli on the screen into two categories. Depending on the counterbalancing order, one task (i.e. bar vs. dot) constituted the symmetric condition whereas the other task constituted the asymmetric condition. Participants had to learn to classify the bar stimuli into category 1 and 2 and the dots task into category A and B. Category 1 or category A represented the high payoff category (i.e. 400 points if correct) depending on which task (bar or dot) was used for the asymmetric condition according to the counterbalancing order (see below). It was decided not to counterbalance which category offered the higher payoffs in the asymmetric task since results from the previous study showed that such counterbalancing was probably not necessary.

Participants received feedback for each response they made. Contrary to the previous study, the feedback message informed them only about how many points they earned but offered no information about how many they could have won on that trial. It was suspected that receiving both types of information may have been perceived as a form of ‘accuracy’ feedback that may have, subsequently, hindered the shift from the optimal accuracy criterion to the optimal reward criterion in the asymmetric task.

Additionally, when participants earned 100 points and were informed that they had the potential to have won 200, 300 or 400 points on that trial, they could have experienced the feedback message as negative feedback and therefore activated the punishment system (FFFS). If this were the case, this could have, subsequently, affected the BAS output (Corr, 2004a; Smillie et al., 2006; Pickering, 2008).

9.3.3.3 Reversal learning phase

In the reversal learning phase, the tasks were presented in the same fashion as in the learning phase although the contingencies had changed. Hence, participants had to withhold from using the old response strategy and switch to a new and optimal strategy. In the symmetric task, it was the categories' location that had been reversed. In fact, during the reversal phase category A/1 stimuli had a mean location equal to 375 pixels and category B/2 stimuli a mean of 275 (cf. 275 for A/1 and 375 for B/2 in the learning phase).

In contrast, in the asymmetric condition the high payoff category was switched, so that category B/2 became the high payoff category (i.e. 400 points for correct responses) and category A/1 the low payoff one (i.e. 200 points for each correct trial).

Participants were not informed about the presence of the reversal learning phase or the need to switch rules. The two learning phases were presented as one task and, therefore, there was no break between the two phases. Hence, the task presentation and the feedback messages were kept identical as in the first phase of the task.

During both learning phases, participants received visual feedback for each response they made. In contrast to the feedback used in the previous study no auditory feedback was given in the present study. This was done in order to ensure that performance was reward-, rather than accuracy-, driven. Indeed, the use of auditory feedback may have been interpreted as a source of accuracy feedback by participants. Following this logic, the visual feedback did not offer any form of accuracy signal and only informed participants on the amount of points they had earned following their response (cf. the amount of both potential and actual points as in the previous study). Feedback was presented for 300 msec and followed by 100-msec ITI.

9.3.4 Counterbalancing

Counterbalancing (CB) across participants was necessary to control for confounding variables, such as bias produced by the task representation (i.e. horizontal vs. vertical) or possible key-mapping bias. Hence, it was decided to counterbalance across participants which task was run as the asymmetric or the symmetric condition. There were only two possible alternative CB orders, i.e. asymmetric-dots and symmetric-bars (CB1) vs. asymmetric-bars and symmetric-dots (CB2). Participants were automatically allocated to one of the two CB orders by the computer depending on the odd or even-numbered ID code they had been allocated to.

9.3.5 Personality measures

In line with the previous study, the main personality measures included were the four personality components extracted in the principal component analysis (PCA, see chapter 4 for details). Participants completed several personality measures, which were: the Eysenck Personality questionnaire revised (EPQ-R), the Oxford-Liverpool inventory of feelings and experiences scale (OLIFE), the sensation seeking scale (SSS), the schizotypal personality questionnaire (SPQ), the big five inventory (BFI) and the BIS/BAS questionnaire. Four personality components were extracted after running a PCA with Varimax rotation on these scales. The four components extracted are: Extraversion (E), Neuroticism (N), Positive Schizotypy (PS) and impulsivity-antisocial (ImpAss). The PCA was run on 232 participants, who constitute the overall number of participants tested.

The components scores used in the present study are the corresponding components scores extracted in the overall PCA involving 232 participants.

Additional measures of impulsivity were also collected in this study. Gray suggested that the personality trait corresponding to the BAS was impulsivity. It was, hence, decided to collect measures of the Dickman Impulsivity Inventory (DII; Dickman, 1990. See chapter 5 for details). The DII captures two sub-types of impulsivity (which are related to reward-sensitive and rash impulsivity). In this way, it is possible to explore how different types of impulsivity relate to performance on the decision-making task and to the four personality components as discussed in chapter 4. The extracted components include measures that are supposed to be related to the BAS output (e.g. extraversion, impulsivity and positive schizotypy).

9.3.6 Procedure

The study method and procedure were approved by the Psychology Department Ethics Committee (DEC) at Goldsmiths, University of London. Testing was conducted in a testing room in the department. The experimental procedure in the present study resembled the procedure adopted in the previous study, since it aimed to replicate and extend those findings. In particular, all participants were given a promotion focus as they were told that they could win two entries into a £25 draw. They could earn an entry in each task, depending on their performance. Indeed, in order to win the entry, they had to exceed a certain amount of points (i.e. performance criterion; table 9.1). Participants could earn one entry in each task so the criteria were calculated across the overall 160 trials which constituted each task. One draw was held for all 32 participants.

Prior to performing on the experimental task, participants carried out a practice session that consisted of 160 trials. The 160 trials contained equal amount of trials from the two tasks that were presented, like in the experimental task, in an intertwined fashion. A set of written instructions was presented to participants on the screen before they carried out the practice session. The instructions closely resembled the ones presented in the study described in the previous chapter. However, the description of the two tasks and the necessary key presses had to be edited to account for the changes in the task. They were presented with written instructions on the computer screen. Participants were informed that their task was to learn, by trial and error, to classify different stimuli into category A and category B, or category 1 and category 2 depending on the task (i.e. dot and bar, respectively).

Initial instructions were presented to remind participants about the structure of the task and to inform them that during this session they had the possibility of earning up to two entries into the £25 lottery, depending on their performance. Following each response, participants earned points (described earlier). All the points that participants won were added up on a vertically oriented "point meter" displayed on the right-hand side of the computer screen. The meter consisted of a 765 pixel tall x 50 pixels wide rectangle and it was set to zero at the beginning of each task. The performance criterion was presented as a horizontal line across the meter and was labelled 'Bonus'. The performance criterion for each task was set at 80% of the points that the optimal classifier would obtain over the overall trials of each task (i.e. 160), relative to 0% accuracy. Each task offered the chance to win an entry into the lottery. Thus, they had different criteria depending on the payoff matrix structure (i.e. asymmetric vs. symmetric; table 9.1). The point meter was updated after each response at a rate of 1 pixel per 50 points. The part of the meter that changed flashed three times to stress the idea that the number of points was increasing.

The task consisted of 4 blocks of 80 trials and at the end of each block, the task stopped to give participants the chance to take a short break if they needed a break. The break was not compulsory and participants could choose to carry on with the task but if they decided to take a break, its length was totally subjective. As in the previous study, at the end of the two tasks, participants were debriefed by a message which informed them on the number of lottery tickets they had earned (i.e. 0, 1 or 2).

9.3.7 Data analysis

A variation of the formal model implemented in the previous study was developed in order to fit the data and to obtain the actual criterion location (cf. calculated 'C' using signal detection theory, SDT). As mentioned in the previous chapter, the use of the formal model allows us to identify and exclude those participants who were not using the appropriate uni-dimensional rule based on the relevant dimension.

As already noticed, the task comprised two task conditions related to the structure of the payoff matrix (asymmetric vs. symmetric payoff matrix). Each condition consisted of 160 trials each. The first half of each task constituted the learning phase whereas the second half of the task represented the reversal learning phase of the task. Hence, each half of the two tasks can be perceived as a separate phase. The stimuli had been randomly generated to have equal number of trials for each task condition across the two phases. However, within each learning phase there was an uneven number of trials of each task (bars vs. dots). In the initial learning phase (first 160 trials), there were 81 trials of the vertical dots and 79 trials of the horizontal bars task. The opposite was true for the second half of the task (i.e. reversal learning). Depending on the CB, each task was associated with either the asymmetric or the symmetric task. Regardless of task type/condition it was decided to include all trials up to the 79th so that the two conditions could be comparable on the amount of learning.

In order to maintain symmetry with the procedure adopted in the previous study, it was decided to exclude the first 25 trials since they only represent familiarisation (Markman, et al., 2005). Thus, the analyses were only run on 54 trials, which are referred to as the 'last' trials. No end trials were removed since, due to the small number of trials, no (learning) ceiling effect was expected to be reached by the 79th trial , and this was supported by scores on the proportion of correct (pc) responses made which were below maximum accuracy in all four halves of the task (see table 9.2 below).

Finally, the results obtained in the empirical study were further explored by simulating the data through the neural model. As mentioned in the previous chapter, the neural model may help explain the human behaviour observed in the present study from a neurobiological perspective.

9.4 Results

9.4.1 Human data

This result section explores participants' performance during the two task conditions across the two learning phases. The results section is constituted by two sub-sections which summarise the results of each learning phase separately (learning vs. reversal phase). Participants' performance is mainly indexed by their decision criteria across the two conditions which were extracted using the formal model for the two learning phases.

9.4.1.1 Preliminary analysis

9.4.1.1.1 Proportion of correct responses

Preliminary analyses were conducted on the proportion of correct (pc) responses in the two task conditions. Hence, as for the rest of the analyses reported below, only the 'last' 54 trials were include in the proportion of correct responses analysis. The analysis showed that overall pc scores were above chance levels across the two payoff conditions and in the two learning phases (table 9.2).

Table 9.2. Mean and standard deviation values of the proportion of correct responses in the two task across the two learning phases including all participants (pc = proportion of correct responses, asy = asymmetric; sym = symmetric; 1h = learning phase; 2h = reversal phase)

	pc_asy_1h	pc_sym_1h	pc_asy_2h	pc_sym_2h
Mean	0.59	0.61	0.59	0.53
SD	0.09	0.09	0.08	0.094

The fact that participants performed above chance is not surprising since they had extended training during the practice trials and accuracy analysis showed that in the last 54 (out of 80 practice) trials, participants had achieved pc scores above chance level in both tasks [i.e. dots and bars; dots: mean = 0.62 and s.d. = 0.10; bars: mean = 0.59 and s.d. = .08]. Further results showed that 81.3% of the sample performed above chance in the learning phase for both the symmetric and asymmetric condition. In contrast, in the reversal phase of the asymmetric condition 84.4% of the sample scored above chance whereas only 62.5% of the sample scored above chance in the symmetric condition.

Those participants who were identified by the formal model as guessing or using an alternative strategy were removed. Following exclusion of these participants, the overall sample size included in the analysis was equal to 17 in the learning phase whereas it was equal to 7 in the reversal phase (see modelling sections). The average proportion of correct (pc) responses increased numerically for all conditions, but for the reversal phase of the symmetric task (table 9.3). In fact, in the symmetric task, participants performed at chance level in the reversal phase.

Table 9.3. Mean and standard deviation values of the proportion of correct responses in the two task across the two learning phases including only those participants that implemented the relevant uni-dimensional strategy (asy = asymmetric; sym = symmetric; 1h = learning phase; 2h = reversal phase)

	pc_asy_1h	pc_sym_1h	pc_asy_2h	pc_sym_2h
Mean	0.65	0.66	0.64	0.50
S.D.	0.05	0.06	0.07	0.12

Paired-sample t-tests indicated that the difference in accuracy scores between the asymmetric and the symmetric condition was statistically significant in the second half of the task [t(16) = 2.11, p= .051]. Results also showed that during the symmetric task, the difference in pc scores across the two learning phases was significant [t(16) = 3.17, p =.002] whereas it was just a trend in the asymmetric task [t(16)=1.7, p = .11, 2-tailed]. A 2x2 mixed-design ANOVA with CB order (i.e. bars vs. dot) as a between-Ss IV and payoff condition as the within-Ss IV showed that CB did not affect accuracy scores. In fact, the CB main effect was non-significant [F(1,15) = 1.88, ns] and, similarly, the CB*payoff interaction was also non-significant [F(1,15) = 1.75, ns].

9.4.1.1.2 Reaction time scores

Table 9.4 summarises the reaction time (RT) scores in the two payoff conditions across the two learning phases. The RT scores are very similar across phases. However, a paired-sample t-test indicated that participants were statistically faster at responding during the reversal phase than the learning phase under asymmetric payoffs [t(31) = 2.27, p = .030]. This is explained by the fact that even though the payoff structure was switched, the categories and the response structure were identical in the two phases. Hence, the extended practice may have enhanced speeding up responding in the second half of the task.

Table 9.4. RT scores in the two task conditions across the two learning phases (asy = asymmetric; sym = symmetric; 1h = learning phase; 2h = reversal phase)

	Asy 1h	Sym 1h	Asy 2h	Sym 2h
RT mean (s.d.)	1.46 (.49)	1.47 (.51)	1.29 (.36)	1.38 (.42)

In contrast, RTs were equivalent across the two phases under symmetric payoffs [t(31) = 1.00, ns].

In order to further investigate RT scores in the asymmetric condition, a difference score was calculated by subtracting the RT scores in the second half of the score from the RT scores in the first half of the task (i.e. RT difference scores) and it was correlated with the four personality components. Results indicated that there was a trend for a correlation between positive schizotypy and the RT difference score (r = .31, p = .087). All other correlations were close to zero (ps > .2). As visually represented in figure 9.1, participants with high scores on the positive schizotypy scale were the ones that showed the greatest RT reduction in the second half of the asymmetric task.

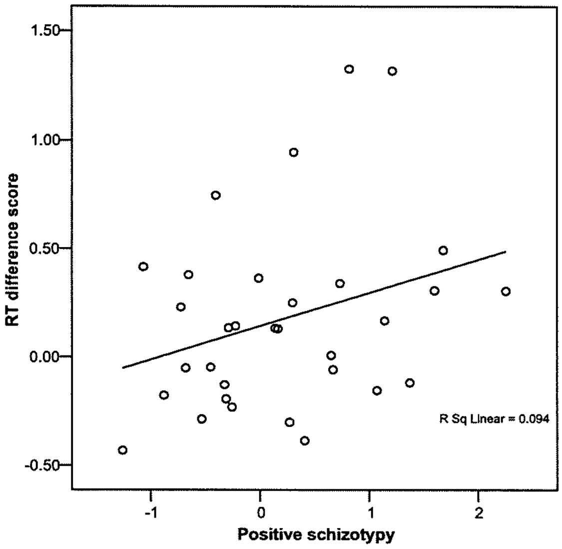


Figure 9.1. Scatterplot that indicates the relationship between positive schizotypy and the RT difference measure in the asymmetric task (axis: x = standardised PS scores; y = RT difference)

9.4.1.1.3 Points earned

Participants gained points throughout the task and aimed to reach the performance criterion to earn entries into the £25 lottery. Table 9.5 shows the amount of points won by the overall sample (N = 32) and the points earned by those participants who had used the appropriate uni-dimensional rule, as indicated by the formal model. For symmetry with the rest of the analysis, the point scores are calculated across the last 54 trials.

Table 9.5. Mean and s.d. scores on the amount of points earned in the two tasks across the two learning phases by the overall sample and those participants who used the appropriate unidimensional rule (UD-users)

	Asy 1h	Sym 1h	Asy 2h	Sym 2h
Overall sample	11459.4 (910.8)	11975.0 (997.7)	12009.4 (1265.8)	1110.0 (1016.0)
UD-users (N= 17)	12041.2 (512.4)	12494.1 (629.0)	12158 (1219.7)	11153.0 (1273.8)

Those participants who implemented the appropriate uni-dimensional strategy gained more points than those who were guessing or using alternative strategies. Further t-tests were run to explore the pattern of earned points among those participants who implemented the appropriate strategy. The results showed that participants earned significantly more points in the symmetric task than the asymmetric task during the learning phase [t(16) = -2.13, p = .049] whereas the opposite pattern was a borderline trend in the reversal phase [t(16) = 2.08, p = .054]. The opposite pattern observed in the reversal phase may be due to the fact that participants gained significantly less points in the symmetric task in the reversal phase compared to the learning phase of the task [t(16)= 3.61, p = .002]. This may be due to the fact that the intra-dimensional switch applied in the reversal phase of the symmetric task required learning of new response key mappings rather than just further fine tuning of a decision boundary. In contrast, the difference in points earned across the two phases of the asymmetric condition was non-significant [t(16) = -.34, ns].

9.4.1.2 Learning phase

9.4.1.2.1 Model fitting

The stimuli presented in the two task conditions consisted of dots or bars that varied in one dimension, i.e. location on the screen. Location was the relevant dimension and participants were expected to develop a uni-dimensional rule based on location in order to successfully solve the categorisation problem. Hence, a uni-dimensional model which used location as the relevant dimension was fitted to the data together with a guessing model. The two formal models were used to fit the data collected in the asymmetric and the symmetric task conditions. Figure 9.2 illustrates the step-like learning functions typical of the data that was well-fitted by the uni-dimensional model (figure 2a) and the distribution typical of data fitted by the guessing model (figure 2b). Moreover, the goodness of each model was compared to the goodness of fit of the saturated model with no free parameters.

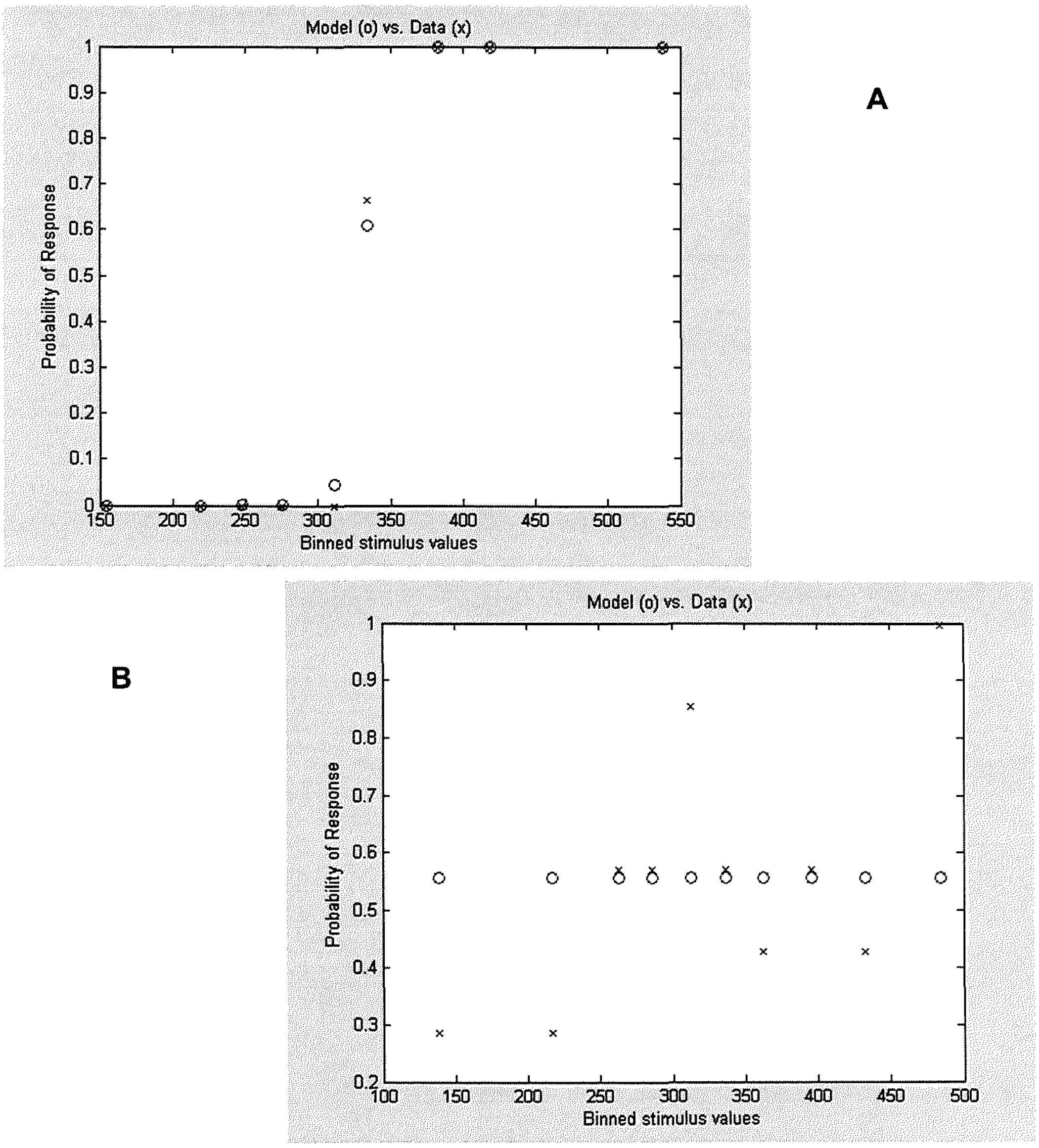


Figure 9.2. It illustrates the distribution typical of data that was well-fitted by the uni-dimensional model (figure a) and the distribution of data fitted by the guessing model (figure b). The red stars represent the actual data whereas the blue circles represent the model fitting.

1 – Fitting through a guessing model

The guessing model describes the situation where participants were randomly assigning the stimuli to one of the two categories, regardless of the location of the stimuli on the screen. This model has only one free parameter (guessing probability for category A, from 0 to 1). Results showed that the guessing model was not significantly worse than the saturated model for 8 cases out of the asymmetric and 7 cases out of the symmetric task conditions. However, none of the cases coincided.

2 – Fitting through a uni-dimensional rule model based on spatial location

The present model describes the situation where participants select spatial location as the relevant dimension and place a criterion on this dimension to categorise the dots. This family of models includes the optimal rule since location on the screen was the relevant dimension in the task. The model was applied to data collected across both payoff conditions.

The formal model showed that for 21 (out of 32) participants the uni-dimensional model was not significantly worse than the saturated model in the asymmetric task condition. Those 21 participants showed a criterion mean equal to 351.27 (47.37) and a noise level equal to 95.51 (84.52). This decision criterion value was statistically different from the optimal accuracy criterion [i.e. 325; $t(20) = 2.54$, $p = .010$]. Additionally, the formal model also showed that for 24 participants the uni-dimensional model was not significantly worse than the saturated model in the symmetric condition. These fitted cases placed their criterion around 333.98 pixels (66.85), which is not statistically different from the optimal accuracy criterion [$t(23) = .66$, ns]. Moreover, they showed a mean noise level equal to 97.41 (80.55).

9.4.1.2.2 Comparison of the two models

As noted above, the uni-dimensional model fits 21 and 24 cases in the asymmetric and the symmetric task conditions, respectively. However, 8 of the 21 fitted data sets collected in the asymmetric task were also fitted by the guessing model and 6 of the 24 fitted data sets from the symmetric condition were also fitted by the both model.

It was decided to compare the two models in order to assess which one offered the best fitting for these participants. Comparisons were made by weighing the Akaike Information Criterion[15] (AIC; Dayton, 2003). The AIC score was calculated on the basis of the free parameters and it is an estimate of the goodness of fit. It penalises the model with extra free parameter so that the lower the AIC score, the better the fitting (i.e. closer to the saturated model; Maddox, Ashby & Bohil, 2003).

Results indicated that the uni-dimensional model offered a better fit than the guessing model for all but two cases. In fact, a data set from one participant was best fitted by the guessing model in the asymmetric task and another data set from the symmetric task was also best fitted by the guessing model. These two cases were excluded from the final sample.

Further analyses showed that for 17 participants, the UD model was non-significantly worse than the saturated model in both conditions. These 17 participants belonged equally to the two CB orders; in fact 8 performed under CB1 (i.e. asymmetric task consisted of dots that varied in location on a vertical dimension and the symmetric task consisted of a bar varying in length over a horizontal dimension) while 9 participants performed under CB 2. Across these 17 cases, the criterion location was equal to 340.38 (30.93) and 336.27 (37.43) in the asymmetric and the symmetric condition, respectively. The noise level was equal to 65.68 (46.76) in the asymmetric and 77.86 (69.85) in the symmetric condition. Paired-sample t-tests indicated that, across the two payoff

---

[15] $AIC = 2r - 2\ln L$, where r is the number of free parameters and L is the log likelihood of the model (Maddox et al., 2003)

conditions, neither the criteria were significantly different [t(16) = .320, ns] nor the noise levels [t(16) = -.79, ns], among those participants who used the appropriate strategy.

In conclusion, these results suggest that 47% of the sample failed to implement the relevant uni-dimensional strategy but they were not guessing either. They will be referred to as 'non-learners' to indicate that they failed to learn to use the appropriate strategy but not that their performance was at or below chance (see below). These participants may have been using alternative strategies that ensured above chance, although sub-optimal, performance.

As in the previous study, some participants reported having treated the distance between the last stimulus on the screen and the new stimulus as the relevant dimension to formulate their decisions. Additionally, some participants reported having used the sequence of 'dots' and 'bar' trials to formulate a strategy. Hence, the sort of rule they developed could have been along the following lines: 'if a 'bar' trial follows two 'dot' trials then the stimulus belongs to category A, otherwise it belongs to category B'. Once again this occurred despite specific warnings in the instructions that such rules and strategies would not be helpful.

9.4.1.2.3 Correlation between the strategy used and personality components

Point-biserial correlations were run in order to assess whether there was a relationship between any of the personality components and the strategy used. The variable that codes the type of strategy implemented was labelled strategy used and coded as 1 the use of a uni-dimensional strategy and as 0 guessing or using an alternative strategy. In the asymmetric task, positive schizotypy was found to be significantly correlated with the type of strategy implemented by participants (r = -.38, p = .03). Indeed, those participants who implemented the appropriate uni-dimensional strategy scored lower on the standardised PS component (mean = -.024 and s.d. = .78) than those who were guessing or using an alternative strategy (mean = .658 and s.d. = .89). No correlation was significant between the strategy used in the symmetric task and positive schizotypy (r = -.005, ns) or any of the other personality components (p >.63).

9.4.1.3 Behavioural data

Some preliminary analyses were carried out to explore the behavioural data and observe where participants placed their decision criteria under the two payoff matrix conditions. Only data from those participants who had implemented the appropriate dimensional strategy was included (N = 17). The mean and standard deviation values for the (fitted) criteria under both payoff conditions are presented in table 9.6. A 1-sample t-test indicated that the criterion location just failed to be significantly different from the optimal accuracy criterion (325) in the asymmetric condition [t(16) = 2.05, p = .057] but the criterion did not differ from the optimal accuracy level in the symmetric condition [t(16) = 1.24, ns].

Table 9.6. Fitted criteria for the asymmetric (asy) and the symmetric (sym) conditions across those participants who used a uni-dimensional strategy (UD-users) and the overall sample

	Criterion asy	**Criterion sym**
UD-users	340.38 (30.9)	336.27 (37.4)
Overall sample	346.40 (51.7)	330.29 (68.8)

These results indicate that the overall sample showed a numerically greater bias than the bias showed by those participants who implemented a uni-dimensional strategy (UD-users). The fitted criterion of the overall sample was significantly different from the optimal accuracy criterion [t(31) = 2.34, p = .03]. Nonetheless, the greater bias might have been determined by the use of alternative (non-dimensional) strategies, which may have relied on the implicit system more strongly than on the explicit system. Indeed, participants who did not use a uni-dimensional strategy showed a pattern of responding that resembled the data simulated by the purely RPE-based implicit model (see non-learners section).

A mixed-design ANOVA was carried out with payoff matrix as the main IVs with two levels (symmetric vs. asymmetric) and it was run as a within-subject factor. The main DV was the decision criterion. Counterbalancing order was also included in the analysis as a between-Ss factor in order to assess whether it affected performance. Results from a mixed-subjects ANOVA indicated that neither the main effect of payoff matrix [F(1,15) = .07, ns] nor the matrix*CB interaction [F(1,15) = .95, ns] were significant.

9.4.1.3.1 Performance and personality

The analyses reported below aimed to establish the relationship between the four personality components extracted through the PCA (chapter 4) and performance during the two task conditions, in particular the decision criterion.

Preliminary correlations were run to assess the relationship between the four personality components and the criterion location as well as the criterion shift score (table 9.7). In the asymmetric condition, participants are expected to place their criterion away from the optimal accuracy criterion (325) closer to the optimal reward criterion (i.e. above 325). Hence, the criterion shift score was obtained by subtracting the criterion location in the symmetric condition from the criterion location in the asymmetric condition and, therefore, it should on average be positive if the expected response bias is occurring in the asymmetric condition.

Table 9.7. Correlations between the four personality components and the criteria set in the two payoff conditions (a = asymmetric; s = symmetric) and their difference score (shift)

		Criterion_a	Criterion_s	Criterion shift
E	Pearson Correlation	-.262	.092	-.218
	Sig. (2-tailed)	.310	.725	.400
	N	17	17	17
N	Pearson Correlation	.378	-.279	.419
	Sig. (2-tailed)	.134	.278	.094
	N	17	17	17
PS	Pearson Correlation	-.149	.059	-.129
	Sig. (2-tailed)	.569	.823	.623
	N	17	17	17
ImpAss	Pearson Correlation	-.415	.417	-.539
	Sig. (2-tailed)	.097	.096	.026
	N	17	17	17

Results showed that there was a trend for a negative correlation between scores on the ImpAss component and the criterion decision placed during the asymmetric task (r = - 0.42, p = .097). Moreover, there was a significant trend for a positive correlation between ImpAss scores on the decision criterion in the symmetric task (r = .42, p = .096). These results suggest that the impact of ImpAss on performance varies depending on the payoff manipulation. Finally, there was a negative correlation between ImpAss scores and the shift in criterion across the two conditions (r = -.54, p = .026). There was also a trend for a positive correlation between neuroticism and the shift in bias (r = .42, p = .094).

The scatterplot, reported below (figure 9.3), displays the negative relationship between scores on the ImpAss component and the criterion shift. Participants who scored high on the ImpAss components showed the smallest criterion shift (i.e. small bias in the asymmetric task)

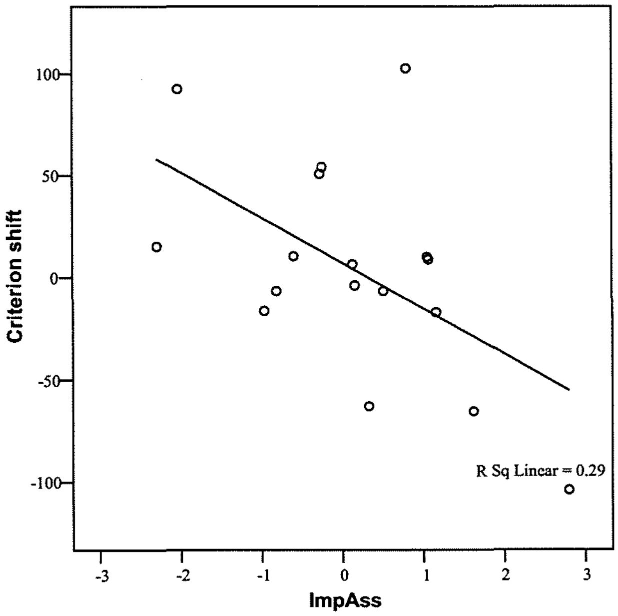


Figure 9.3. The scatterplot summarises the negative correlation between ImpAss scores and the criterion shift values

9.4.1.4 Further analyses

In order to gain a better understanding of the two (near-) significant correlations, the median split technique was implemented on the relevant personality components, in order to transform them into binary valued factors coded as 1 for low scores and as 2 for high scores. Two independent mixed-design ANOVAs were run to further explore these relationships, the binary-valued personality factors were entered as between-Ss factors in the ANOVAs.

9.4.1.4.1 ImpAss

Due to the loss of power caused by transforming the ImpAss component into a binary-valued factor and the small sample size, the interaction between ImpAss and payoff was only a trend [F(1,15) = 2.87, p = .11]. Figure 9.4 shows the relationship between the ImpAss factor and payoff condition. The figure indicates that participants with low scores on the ImpAss measures show a greater bias during the asymmetric payoff condition compared to those participants with high scores on the ImpAss traits[16].

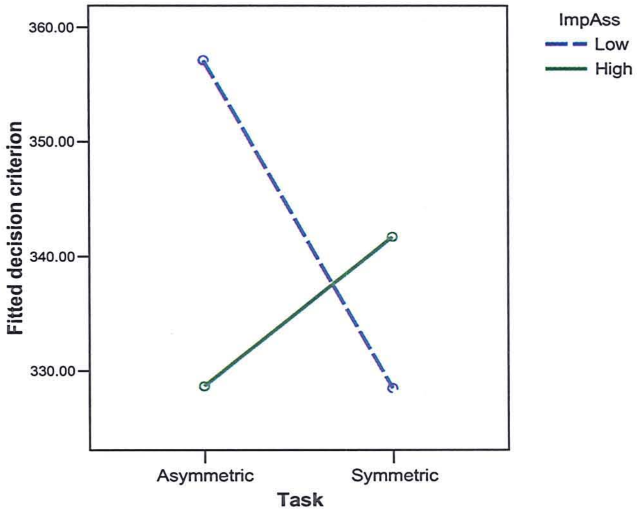


Figure 9.4. Relationship between the decision criterion across the two task conditions and scores on the ImpAss component

---

[16] Power analyses carried out using the G*Power program with n = 7 and an expected small to medium effect size (d = .3) showed that the actual test's power for a repeated measure t-test was equal to 0.1741, which is very low. However, even with low power there was a trend for the low ImpAss individuals to show a positive bias towards the high payoff category under the asymmetric condition compared to the symmetric condition [t(6) = 1.96, p = .098, 2-tailed].

An independent t-test showed that the difference in criterion placement during the asymmetric task across low and high ImpAss individuals just failed to be significant [t(15) = 2.05, p = .058].

9.4.1.4.2 Neuroticism

Despite the loss of power caused by recoding the neuroticism component into a binary-valued factor, the neuroticism by payoff interaction was significant [F(1,15) = 6.82, p = .020]. Figure 9.5 shows the relationship between the neuroticism (N) factor and the payoff condition. Participants with high scores on the N component showed a shift in bias during the asymmetric task condition and no bias during the symmetric task. In contrast, individuals with low scores on the neuroticism component showed the opposite pattern, i.e. a bias in the symmetric task and no bias in the asymmetric task condition[17].

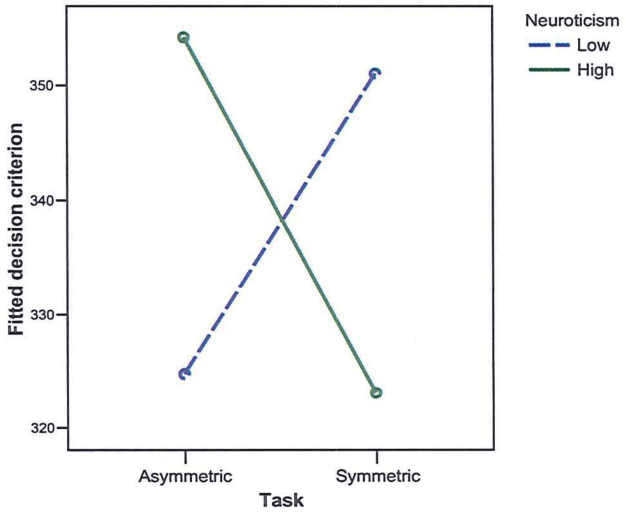


Figure 9.5. Relationship between the decision criterion across the two task conditions and scores on the neuroticism component

---

[17] Even with very low power, post-hoc tests on neuroticism showed that there was a trend for a difference in criteria between the asymmetric and the asymmetric task for participants with both high [t(8) = 1.82, p = .104, 2-tailed] and low [t(7) = -1.97, p = .089, 2-tailed] scores on the neuroticism component.

9.4.1.4.3 Dickman's impulsivity inventory

Correlations were run to assess the overall relationship between the four personality components extracted in the PCA (especially extraversion and impulsivity) and the two impulsivity scales of the Dickman impulsivity inventory (functional and dysfunctional impulsivity; table 9.8).

Table 9.8. Correlation between the four personality components and functional and dysfunctional impulsivity

		E	N	PS	ImpAss
FI	Pearson Correlation	.441	-.552	-.242	.131
	Sig. (2-tailed)	.011	.001	.183	.476
	N	32	32	32	32
DI	Pearson Correlation	.325	.038	.342	.371
	Sig. (2-tailed)	.070	.838	.055	.036
	N	32	32	32	32

Results indicated that functional impulsivity was significantly and positively correlated with extraversion ($r = .44$, $p = .01$) and negatively correlated with neuroticism ($r = -.55$, $p = .001$). In contrast, dysfunctional impulsivity was significantly correlated with the ImpAss component ($r = .37$, $p = .04$) while its positive correlation with positive schizotypy just failed to be significant ($r = .32$, $p = .06$) and the correlation with extraversion showed only a trend ($r = .33$, $p = .07$). Hence, these results support the view that extraversion is a better measure of reward-sensitivity (impulsivity) whereas ImpAss measures are better measures of rash impulsiveness (Smillie and Jackson, 2006).

The significant correlation between dysfunctional impulsivity and the ImpAss component is in line with the literature that suggests that they measure the same type of impulsivity (i.e. rash impulsivity; Dawe, 2004; Smillie and Jackson, 2006). Hence, it was decided to calculate a composite score for measures on the standardised ImpAss component and the dysfunctional impulsivity (DI) factor. This was done in order to further explore the relationship between impulsivity and behaviour (i.e. criterion placement). This composite impulsivity factor was obtained by transforming the DI scores into z scores and, subsequently, averaging these scores with the scores on the standardised ImpAss component {i.e. [Z(ImpAss)+Z(DI)]/2}. Further correlations were run between the new composite impulsivity score, the decision criteria and shift score for those participants who had used the appropriate uni-dimensional strategy (table 9.9).

Table 9.9. Correlations between the composite impulsivity scores and the criteria set in the two payoff conditions (a = asymmetric; s = symmetric) and their difference score (shift)

		Criterion_a	Criterion_s	Criterion shift
Impulsivity composite score	Pearson Correlation	-.469	.238	-.443
	Sig. (2-tailed)	.057	.358	.075
	N	17	17	17

Results show that the negative correlation between the composite impulsivity score and the criterion location in the asymmetric condition just failed to be significant (r = -.47, p = .06) whereas the correlation between impulsivity and the criterion location in the symmetric task was non-significant (r = .24, ns). Nonetheless, the correlation between impulsivity and the criterion shift showed a trend (r = -.44, p = .075). Hence, the composite score confirms the results reported above and it shows a stronger correlation between impulsivity and criterion location in the asymmetric task. Impulsivity does not seem to moderate the criterion location in the symmetric condition. The scatterplot displays the negative correlation between impulsivity and the shift score (figure 9.6).

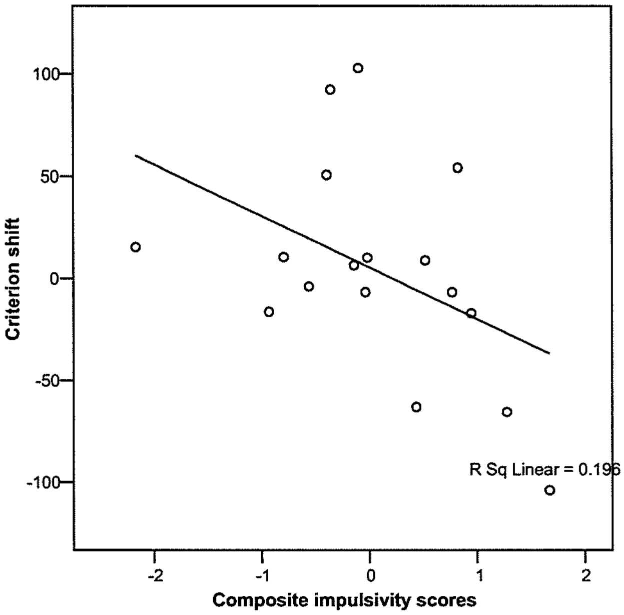


Figure 9.6. The scatterplot summarises the negative correlation between the composite impulsivity score scores [(ImpAss + DI)/2] and the criterion shift values (when the criterion shift score is equal to 0, the criterion was placed near the optimal accuracy criterion)

9.4.1.4.4 Correlations between criterion location measures, personality and the amount of points earned during the first phase of the task

A set of correlations was carried out to assess the relationship between the fitted criterion measures and the number of points earned during the two task conditions. This set of correlations did not identify any significant correlation between the decision criteria set during the asymmetric task and the number of points earned (p >.22). However, there was a trend for the correlation between impulsivity and the number of points earned during the symmetric task [r = .440, p = .08], which suggests that impulsive individuals may have been more accurate than their counterparts and, therefore, earned more points.

9.4.1.4.5 Non-learners

Results showed that non-learners had an accuracy level equal to 0.54 (.08) and 0.55 (.10) in the asymmetric and the symmetric task, respectively (cf. UD-users: asymmetric= .65, s.d. = .05; symmetric = .66, s.d. = .06]. The accuracy level of non-learners was significantly above chance in the symmetric task [t(14) = 2.21, p = .04] whereas in the asymmetric condition, accuracy was just a non-significant trend above chance [t(14) = 1.76), p = .10]. However, an independent t-test showed that non-learners achieved significantly lower accuracy scores than UD-users in the asymmetric [t(30) = -5.1, p<.001] and the symmetric task [t(30)= -3.7, p<.001]. Moreover, a chi-square test showed that non-learners earned significantly fewer tickets overall than learners ($\chi^2(2) = 12.42$, p = .002; table 9.10). None of the personality components was found to be correlated with the accuracy scores (ps >.2).

Table 9. 10. Number of tickets earned by learners vs. non-learners

		Tickets earned		
		0	1	2
Learners	**Yes**	1	8	8
	No	7	8	0

Overall, these results indicate that the use of an alternative strategy (cf. appropriate strategy) hindered performance compared to the use of the appropriate strategy, although it ensured accuracy levels above chance.

Finally, the results indicated that non-learners showed a greater response bias and higher noise levels than those participants who implemented the appropriate strategy (i.e. learners; table 9.11).

Table 9.11Criteria and noise levels shown by those participants who implemented a uni-dimensional strategy (UD-users; N = 17) and those who implemented an alternative strategy (non-UD users; N = 15) on the asymmetric (a) and the symmetric (s) conditions

	Criterion a	Noise a	Criterion s	Noise s
Non-UD users	355.22 (66.8)	278.93 (57.3)	323.52 (93.8)	225.52 (96.7)
UD users	340.38 (30.9)	65.68 (46.8)	336.27 (37.4)	77.86 (69.8)

The learning pattern displayed by non-learners resembles the learning pattern (i.e. flat learning curves) simulated using the purely implicit model (see chapter 8). These observations suggest that non-learners may be more strongly relying on the implicit system, rather than the explicit module, to solve the categorisation problem. Therefore, non-learners might be using an alternative strategy that relies on procedural learning alone which would, subsequently, account for the strong response bias (and lower accuracy) observed in this sub-sample. Indeed, simulations run with low rulebias values (e.g. 0.3), which reduced the impact of the explicit module on learning, produced a strong response bias, high noise levels and flat learning curves (see chapter 8).

9.4.1.5 Reversal learning phase

9.4.1.5.1 Model fitting

During the reversal learning phase, participants performed on the same tasks they had performed on during the first phase, although they had to learn a new strategy to perform optimally on the tasks. In fact, in the symmetric task, categories and responses were simply swapped (intradimensional shift) while the payoff structure stayed the same; whereas in the asymmetric task the high and low payoff categories were simply reversed (i.e. payoff shift) while the categories' location stayed the same. Hence, the same two formal models were fitted to the data, i.e. a uni-dimensional model which used location as the relevant dimension and a guessing model. As in the previous section, the two formal models were run to fit the data collected in the two task conditions.

1 – Fitting through a guessing model

Results showed that the guessing model was not significantly worse than the saturated model for 9 and 5 cases in the asymmetric and the symmetric task conditions, respectively.

2 – Fitting through a uni-dimensional rule model based on spatial location

Results indicated that during the asymmetric task, the 1-dimensional model was not significantly worse than the saturated one for 20 cases out of the total 32. Fifteen of these 20 cases had learnt the appropriate rule during the first half of the task. During the symmetric task, the 1-dimensional model was not significantly worse than the saturated model for 9 cases. Seven of these 9 cases had previously used the appropriate rule in the first phase of the task and, subsequently, showed the appropriate rule shift. Moreover, formal modelling was implemented to assess whether some participants were using a non-switch dimensional rule. Results showed that for six participants the non-switch dimensional model was non-significantly worse than the saturated model. Thus, these six participants had used the appropriate strategy in the learning phase and continued using it during the reversal phase. However, five of the six participants were able to implement the appropriate rule-switch strategy in the asymmetric task condition. Despite the fact that they had failed to show the appropriate rule-switch in the symmetric task, they were retained in the final sample since they were adopting a uni-dimensional rule.

9.4.1.5.2 Comparison of the two models

Results indicated that the uni-dimensional (UD) model was not significantly worse than the saturated model for 20 cases in the asymmetric task condition. However, 9 of the 20 fitted data sets collected in the asymmetric task were also fitted by the guessing model. In contrast, the UD model was not significantly worse than the saturated model for 15 cases in the symmetric task condition but three cases were also fitted by the guessing model.

It was decided to compare the two models in order to assess which one offered the best fitting for these participants. Comparisons were made by weighing the AIC values of the two models against each other, for those cases where the guessing and the UD model were both non-significantly worse than the saturated model. The model with the lowest AIC value was selected and identified as the best fitting (Dayton, 2003; Maddox et al., 2006). As in the previous section, results indicated that the UD model offered a better fit than the guessing model for all overlapping cases in the symmetric condition. However, four cases were found to be better fitted by the guessing model than the UD model for the asymmetric condition. Hence, the uni-dimensional model fitted 16 cases in the latter condition.

9.4.1.5.3 Model fitting - summary

Overall, the formal model indicated that the uni-dimensional model was non-significantly worse than the saturated model for 16 participants in the asymmetric task and for 15 participants in the symmetric task.

In the asymmetric condition, out of the 16 participants who implemented the appropriate strategy thirteen had used the appropriate uni-dimensional strategy in the learning phase of the task. In contrast, three participants who had not implemented the correct strategy in the learning phase were able to do so in the reversal phase (table 9.12).

Table 9. 12. Proportion of learners and non-learners across the two phases of the asymmetric task

		Learners reversal	
		Yes	**No**
Learners original learning	**Yes**	13	8
	No	3	8

It was decided to create a variable that coded whether the 15 cases fitted by the UD model also showed the appropriate switch in response bias during the reversal phase. The variable coded 1 all those criterion values that were below 325 and 0 those that were above this value (i.e. old response bias). Eleven of the 16 participants fitted by the uni-dimensional model showed the appropriate shift in bias (i.e. placed their criterion below 325) while 5 failed to show the shift. Moreover, results showed that among the 13 participants who had used a dimensional strategy in both learning phases of the asymmetric task, nine showed the appropriate shift in bias. In addition, all but one of the participants, who adopted the uni-dimensional rule only in the reversal phase, showed the appropriate (switched) response bias. These 16 cases showed an asymmetric criterion mean equal to 300.06 (55.17) and a noise level equal to 106.35 (85.01).

For the symmetric condition, the formal model identified 15 that were fitted non-significantly worse by the uni-dimensional model than the saturated model. Thirteen of these cases had also used the appropriate strategy in the learning phase (table 9.13).

Table 9. 13. Proportion of learners and non-learners across the two phases of the symmetric task

	Learners reversal		
		Yes	**No**
Learners original learning	**Yes**	13	11
	No	2	6

Among these 15 participants, the criterion mean was equal to 332.65 (94.53) and the noise level was equal to 177.88 (99.12). The noise level was high and, following a data inspection, it was detected that seven cases that were well-fitted by the formal model ($p > .05$) had very high noise levels. Indeed, four had a noise level equal to 300 (i.e. the upper limit) whereas three cases had noise levels above 200 (i.e. 209.30, 235.73 and 298.90). Once individuals with such high noise levels were removed, the retained sample ($N = 8$) showed a symmetric criterion equal to 336.22 (44.24) and a lower noise level (mean = 100.88, s.d. = 59.56).

Finally, analysis showed that 7 cases applied the appropriate uni-dimensional rule in both task conditions. They showed a criterion equal to 300.21 (s.d.= 48.23) and 330.70 (s.d.= 44.72) in the asymmetric and the symmetric condition, respectively. The noise level in the asymmetric condition is equal to 85.41 (s.d.= 67.82) and in the symmetric task 100.84 (s.d.= 64.33).

9.4.1.6 Preliminary analysis

9.4.1.6.1 Correlation between the strategy used and personality components

As done for the data collected during the learning phase, point-biserial correlations were run in order to assess whether there was a relationship between any of the personality components and the strategy used. The variable that codes the type of strategy implemented was labeled strategy used and coded as 1 the use of a uni-dimensional strategy and as 0 guessing or the use of other strategy. None of the correlations was found to be significant ($p > .20$).

9.4.1.6.2 Correlation between criterion location measures and the amount of points earned during the reversal learning phase of the task

As expected, the correlation results showed that there was a significant positive correlation between the amount of points earned in the asymmetric condition and the decision criterion implemented ($r = .63$, $p = .009$) for those 16 participants fitted by the uni-dimensional model. No significant correlation was found in the symmetric condition ($r = .21$, ns) for those 8 participants fitted by the formal model.

9.4.1.6.3 Individual differences and rule-switch during the symmetric task

After removing those participants who were guessing or using an alternative strategy in the second half of the symmetric task (total $N = 15$), correlations were run between the four personality components and the strategy used. A variable was created that coded whether participants had implemented the appropriate rule-shift ($N = 9$) or whether they persevered using the old rule ($N = 6$). This variable coded the appropriate switch as 1 and no switch as 2. During this learning phase participants had to learn to reverse their strategy by switching the previously learnt category-location association. Results showed that there was a negative correlation between extraversion and the rule applied ($r = -.54$, $p = 0.04$). Indeed, those individuals who showed the appropriate switch scored higher on the extraversion component (mean = 0.17 and s.d. = .89) than those participants who failed to apply the appropriate switch (mean = -1.33 and s.d. = 1.61).

The same extraversion by strategy correlation was observed when including only those 8 participants that had noise levels equal or below 200 ($r = -.72$, $p = .046$). This correlation indicates that extraverts were more likely to show the appropriate rule-switch than introverts. Therefore, it seems that the modulating impact of extraversion is strong enough to be detected in a small sample size despite the greatly reduced power.

9.4.1.6.4 Individual differences and rule-switch during the two learning phases

A variable was computed in the SPSS syntax, to code the strategy used by participants across the asymmetric task. The variable coded those participants who had used the appropriate uni-dimensional strategy in the first half of the task as 1, those who had used the appropriate strategy in both phases of the task as 2 and those who had not used the appropriate strategy at any stage as 0. Table 9.12 shows that, out of the 21 participants who had used the appropriate strategy in the first half the asymmetric task, 13 were able to show the appropriate shift in response bias whereas 8 failed to implement the appropriate shift. Correlations were run in order to establish whether there were individual differences between those participants who used the appropriate strategy in both learning phases and those who only used it in the first half of the task. Hence, those participants coded as 0 were excluded from the analysis. Results showed that there was no significant correlation between the strategy used and the any of the four personality components (Ps > .63). Hence, the present results suggest that personality did not mediate performance variations across the two phases of the asymmetric task.

A repeated-measure t-test showed that the shift in bias across the two learning phases of the asymmetric task was significant among those participants who had implemented the appropriate strategy in both phases [t(12) = 2.96, p = .012]. Indeed, participants showed a higher bias in the first half (mean = 337.67, s.d. = 29.26) and a reversed bias in the switch phase of the task (mean = 291.45, s.d. = 53.05). Subsequently, a shift score was calculated by subtracting the criterion scores in the second half of the task from the criterion cores in the first half of the task. Correlations were run between the shift scores and the four personality components. However, none of the personality components was found to correlate with the shift score (p >.30). No personality component enhanced the appropriate shift in bias from the learning phase to the reversal phase.

9.4.1.7 Behavioural data

It was decided to run some exploratory analysis only including those participants who implemented the optimal uni-dimensional strategy in both task conditions (N = 7).

9.4.1.7.1 Correlations between personality and decision criteria shown in the two task conditions

Preliminary correlations were run to assess the relationship between the four extracted personality components and the criterion location measures as well as the criterion shift score. The criterion shift score was calculated by subtracting the criterion in the asymmetric condition from the criterion in the symmetric condition, since participants are expected to place their criterion below the optimal accuracy bounds (325); thus, the shift scores should on average be positive.

Results showed that there was a trend for the correlation between neuroticism and the criterion location set in the asymmetric task (r = .69, p = .086). This correlation suggests that there is a trend for stable, compared to neurotic, participants to show the appropriate shift in bias.

Owing to the small sample size, it was not possible to further explore the data from the reversal phase. However, it was decided to explore how performance varied in the task conditions across the two learning phases. Hence, a repeated measure ANOVA was run with learning phase and task condition entered as repeated measure factors with two levels each (pre vs. post-shift and asymmetric vs. symmetric). The ANOVA is run on the 17 participants who were found to use the appropriate uni-dimensional strategy in both tasks during the pre-shift phase of the task. Results showed that there was a significant main effect for learning phase [F(1,16) = 13.00, p = .002] while there was only a trend for the main effect of task condition [F(1,16) = 3.10, p = .097]. Moreover, there was a significant interaction between task phase and task condition [F(1,16) = 6.50, p = .021], which suggests that criterion location varied across the task phases in relation to the task condition (figure, 9.7).

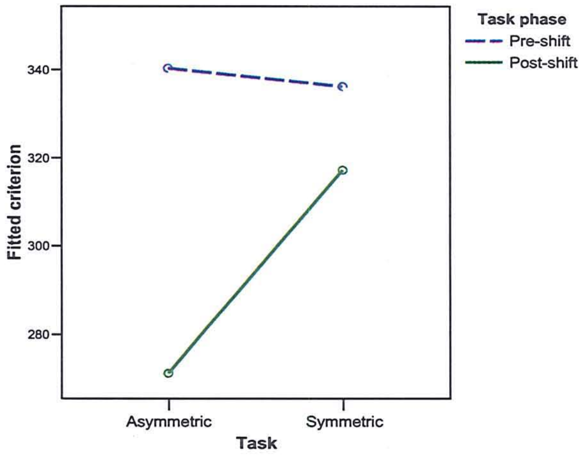


Figure 9. 7. Criterion location in the two task conditions across the learning phases

A paired sample t-test indicated that, among these 17 participants, the criterion shift across the two learning phases was statistically significant in the asymmetric task [t(16) = 3.89, p = .001]. In the post-shift phase participants placed their criterion around 271.37 (64.43) and showed a noise level equal to 162.12 (116.82). The noise level is higher than in the pre-shift phase since 5 of the participants included were not significantly fitted by the uni-dimensional model and they these cases had noise levels equal to 300.

In contrast, the criterion location set in the symmetric task was not statistically different across the two task conditions [t(16) = 1.44, ns]. In the post-shift phase, the 17 participants placed their criterion around 317.23 (59.73) and showed a noise level equal to 202.52 (107.41). The reason why on average these participants showed a high noise level is due to the fact that 6 cases were not fitted by the formal model in the post-shift phase of learning and had noise levels equal to 300.

A 2*2 ANCOVA was conducted in order to assess whether decision criterion was affected by individual differences by entering the personality components as covariates. Due to the small sample (N = 17), each personality factor was entered separately. Results showed that there was a non-significant trend for the correlation between ImpAss and task condition [F(1,15) = 4.17, p = .059]. Moreover, there was a significant main effect of phase [F(1,15) = 11.83, p = .004] and the 2-way interaction between task condition and phase was also significant [F(1,15) = 6.45, p = .023]

A median-split technique was applied to the ImpAss component and the binary-coded factor was entered in the ANOVA as a between-Ss factor. Results showed that the task*ImpAss interaction was still a non-significant trend [F(1,15) = 4.01, p = .064]. This interaction is graphically represented in figure 9.8.

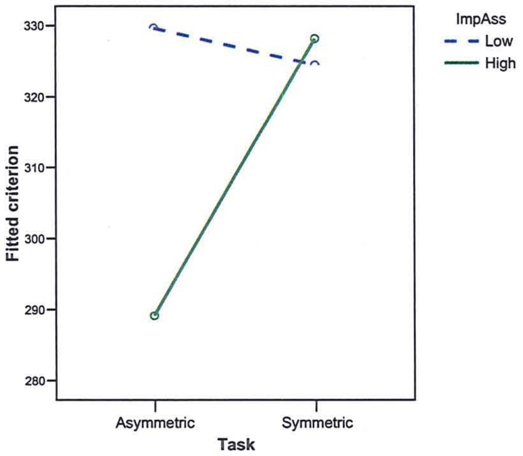


Figure 9.8. Criterion location across the two task conditions for high and low impulsive individuals

9.4.1.7.2 Non-learners

Results showed that non-learners had an accuracy level equal to 0.58 (.08) and 0.54 (.07) in the asymmetric and the symmetric task, respectively. The accuracy level was above chance levels in the asymmetric task [t(24) = 5.13, p <.001] and the symmetric task [t(24) = 2.89, p =.008]. Correlations were run to establish any possible relationship between accuracy scores and the personality components. There was a non-significant trend for accuracy scores in the asymmetric task to positively correlate with functional impulsivity (r = .39, p = .056) and extraversion (r = .38, p = .064).

Non-learners showed the appropriate switch in bias during the asymmetric task (mean = 303.38 and s.d. = 91.3) while they placed their criterion close to the optimal accuracy bound during the symmetric task (mean = 325.03 and s.d. = 82.1). Correlations were run between the personality components and the criteria and they showed a trend for the correlation between functional impulsivity (FI) and the criterion set in the asymmetric task (r = -.35, p = .08).

The correlation results also showed that FI was positively correlated with extraversion (r = .47, p = .02) and negatively correlated with neuroticism (r = -.50, p = .01).

Overall results from both learning phases indicated that the use of an alternative strategy allowed participants to show accuracy levels above chance and show the appropriate response bias, which maximised winnings, in the asymmetric condition. This suggests that those participants who did not learn to implement the optimal strategy must have used an alternative and reasonably effective strategy. This might have been mostly dependent on the implicit system and, therefore, not captured by the formal modelling (following earlier considerations).

9.4.2 Neural model simulations

Owing to the fact that participants had experienced an extensive practice session which consisted of 160 trials in total, 80 trials from each task (i.e. bars and dots), the simulations included the practice trials. This was done to account for the training effect of the practice trials on the experimental session. Indeed, learning was expected to occur over the practice trials when participants should have been able to develop the appropriate strategy to solve the task. Learning during the practice session should be reflected by changes in the learning weights in the neural model.

Simulations were run separately for each payoff condition (asymmetric vs. symmetric). Thus, the simulations were run across 240 trials which included the 160 trials of the experimental task and 80 practice trials. The practice trials corresponded to the task in the experimental session (bars vs. dots). The stimulus sequence used in the simulations corresponded to the sequence experienced by the actual participants during the experiment.

A parameter set that was able to simulate the behavioural data was identified by trial-and-error which was able to fit those parameters to the requirements of the present task. In order to successfully simulate the behavioural data the **scaler** parameter was set equal to 2, **rulebias** was set equal to 0.65 and the reward prediction signal (**rptonic**) was set equal to 0.5. The level of perceptual noise (**pnoise**) was set equal to 50 as for the simulation reported in the previous study.

During the practice trials, participants received accuracy feedback informing whether their responses were correct or incorrect. The accuracy feedback signals were coded into **rf** values equal to 1 for correct responses and 0 for incorrect responses. In contrast, during the experimental task the **rf** values were coded as 1, 0.75, 0.5 and 0.25 which represent a linear transformation of the actual reward values, i.e. 400, 300, 200 and 100, respectively.

The **rf** multiplier was set equal to 0.5 (i.e. rf = rf.*0.5) throughout the simulations. Thus, the actual **rf** values were equal to 0.5, 0.375, 0.25 and 0.125 (for 400, 300, 200 and 100 points, respectively) in the experimental task and 0.5 and 0 (for correct and incorrect responses, respectively) in the practice trials. The parameter values used in the simulations reported below are summarised in table 9.14.

Table 9. 14. Parameters' values implemented in the simulations

Symbol	Description	Value
rptonic	**Reward prediction tonic value**	0.5
rulebias	**Strength of explicit module**	0.65
m	**Scales the rf values**	0.5
Pnoise	**Perceptual and criterial noise**	50
b_1 and b_2	**Learning rates of the RP unit**	0.05*scaler1
b_3 and b_4	**Learning rates of the SR unit**	0.05*scaler2
Scaler1	**It scales the learning rate of RP unit**	2
Scaler2	**It scales the learning rate of RP unit**	2

The parameter values reported in table 9.14 were used at the start of the simulations (i.e. first trial of the practice session) and the learning weights were updated following each trial. Simulations of the experimental session would, therefore, implement the updated learning weights which should reflect the learning occurred over trials (as expected to occur in the human data).

9.4.2.1 Preliminary simulations of the behavioural data

Preliminary simulations were run to check that the model and the parameter set could capture criterion and noise mean values close to the ones observed in the empirical data. The simulated criteria and noise levels for the asymmetric and symmetric conditions are reported in table 9.15 together with the fitted criteria shown by human participants.

Table 9. 15. Simulated and actual (human) criteria and noise levels obtained in the asymmetric and the symmetric condition (Crit = criterion; a = asymmetric, s = symmetric)

	Crit_a	**Noise_a**	**Crit_s**	**Noise_s**
Simulated (N = 300)	339.59 (13.7)	85.62 (19.2)	320.32 (13.3)	73.81 (14.7)
Human (N = 17)	340.38 (30.9)	65.78 (46.8)	336.27 (37.4)	77.86 (69.8)

The results reported in table 9.15 show that the neural model was able broadly to capture the mean of the human data.

9.4.2.2 Preliminary analyses to simulate the relationship between behaviour and individual differences

Simulations were separately run for the two payoff sub-conditions (i.e. asymmetric vs. symmetric) and the parameter set presented above was kept constant across all the simulations reported in the following sections. These initial simulations capture a BAS variation that ranges uniformly from 0 to 2.

9.4.2.3 Individual differences on the reward cell

The correlation between the criterion and the personality variance was non-significant in the asymmetric condition task (r = 0.08, ns; figure 9.9). Once again there was a clear tendency for non-linearity in the relationship, as observed in chapter 8.

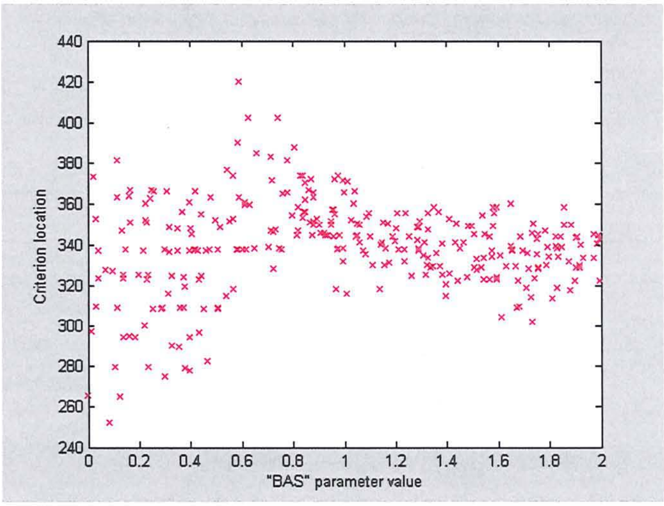


Figure 9.9. Scatterplot of the asymmetric condition for 300 simulated subjects with random uniform BAS parameter variation

In contrast, there was a weak correlation in the symmetric condition (r = .17, p = .003). As observed in chapter 8, the relationship between the BAS variation and the fitted criterion is curvilinear. Indeed, the relationship between the BAS and performance is negative for those 'individuals' with BAS scores between 0.7 and 1.2, whereas it is flat above 1.2 and positive for those with BAS scores below 0.7.

Thus, the overall positive correlation presumably derives from the fact that there are slightly higher criteria observed for those individuals with "BAS" parameter values below 0.5 (figure 9.10)

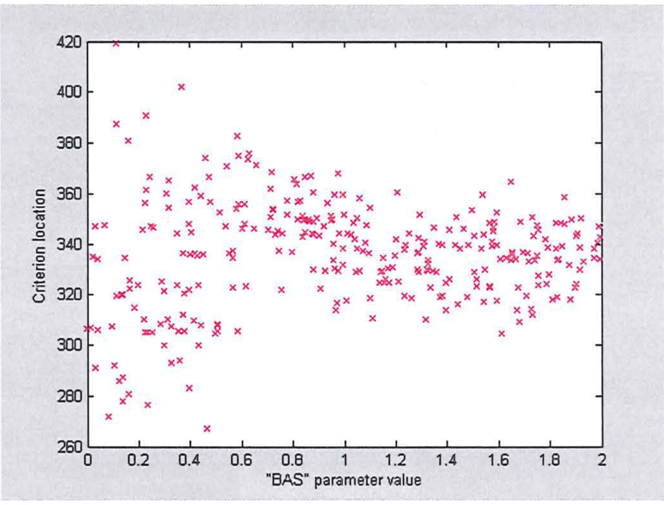


Figure 9.10. Scatterplot for the symmetric condition for 300 simulated subjects with random uniform BAS parameter variation

9.4.2.4 Individual differences on the RPE signal acting on the actor synapses of the SR cell

Simulations of the asymmetric condition showed that criterion was positively correlated with personality ($r = .64$, $p < .001$). Instead, there was no significant correlation in the symmetric condition ($r = .07$, ns). Figure 9.11 visually represents these relationships.

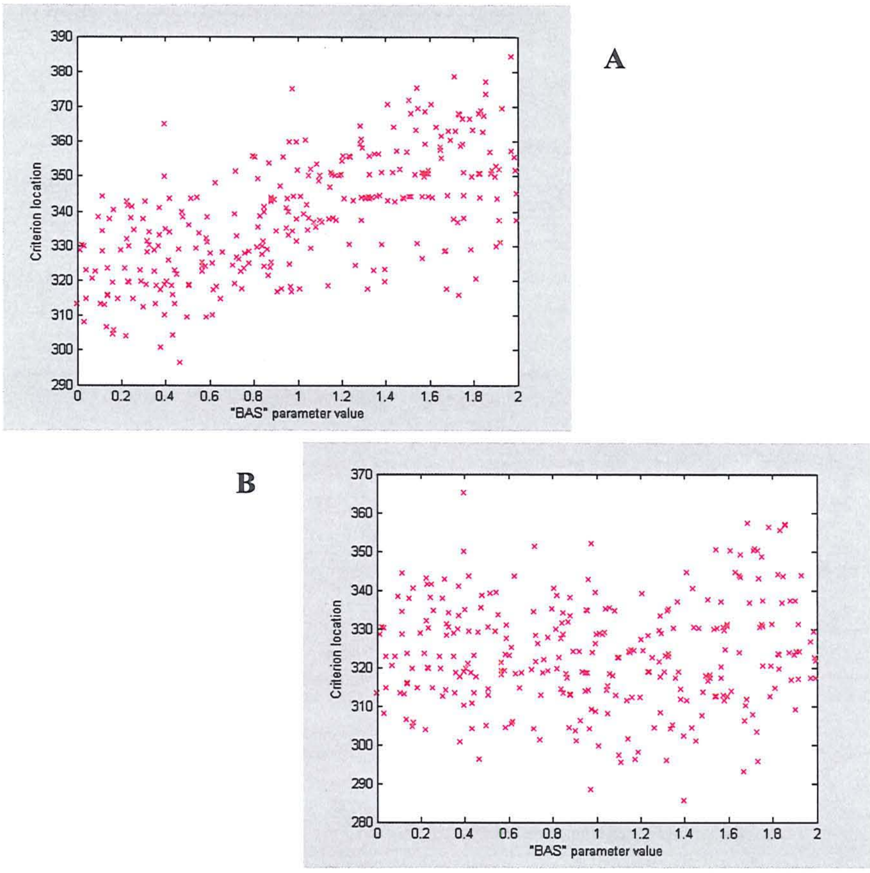


Figure 9.11. Scatterplot for the BAS*trait relationship in the asymmetric condition (figure A) and the symmetric condition (figure B), for 300 simulated subjects with random uniform BAS parameter variation

9.4.2.5 Individual differences on the RPE signal acting on the synapses of the critic (RP) cell

In line with the results obtained in the previous chapter, the simulations showed that there was no significant correlation between BAS variation and the criterion placement under the asymmetric condition (r = -.08, ns) or the symmetric condition (r = .02, ns).

9.4.2.6 Individual personality simulations

The preliminary simulations above indicate that the model with personality acting as a multiplier on the RPE signals impacting on the SR cell is the best candidate to simulate and explain the human data. These results are in line with the results and observations in chapter 8. Thus, further simulations were run using this model to simulate the human data with individual values of personality variance. Simulations were only run to simulate the data from the learning phase since, due to the small sample size, the behaviour data from the reversal phase were inconclusive.

9.4.2.7 Personality on the SR cell

Taking figure 9.11 into consideration, it seems that it should be possible to capture criterion and noise levels close to the ones observed in the human data by considering a normal distribution of the BAS variance with a mean value of 1 and a standard deviation of 0.25. These values should also be able to simulate the trait*criterion relationship, observed in the asymmetric condition.

Three hundred and forty simulations were carried out and then sub-divided into twenty groups of seventeen 'participants', in order to obtain groups with the same samples size of the human data analysed in the learning phase.

In the overall sample (N = 340), the correlation between personality and criterion was significant (r = .51, p <.001; figure 9.12).

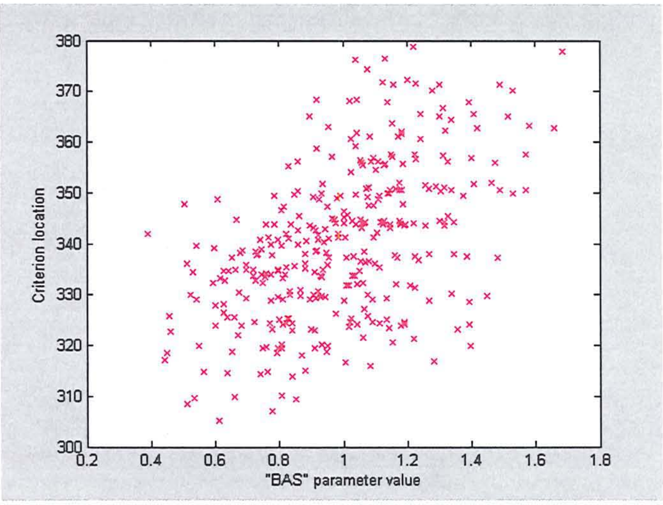


Figure 9.12. Scatterplot of the relationship between personality and criterion in the asymmetric condition (N = 340)

Nonetheless, the results in the twenty sub-groups are of greater interest since they consist of the same sample size as the human data and, therefore, have similar power levels. Across the 20 sub-groups, the correlation coefficients ranged between 0.22 and 0.82, with a mean value equal to 0.5 (s.d. = 0.17).

The correlation was significant for twelve of the twenty sub-groups with correlation coefficients ranging from 0.47 to 0.82 (mean = 0.6, s.d. = 0.11). In three of the sub-groups, the correlation showed a non-significant trend [group 1: r = 0.45, p =0.07; groups 2: r = 0.42, p =0.09; group3: r = 0.42, p = .101]. However, the correlation was non-significant in five of the sub-groups. Overall, these results indicate that despite the small sample size, it might be possible to detect the correlation between the trait parameter and the criterion in an empirical study. Additionally, the criterion and noise mean values were close to the scores in the human data. Indeed, the criterion mean ranged between 333.13 (13.61) and 345.21 (16.21) and the noise level ranged between 69.74 (26.52) and 96.03 (23.72). The overall criterion and noise level values are reported in table 9.16 (together with the actual scores).

Table 9. 16. Simulated and actual criterion in the asymmetric condition

	Criterion	**Noise**
Simulated (N= 340)	340.14 (14.9)	**85.58 (25.4)**
Human (N = 17)	**340.38 (30.9)**	**65.78 (46.8)**

These results across the sub-groups support the findings obtained in the empirical data.

9.5 Discussion

The present study replicated the behavioural data obtained in the study reported in chapter 8 and the results obtained by Markman and colleagues (2005). Indeed, under asymmetric payoffs participants showed a tendency to place their decision criterion away from the optimal accuracy criterion towards the optimal reward one. Moreover, the study showed that the response bias was mediated by individual differences and, more specifically, impulsivity was found to be negatively associated with the bias. These results suggest that impulsivity is the underlying personality trait of the BAS system, at least if one assumes that the response bias in this task is a BAS-mediated phenomenon. The model simulations are a weak test of the latter assumption since the model simulates a positive relationship between response bias and the parameters within the biologically-inspired model, which is built so as to reflect probable features of the BAS.

Results further identified the presence of a trend that indicated that high scores on the neuroticism (N) component were positively associated with a shift in bias across the two conditions. In particular, individuals with high scores on the N component showed a response bias in the asymmetric condition and no bias in the symmetric condition. The opposite pattern was observed for low N individuals. Since the two tasks were intertwined, it is possible that some participants may have not been able to distinguish which task offered symmetric payoffs. Thus, it is possible that low N participants were less likely to identify which task had an asymmetric payoff matrix. Neuroticism has been identified as the underlying personality trait of the conflict resolution system (i.e. BIS; Corr, 2006) and, therefore, these results indicate that the BIS system became active during performance on the task. Indeed, feedback signals that indicated participants had earned fewer points than 400 may have been perceived as 'punishing' which would have activated the FFFS system. Simultaneous activation of the BAS and the FFFS could have caused the activation of the BIS system in order to solve the conflict. Individuals with a more reactive BIS system (high N) may have been more sensitive to the emotional/motivational factors (i.e. feedback signals) and this could have enhanced learning, altering criterion placement.

In the study reported in the previous chapter, neuroticism was not found to mediate performance during the task. Hence, it is possible that in the previous task receiving a message about the amount of potential winnings, together with the amount of actual points won, may have been perceived as more rewarding. However, this does not seem especially plausible and goes against the opposing reasoning advanced earlier; that reasoning led to the change in procedures used in the present study. Nonetheless, the payoff structure implemented in the present study (together with the intertwined design of the task) may have rendered it more difficult to identify which task was the asymmetric task. It is possible that having presented participants with information on both the actual and the potential points might have helped participants to distinguish between the two tasks.

Simulations run using the neural model were able to capture the relationship between a plausible BAS parameter (i.e. simulated underlying biology) and performance and, therefore, they validate to some extent the results obtained in the present study. The results showed that impulsivity (i.e. a BAS-trait candidate) mediated the response bias in the asymmetric task. The simulations also offer further insight into the empirical data. In fact, the human data was best simulated by the sub-model which has 'BAS-variation' acting on the RPE projections affecting the synapses on the stimulus-response actor cell. Thus these results suggest that the BAS may be characterised by responsivity to the RPE signals rather than to reward *per se*. In particular, the data seem to indicate that individuals who score low on impulsivity are more reactive to RPE signals (i.e. which might be described as high BAS-sensitivity). This was indicated by the fact that low impulsives showed greater response biases than their high-scoring counterparts. In other words, these results seem to indicate that high BAS individuals show the greater learning due to high sensitivity to the reward prediction signal which guides SR learning. Thus, consistent with some biological accounts of RST, BAS reactivity is associated with greater DA receptor reactivity to reward-related (in this case, RPE) signals. However, contrary to RST, this sensitivity is in response to RPE rather than reward *per se* and it is related to low (cf. high) scores on impulsivity.

The neural model is a reward-only model and, therefore, it cannot capture the modulatory effect of neuroticism observed in the human data, assuming that the neuroticism effect is related to perceived losses during the task. However, the neuroticism effect observed in the human data may have hindered the BAS*behaviour relationship from being observed in the empirical data.

The simulations produced reasonably strong associations (i.e. correlations of about 0.5). However, the corresponding real personality*behaviour correlations will have also been weakened by the series of factors previously mentioned (see chapter 8). In fact, the neural model simulates the relationship between the biology (i.e. effects of midbrain DA activity) and behaviour, which is usually stronger than the association between personality and behaviour (Munafo et al., 2003; Cohen et al., 2005). Finally, empirical studies also include external variations (not present in the simulations) that add noise and variance to the data and, subsequently, it may further weaken the phenomenon under study. This implies that the real personality*behaviour correlations would be expected to be considerably weaker than the simulated biology*behaviour correlation of around 0.5, although it was still possible to observe a significant relationship in a sample of only 17 participants.

One of the main limitations to drawing conclusions from the present study was the small sample size, and low retention, as many participants failed to learn to perform on the task. In particular, the data from the reversal learning phase indicated that only a very small number of participants learnt to perform on the task by applying the appropriate switched strategies. Thus, the data from this phase are inconclusive due to small sample size (i.e. low power). This does not allow one to fully explore the dataset to detect any possible relationship between personality and behaviour (i.e. response bias).

The data showed that individuals with low scores on the positive schizotypy (PS) component were significantly more likely to use the appropriate uni-dimensional strategy (cf. guessing/other dimensional) than their high-scoring counterparts during the learning phase. Owing to the overlap between the ImpAss and the PS components, these results are in line with the findings obtained by Pickering (2004) who showed that impulsivity measures (e.g. EPQ-P) were positively correlated with performance on a rule-based CL task with stimuli varying on continuous dimensions. Moreover, current results showed that individuals who scored high on the impulsivity component showed a non-significant trend to earn more points in the symmetric condition than their low counterparts in the symmetric condition. These results suggest that high scorers on the ImpAss component led to superior performance (i.e. greater accuracy levels) and these conclusions are in line with previous results obtained by Pickering (2004) and Tharp (2007) who found that high ImpAss individuals perform better than low ImpAss individuals during simple rule-based tasks. These researchers concluded that ImpAss is associated with efficient performance during uni-dimensional RB tasks but, on the contrary, they show inefficient performance during more complex (i.e. conjunctive) tasks and/or rule switch learning. Tharp suggested that this is due to generally low cognitive flexibility typical of high ImpAss individuals.

In the reversal phase, extraversion was found to significantly enhance learning on the symmetric payoff task condition. These results suggest that extraverts were able to switch their strategy once contingencies on the task had changed. These results seem to be counterintuitive to the literature that indicates that extraverts are impaired at inhibiting a dominant response (previously associated with reward) due to their hypersensitivity to reward (Patterson et al., 1987; Avila, 2001). However, the present findings support the view that enhanced rule-switching behaviour is due to greater cognitive flexibility (Pickering, 2004; Tharp, 2007). As previously mentioned, there is a great overlap between the regulatory focus theory and RST. For example, a high functioning BAS is characterised, just like the promotion focus, by a greater sensitivity to reward and a tendency to approach potential reward-related goals. Hence, it is possible that the greater cognitive flexibility shown by extraverts in the present study was determined by the regulatory fit between the experimental manipulation and the participant's chronic 'promotion' focus (i.e. high BAS individuals). Maddox and colleagues (2006) have suggested that regulatory fit leads to greater cognitive flexibility, which could have facilitated rule-switching in the present study. These results are in line with a study conducted by Tharp (2007) where extraversion was found to facilitate learning during a conjunctive RB task with a gain-only payoff matrix. The regulatory fit theory may account for the fact that the present results do not replicate those obtained by Patterson and colleagues (1993) and Avila's that used a mixed-payoff matrix and, therefore, did not lead to a regulatory fit.

Overall, the results are not conclusive, due to the small sample size and the small size effect (i.e. response bias). Indeed, just above half of the participants implemented the appropriate dimensional rule and performed above chance. This may be due to the increased difficulty added to the task by its intertwined structure which may have required additional working memory effort. Similarly, the small number of trials may have not allowed participants to show the appropriate response bias to maximise winnings. Moreover, the small effect size may have been due to the task pre-exposure during the practice trials. This should have encouraged participants to use the appropriate uni-dimensional strategy and placed their criterion close to the optimal accuracy criterion as accuracy feedback is symmetric across all categories. Originally, the practice session had been included to ensure the activation of the explicit system during the experimental categorisation task. In this way, it would have been possible to explore the activation of the implicit system over trials and how it mediated performance (i.e. response bias) under asymmetric payoffs. However, the practice trials may have strengthened the activation of the explicit system while dampening the activation of the implicit system and, subsequently, led to a smaller response bias. Hence, these results may suggest that practice trials may have been detrimental to exploring the behaviour under evaluation (i.e. response bias); nonetheless, inclusion of these trials ensures that it is the activation of the implicit system that leads to the response bias. However, future studies should probably not include such an extensive practice in order to obtain a greater effect size.

Following these considerations, it seems necessary to replicate the present behavioural and psychometric findings in a bigger sample. In order to ensure learning and, therefore, greater retention of datasets for the analysis, it is necessary to render the task simpler. Owing to the results obtained in this and the previous chapter, inclusion of two task conditions (asymmetric vs. symmetric) may hinder performance: the results may have been affected by either order effects and/or overly-complicated task structure (i.e. the simultaneous requirement to perform the two conditions). Therefore, it may be necessary to have participants perform only on the asymmetric task, since it has already been established that an asymmetric payoff matrix leads to a response bias in comparison to a symmetric payoff structure.

A follow-up study would also allow us to test the validity and efficiency of the neural model at predicting human behaviour on a similar categorical task as a result of the biology-behaviour-trait relationship. The follow-up study reported in the next chapter aims to further replicate the results obtained in this chapter as well as further testing of the neural model. The task explores how categorisation is mediated by a variety of payoffs of different magnitudes.

The current model suggests that learning in the 3-factor rule is highly dependent on the phasic firing of the mesolimbic and the nigrostriatal DA cells and their projections into the caudate, which reinforce the active and correct SR synapses. BAS functioning is also believed to vary depending on the DA activity of these systems (Pickering and Gray, 2001), therefore individual differences in BAS functioning should affect the rate of learning (i.e. the formation of SR associations). Hence, the general assumption is that individuals with a highly reactive BAS should learn SR associations in a more efficient way (through faster development of stronger synaptic connections) than those with a less reactive BAS. Learning relies on the reward prediction signals and the magnitude of the response they induce at their target synapses. The simulations in the present chapter suggest that the BAS is responsive to RPE signals (cf. reward *per se)* and, therefore, it is possible that BAS variations may affect learning as a function of RPE magnitude. The following study attempts to address this issue more directly by manipulating RPEs over a wider range of values.

Chapter 10

Motivation and personality in response to several payoffs varying in magnitude

10.1 Abstract

The present study aimed to extend the findings obtained in the previous chapter. The study explored how individual differences mediated decision bound placement under an asymmetric matrix with several payoff magnitudes. Additionally, it aimed to assess the efficacy of the neural model at simulating human data. Results showed that impulsivity mediated learning and, in particular, that participants who scored low on impulsivity showed greater criterion shifts than their high-scoring counterparts. Consistent with the results of the previous chapter, the neural model results suggested that learning was dependent on the dopaminergic reward prediction error (RPE) signals that are responsible for the formation of SR association (according to the 3-factor rule; Pickering and Gray, 2001). Moreover, RST postulates that DAergic activity is the substrate of the behavioural activation system (BAS; Gray, 1987). Therefore, the present results suggest that BAS is responsive to DA-RPE signals rather than to reward *per se* as originally postulated by RST.

10.2 Introduction

Data from the two previous studies indicated that participants tend to over-classify probabilistic stimuli as members of the high payoff category during an RB category learning task with an asymmetric payoff matrix. Moreover, the previous studies indicate that this response bias was mediated by individual differences. In particular, the study reported in chapter 9 identified impulsivity as the personality trait that affects the magnitude of the bias. Contrary, to RST prediction it seems that high BAS activation (of processes responsible for the bias) is associated with low scores on impulsivity. However, due to the small sample size included in the analysis the study needs replication. The present study aims to replicate these findings.

The neural model simulations offered a further insight into the human data. In fact, simulation results indicated that both the explicit system and the implicit system were active during performance on the simple RB CL task and they mediate different aspects of learning. According to COVIS predictions, the explicit system was involved in solving the task in general terms by applying a simple and easily verbalised uni-dimensional rule, which was probably developed fairly early on during the practice trials. In contrast, the implicit system was found to mediate SR learning and, therefore, be responsible for the response bias observed under an asymmetric payoff matrix. In other words, the explicit system determines a boundary that is needed, to solve the categorisation problem, but sets it at a perceptually symmetric position; the implicit system, by contrast, gradually adjusts that boundary via reward-based learning in a biased fashion which helps to maximise winnings under asymmetric payoffs. The simulations showed that BAS variation could lead to individual differences on this task by affecting SR learning, which is guided by RPE-DA signals. In other words, BAS variability might reflect responsivity to RPE-DA signals (cf. responsivity to reward *per se*). In particular, the model showed that simulated high BAS individuals were more responsive to RPE signals and, therefore, showed faster and greater SR learning, as indexed by the response bias.

Therefore, high BAS activation is associated with increased effectiveness of the RPE-DA projections from the SNpc to the medium spiny neurons in the striatum which, subsequently, strengthen the appropriate SR association in corticostriatal synapses. Indeed, DA firing has been identified as one of three fundamental factors involved in SR learning, according to the 3-factor learning rule (Schultz, 1998; Pickering and Gray, 2001). This model accounts for DA-driven synaptic plasticity in the striatum (Pickering and Gray, 2001; Schultz, 2002; Corr 2006; see chapter 2). The COVIS model also assumes that this same 3-factor model is the substrate of the implicit system (Pickering, 2004).

Shohamy and colleagues (2008) have reviewed behavioural, computational and neuro-imaging data on probabilistic learning (in particular, the so-called 'weather-prediction' task) and have concluded that optimal learning improves over trials. More specifically, there is evidence that indicates that performance seems to be initially driven by the explicit system (i.e. MTL), which implements simple strategies; whereas over time the implicit system becomes more dominant as the appropriate SR associations are strengthened. In particular, the evidence reviewed by these authors indicates that incremental learning was driven by the basal ganglia and it was dependent on strengthening of the SR association through a DAergic error-correcting, phasic signal (i.e. RPE signals). Therefore, these data support the simulation results obtained in the previous chapters that indicate that the two systems may be involved at different stages, and in different aspects, of learning.

One of the main aims of the study was to replicate and extend the data obtained in the previous studies. In order to replicate the results, several changes were applied to the task design in order to facilitate performance. Indeed, the tasks implemented in the previous two chapters proved quite difficult as indicated by the low proportion of participants performing above chance using the appropriate, dimensional strategy. Therefore, it was decide to lower the difficulty of the task and enhance learning by increasing the discriminability of the categories (d prime) used in the task. Stimuli were generated so that the task had a d prime equal to 2[18] (cf. d' = 1 in the previous studies). By doing this, it is expected that a greater proportion of participants should learn to perform using an appropriate rule in the task and, therefore, the final sample size should be of a reasonable size (i.e. power).

Additionally, the task implemented in the previous chapter may have been difficult because participants performed simultaneously on the two tasks, which were presented in an intertwined fashion. This may have proved demanding on working memory and hindered overall performance. However, running the two conditions separately raised order effect issues (chapter 8). Hence, it was

---

[18] The d prime (d') value in this task indicates the discriminability of 'adjacent' categories; see below

decided to lower the level of difficulty of the task by having participants perform only in the asymmetric payoff condition. There was no need to include the symmetric payoff matrix condition, since the two previous studies have already shown that asymmetric payoffs lead to a response bias in relation to symmetric payoffs.

Finally, in order to increase the likelihood of detecting the predicted association between behaviour and a 'trait' measure of BAS variation, it was decided to include a behavioural measure of impulsivity together with the usual battery of questionnaires. Following RST predictions and results from the previous study, trait impulsivity seems the best BAS-trait candidate and it should be the BAS measure that best predicts performance on the task. The decision to include a behavioural measure of impulsivity (i.e. an endophenotype) was guided by the fact that behavioural measures are expected to be more reliable indexes of individual differences than self-reported, psychometric measures of impulsivity (Lee, et al., 2007). Indeed, self-report measures are limited by response bias induced by participants' interpretation of the questionnaire items and by the fact that the self-report measure may not be capturing the facet of impulsivity assessed in the study (e.g. reward evaluation, or behavioural inhibition; Eisenberg et al., 2007). Indeed, Eisenberg and colleagues (2007) found that DA activity, as indexed by the DRD2 and the DRD4 genetic polymorphisms, was associated with greater delay discounting (i.e. impulsivity). In contrast, there were weak effects on questionnaires: a trend main effect of the genetic variance on the total SSS score, but no effect on any of the Barratt Impulsive Scale (BIS) or the Eysenk Impulsivity Questionnaire (EIQ) scales. Nonetheless, the delay discounting (DD) scores were positively correlated with the BIS motor impulsivity and EIQ impulsivity subscales. Hence, these results confirm that the personality measures indexed impulsivity but the behavioural measure of impulsivity (i.e. delay discounting) was most strongly correlated with the genetic variance (i.e. biological mechanism).

The behavioural index of impulsivity included in the present study was the delay discounting task, which is a standardised measure of impulsivity (Monterosso and Ainslie, 1999; Eveden, 1999; Dawe, 2004). Indeed, impulsivity is often operationally defined as the tendency to choose small but immediate rewards over larger but delayed rewards (this is delay discounting; Bickell et al., 1999; Field et al., 1999; Richards et al., 1999). High impulsive individuals tend to discount delayed rewards more than low impulsive individuals. There is empirical evidence that supports the claim that delay discounting is a behavioural measure of trait impulsivity (Eisenberg et al., 2007). Richards and colleagues (1999) have found that the hyperbolic discounting coefficient is positively correlated with standardised impulsivity measures (i.e. EPI-impulsivity and extraversion as well as SSS-disinhibition).

There is also evidence that indicates that the decision-making processes involved during performance on the delay discounting task resemble the processes involved in reward-related learning and, more importantly, relies on reward prediction error (Daw and Doya, 2006; Kable and Glimcher, 2007; Murray, et al., 2007; Kalenscher and Pennartz., 2008). Indeed, Kalenscher and Pennartz (2008) describe hyperbolic delay discounting as constant updating of reward values over trials following Pavlovian learning. Moreover, these authors suggest that with practice participants learn to predict future (expected) rewards in relation to reward-related cues. This reward representation is constantly updated and future, predicted rewards are constantly discounted throughout learning. If there is a mismatch between the actual and the predicted reward, a reward prediction error can be computed (equation 10.1)

$$\boldsymbol{\delta_t = r_t - \gamma V_t - V_{t-1}} \qquad \textbf{Eq.10. 1}$$

where $\boldsymbol{\delta_t}$ is the error term at time **t**, , and γ is the delay discounting term, $\mathbf{r_t}$ is the actual reward a time **t**, whereas $(\boldsymbol{\gamma V_t - V_{t-1}})$ represents the predicted reward. As in the operational learning formulae reported in chapter 2, the reward prediction error term acts as a reinforcement signal according to the 3-factor rule. Following these considerations, the authors proposed a computational model that explains delay discounting in terms of implicit learning. Indeed, they suggest that RPE-DA signals strengthen SR learning and, therefore, guide decision-making processes according to equation 10.2:

$$\Delta \mathbf{w}_{ij,t} = \mathbf{c}\ \delta_t\ \mathbf{a}_{i,t}\ \mathbf{a}_{j,t} \qquad \textbf{Eq.10. 2}$$

where $\Delta \mathbf{w}_{ij,t}$ denotes the strength of the synapse between the pre- (**j**) and post-synaptic (**i**) neuron and **c** is the learning constant. In other words, the computational model suggests that delay discounting is mediated by DA-firing changes in the striatum.

Probabilistic discounting, which is equivalent to delay discounting, has been found to be mediated by dopamine and, in particular, RPE-DA signals (Richards, et al., 1999; Wade et al., 2000; Pessiglione et al., 2006; Murray et al., 2007; Kalenscher et al., 2008). For example, Pessiglione and colleagues (2006) have observed that during performance on a probabilistic task, L-Dopa administration (cf. haloperidol) was associated with increased responding to high-probability stimuli under a gain-only matrix. In contrast, L-Dopa administration did not affect learning under a loss-only matrix. Additionally, increased RPE signals in the ventral striatum were observed following L-Dopa (cf. haloperidol) administration. RPE-DA signals can also control learning during a reward-mediated learning task of the type implemented in the present and the two previous chapters, as indicated by the neural model and by previous studies (Hollerman and Schults, 1998; Tobler et al., 2005; Abler et al., 2006; Juckel et al., 2006a). In particular, it has been found that reward-dependent learning is mediated by DA activity in the striatum primarily through D1 receptors (Beninger and Miller, 1998; Frank et al., 2007; Klein et al., 2007)

According to the evidence, it was expected that a behavioural measure of impulsivity such as delay discounting (i.e. an endophenotype for trait impulsivity) should be a more reliable and specific measure of impulsivity than a self-report personality questionnaire (i.e. the phenotype). Thus, it should be more strongly correlated with behavioural responding on the CL task (i.e. bias) than self-reported psychometric measures of impulsivity (Eisenberg et al., 2007; Lee, et al., 2007). There is extensive evidence that shows that the specific biological markers (such as those simulated in the model) are correlated relatively weakly to individual differences on trait measures (explaining at most 10% of the variance; Bau and Salzano, 1995; Blum et al., 2006; Reuter et al., 2006). Subsequently, it can be assumed that there is a weak association between the personality and task behaviour as a direct result of their mutual association with the underlying biology. Indeed, the impulsivity*task correlations observed in the human data reported in the previous chapters were quite weak, despite the fact that the simulated behaviour*biology correlation was strong (up to r = .70). Therefore, it was assumed that a behavioural measure of impulsivity might be closer to the (BAS-mediated) task behaviour assessed in the study.

In order to extend the findings from the previous study, the present study aims to explore the impact of several payoff magnitudes on learning. The previous studies show that participants showed a response bias when presented with two categories that offered asymmetric payoffs. In contrast, the present study explores how several payoffs of differing magnitude affect learning and, possibly, response bias.

RST suggests that learning according to the 3-factor model is highly dependent on the phasic firing of the mesolimbic and the nigrostriatal DA cells and their projections into the caudate that reinforce the active and correct SR synapse. BAS functioning is also believed to vary depending on the DA activity of these systems and, therefore, individual differences in BAS functioning should affect the rate of SR learning (Pickering and Gray, 2001). Hence, RST predicts that high BAS individuals should learn more efficiently than low BAS individuals. Learning relies on the reward prediction signals and their magnitude. The simulations from the previous chapter suggest that the BAS is responsive to RPE signals (cf. reward *per se*). Therefore, it is possible that BAS variations may affect learning as a function of RPE magnitude, which represents the teaching signal that guides learning. Following these considerations, high BAS individuals are expected to show superior learning than low BAS individuals thanks to their greater sensitivity to RPE signals.

The use of several payoffs varying in magnitude also offers the chance to test the efficacy of the neural model to simulate the human, behavioural data (i.e. response bias) in response to different payoffs and the impact of individual variation in BAS-reactivity to a variety of payoffs.

10.2.1 Study aims

The present study aims to extend the findings obtained in the previous studies and to assess how BAS variation mediates response bias in relation to a greater variety of payoff magnitudes. Moreover, the study aims to explore the impact of payoff matrix switch in the second half of the study. More specifically, the study explores the impact of switching from an asymmetric payoff matrix to a symmetric matrix, where correct responses earn either greater or smaller gains than any of the correct responses in the learning phase. Finally, the present study also has as a main goal to test the validity of the two individual differences sub-models and further assess which one is best at capturing the human data and, therefore, interpret the results.

10.3 Method

10.3.1 Participants

Participants were an opportunity sample recruited at Goldsmiths. The sample consisted of 60 participants, 18 males and 42 females (70% female). Fifty-four participants (i.e. 90% of the sample) were right-handed. All participants were aged between 18 and 39 (mean age = 22.9; s.d. = 4.9).

The majority of the sample (62%) was recruited through the research participation scheme run by the psychology department and they all received 4 course credits for their participation. The remaining 38% of the sample comprised students from other departments within Goldsmiths. These participants received £12 for their participation. Participants were guaranteed confidentiality. Testing consisted of one sitting that lasted approximately 2 hours.

10.3.2 Design

The aim of the present experiment was to assess the impact of payoffs of varying size during learning on a CL task. Therefore, an asymmetric payoff matrix was used where payoffs varied in magnitude.

In particular, the study was interested in assessing the impact of reward magnitude on the decision criteria set by participants during a CL task that required participants to categories stimuli into four categories. The task consisted of two phases: a learning phase and the subsequent "payoff-shift" phase. In the learning phase, the primary independent variable was payoff matrix which had four within-subjects levels that characterised the magnitude of the payoffs earned for correct responses (i.e. highest, high, low and lowest). The different payoff values offered by the four different categories are summarised in table 10.1.

Table 10.1. Payoff values used for correct and incorrect responses in the two learning phases of the task

Phase->	**Learning**	**Shift condition highest**	**Shift condition lowest**	**All phases**
Payoff ↓	**Correct**	**Correct**	**Correct**	**Incorrect**
Highest	400	467	133	100
High	333	467	133	100
Low	267	467	133	100
Lowest	200	467	133	100

In the payoff-switch phase, the payoff matrix was altered in such way that correct responses to any of the four categories offered the same amount of points (i.e. symmetric payoffs). The main IV in the second half of the task was payoff structure and it was run as a between-subjects factor with 2 levels (i.e. higher vs. lower). Indeed, one matrix condition offered more points for correct responses than any of the categories in the first task (467 points) while the other matrix condition offered fewer points than any of the categories in the first task (133 points).

10.3.3 Task and apparatus

The task was run on a Mesh PC and a Mitsubishi 21" monitor with 1024 x 768 pixel resolution in an artificially lit room. Each stimulus was computer generated by using Matlab routines from Brainard's (1997) Psychophysics Toolbox. Stimuli were flashed on a black background that filled the entire screen.

The category learning task required participants to learn to categorise small circular dots (14 pixels in diameter, approximately .525 cm) as members of one of the four categories. The stimuli appeared within a 600*600 pixels display box that delineated the attention space participants should focus on. The dots always appeared along an imaginary, horizontal 600-pixel long line which ran along the horizontal midline of the display box.

The relevant dimension that determined category membership was location of the dot across the imaginary horizontal line. The stimuli had to be classified as members of category A, B, C or D. The stimuli for the four different categories were generated by sampling from four independent but overlapping normal distributions, thus making the task probabilistic. In the two previous studies the mean and standard deviation values (of the categories' distributions) were chosen so that the signal detection discriminability of the two categories (i.e. d prime) was equal to 1. However, results from these studies have shown that a d prime (d') equal to 1 made the task very hard for participants to learn successfully. In fact, a high proportion of participants failed to implement the appropriate uni-dimensional strategy. Hence, it was decided to increase the d' value to render the task easier and, hence, enhance learning. This was also necessary since the present task is harder than the two tasks reported in the previous chapters since participants have to classify the stimuli as members of one of four categories (cf. two categories in the previous studies). The increased numbers of categories was expected to add some requirement for extra cognitive and short-term memory effort.

The d' value was increased from 1 to 2, following the results obtained in a pilot study which tested the impact of a d' of 1.5 and 2 (see appendix 2). Stimuli were generated by sampling from 4 independent but overlapping normal distributions. Hence, in order to develop stimuli with a d' equal to 2, the standard deviation value was chosen to be 50 pixels whereas the distance between the centre of adjacent categories was equal to 100 pixels. The distribution used to generate category A stimuli had a mean pixel location equal to 150, the one used to generate category B stimuli had a mean equal to 250, the one for category C stimuli had a mean equal to 350 and, finally, the one for category D stimuli a mean equal to 450. The dots were drawn, centred at these random, varying pixel positions, measured from the left-hand end of the imaginary vertical line. The stimuli were presented in a fixed quasi-random order to all participants. Each stimulus appeared on the screen until participants responded.

During the task participants were expected to learn to categorise the different stimuli into the four different categories. In order to facilitate learning, participants received feedback for each response they made and the feedback they received was in the form of points. Indeed, as soon as they responded they received a message on the screen below the stimuli which told them how many points they earned and how many they could have earned, as in the task reported in chapter 8. Indeed, results obtained in chapter 9 indicated that neuroticism (N) scores mediated performance in the learning phase. N may have been found to mediate performance because earning 100 points may have been perceived as punishing compared to 200, 300 and 400 points earned for correct responses. However, this N effect was not found in the first study. This may have been due to the fact that, in chapter 8, the presentation of the potential points that could have been won, together with the actual points won, may have attenuated the 'punishing' effect of earning 100 points. Indeed, the perspective of a potential, greater earning may have been perceived as rewarding. It is possible that including both types of information is perceived as more rewarding and, therefore, may lead to a pure BAS activation. Additionally, information on actual and potential rewards may be more informative and enhance learning. In conclusion, following results from the previous studies, it appears that both auditory and visual feedback may enhance learning without impacting on the response bias. Thus, participants were told how many points they had won and how many they could have won since both type of information seem to be useful for learning.

The participants made their responses by pressing some specific keys on the keyboard. They had to press the 'up arrow' to make an A response, the 'down arrow' key to make a response D, a 'right arrow' key response for category B and, finally, a 'left arrow' response for category C. The choice of this key was guided by the fact that this key mapping would have allowed making responses just by using 'the index finger of the dominant hand' to avoid any possible handedness effect.

The corresponding keys were covered by a sticky label marked as A, B, C and D as necessary. This was done in order to make it easier for participants to remember which key corresponded to which category. Visual feedback informed participants on the amount of points they had earned as well as the amount of potential points that could have been earned in that trial. Feedback was presented for 300 msec and followed by 100-msec ITI.

Participants received both visual and auditory feedback for each response they made. The auditory feedback was represented by the sound of a ringing cash register for correct responses, whereas for incorrect responses it was an unpleasant buzzer.

As mentioned in the design section, the task consisted of two phases: the learning phase and the payoff-switch phase; each phase constituted a separate (sub-)task. The task comprised 240 trials in total; the learning phase consisted of 180 trials (45 trials per category) whereas the switch phase consisted of 60 trials (15 trials per category). The two phases were presented as two separate tasks and each task offered participants the chance to earn one entry into the lottery, depending on their performance.

10.3.3.1 Learning phase

The learning phase aimed to assess the impact of the asymmetric payoff matrix on learning during an RB category learning task and, therefore, extend previous findings. As mentioned in the design section, during the learning phase, participants were exposed to the asymmetric payoff matrix which offered different amount of points for correct responses to the different categories (table 10.1). Incorrect responses to any category earn participants 100 points. The structure of the asymmetric payoff matrix was counterbalanced across participants to account for potential artefacts determined by the stimulus location on the screen (see below for details).

10.3.3.2 Payoff-switch phase

The switch phase aims to explore possible behavioural changes (change in response bias) during the task as a result of a switch in the structure of the payoff matrix. This phase consisted of only 60 trials and it represents an exploratory analysis of the potential carry-over effect from a task with asymmetric payoffs to one that delivers symmetric payoffs.

A symmetric payoff matrix was implemented in the switch phase so that all categories offered the same amount of points for correct responses (table 10.1) As in the learning phase of the task, participants earned 100 points for incorrect responses. The switch phase had two switch conditions (higher vs. lower) and participants were randomly allocated to either condition depending on their experimental number code. Under the lower payoff matrix correct responses to any category earned participants 133 points, which represents the lowest amount of points on offer compared with the points previously earned during the learning phase. On the contrary, under the higher payoff matrix participants earned 467 points for correct responses regardless of category type, which is the highest amount of points on offer compared to the points earned in the learning phase. Incorrect responses still earned 100 points as in the learning phases.

10.3.4 Counterbalancing

Only two forms of counterbalancing (CB) were used across participants. The first type of CB was used to avoid any possible effect of stimulus location interacting with payoff magnitude. Indeed, participants may have found it easier to learn to correctly respond to the categories close to the edge of the box (A and D) and may have over-classified ambivalent stimuli from categories B and C as members of the two edge categories (categories A and D). In other words, participants may have shown a tendency to over-classify ambiguous stimuli as members of the near-the-edge categories independently on their payoff and, possibly, this would reduce the effect of the asymmetric matrix on the response bias. Hence, the first type of CB aimed to counteract any possible location-bias. This was achieved by having two alternative matrix structures (PO matrix1 and PO matrix2) counterbalanced across participants, while keeping constant the location of the categories across participants. The first 30 participants performed under CB1 (POmatrix1) and the last 30 under CB2 (POmatrix2). The points earned for correct responses to the different categories are presented below in table 10.2 for the two different PO matrices.

Table 10.2. Payoff matrix of the four categories across the two counterbalancing orders

	Cat A (150 pixels)	**Cat B (250 pixels)**	**Cat C (350 pixels)**	**Cat D (450 pixels)**
PO1	333	200	400	267
PO2	200	333	267	400

The second CB was used in the payoff-switch phase in order to ensure random allocation of participants to either of the two switch conditions. Participants were automatically allocated to one of the two switch CB condition by the computer depending on the ID code they had been allocated. Participants with an odd ID code experienced CBA (higher: 467 points) whereas those with an even ID code were allocated to CBB (lower: 133 points). Thus, there were two crossed counterbalancing conditions (i.e. 1/2; A/B).

10.3.5 Personality measures

Participants completed several personality measures, which were: the Eysenck Personality questionnaire revised (EPQ-R), the Oxford-Liverpool inventory of feelings and experiences scale (OLIFE), the sensation seeking scale (SSS), the schizotypal personality questionnaire (SPQ), the big five inventory (BFI) and the BIS/BAS questionnaire. Four personality components were extracted after running a PCA with Varimax rotation on these scales (see chapter 4 for details). The four components extracted are: Extraversion (E), Neuroticism (N), Positive Schizotypy (PS) and impulsivity-antisocial (ImpAss). The PCA was run on 232 participants, who constitute the overall number of participants tested. The components scores used in the present study are the corresponding scores extracted in the overall PCA involving 232 participants.

Additional impulsivity measures were also included to assess their impact during the CL task and to explore their relationship with other personality components which, as described in the PCA chapter, are expected to be related to the BAS output. The additional impulsivity measures were the delay discounting (DD) task and the Dickman Impulsivity Inventory (DII; Dickman, 1990; see chapter 5). In the present study, the Chronbach's alpha for the FI scale was 0.73 and the alpha value for the DI scale was 0.86, which indicates that the two scales have a reasonable internal consistency. The instructions and procedure reported in chapter 5 for the DD task were also implemented in the present study (thus, consult chapter 5 for details).

10.3.6 Procedure

The study method and procedure were approved by the Psychology Department Ethics Committee (DEC) at Goldsmiths, University of London. Testing was conducted in a testing room in the department. The procedure used in the present study closely resembled the procedure followed in the previous two studies. The instructions and the task layout followed the format used in the other studies although they allowed for the different requirements and structure of the task (i.e. 4 categories).

All participants were given a so-called promotion focus. Indeed, they were told that they could win two entries into a £25 draw per task. In order to win an entry, they had to exceed a certain amount of points (i.e. performance criterion; table 10.3) for the learning or switch phases of the task (considered separately). One draw was held for all 60 participants.

The performance criterion in the learning phase was the same for the two CB groups (i.e. 37100). Nonetheless, in the switch phase the performance criterion varied depending on the CB condition participants performed in. The performance criterion for the two phases are summarised in table 10.3.

Table 10.3. Performance criteria in the different phases of the task and counterbalancing condition (CB)

	Learning Phase	**Switch phase**	
		CB1 (467 points)	**CB2 (133 points)**
Performance Criterion	37100	17200	7000

Participants read the instructions on the computer screen and were encouraged to ask questions for clarification if needed. Once participants reported being ready to start, the experimenter left the room only to return at the end of both tasks. The symmetric payoff matrix condition always followed the learning phase with the asymmetric payoffs. No practice trials were included in the present study since modelling results from the previous study suggest that extended practice may have hindered the activation of the implicit system and, therefore, the response bias.

Instructions informed participants that a series of stimuli represented by circular dots would be individually presented on the screen and that they would need to learn, by trial and error, to classify them into category A, B, C or D by pressing the corresponding keys.

As in the previous studies, points earned were added up on a vertically oriented "point meter" displayed on the right hand of the computer screen. The meter consisted of a 765 pixel tall x 50 pixels wide rectangle and it was set to zero at the beginning of each task. The performance criterion was presented as a horizontal line across the meter and was labelled 'Bonus'. Contrary to the two previous studies, where the performance criterion for each condition was set at 80%, in the present study it was necessary to set the performance criterion at 67% of the points that the optimal classifier would obtain over the overall trials of each task (i.e. 180 and 60), relative to 0% accuracy. This was a necessary measure since, due to the extra effort required by the task as a result of having four categories, the pilot study showed that even with a d' equal to 2 the task was too hard for participants to obtain 80% of the points earned by the optimal classifier. In the pilot study participants reached 67% accuracy levels and, therefore, this level was used in the actual task. Indeed, 67% was considered to be a more realistic target that would have prevented participants to experience frustration but rather maintained a promotion focus. The 67% accuracy level was implemented for the learning phase and for the two payoff-switch conditions, although for the latter it was obviously calculated on a far smaller number of trials.

The region above the criterion line was labelled 'Yes' whereas the one below it was labelled 'No', which indicated whether the participant had won the entry or not. The meter was updated following each response. In the learning phase, it was updated at a rate of 1 pixel per 52 points regardless of CB order. In the switch phase, the meter was updated at a rate of 1 pixel per 37 points in CB1 and a rate of 1 pixel for 15 points in CB2. The part of the meter that changed flashed three times to stress the fact that points had been earned.

The learning phase task consisted of 2 blocks of 80 trials each. Participants had the chance to take a break at the end of the first block and before starting performance on the second task. The instructions presented to participants during the inter-block and inter-task breaks closely resembled those used in the original study (chapter 8; see Appendix 3 for the instructions).

10.3.7 Data analysis

Each data set was fitted using a formal model that could extract the actual criterion location which, as previously mentioned, is a more useful measure of decision bound than the criterion calculated following SDT formulae. The formal model was applied to the last 160 trials of the learning phase. Indeed, it was decided to exclude the initial 20 trials since they constitute practice/familiarisation trials and, therefore, their inclusion would include noise in the analyses. Despite the fact that the first 20 trials were excluded, similar numbers of trials of the four categories were included. Formal fitting was not meaningful for the reversal phase due to the small number of trials on which the fitting would have been based. As already noted, this phase of the study was largely exploratory anyway.

As in chapter 5, prior to analysing the delay discounting data, all delay intervals were transformed into months [e.g. 1-week = (1/30)*7 = 0.23]. Subsequently, data obtained from each participant was individually fitted by both the exponential and the hyperbolic decay fitting model using non-linear regression in SPSS.

The hyperbolic function assumes that discounting decreases in proportion to the time delay, in particular discounting is greater with short time delay and smaller as the time delay increases (equation 10.3):

$$\mathbf{V_d = V_i/[1+(k*D)]} \qquad \textbf{Eq.10. 3}$$

In contrast, the exponential function assumes that the value of the reward decreases by a fixed amount constantly over time (equation 10.4):

$$\mathbf{V_d = V_i *exp^{-kD}} \qquad \textbf{Eq.10. 4}$$

In both equations $\mathbf{V_d}$ is the value of the delayed reward, $\mathbf{V_i}$ represents the value of the immediate reward (i.e. 1000), **D** represents the delay and **k** is the discounting rate constant (i.e. 0.1). Consistent with previous studies, the hyperbolic function offered a better fit to the data than the exponential function (Monterosso and Ainslie, 1999; Richards, et al., 1999; Kalenscher et al., 2008).

10.4 Results

The present results explore participants' performance during the two phases of the 4-category category learning task (i.e. learning and payoff switch phase). Hence, the section is subdivided into two parts, a first section which describes the main results obtained during the learning phase and a second section which describes the secondary results for the switch phase.

10.4.1 Learning phase

10.4.1.1 Preliminary analyses

10.4.1.1.1 Proportion of correct responses

Preliminary analyses were conducted to assess the accuracy level achieved by participants during the initial learning phase of the task. Accuracy was indexed by the proportion of correct responses across the four categories. The analysis indicates that the proportion of correct responses was above chance levels (i.e. 25%) across the four categories and regardless of payoff matrix condition (table 10.4). Indeed, a mixed-design ANOVA on accuracy indicated that the interaction between PO matrix condition and the category type was non-significant [F(3,56) = .11, ns].

Table 10.4. Mean and standard deviation values of the proportion of correct across the four categories in the two payoff conditions (N = 60)

	Proportion of correct responses			
	Cat A	**Cat B**	**Cat C**	**Cat D**
PO matrix 1	0.62 (.16)	0.46 (.18)	0.48 (.18)	0.55 (.18)
PO matrix 2	0.60 (.25)	0.44 (.19)	0.47 (.19)	0.56 (.22)

Table 10.4 also indicates that participants were more accurate when responding to category A and category D stimuli than stimuli belonging to category B and C. Indeed, during briefing several participants verbally reported that they were quite confident in their categorisation of stimuli A and D. This may be due to the fact that the near the edge 'location' of these stimuli made them more discernible (e.g. 'category A stimuli are the ones closest to the left-hand edge of the screen'). Additionally, stimuli from category A and D overlapped with only one other category (B and C, respectively) whereas stimuli from categories B and C probabilistically overlapped with two other categories (i.e. A and C vs. B and D, respectively). Indeed, the main effect of category was found to be significant [F(3,56) = 7.50, p <.001] and the contrasts showed that the proportion of correct responses was statistically greater for category A than category B stimuli [F(1,58) = 21.5, p<.001] and, similarly, participants made significantly more correct responses to category D than category C stimuli [F(1,58) = 5.5, p = .023].

A similar pattern of responding was observed among participants who were using a uni-dimensional strategy (as identified by the formal model; see below). Table 10.5 reports accuracy level for those participants implementing a dimensional strategy regardless of payoff (PO) matrix conditions since previous analysis showed that accuracy was not mediated by payoff condition.

Table 10.5. Mean and standard deviation values of the proportion of correct responses across the four categories for those participants (N = 38) that implemented a uni-dimensional strategy

	Proportion of correct responses			
	Cat A	**Cat B**	**Cat C**	**Cat D**
Mean	0.71	0.51	0.52	0.67
(s.d.)	(.12)	(.19)	(.17)	(.09)

10.4.1.1.2 Points earned

Participants gained points throughout the task and aimed to reach the performance criteria to earn entries into the £25 lottery. Participants could earn one entry in each task. In order to earn an entry during the learning phase of the task, they had to earn 37100 points. The overall sample (N = 60) earned 35581 (s.d. = 4915) points on average, whereas those participants (N = 38) who had use a uni-dimensional rule to perform on the task earned an average of 38004 (s.d. = 3794) points [cf. those guessing/using an alternative strategy earned 31176 (s.d. = 3400) points on average]. An independent sample t-test showed that those participants who implemented a dimensional strategy earned significantly more points than those who were guessing/using an alternative strategy [t(58) = -6.97, p <.001]. This observation accounts for the fact that 26, of the 27 participants who earned an entry during the learning phase of the task, had implemented a dimensional rule.

10.4.2 Model fitting

As described in the method section, the stimuli presented in the task were dots that varied on their horizontal location across the screen. Therefore, the relevant dimension was spatial location and participants were required to use this dimension to perform effectively on the task. As in the previous experiments, participants were expected to use a dimensional rule that allows them to split the perceptual space into 'categorical space'. Since the present task consists of four categories, participants have to set three decision criteria in order to split the perceptual space. Two models were implemented to fit the behavioural data: a uni-dimensional model and a guessing model.

The uni-dimensional model has four free parameters which are the three decision boundaries between categories and a noise parameter. The uni-dimensional model assumes that participants used spatial location as the relevant dimension. The guessing model has three parameters and predicts that participants were guessing their responses across all four categories. The three free parameters are the probability of guessing category A, B and C (guessing of category D is not free to vary and is given by the probabilities of the other 3 categories).

The two models were applied to each participant's data for the last 160 trials. By using maximum likelihood methods each model estimated its parameters. Each model is initially compared to a saturated model which has no free parameters.

The goodness of fit of each model was compared to the goodness of fit of the other models by comparing the Akaike Information Criterion[19] (AIC; Dayton, 2003; Motulsky and Christopoulos, 2004). The AIC score was calculated on the basis of the free parameters and it is an estimate of the goodness of fit. It penalises the model with extra free parameter so that the lower the AIC score, the better the fitting (i.e. closer to the saturated model; Maddox, Ashby and Bohil, 2003).

10.4.2.1 Fitting results

Results showed that the guessing model was significantly worse than the saturated model for all cases, whereas the dimensional model was not significantly worse than the saturated model for 5 cases of the whole sample.

However, by exploring the individual plots it became evident that some participants were well-fitted by the uni-dimensional model (even though this fit

---

[19] $AIC = 2r - 2\ln L$, where r is the number of free parameters and L is the log likelihood of the model (Maddox et al., 2003)

was significantly worse than the fit offered by the saturated model). Hence, further observations and analyses were carried out using these models.

1 – Fitting through a uni-dimensional model

Following data-fitting via the formal model, it was possible to notice that some participants were well-fitted by the uni-dimensional (UD) model. Those data files that were fitted by the uni-dimensional model also had 'noise' values below 150 (cf. bad fitting cases with noise above 150; see figure 10.1). Therefore, it was decided to use noise as the cut-off inclusion criterion and to include into future analysis those participants with noise levels equal or below 150 pixels. Thirty-eight of the 60 participants met this requirement and they were classified as using a UD strategy.

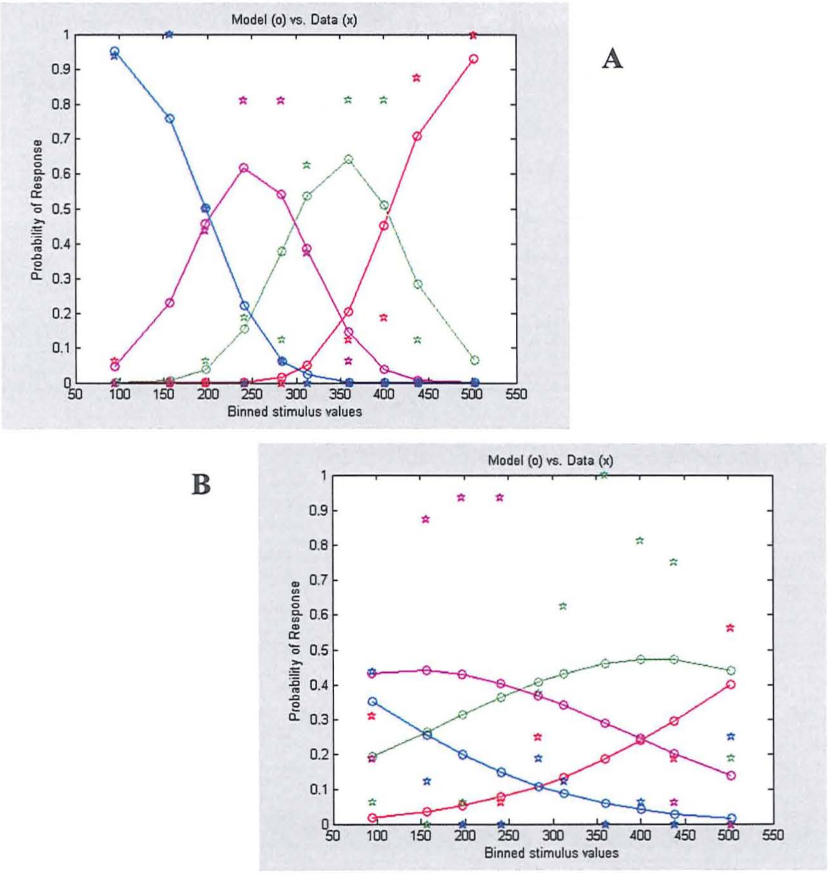


Figure 10.1. It illustrates an example of good fit (figure A) and example of bad fit (figure B) of the data through the use of a uni-dimensional model. The stars represent the actual responses and the circles (unified by the lines) the model fitting for the four categories (blue: category 1, pink: category 2, green: category 3 and red: category 4)

2 – Fitting through a guessing model

For those 38 participants who were found to be well fitted by the uni-dimensional model, it was decided to compare the AIC value of the best fitting model with the AIC value of the guessing model in order to assess whether the data were better explained by the guessing model. In order to do so, the AIC values of guessing model were compared to the AIC values of the uni-dimensional model. The AIC comparison indicated that the UD model offered a better fit than the guessing model in all 38 cases.

For those 22 participants who were not found to be fitted by the uni-dimensional, AIC comparisons were also held to establish whether the guessing model may have fitted these data and accounted for those participants' performance on the task. The AIC values of the guessing model were separately compared to the AIC values of the UD model and these comparisons showed that 16 of the 22 cases, not well fitted by the UD model, were better-fitted by the guessing model (i.e. AIC values were lower for the guessing model than for the UD model). Figure 10.2 illustrates an example of good fitting through the guessing model.

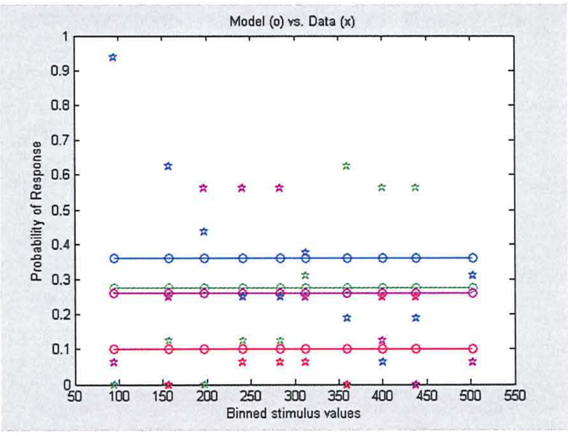


Figure 10.2. It illustrates an example of good fit of the data through the use of a guessing model. The stars represent the actual responses and the circles (unified by the line) the model fitting for the four categories (blue: category 1, pink: category 2, green: category 3 and red: category 4)

Overall, the model fitting results indicate that 38 participants were using a dimensional rule whereas 16 were guessing and six cases were not fitted by any of the models tested in this chapter. These participants may be referred to as non-learners since they fail to learn to perform appropriately on the task. Nonetheless, it is possible that they may have been using alternative strategies considering participants' self-report and accuracy levels among non-learners (see below).

10.4.2.2 Strategy used and personality correlation

Point-biserial correlations were run in order to assess whether there was a relationship between any of the personality components and the strategy used, i.e. uni-dimensional vs. guessing/other strategy. The variable that codes the type of strategy implemented was labelled strategy used and coded as 1 for a uni-dimensional strategy and as 2 for guessing/other. No correlation was found to be significant (Ps > .24).

10.4.2.3 Behavioural data

Some analyses were carried out to explore the behavioural data and observe where participants placed their decision criteria during the learning phase of the task.

The mean and standard deviation values for the (fitted) criteria for those participants implementing a UD strategy (N = 38) are presented in table 10.6. The criteria do not seem to differ across the two feedback conditions.

Table 10.6. Mean and standard deviation of the three criteria fitted by the uni-dimensional model across the matrix conditions (N = 38)

	Criterion 1	Criterion 2	Criterion 3	Noise
PO Matrix 1	182.97 (26.4)	300.13 (14.0)	420.84 (23.8)	80.27 (31.7)
PO Matrix 2	191.77 (27.3)	303.21 (21.0)	413.31 (20.7)	79.62 (33.3)

A mixed-design ANOVA confirmed this initial observation and indicated that the criterion locations did not differ across the two PO matrix conditions, as the PO condition*criterion interaction was non-significant [F(2,35) = 1.08, ns]. Hence, it was decided to merge the data from the two PO matrix conditions in the following analyses. Table 10.7 represents the overall criteria and noise levels regardless of payoff condition.

Table 10.7. Mean and standard deviation of the three criteria and noise levels, regardless of payoff condition

	Criterion 1	Criterion 2	Criterion 3	Noise
Mean	187.60	301.75	416.88	79.93
(s.d.)	(26.90)	(17.83)	(22.25)	(31.60)

For optimum accuracy, in the present task the criteria should be set at 200, 300 and 400 pixels, respectively[20]. One-sample t-tests were run to assess whether the fitted criteria were statistically different from the optimal accuracy criteria. Results showed that criterion 1 and 3 were significantly different from the optimal accuracy criterion [200 and 400, respectively; criterion 1: t(37) = =-2.84, p = .007; criterion 3: t(37) = 4.68, p<.001]. In contrast, criterion 2 was not found to be significantly different from the optimal accuracy criterion [i.e. 300; t(37) = .61, ns].

Further analyses were carried out to assess the relationship between the criteria and personality traits among those participants who implemented the appropriate UD rule.

10.4.2.4 Correlations between the four personality components and the criteria set across the task

A few exploratory correlations were run between the four personality components extracted in the PCA (chapter 4) and the criteria set during the task (table 10.8). The correlations showed that positive schizotypy was negatively

---

[20] The optimal accuracy criteria are determined by calculating the distance midpoint between mean-pixel location of adjacent categories; e.g. the accuracy criterion between category A and category B is equal to the mean score of their mean distribution value [i.e. (150+250)/2 = 200].

correlated with criterion 2 scores (r = -.35, p = .03). No other correlation was significant (ps > .14).

Table 10.8. Correlations between the criteria fitted by the uni-dimensional model and the four personality components (E = extraversion, N = Neuroticism, PS = positive schizotypy and I = ImpAss)

		E	N	PS	I
Criterion 1	Pearson Correlation	.064	.003	-.242	-.088
	Sig. (2-tailed)	.703	.985	.142	.601
	N	38	38	38	38
Criterion 2	Pearson Correlation	-.139	-.075	-.347	-.154
	Sig. (2-tailed)	.406	.656	.033	.357
	N	38	38	38	38
Criterion 3	Pearson Correlation	.099	.120	-.133	.157
	Sig. (2-tailed)	.555	.475	.425	.348
	N	38	38	38	38

10.4.2.5 Correlations between the criteria set across the task and the impulsivity measures

Further correlations were run between the three criteria and the impulsivity measures collected in the present study, i.e. measures of functional and dysfunctional impulsivity (Dickman's scale) and delay discounting measures as indexed by the hyperbolic k scores (see data analysis section).

Preliminary correlations were run to assess the relationship between the four personality components extracted in the PCA and the impulsivity measures collected in the present study [i.e. DII's functional impulsivity (FI) and dysfunctional impulsivity (DI) measures; table 10.9].

Table 10.9. Correlation between the PCA-extracted personality components and the DII measures (FI = functional impulsivity; DI = dysfunctional impulsivity; E = extraversion, N = Neuroticism, PS = positive schizotypy and I = ImpAss)

		E	N	PS	I
FI	Pearson Correlation	.514	-.146	-.100	.351
	Sig. (2-tailed)	.001	.381	.551	.031
	N	38	38	38	38
DI	Pearson Correlation	.206	.258	.347	.312
	Sig. (2-tailed)	.215	.118	.033	.056
	N	38	38	38	38

Results show that extraversion was positively correlated with functional impulsivity (r = .51, p = .001) and so was the ImpAss component (r = .35, p = .031). In contrast, dysfunctional impulsivity was correlated with the positive schizotypy component (r = .35, p = .033) and its correlation with ImpAss just failed to be significant (r = .31, p = .056).

The hyperbolic k scores were highly positively skewed and contained a few possible outliers, hence it was decided to run correlations between the scores and the criteria using Spearman's rho. Delay discounting scores were found to be positively correlated with dysfunctional impulsivity (r = .37, p = .035).

The correlations assessing these impulsivity measures and the criteria scores showed a non-significant trend for a negative correlation between criterion 2 and dysfunctional impulsivity (r = -.28, p = .09). None of the other relationship with FI or DI was significant (ps > .23). There was also a significant negative correlation between delay discounting scores and criterion 1 (in the criterion used between category A and category B; rho = -.36, p = .026; figure 10.3).

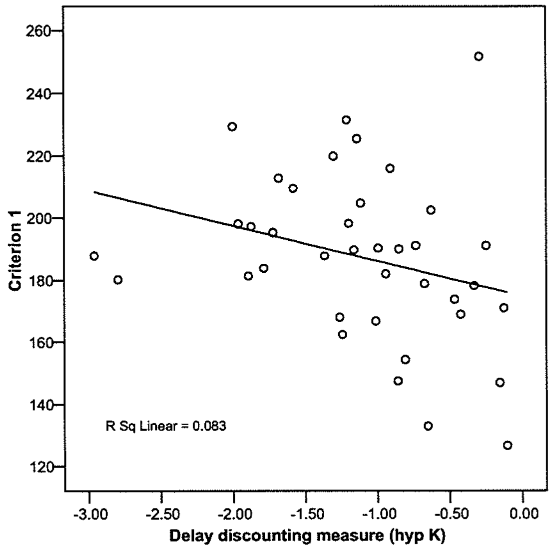


Figure 10.3. The scatterplot shows the relationship between delay discounting scores and criterion 1.

Results also showed that there was a non-significant trend for a positive correlation between the delay discounting measure and criterion 3 (rho = .28, p

= .09). The correlation between criterion 2 and the delay discounting measure was non-significant (r = -.10, ns).

A difference scores was computed, subtracting criterion 1 scores from criterion 3 scores, and it was correlated with the personality components. Spearman's Rho identified a strong positive correlation between the difference score and the hyperbolic K scores (rho = .43, p = .007; figure 10.4). No other correlation was found to be significant (ps > 0.3).

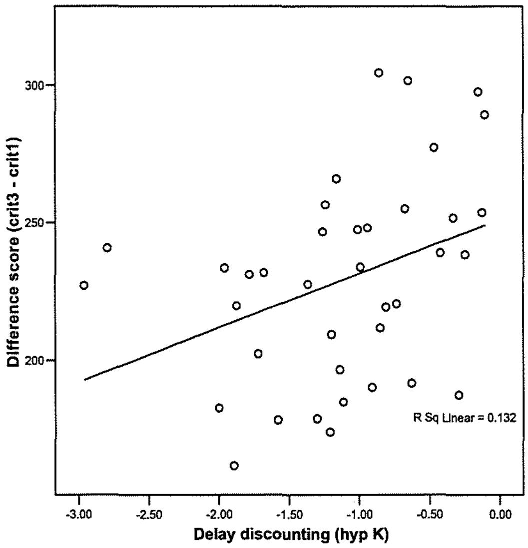


Figure 10.4. The □catterplot represents the relationship between delay discounting scores and the difference score

The hyperbolic k scores were recoded by applying a logarithm transformation in order to render the data more normally distributed. This transformation allows one to further explore the present results using parametric methods (e.g. multiple regression, ANCOVA). Simple contrasts were run to compare criterion 1 and criterion 3 in relation to the delay discounting measure (i.e. the recoded hyperbolic K).

10.4.2.5.1 Delay discounting (DD)

A mixed-design ANCOVA was run in order to explore the relationship between the impulsivity delay discounting scores and the criterion locations set during the learning phase of the task. The recoded hyperbolic K scores were entered as a covariate while criterion was entered as repeated measure factor. Results showed that the 2-way interaction between the criteria and the impulsivity measure (i.e. DD scores) was significant [F(1.5, 53.11) = 4.37, p = .027].

10.4.2.6 Non-learners

This section considers data from those participants who failed to implement the appropriate uni-dimensional strategy (i.e. non-learners). Results showed that non-learners had accuracy levels equal to 38% (s.d. = .100) which was significantly above chance (0.25; t(21) = 6.2, p <.001). This observation indicates that despite the fact that they were not using the optimal strategy, they may have been implementing an alternative strategy effective enough to score above chance level.

These results suggest that 16 of the 22 learners who were reasonably well fit by the guessing model cannot have been simply guessing (or else they would have achieved accuracy scores around 25%). This is explained by the fact that the guessing model did not offer as good a fit as the saturated model, but it offered a better fit than the dimensional model (according to AIC scores). However, overall non-learners earned fewer lottery entries than learners [$\chi^2(2) = 31.29$, p <.001; table 10.10].

Table 10.10. Number of lottery tickets earned by learners (N = 38) and non-learners (N = 22)

		Tickets earned		
		0	**1**	**2**
Learner	**Yes**	0	12	26
	No	7	15	0

Non-learners tend to earn 14 tickets in the reversal phase (14/16 vs. 1/16 in the learning phase). This suggests that the extended practice on the CL may have allowed non-learners to show levels of accuracy high enough to earn an entry into the lottery in the reversal phase (i.e. last 60 trials).

Correlations were run to establish the relationship between the guessing parameter values (g_1 – g_3) and the different personality traits. Results showed that there was a significant correlation between the delay discounting measure and g_3 (i.e. the guessing parameter between categories C and D; rho = .46, p <.04). Similarly, there was a negative trend for the correlation between the DD measures and g_1 (i.e. the guessing parameter between categories A and B; rho = -.42, p <.06). Finally, there was a positive correlation between g_2 and positive schizotypy (r = .53, p = .01). It is of particular interest that the same personality traits that were found to mediate criterion location among learners mediate performance among non-learners. Indeed, it strengthens the assumptions that non-learners might have been using an alternative strategy that might have been mediated by DAergic mechanism. Indeed, as mentioned earlier, the guessing model offers a significantly worse fit than the saturated model and, therefore, it is possible that alternative processes may have been mediating performance.

Moreover, accuracy scores, indexed by the proportion of correct responses, were positively correlated with functional impulsivity (r = .48, p = .02) and its negative correlation with positive schizotypy was just significant (r = -.43, p = .048). In contrast, there was no significant correlation between accuracy and personality among learners (ps >.34).

10.4.3 Payoff switch phase

It was not possible to fit the data through the formal model due to the small number of trials included in the reversal phase, which contained 15 stimuli of each category. Hence, overall performance levels (e.g. accuracy) were explored among those 38 participants who had successfully implemented a uni-dimensional strategy in the learning phase. There were equal numbers of participants from the two payoff switch condition (higher and lower).

10.4.3.1 Preliminary analyses

10.4.3.1.1 Proportion of correct responses

Preliminary analyses were conducted to assess the accuracy level achieved by participants during the payoff switch phase of the task. Accuracy was indexed by the proportion of correct responses across the four categories (table 10.11). A mixed-design ANOVA on accuracy showed that there was a significant main effect of category type [F(3,34) = 29.07, p <.001]. In fact, as in the learning phase, participants showed greater accuracy for stimuli that appeared closer to the external edges of the box (i.e. stimuli from categories A and D). These results can be explained by the same rationale presented earlier. Additionally, there was a significant main effect of switch condition [F(1,36) = 4.35, p = .044]. Indeed, participants in the higher reversal condition showed greater accuracy (.72, s.d. = .06) than those in the lower condition (.67, s.d. = .08).

Table 10.11. Mean and standard deviation values of the proportion of correct across the four categories in the two payoff conditions

	Proportion of correct (pc) responses			
	Cat A	**Cat B**	**Cat C**	**Cat D**
Switch_higher **(N = 19)**	0.82 (.10)	0.65 (.16)	0.63 (.13)	0.78 (.06)
Switch_lower **(N = 19)**	0.85 (.09)	0.56 (.21)	0.56 (.16)	0.73 (.16)

The ANOVA also indicated that the interaction between switch condition and the category type was non-significant [F(3,56) = .28, ns]. Thus, the accuracy levels across the two conditions were merged (table 10.12).

Table 10.12. Proportion of correct responses for the four categories regardless of payoff switch condition

	Proportion of correct (pc) responses			
	Cat A	Cat B	Cat C	Cat D
Mean	0.84	0.61	0.59	0.75
(s.d.)	(.10)	(.19)	(.15)	(.12)

The accuracy levels across the four categories were found to be significantly above chance level (i.e. 0.25; 1-sample t-tests: ts >11.7, ps<.001).

10.4.3.1.2 Points

Owing to the high accuracy levels achieved in the switch phase of the task, fifty-two participants out of the whole sample earned one entry in this second half of the task. Participants allocated to the higher switch (i.e. 467 points won for each correct response and 100 for incorrect responses) condition had to earn 17200 points to earn an entry whereas those allocated to the lower switch condition (i.e. 133 points earned for correct responses and 100 for incorrect ones) had to earn 7000 points to earn an entry into the £25 lottery. On average, participants performing under the higher switch condition earned 21916 points (s.d. = 1412) while those under the lower switch condition earned 7336 (s.d. = 154). All 38 participants earned an entry in the lottery during the reversal phase.

10.4.3.2 Behaviour and personality across the task

It was decided to assess the impact of payoff switch on accuracy and, therefore, accuracy levels between the last 60 trials of the learning phase and the 60 trials of the reversal phase were analysed across all four categories. Accuracy was indexed by the proportion of correct responses. A mixed-design ANOVA was run with phase as a repeated measure factor with two levels (learning vs. reversal), category as a repeated measure factor with four levels and switch condition as an independent factor with two levels. Overall, results indicated that there was a significant main effect of phase [F(1,36) = 30.34, p <.001], which indicates that accuracy improved over trials (i.e. in the reversal phase; table 10.13).

Table 10.13. Accuracy levels across the four categories in the last 60 trials of the learning phase and the reversal phase

	Cat A	**Cat B**	**Cat C**	**Cat D**
Learning last 60 trials	.78 (.14)	.51 (.22)	.55 (.21)	.65 (.17)
Reversal phase	.85 (.10)	.61 (.17)	.60 (.18)	.74 (.09)

Similarly, there was a significant main effect of category [F(1,36) = 29.40, p <.001] since accuracy for categories A and D was greater than for categories B and C (table 10.13). The main effect of switch condition was also significant [F(1,36) = 5.63, p = .023]. The latter result indicates that, even though accuracy levels are higher in the reversal phase, the proportion of correct responses is higher for categories A and D, rather than B and C, in the reversal phase as in the learning phase. None of the interactions was found to be significant (ps >. 4).

Each personality component was entered as a covariate into several ANCOVAs that explore the impact of individual differences on accuracy. The ANCOVAs had two repeated measure factors, i.e. phase (two levels) and category (four levels). However, none of the personality component was found to mediate learning significantly (Fs <.3, ps > .2).

It was then decided to assess how performance (i.e. accuracy) improved across the overall task. The learning phase was broken down into three blocks of 60 trials each and data from the reversal phase was also included and it constituted the fourth block. A mixed-design ANOVA was run with phase as a repeated measure factor with four levels (i.e. each 60-trial block), category as a repeated measure factor with four levels and switch condition as an independent factor with two levels. The main effect of phase [F(2.25, 80.95) = 48.69, p <.001] and category [F(2.29, 82.41) = 19.26, p <.001] were significant. The switch condition factor did not seem to affect performance as none of the interactions was significant (ps > .2). Thus, switch condition was excluded from further analysis.

The personality measures were individually entered as covariates into several repeated-measure ANCOVAs. Results showed that neuroticism was the only personality component that affected accuracy scores. Indeed, there was a significant interaction between neuroticism and phase [F(2.3,83.19) = 3.48, p = .03] whereas there was a non-significant trend for the 3-way interaction between category, phase and neuroticism [F(6.9, 247) = 1.99, p = .06]. The contrasts indicated that the interaction involved the linear trend [F(1,36) = 6.82, p = .013]. Therefore, the linear trend components were extracted and correlated with scores on the neuroticism component. The correlation showed a negative relationship between these two measures (r = -.38, p = .02). These results suggest that participants with low scores on neuroticism showed better learning improvements across the four learning blocks (i.e. greater accuracy increases) than their high scoring counterparts.

10.4.4 Neural model simulations

10.4.4.1 Preliminary simulations of the behavioural data

In the previous chapters, it has been shown that the neural model is able to capture the human behaviour including individual differences observed in the actual experiments. These results not only support the empirical findings but also the validity of the CL task to capture the behaviour under study (i.e. response bias) as well as the relationship between behaviour and personality. Initial simulations were conducted in order to assess the validity of the task to portray the behavioural data during the present task.

As in the previous chapters, the simulated data were subsequently fitted by the formal model in order to obtain the decision boundaries. In order to keep the analysis of the simulated data close to the analyses of the empirical data the initial 20 trials were excluded from further analyses since they represent familiarisation and practice trials and, therefore, would add noise to the data. The simulations below all generated 300 simulated 'participants', unless otherwise stated.

Initial simulations attempted to capture decision criteria close to the criteria observed in the human data. These simulations were run without any inclusion of individual differences and were separately run for the two payoff matrix counterbalancing orders. Following preliminary simulations (see appendix 4), it was found that it was necessary to adopt two values of perceptual noise (cf. original noise level equal to 50) in order simulate the higher accuracy scores, observed in the human data, for categories A and D (cf. categories B and C). The two **pnoise** values were set equal to 40 and 60 for the stimuli at the edge and in the middle of the display, respectively.

Additionally, the simulations indicated that participants had initially placed their decision bounds close to the perceptual criteria (i.e. 150, 300 and 450 pixels) and learning was characterised by the shift away from the perceptual bounds and closer to the optimal accuracy bounds i.e. 200 300 and 400 (figure 10.5).

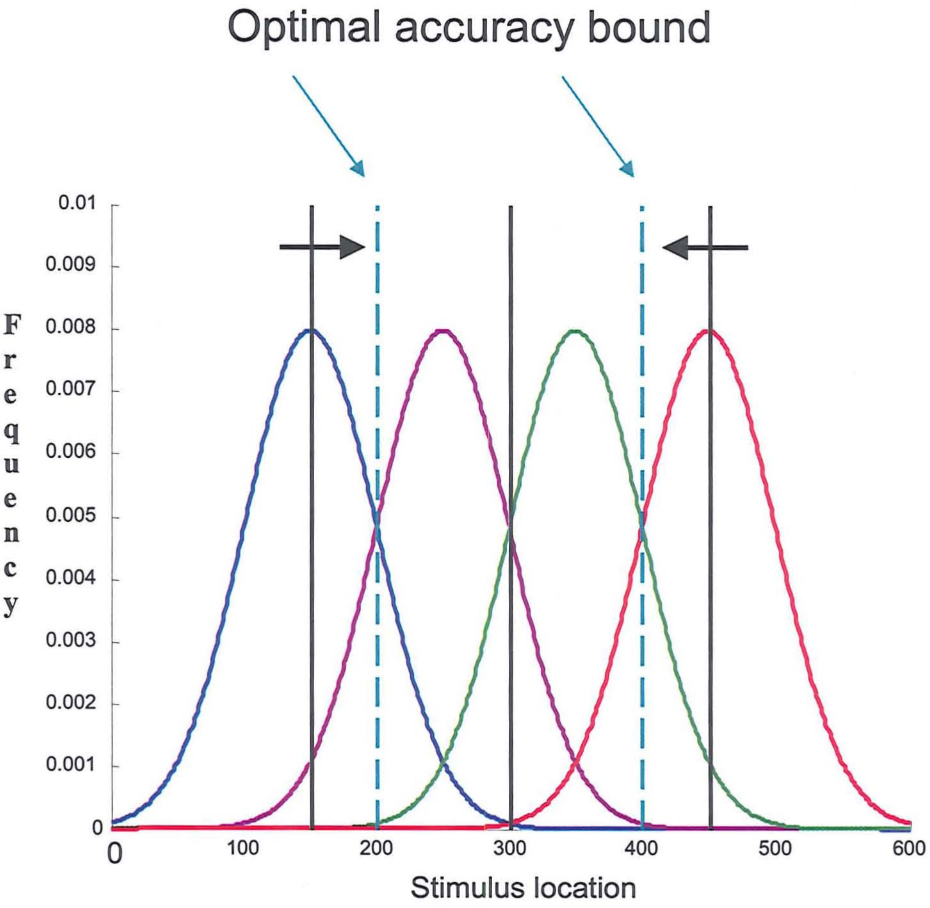


Figure 10.5. Criterion shift from the perceptual bounds (black lines) towards the optimal accuracy bounds (cyan lines)

Finally, the preliminary simulations indicated that partially asymmetric payoff matrices had to be implemented, in order to simulate the human data. The partially asymmetric matrices recoded the reinforcement values (points) following correct categorisations into an **rf** value of 0.75 apart from the 400 points, which were coded into an **rf** value of 1 [i.e. PO matrix 1: 0.75; 0.75; 1; 0.75; while PO matrix 2: 0.75; 0.75; 0.75; 1][21]. Indeed, participants verbally reported not having paid attention to the 'point' feedback; although they reported having noticed the 400 points more frequently than the other payoffs so it was assumed that the 400 points particularly may have affected learning and criterion placement. Additionally, the reward prediction (**rptonic**) parameter was set equal to 0.5 and the **scaler** equal to 2.

---

[21] These **rf** values code winnings received following correct responses for categories A to D, whereas an **rf** value equal to 0.25 was used to code winnings for incorrect responses.

The simulations obtained using these parameter values were able to capture the criteria and noise level observed in the human data (table 10.14).

Table 10.14. Human and simulated criteria (s.d.) under the revised asymmetric matrices across the two payoff matrices (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Criterion 1	Criterion 2	Criterion 3	Noise
Sim_PO1 **(N = 300)**	178.98 (10.7)	300.02 (8.2)	432.71 (11.7)	82.79 (9.5)
Sim_PO2 **(N=300)**	180.33 (10.3)	303.35 (7.7)	418.02 (12.0)	79.38 (8.9)
Hum_PO1 **(N = 18)**	182.97 (26.4)	300.13 (14.0)	420.84 (23.8)	80.27 (31.7)
Hum_PO2 **(N = 20)**	191.77 (27.3)	303.21 (21.0)	413.31 (20.7)	79.62 (32.3)

Additionally, these matrices also simulated the accuracy asymmetry for the edge stimuli that were observed in the human data (table 10.15). Therefore, these **rf** values were selected to be used for all simulations reported below.

Table 10.15. Simulated and human accuracy level under the revised asymmetric matrices (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Cat A	Cat B	Cat C	Cat D
Sim_PO1	0.67	0.59	0.59	0.59
Sim_PO2	0.67	0.60	0.52	0.65
Hum_PO1	0.68	0.53	0.54	0.66
Hum_PO2	0.73	0.48	0.50	0.68

The model was also able to capture the step-like learning curve observed among those human participants who had implemented the appropriate, UD rule (figure 10.6a). In contrast, a purely RPE-driven, implicit system captured the learning pattern observed among those participants who failed to implement the optimal rule (figure 10.6b). Similar simulation results had been observed when simulating the 2-category task implemented in the previous two chapters (see chapter 8).

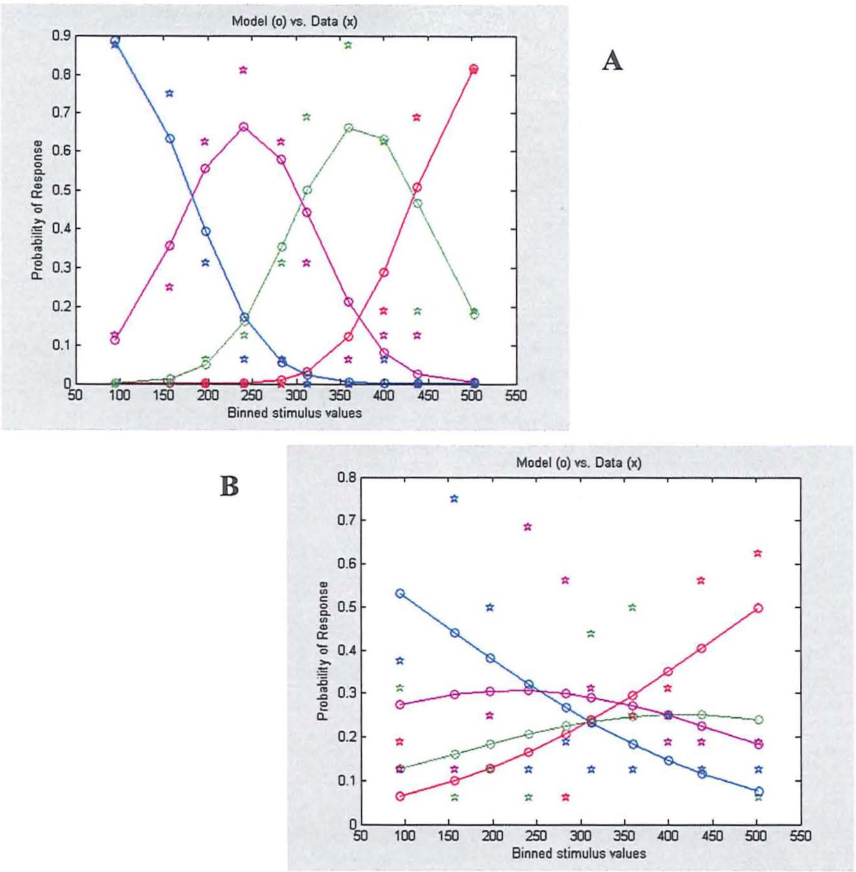


Figure 10.6. Figure A represented the simulated, step-like learning function obtained with the revised neural model whereas figure B represented the learning function simulated via the implicit, RPE-driven model alone. The stars represent the actual responses and the circles (unified by the lines) the model fitting for the four categories (blue: category 1, pink: category 2, green: category 3 and red: category 4)

Further simulations were run to explore the impact of individual differences (i.e. along the "BAS" parameter) on criterion placement. The parameter values used for the next simulations, are the values that were able to capture the decision bounds described above (table 10.16).

Table 10.16. Parameters' values implemented in the simulations

Symbol	Description	Value
rptonic	**Reward prediction tonic value**	0.5
rulebias	**Strength of explicit module**	0.4
m	**Scales the rf values**	1
Pnoise	**Perceptual and criterial noise**	40/60
b_1 and b_2	**Learning rates of the RP unit**	0.05*scaler1
b_3 and b_4	**Learning rates of the SR unit**	0.05*scaler2
Scaler1	**It scales the learning rate of RP unit**	2
Scaler2	**It scales the learning rate of RP unit**	2

10.4.4.2 Preliminary analyses to simulate the relationship between behaviour and individual differences

In the two previous chapter, it was observed that two models were best at capturing the trait*behaviour relationship. One model has BAS variation acting on the reward input to the DA cell, although nonlinearities made this model less likely to account for linear relationships observed in real human data. The other model has BAS variation acting as a multiplier on the reward prediction error (RPE) signal affecting learning at synapses in the SR pathways (on the actor units). Thus, in the present chapter several simulations were run using those two models to simulate the empirical data. Additionally, further simulations were run using the model with BAS acting on the RPE projections to the critic unit in order to assess whether this model could also account for the human data (although, it had previously failed previously to do so). Simulations were separately run for the two payoff sub-conditions (i.e. POmatrix 1 and POmatrix 2). These initial simulations capture a BAS parameter variation that ranges randomly and uniformly from 0 to 2.

10.4.4.2.1 Individual variance on the reward input to the DA cell

Results showed that the simulated criteria and, in particular, the noise levels were not particularly close to those observed in the human data (table 10.17).

Previous simulations had 'BAS' values set at 1 for all simulated participants, adding BAS variation in the range 0-2 adds variance to the simulation.

Table 10.17. Simulated criteria (s.d.) with BAS on the reward cell compared to the human data (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Criterion 1	Criterion 2	Criterion 3	Noise
Sim_PO1	161.24 (33.3)	300.88 (10.6)	446.46 (35.4)	115.56 (80.80)
Sim_PO2	163.23 (31.8)	302.70 (10.8)	434.27 (34.2)	110.95 (79.0)
Hum_PO1 **(N = 18)**	182.97 (26.4)	300.13 (14.0)	420.84 (23.8)	80.27 (31.7)
Hum_PO2 **(N = 20)**	191.77 (27.3)	303.21 (21.0)	413.31 (20.7)	79.62 (32.3)

Overall, simulations under both PO matrices indicate that the stronger correlations are between the BAS variance and criteria 1 and 3 whereas the BAS*criterion 2 relationship are much weaker.

POmatrix 1 (points: 333 200 400 267 for categories A to D, respectively)

Results showed that there was a significant, positive correlation between criterion 1 and the BAS variation ($r = .74$, $p <.001$) and a negative correlation between criterion 3 and the BAS ($r = -.72$, $p <.001$). However, the criterion 2* BAS variation correlation was non-significant ($r = 0.07$, ns). All these relationship are visually represented in figure 10.7. Panel 'a' and panel 'c' show a strongly curvilinear relationship as observed in the previous chapters for this model.

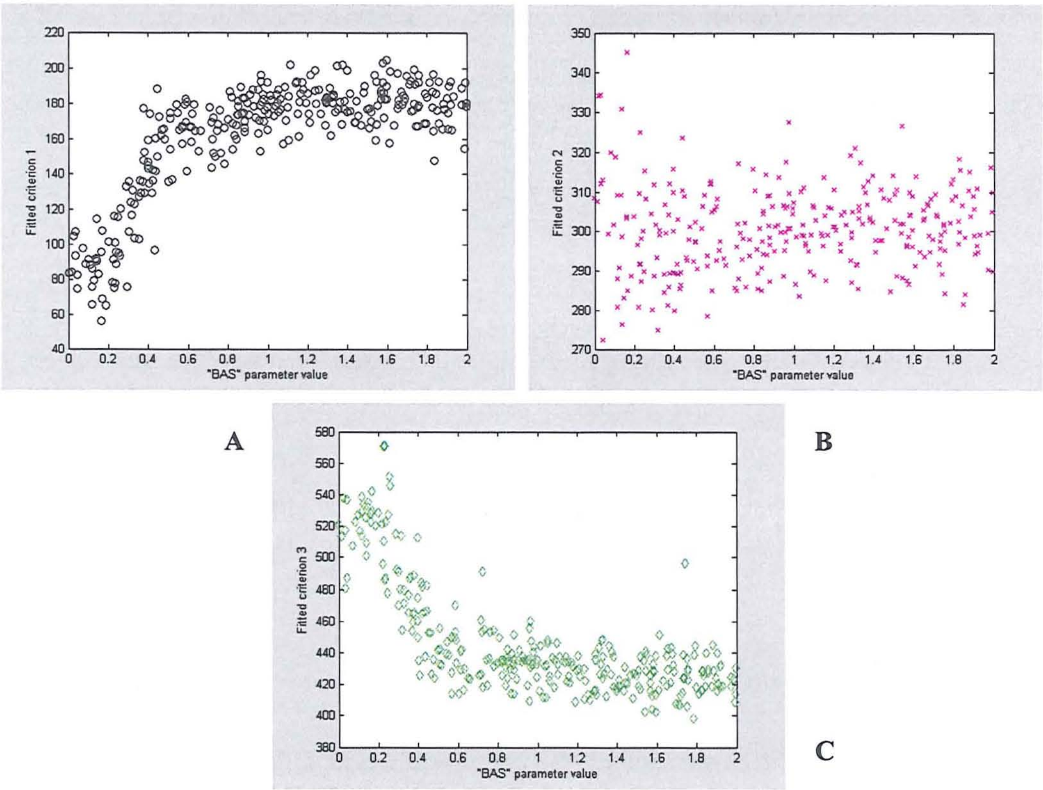


Figure 10.7. Scatterplots indicating the relationship between "BAS" variation and the fitted criteria under POmatrix 1 (N = 300). Figure A illustrates the relationship between criterion 1 and BAS variation, figure B the non-significant BAS*criterion 2 relationship and figure C the relationship between criterion 3 and BAS variation

As observed in the human data, the simulated data showed that the uni-dimensional model was non-significantly worse than the saturated model for 37 'participants' out of the whole sample (N =300). Thus, approximately 12% of the sample was well-fitted by the uni-dimensional model. This ratio resembles the ratio observed in the human data where the uni-dimensional model was non-significantly worse than the saturated model for 3 out of the 30 (10%) participants who performed under PO matrix 1. It was decided to apply the same exclusion criterion applied to the human data and, therefore, to exclude data sets with noise levels above 150.

After applying this exclusion criterion, the sample was equal to 231 (77% of the actual sample) and the mean criterion values were close to the values obtained in the whole sample; indeed, criterion 1 = 176.59 (s.d. = 12.8), criterion 2 = 300.99 (s.d. = 8.8) and criterion 3 = 430.22 (s.d. = 13.8). The noise level was equal to 75.96 (27.8). The correlations were weaker than those observed in the unfiltered simulations (BAS*criterion 1: r = .41, p <.001; BAS*criterion 2: r = .15, p = .03 and BAS*criterion3: r = -.34, p <.001).

POmatrix 2 (points: 200 333 267 400 for categories A to D, respectively)

Results showed that there was a significant, positive correlation between criterion 1 and the BAS variation (r = .70, p <.001) and a negative correlation between criterion 3 and BAS variance (r = -.65, p < .001). Finally, the BAS (variation) *criterion 2 correlation was significant but very weak (r = .13, p = .02). All these relationship are visually represented in figure 10.8. As in the previous simulations, the relationships reported in panels 'a' and 'c' are strongly non-linear.

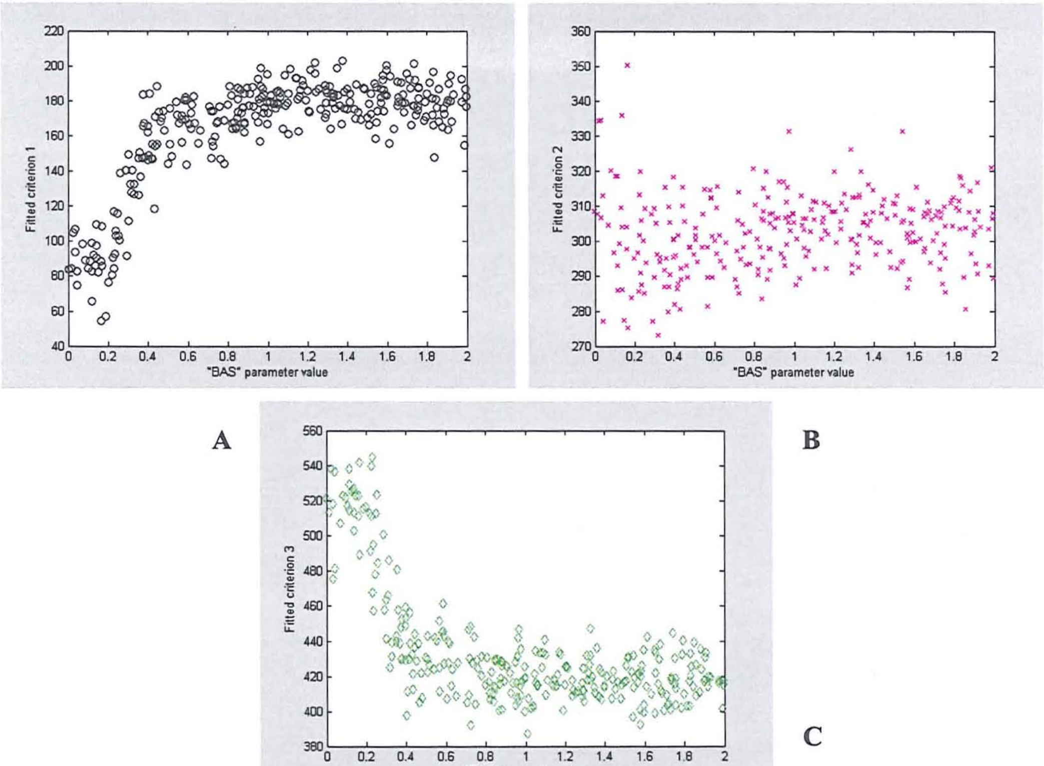


Figure 10.8. Scatterplots indicating the relationship between "BAS" variation and the fitted criteria under POmatrix 2 (N = 300). Figure A illustrates the relationship between criterion 1 and BAS variation, figure B the non-significant BAS*criterion 2 relationship and figure C the relationship between criterion 3 and BAS variation

As observed in the human data, the simulated data showed that the uni-dimensional model was non-significantly worse than the saturated model for 37 'participants' out of the whole sample (N =300). Thus, approximately 16% of the sample was well-fit by the uni-dimensional model. This ratio is slightly higher than the ratio observed in the human data where the uni-dimensional model was non-significantly worse than the saturated model for 2 out of the 30 (7%) participants who performed under PO matrix 2. It was decided to apply the same exclusion criterion applied to the human data and, therefore, to exclude data sets with noise levels above 150.

Following the exclusion of those simulate participants, the sample was equal to 234 (78% of the actual sample) and the mean criterion values were close to the values obtained in the whole sample; indeed, criterion 1 = 177.14 (s.d. = 11.8), criterion 2 = 300.28 (s.d. = 8.9) and criterion 3 = 419.84 (s.d. = 12.6). The noise level was equal to 73.82 (28.4). The correlations were weaker than those observed in the unfiltered simulations (BAS*criterion 1: r = .33, p <.001; BAS*criterion2: r = .16, p = .02 and BAS*criterion3: r = -.18, p = .005).

10.4.4.1.2 Individual variance on the RPE signal projections to the synapses in the SR pathway

The simulated criteria under this sub-model also offer a close fit, especially in terms of noise, to the empirical data under both payoff manipulations (table 10.18).

Table 10.18. Simulated criteria (s.d.) with BAS variance on the SR cell compared to the human data (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Criterion 1	Criterion 2	Criterion 3	Noise
Sim_PO1	176.42 (15.8)	300.76 (8.3)	433.75 (14.2)	79.29 (16.9)
Sim_PO2	177.54 (16.1)	303.52 (8.2)	422.60 (16.5)	76.91 (15.5)
Hum_PO1 **(N = 18)**	182.97 (26.4)	300.13 (14.0)	420.84 (23.8)	80.27 (31.7)
Hum_PO2 **(N = 20)**	191.77 (27.3)	303.21 (21.0)	413.31 (20.7)	79.62 (32.3)

POmatrix 1 (points: 333 200 400 267)

The simulation results showed that the correlation between criterion 1 was positively correlated with BAS variation (r = .78, p <.001) while criterion 3 was found to be negatively correlated with the BAS (r = -.58, p <.001). The correlation between criterion 2 and the BAS variance was significant but weak (r = .17, p = .002). Figure 10.9 presents the actual scatterplots that describe these relationships, whcih are linear with little evidence of non-linearity.

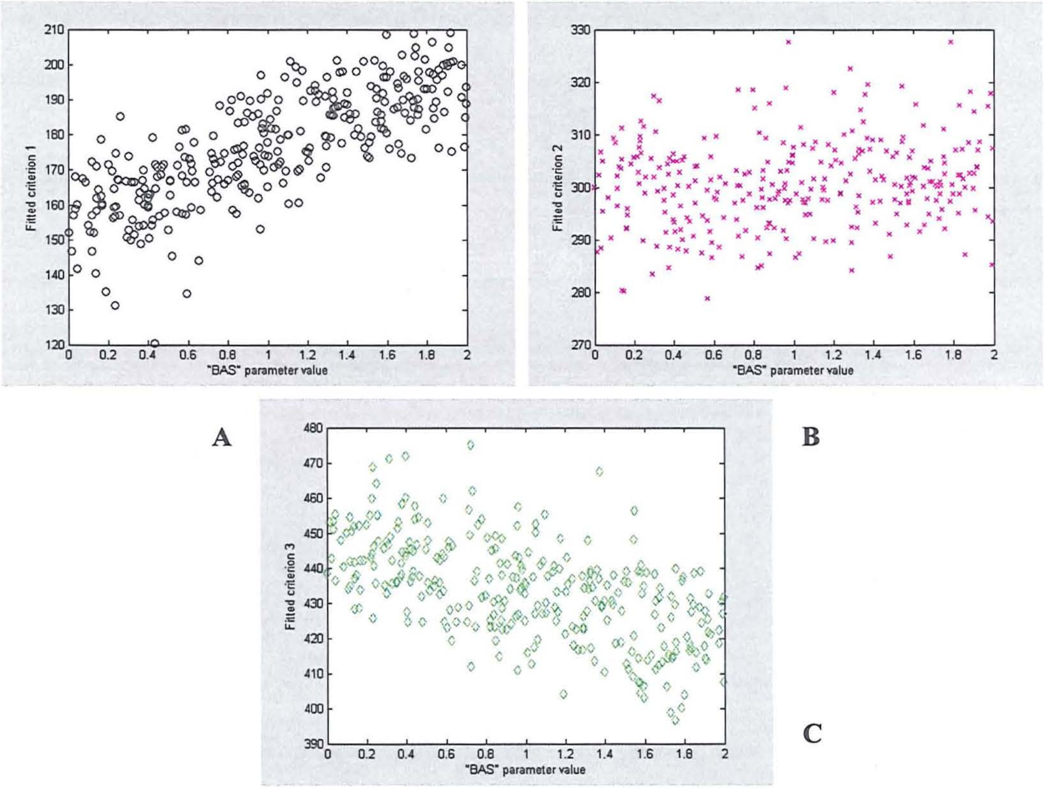


Figure 10.9. Scatterplots indicating the relationship between "BAS" variation and the fitted criteria under POmatrix 1 (N = 300). Figure A illustrates the relationship between criterion 1 and BAS variation, figure B the non-significant BAS*criterion 2 relationship and figure C the relationship between criterion 3 and BAS variation.

As observed in the human data, the simulated data showed that the uni-dimensional model was non-significantly worse than the saturated model for 12 'participants' out of the whole sample (N =300). Thus, approximately 4% of the sample was well-fit by the uni-dimensional model. This ratio is roughly equivalent to the ratio observed in the human data where the uni-dimensional model was non-significantly worse than the saturated model for 3 out of the 30

(10%) participants who performed under PO matrix 1. It was decided to apply the same exclusion criterion applied to the human data and, therefore, to exclude data sets with noise levels above 150. Results showed that all of the simulated 'participants' had noise levels below 150, and so the correlations were not weakened by any exclusions.

POmatrix 2 (points: 200 333 267 400)

Results showed that under POmatrix 2, there was a strong correlation between criterion 1 and the individual variance on the BAS parameter (r = .79, p <.001) and a weaker but still significant BAS*criterion 3 correlation (r = -.66, p <.001). The BAS*criterion 2 correlation was significant but very weak (r = .28, p <.001). All these relationships are illustrated in figure 10.10. Once again the simulated relationships are linear.

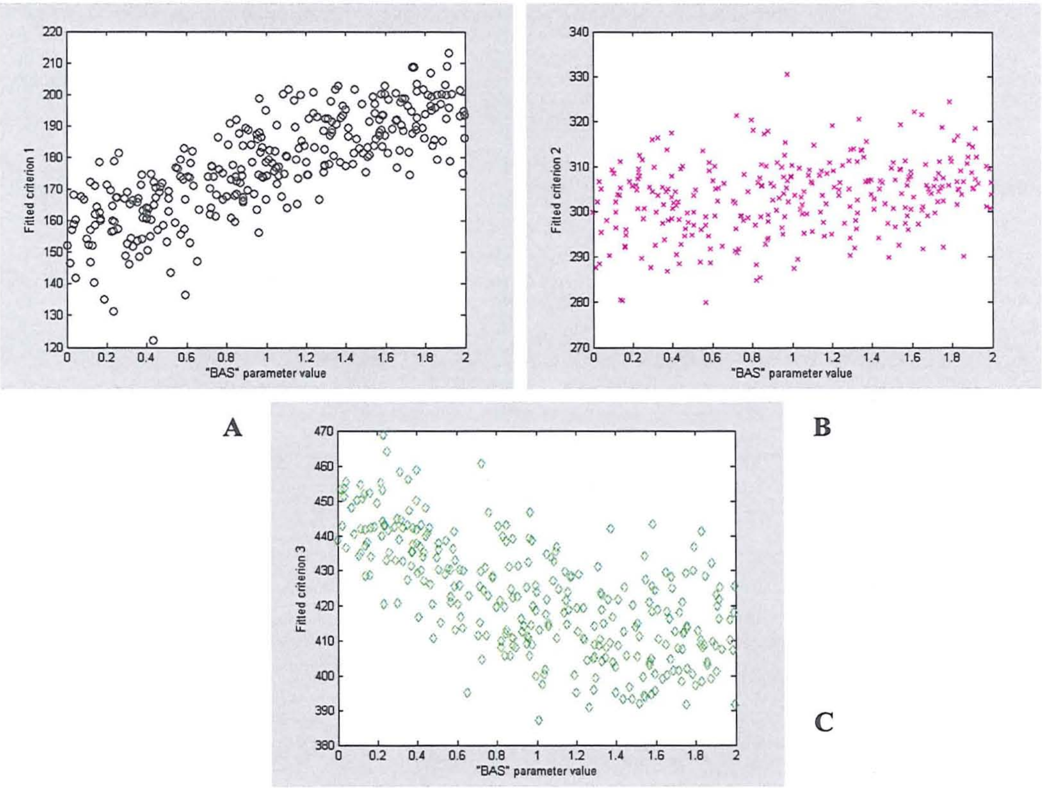


Figure 10.10. Scatterplots indicating the relationship between "BAS" variation and the fitted criteria under POmatrix 2 (N = 300). Figure A illustrates the relationship between criterion 1 and BAS variation, figure B the non-significant BAS*criterion 2 relationship and figure C the relationship between criterion 3 and BAS variation

As observed in the human data, the simulated data showed that the uni-dimensional model was non-significantly worse than the saturated model for 12 'participants' out of the whole sample (N =300). Thus, approximately 4% of the sample was well-fit by the uni-dimensional model. This ratio is roughly equivalent to the ratio observed in the human data where the uni-dimensional model was non-significantly worse than the saturated model for 2 out of the 30 (7%) participants who performed under PO matrix 2. It was decided to apply the same exclusion criterion applied to the human data and, therefore, to exclude data sets with noise levels above 150. Results showed that all of the simulated 'participants' had noise levels below 150, and so the correlations were not weakened by any exclusions.

These preliminary results support the efficacy of the task to examine both behavioural responses to asymmetric payoffs and the relationship between behaviour and inter-individual variation. Indeed, the simulated data shows that there seems to be a strong relationship between the simulated biology and the behaviour. The fact that a strong biology*behaviour association is simulated suggest that the task may be able to uncover the personality*behaviour relationship. Nonetheless, as discussed above, the trait*behaviour correlation would be weaker due to the generally weaker association observed between task measures and personality (Franken and Muris, 2006; Pardo et al., 2007).

10.4.4.1.3 Individual variance on the RPE signals projecting to the synapses on the reward prediction (RP) cell

Simulations conducted with this sub-model showed a good fit to the human data (table 10.19) although the model was not able to capture the relationship between the BAS and the criteria. Indeed, the correlations were mostly weak and non-significant.

Table 10.19. Simulated criteria (s.d) with BAS on the RP cell compared to the human data (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Criterion 1	Criterion 2	Criterion 3	Noise
Sim_PO1	179.08 (10.7)	299.87 (8.9)	433.38 (11.5)	81.53 (15.0)
Sim_PO2	180.16 (10.4)	303.72 (8.1)	417.64 (12.6)	78.53 (13.5)
Hum_PO1 **(N = 18)**	182.97 (26.4)	300.13 (14.0)	420.84 (23.8)	80.27 (31.7)
Hum_PO2 **(N = 20)**	191.77 (27.3)	303.21 (21.0)	413.31 (20.7)	79.62 (32.3)

POmatrix 1

These simulations showed that the correlation between BAS and criterion 1 was significant but weak (r = .20, p <.001) and so were the correlations between BAS variation and criterion 2 (r = .36, p <.001) and criterion 3 (r = -.20, p <.001). The scatterplots (figure 12.11) indicate that the relationships between BAS variation and criteria 1 and 3 are very weak although it shows the same patterns simulated by the model where BAS mediates the RPE projections to the SR pathway. However, an effect on criterion 2 has not been observed in the previous simulations.

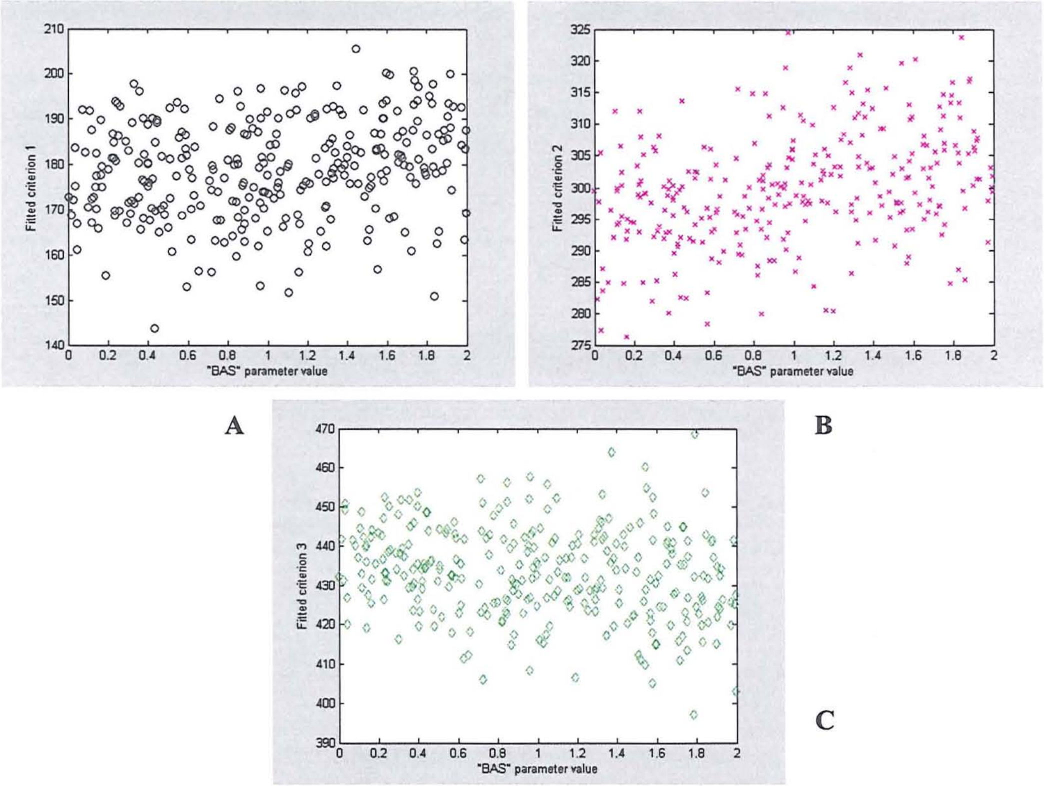


Figure 10.11. Scatterplots indicating the relationship between "BAS" variation and the fitted criteria under POmatrix 1 (N =300). Figure A illustrates the relationship between criterion 1 and BAS variation, figure B the non-significant BAS*criterion 2 relationship and figure C the relationship between criterion 3 and BAS variation

As observed in the human data, the simulated data showed that the uni-dimensional model was non-significantly worse than the saturated model for 9 'participants' out of the whole sample (N =300). Thus, approximately 3% of the sample was well-fit by the uni-dimensional model. This ration is roughly equivalent to the ratio observed in the human data where the uni-dimensional model was non-significantly worse than the saturated model for 7% of the sample who performed under PO matrix 1. It was decided to apply the same exclusion criterion applied to the human data and, therefore, to exclude data sets with noise levels above 150. As observed in the simulations using model 2, results showed that all of the simulated 'participants' had noise levels below 150.

POmatrix 2

Similarly, under PO matrix 2 the correlation between criterion 3 and BAS variance was weak but significant (r = .23, p <.001) and so was the correlations between BAS variance and criterion 2 (r = .14, p = .02). Criterion 1 was not significantly correlated with individual differences (r = .09, ns). These relationships are reported in figure 10.12.

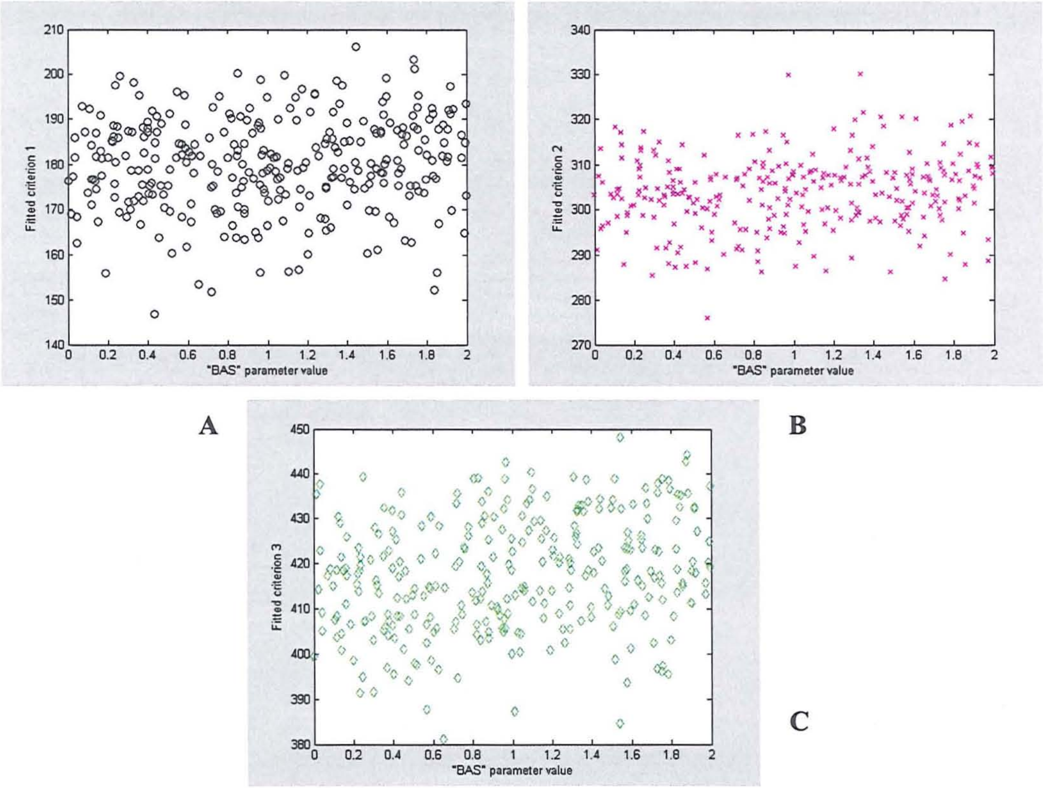


Figure 10.12. Scatterplots indicating the relationship between "BAS" variation and the fitted criteria under POmatrix 2 (N = 300). Figure A illustrates the relationship between criterion 1 and BAS variation, figure B the non-significant BAS*criterion 2 relationship and figure C the relationship between criterion 3 and BAS variation

As observed in the human data, the simulated data showed that the uni-dimensional model was non-significantly worse than the saturated model for 1 'participants' out of the whole sample (N =300). It was decided to apply the same exclusion criterion applied to the human data and, therefore, to exclude data sets with noise levels above 150. As observed in the simulations using model 2, results showed that all of the simulated 'participants' had noise levels below 150.

Thus, as in the two previous studies, the latter sub-model is not able to simulate the behaviour*BAS relationship. In contrast, the sub-model with personality acting as a multiplier on the reward input to the DA cell (model 1) and the one with personality as a multiplier on the RPE signals projecting to the SR cell (model 2) are the best candidates to explain the human data. However, model 1 captures a curvilinear relationship between BAS variation and the criteria (cf. the linear relationship in the human data) and simulates noise levels higher than the levels observed in the human data. Thus, it reinforces the earlier view that model 2 might represent the best candidate to explain the human data and it is assessed in the next section.

10.4.5 Simulations using normally distributed 'BAS' variance

10.4.5.1 Learning phase

In the preliminary simulations, sub-model 1 and 2 appeared to be the best candidates to capture the behavioural data obtained in the present study. However, model 1 seems the less good candidate owing largely to the non-linearity of the simulated BAS*behaviour relationships. In this section, the two models will be further tested to assess whether they can efficiently simulate the present data using normally distributed simulated personality variance. Additionally, these extra analyses may identify the model that best explains the human data.

In order to assess the efficacy of the model at simulating the empirical data, it was decided to simulate 10 subgroups of 40 participants each in order to reproduce a sample size similar to the human sample (N=38), whose data was reported in the result section. Each sub-group should contain roughly equal numbers of simulated 'participants' from each payoff condition as in the human data (PO matrix 1 = 18 vs. POmatrix2 = 20). Thus, it was necessary to simulate a total of 400 data sets (200 in each payoff condition).

10.4.5.1.1 Individual differences on the reward input to the DA cell (model 1)

By exploring the scatterplots presented above (figures 1 and 2), it can be seen that there are quite strong but non-linear relationships between personality and criterion 1 as well as personality and criterion 3 under both payoff matrices. These relationships are however roughly linear for simulated individual differences corresponding to BAS parameter values below 0.5; whereas values above 0.5 are associate to flat, non-significant relationships.

Hence, it was decided to conduct the simulations for individuals with a BAS parameter mean equal to 0.4 and a standard deviation equal to 0.1 (values of the BAS parameter falling at or below zero were set to 0). The s.d. value cannot be increased much beyond this as, at the upper end, the values start to involve more and more of the flat portion of the simulated BAS-behaviour curve and, thus, are not effective as a means of adding individual differences. The model was not able to capture the "noise" values shown by human participants (table 10.20). In this case, noise is the value of interest to evaluate the goodness of the model fitting. The higher a participant's noise value the less sharp the implemented category boundaries. Obviously, by moving from the uniform distribution simulations (mean =1, range 0-2), to a normal distribution with a mean of 0.4, chosen so as to fall on the most linear portion of the BAS-behaviour simulation curve, this change of mean may have an effect on the quality of the simulations. Additionally, a BAS mean value equal to 0.4 causes the strength of the reward effect to be low (0.4 vs. 1 on average).

Table 10.20. Simulated criteria for model 1 using normal distributed variance (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Criterion 1	Criterion 2	Criterion 3	Noise
Sim_PO1	144.05	298.49	467.84	181.46
(N = 200)	(20.8)	(11.0)	(22.5)	(30.0)
Sim_PO2	154.26	297.10	437.72	163.86
(N = 200)	(17.8)	(9.6)	(18.1)	(23.8)
Hum_PO1	182.97	300.13	420.84	80.27
(N = 18)	(26.4)	(14.0)	(23.8)	(31.7)
Hum_PO2	191.77	303.21	413.31	79.62
(N = 20)	(27.3)	(21.0)	(20.7)	(32.3)

As for the human data, it was decided to exclude from further analysis those participants that had noise levels above 150 pixels. After applying this exclusion criterion, ninety simulated participants only were retained (29 from PO matrix 1 vs. 61 from PO matrix 2). The criteria and noise levels of these participants are reported in table 10.21, which shows that the simulated noise levels were still higher than the human noise levels.

Table 10.21. Simulated criteria for model 1 for those participants with noise levels below 150 pixels (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Criterion 1	Criterion 2	Criterion 3	Noise
Sim_PO1	164.41	299.45	446.66	139.49
(N = 29)	(10.1)	(8.6)	(10.2)	(6.5)
Sim_PO2	167.56	298.44	424.53	138.00
(N = 61)	(10.6)	(8.9)	(13.2)	(8.0)
Hum_PO1	182.97	300.13	420.84	80.27
(N = 18)	(26.4)	(14.0)	(23.8)	(31.7)
Hum_PO2	191.77	303.21	413.31	79.62
(N = 20)	(27.3)	(21.0)	(20.7)	(32.3)

Owing to the low retention rate and the simulated high noise levels, the model does not seem good at capturing the human data and, therefore, not a good candidate to explain human behaviour. Thus, no further analyses were conducted on the data simulated using this model.

10.4.5.1.2 Individual differences in the effect of the RPE signals projecting to synapses on the SR cell (model 2)

By exploring the scatterplots presented above (figures 10.3 and 10.4), it was noticed that there were quite strong relationships between BAS variation and criterion 1 as well as personality and criterion 3 under both payoff matrices. The figures also suggest that a normal distribution with a mean of 1.1 could be used to capture the human data. In order to maximise individual differences, a large standard deviation relative to the mean was adopted. In fact, a standard deviation equal to 0.25 was used to simulate the human data. Again simulated BAS values randomly falling at or below 0 were set equal to 0.

Overall analysis of the total sample showed that these values could capture the mean data (i.e. criteria and noise level; table 10.22) as well as the BAS*behaviour relationship (see below).

Table 10.22. Simulated criteria for model 1 using normal distributed variance (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Criterion 1	**Criterion 2**	**Criterion 3**	**Noise**
Sim_PO1 (N = 200)	181.05 (11.5)	300.35 (8.1)	430.96 (11.8)	83.16 (11.4)
Sim_PO2 (N = 200)	181.96 (11.3)	303.97 (8.0)	417.19 (12.3)	80.04 (11.4)
Hum_PO1 (N = 18)	182.97 (26.4)	300.13 (14.0)	420.84 (23.8)	80.27 (31.7)
Hum_PO2 (N = 20)	191.77 (27.3)	303.21 (21.0)	413.31 (20.7)	79.62 (32.3)

It was decided to apply the same exclusion criterion applied to the human data and, therefore, participants with noise levels above 150 were excluded from further analysis. However, none of the simulated participants had such high noise scores and, therefore, the whole sample was retained (N = 400; 200 for each PO condition).

PO matrix 1

According to the human data, the simulations identified a moderately strong and significant correlation between criterion 1 and BAS variation (r = .46, p <.001) whereas the correlation between BAS and criterion 3 was significant, although a little weaker (r = -.37, p <.001). As expected, criterion 2 did not significantly correlate with personality variation (r = .12, ns). These relationships are represented in figure 10.13.

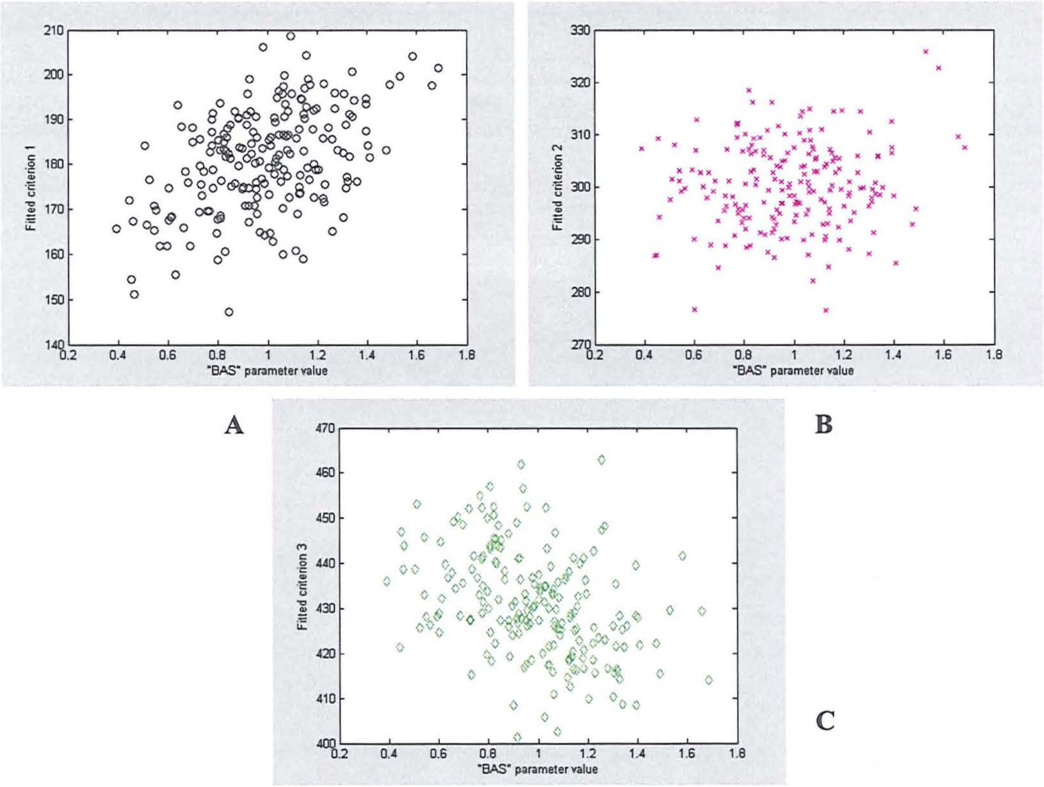


Figure 10.13. Scatterplots indicating the relationship between "BAS" variation and the fitted criteria under POmatrix 1 (N =200). Figure A illustrates the relationship between criterion 1 and BAS variation, figure B the non-significant BAS*criterion 2 relationship and figure C the relationship between criterion 3 and BAS variation

PO matrix 2

According to the human data, the simulations identified a strong and significant correlation between criterion 3 and BAS variation (r = .52, p <.001) whereas the correlation between the BAS and criterion 3 was significant, although weaker (r = -.32, p < .001). There was also a weak, but significant, correlation between

BAS variation and criterion 2 (r = .17, p = .02). These relationships are represented in figure 10.14.

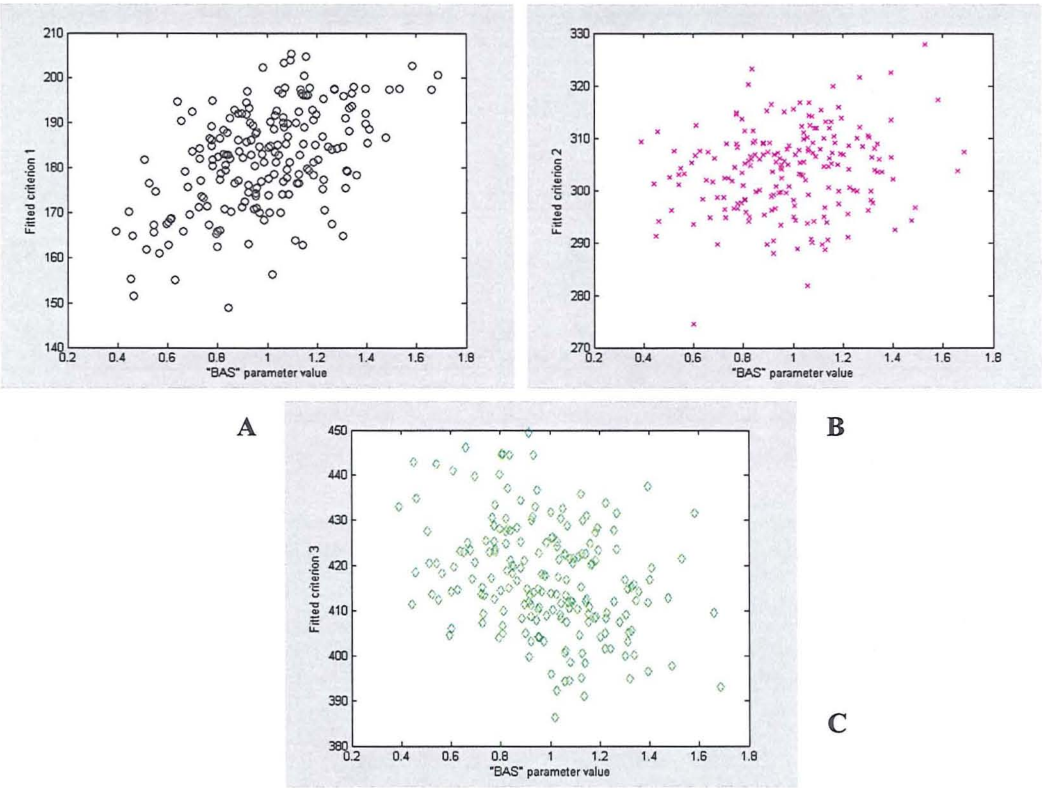


Figure 10.14. Scatterplots indicating the relationship between "BAS" variation and the fitted criteria under POmatrix 2 (N = 200). Figure A illustrates the relationship between criterion 1 and BAS variation, figure B the non-significant BAS*criterion 2 relationship and figure C the relationship between criterion 3 and BAS variation

Further analyses were conducted on the 10 sub-groups within the 400 simulated data sets. Each subgroup consisted of 40 simulated 'participants' and within each subgroup half of the 'sample' was simulated using POmatrix 1 (PO 1) and the remaining half using POmatrix 2 (PO 2). Table 10.23 reports the mean and standard deviation values of the criteria and noise level observed in these 10 samples across the two payoff conditions.

Table 10 23. Simulated criteria and noise level in the 10 sub-group (GPS)

GPS	Criterion 1	Criterion 2	Criterion 3	Noise
1				
PO 1	183.87 (10.0)	301.15 (5.0)	428.31 (8.8)	83.64 (9.7)
PO 2	184.81 (9.1)	304.37 (6.7)	417.24 (7.6)	80.82 (8.5)
2				
PO 1	180.33 (9.7)	298.62 (9.7)	430.31 (11.6)	85.2 (7.0)
PO 2	180.68 (9.9)	302.79 (8.7)	415.08 (11.2)	83.70 (8.9)
3				
PO 1	176.46 (10.4)	298.03 (9.1)	432.09 (13.1)	84.80 (13.4)
PO 2	176.38 (11.1)	299.60 (10.9)	413.83 (10.2)	79.07 (15.3)
4				
PO 1	182.52 (10.8)	297.34 (7.5)	428.16 (12.6)	79.26 (12.4)
PO 2	183.40 (10.7)	301.62 (5.6)	411.86 (12.3)	75.48 (11.5)
5				
PO 1	179.75 (12.6)	303.09 (6.8)	433.68 (10.6)	78.40 (12.9)
PO 2	181.38 (13.0)	306.71 (7.3)	422.61 (11.2)	77.39 (12.9)
6				
PO 1	178.01 (11.2)	300.09 (7.7)	433.05 (12.6)	85.53 (11.4)
PO 2	178.54 (11.7)	304.43 (7.9)	418.80 (13.6)	82.92 (11.4)
7				
PO 1	185.41 (11.7)	301.68 (8.4)	433.87 (10.4)	79.57 (12.0)
PO 2	187.01 (12.0)	305.94 (7.5)	421.73 (13.6)	77.70 (10.8)
8				
PO 1	181.29 (14.1)	300.50 (9.4)	429.77 (14.6)	83.67 (12.2)
PO 2	182.70 (13.1)	304.13 (7.6)	416.24 (12.8)	79.95 (10.2)
9				
PO 1	181.00 (9.7)	304.52 (8.8)	430.00 (12.0)	85.70 (11.9)
PO 2	182.41 (8.2)	308.36 (8.2)	418.26 (14.2)	82.28 (12.7)
10				
PO 1	181.83 (13.1)	298.47 (6.4)	430.37 (12.1)	85.81 (8.4)
PO 2	182.32 (12.4)	301.80 (6.5)	416.26 (12.2)	81.09 (9.7)

Ten separate mixed-design ANCOVAs were run with criteria as a repeated-measure factor and payoff matrix as a between-measure factor, whereas BAS variation was entered as a covariate. The ANCOVAS aimed to assess how different factors affected learning and, in particular, criterion location. Results showed that the main effect of criteria was significant for all 10 subgroups (Fs > 322.1, ps <.001) while the 2-way interaction between BAS and criteria was significant for 9 of the 10 subgroups (Fs > 5.8, ps < .006). The main effect of payoff condition was non-significant in all groups (Fs < 2.1, ps > .16). Moreover, the 2-way interaction between criteria and payoff matrix condition was non-significant for all groups (Fs < .09, ps >.89) and neither was the 3-way interaction between BAS, criteria and PO matrix (Fs <.02, ps > .98). These results indicate that individual variation in RPE sensitivity mediated performance (i.e. criterion location) whereas the payoff matrix had no impact on performance, as observed in the human data.

It was, subsequently, decided to explore the strength of the behaviour*BAS variation relationship in each sub-groups (table 10.24). The correlations were run for each subgroup regardless of payoff condition since the ANCOVA's results showed that this variable did not affect performance.

Table 10.24. Criteria*BAS variation correlations across the 10 sub-groups

	Criterion 1	**Criterion 2**	**Criterion 3**
Subgroup 1	r = .16, ns	r = -.002, ns	r = -.01, ns
Subgroup 2	r = .67, p <.001	r = -.03, ns	r =-.46, p=.003
Subgroup 3	r =.46, p =.003	r = .13, ns	r = -.25, ns
Subgroup 4	r =.35, p = .03	r =.32, p = .04	r = -.33, p= .04
Subgroup 5	r =.66, p <.001	r =.28, p = .08	r = -.22, ns
Subgroup 6	r =.45, p=.004	r =.33, p =.04	r = -.17, ns
Subgroup 7	r =.58, p <.001	r = -.07, ns	r= -.41, p=.009
Subgroup 8	r =.55, p <.001	r =.33, p = .04	r= -.31, p=.051
Subgroup 9	r =.40, p= .010	r = .04, ns	r = -41, p=.008
Subgroup 10	r=.59, p<001	r =.35, p=.03	r = -.25, ns

The simulated data indicated that the strongest correlation was between criterion 1 and BAS variation, which was positive and significant in all but one of the 10 sub-groups. In contrast, the negative correlation between criterion 3 and BAS variation was only found to be significant in four sub-groups and just failed to be significant in another group (r = -.31, p = .051). Finally, criterion 2 was also found to be significantly, although weakly, correlated with BAS in four subgroups and the correlation was a non-significant trend in one further group (r = .28, p .08).

In the analysis of the human data, a difference score was calculated by subtracting criterion 1 scores from criterion 3 scores. This difference score was found to be positively correlated with participants' score on the delay discounting (DD) task (rho = .43, p = .007). Similarly, the simulated data could capture this relationship as indicated by the fact that the simulated personality variance was negatively correlated with the difference score (r = -.53, p <.001) across the overall simulated sample (N = 400). Additionally, this correlation was significant in 9 of the 10 simulated sub-groups (rs > -.41, ps <.009).

10.4.5.1.3 Simulations overview

Overall, the simulations indicated that the sub-model with BAS acting on the RPE projections to the SR cell was the best candidate model to capture and further understand the human data. In fact, the criteria and, especially, the noise levels simulated using model 2 are closer to the values displayed by human participants (cf. model 1). The model can also potentially simulate the correlation observed between criterion 2 and personality (i.e. positive schizotypy) observed in the human data. However, this effect was quite weak and observed only in a third of the sub-groups and was equivalent in size to the human sample. It is, therefore, unlikely that the present study procedure, used with human participants, would be able to detect it. One also needs to consider what it would mean for a specific simulated individual differences effect in the biology to underlie a correlation between delay discounting and some aspects of task performance (for criteria 1 and 3) and underlie a correlation between schizotypal personality and another index of task performance (criterion 2). This would mean that delayed discounting and schizotypal personality would share underlying biology and should correlate. In the present study positive schizotypy and DD scores did not correlated (rho = .08, ns); similarly there is no literature that reports relationships between delay discounting and positive schizotypy. One study by Heery and colleagues (2007) showed that schizophrenic patients, on antipsychotic medication, showed steeper discounting during a DD task than healthy controls. Delay discounting was found to be negatively correlated with negative symptoms whereas no correlation was found between DD scores and positive symptoms.

10.4.5.2 Reversal phase

As mentioned above, owing to the small number of trials included for each category in the reversal phase, it was not possible to use the formal model to fit the data. Thus, the simulated performance levels, indexed by the proportion of correct responses (i.e. accuracy) were analysed. Following the results that indicate that model 1 could not capture the human data, accuracy scores were only simulated using model 2 (BAS variation on the RPE projections to the actor unit).

Simulations of the last 60 trials, which constitute the reversal phase, were run straight after the simulations of the learning phase. During the payoff switch phase, the **rf** values were coded as 1.17 and 0.33 which represent a linear transformation of the actual reward values, i.e. 467 and 133, respectively. In other words, simulations included the total 240 trials that were run as one long session. In this way performance during the reversal phase would be affected by the weight changes occurred during the learning phase. It was decided to simulate 10 subgroups of 40 participants each in order to reproduce a sample size similar to the human sample (N=38) whose data was reported in the result section. Each sub-group should contain equal numbers of simulated 'participants' from each switch condition as in the human data (lower = 19 vs. higher = 19). Thus, it was necessary to simulate a total of 400 data sets (200 in each switch condition).

The simulated accuracy scores are reported in table 10.25 and they are close to the accuracy scores obtained by the human participants in both switch conditions.

Table 10.25. Proportion of correct responses across the four categories in the simulated and the human participants in the two switch conditions (sim = simulated; hum = human)

	Proportion of correct responses			
	Cat A	Cat B	Cat C	Cat D
Sim higher **(N = 200)**	0.89 (.07)	0.66 (.09)	0.53 (.11)	0.68 (.09)
Sim lower **(N = 200)**	0.81 (.08)	0.56 (.10)	0.45 (.10)	0.61 (.08)
Hum higher **(N = 19)**	0.82 (.10)	0.65 (.16)	0.63 (.13)	0.78 (.06)
Hum lower **(N = 19)**	0.85 (.09)	0.56 (.21)	0.56 (.16)	0.73 (.16)

Further analyses were conducted separately on the 10 subgroups in order to assess whether the neural model could simulate the results obtained in the human sample. Thus, 10 mixed-design ANOVAs were run to assess the impact of category type and switch condition on accuracy scores. In line with the human data, results showed that the main effect of category type was significant (Fs> 87.5, ps <.001) in all 10 groups. As in the human data, greater accuracy was achieved for those categories closer to the external edge of the display box (categories A and D). Similarly, the main effect of switch condition was also significant (Fs > 17.7, ps < .001) in the 10 groups. Indeed, overall 'participants' simulated under the high reversal condition showed greater accuracy (0.69 and s.d. = .04) than those simulated under the lower condition (0.61 and s.d. = .06). These results compare with the corresponding figures for the human participants [switch_467: 0.72 (.06); switch_133: 0.67 (.08)]. In line with the human data, the 2-way interaction between category and switch condition was non-significant in all 10 groups (Fs < 1.8, ps >.17). These results replicate the human data and indicate that the neural model is able to simulate the accuracy pattern observed in the human data (as well as the categorical bounds). As for the human data, it was decided to merge the accuracy score for each category across the two switch conditions (table 10.26).

Table 10.26. Simulated and human accuracy scores across the four categories regardless of switch condition

	Proportion of correct (pc) responses			
	Cat A	Cat B	Cat C	Cat D
Simulated **(N = 400)**	0.85 (.09)	0.61 (.11)	0.49 (.11)	0.64 (.09)
Human **(N = 38)**	0.84 (.10)	0.61 (.19)	0.59 (.15)	0.75 (.12)

The accuracy scores for the four categories were above chance levels (i.e. 0.25; 1-sample t-tests: $t > 42.7$, $p < .001$). Overall, these results show that the model was able to capture the proportion of correct responses (i.e. accuracy scores) scores observed in the human data during the reversal phase.

10.5 Discussion

The present behavioural results replicate and confirm the results obtained in the previous chapters. In fact, participants shifted their decision bounds away from the perceptual bounds to maximise performance (i.e. they demonstrate a criterion shift). The use of the neural model offered a further insight into the human data. The simulation results confirmed that the criterion shift processes were mediated by the implicit system whereas the explicit system was involved in solving the categorisation problem by initially setting perceptually-driven decision boundaries and using these as the basis of uni-dimensional rules to systematically classify the stimuli into the different categories. As previously observed, the criterion shift was found to be mediated by individual differences. In particular, results confirm a primary role of impulsivity as indexed by the behavioural measure of the delay discounting. More specifically, low impulsive individuals were found to show a greater shift than their high scoring counterparts for criterion 1 and for criterion 3.

The neural model produced similar relationships between criteria 1 and 3 and the biological mechanism underlying individual differences. Nonetheless, the model captured a trait*behaviour relationship opposite in sign to the human impulsivity data; in fact the model showed that high trait individuals showed the greatest shifts in response bias. Taking the empirical and simulated data into account jointly, the present results suggest that low impulsivity scores might correspond to high trait reactivity (at least as reactivity is captured in neural model 2; i.e. elevated synaptic plasticity in the striatum in response to RPE signals). As expected the simulated relationships were stronger than those observed in the human data, since the biology*task behaviour associations are usually stronger and easier to detect than trait*task behaviour associations (Munafo et al., 2003; Cohen et al, 2005).

The behavioural measure of impulsivity indexed by delay discounting (hyperbolic K) was found to be the only personality measure to relate to criterion shifts significantly and, therefore, to the underlying learning processes responsible for the shifts. These results seem to support the claims that behavioural, compared to self-report, measures are a better index of individual differences (Lee et al., 2007; Eisenberg et al., 2007). Nonetheless, delay discounting (DD) scores were positively correlated with scores on the Dickman's dysfunctional impulsivity (DI) scale although not with the ImpAss component (SSS, OLIFE-impulsivity/non-conformity and EPQ-P). These correlational results indicate that both the DI scores and the DD scores measured a similar construct of impulsivity whereas the ImpAss component may have captured somewhat different or varied facets of the trait. Overall, these results indicated that the delay discounting scores may be a more reliable endophenotype of BAS reactivity and variation than the more commonly employed phenotypic measures of impulsivity. At least, this seems to be true for the relationship with performance on a reward-related CL task.

Moreover, because the delay discounting scores may have been found to be associated with the criteria placed in the CL task, this supports the existing arguments that claim that they rely on the same biological mechanism. Indeed, there is evidence that suggest that decision-making during the DD task is guided RPE-DA signals (Pessiglione et al., 2006; Kalenscher and Pennartz, 2008). The neural model of the thesis also showed that during the CL task implicit learning (i.e. as reflected in criterion placement) is guided by the dopaminergic RPE signals. According to the 3-factor model, RPEs act as a teaching signal that strengthens the SR association to, subsequently, guide responding and decision-making (Pickering and Gray, 2001; Schultz, 2002).

Two neural sub-models showed that BAS variation affected the activation of the implicit system in such a way as to affect the criterion shifts. One sub-model suggested that BAS affects reward processing by acting on the reactivity of the DA cell to reward whereas the second sub-model had BAS mediating SR learning by acting on the cortico-striatal synapses of the actor cell (Joel et al., 2002). The results showed that both sub-models were able to simulate the human data to some extent. In fact, they both captured the criterion shift and the relationship between BAS variation and the criteria. However, the sub-model with BAS acting on the SR cell could better capture the human data (i.e. mean criteria and, particularly, noise levels). Additionally, this model could capture the linear relationship between the criteria and the BAS variance, as observed in the human data. In conclusion, the simulation results indicated that BAS variation is possibly related to responsivity to the dopaminergic teaching RPE signal, which is responsible for the strengthening of SR associations according to the 3-factor rule. The simulations make it less likely that BAS variation is related to responsivity to reward *per se*. This argument holds at least for BAS variations that affect criterion shifts in the present CL task. These findings replicate the results from the previous study where the data could be simulated by the sub-model with the BAS on the SR cell but not the sub-model with BAS on the reward cell. These results offer a new view of RST, revising the assumption of BAS variation as responsivity to reward and reward-related stimuli which, subsequently, lead to approach behaviour (Corr, 2006). This older conjecture was based on theoretical assumptions rather than empirical data. The present data bring theory and data much closer together. Nonetheless, the present findings are not fully conclusive and, therefore, further research is necessary to confirm them.

One of the main limitations with the present study is due to the fact that the perceptual decision bounds do not correspond with the optimal accuracy bounds. As the learning process moves the criteria from the (assumed) initial perceptual, symmetric bounds, the shift observed is small owing to the relatively small number of trials per category and the gradual nature of procedural learning under the 3-factor rule. Thus, one cannot confidently determine whether the shift is moving towards optimum accuracy or optimum winnings as both these optimal bounds are shifted in the same directions away from the initial perceptual bounds. However, the simulations reported in the previous chapters have indicated that shifts away from the optimal accuracy bounds, towards the optimal reward bounds, are simulated by the model.

The simulations suggested that participants might have initially split the display box (600*600-pixel) into 4 equally spaced areas using the perceptual-symmetry bounds (i.e. at 150, 300 & 450 pixels). These perceptually-driven bounds could then be used to formulate their dimensional strategy. However, over trials, and under the influence of implicit learning processes mediated by RPE signals, participants would first have shifted their criterion to a position closer to the optimal accuracy bounds (200, 300 & 400 pixels), which creates four unequally spaced regions. In this way, participants' performance increases accuracy-related winnings. It could be suggested that increasing the number of trials further may have allowed participants to further maximise their winnings by shifting the criterion away from the accuracy bound closer to the optimal reward bound. However, this may not be the case since the neural model indicated that the human data were best simulated by implementing a symmetric, rather than the actual asymmetric, payoff matrix. Participants' self-report support this observation. During debriefing, several participants reported not having paid much attention to the visual feedback (i.e. points earned). Although, they were mostly able to recall the 400-point feedback signals and none of the other point values. Since the task was quite complex and demanding, participants may have relied solely on the auditory feedback in order to limit the effort of processing feedback and direct them to solve the categorisation problem. It is, therefore, possible that participants relied more strongly on the auditory feedback than the visual feedback and this would have not allowed them to notice that different categories earned different amount of points for correct responses. Thus, contrary to arguments raised earlier, based on the results obtained in the 2-category tasks, the inclusion of auditory feedback may have limited the impact of the visually-presented point feedback due to higher cognitive demand of the

4-category task. Future replications should exclude auditory feedback in favour of the visual feedback, presenting both the potential and the actual amount of points (see chapter 9). Moreover, alternative task designs, where the perceptual and the optimal accuracy criteria overlap, may be more suited to detect a response bias towards the optimal reward criterion (see chapter 11 for examples).

Owing to the small number of trials included in the reversal phase, it was not meaningful to extract the criteria values using the formal model. However, overall analysis of performance (i.e. accuracy levels) indicated that learning continued over the trials of the reversal phase. Accuracy was assessed over four blocks of 60 trials (3 from the learning task and last from the reversal task). Accuracy increased in a linear fashion across the four blocks and it was greatest in the last block (i.e. reversal phase). Participants who scored low on neuroticism were found to show greater learning over the whole task than their high-scoring counterparts. Stable individuals may have a chronic promotion focus (cf. prevention focus in neurotics) that fits the induced focus (i.e. low N participants may experience a regulatory fit; Maddox et al., 2006). Regulatory fit is expected to improve cognitive flexibility and, therefore, may have enhanced performance for stable (cf. neurotic) participants.

Overall, these results indicate that participants carried the strategy learnt in the first task over to the second task despite the fact that they were presented as two independent tasks. However, results indicated that participants who performed under the higher switch condition achieved higher accuracy than those participants under the lower switch. This effect was also captured by the simulations. It is possible that by the switch phase, participants had efficiently learnt to solve the categorisation task as a direct result of practice. Indeed, they seem to have applied the same strategy used in the learning phase. Thus, participants may have had more available cognitive resourced to process visual feedback. As participants mostly noticed the 400 points in the learning phase, comparison of the switch (467 vs. 133) points to 400 points may have affected RPE-driven performance on the task. Indeed, receiving 467 points would have produced a positive RPE signal which would have further strengthened the correct SR association (e.g. stimulus on the far left-hand side of the screen belongs to category A).

Nonetheless, the present results confirm that implicit reward-driven learning, as indexed by criterion shifts to maximise performance, is affected by individual differences. Simulated high BAS individuals have been found to show greater learning thanks to their increased sensitivity to the RPE-DA signals which strengthens correct SR associations. Certain impulsivity measures were the ones which reproduced the simulated pattern of BAS-behaviour correlations. Therefore, these observations are in line RST's original assumptions (Gray, 1987), the present study identifies impulsivity as the personality component that mediates learning and, therefore, impulsivity may represent the BAS-trait. However, contrary to the original assumption high BAS activity was found to be associated with low scores on impulsivity. Similar results had been observed in the study in the previous chapter. Nonetheless, further replication is required to confirm these findings.

Chapter 11

Summary and discussion

The thesis offered a detailed review of the literature of category learning (CL), dopamine and the reinforcement sensitivity theory (RST) and drew the links between them. Owing to the evidence that indicates how performance on CL tasks is mediated by dopaminergic functioning, and because the BAS has a dopaminergic nature, the CL paradigm was identified as a useful tool to explore reward-related learning and identify the BAS-related traits (Ashby et al., 1998; Pickering and Gray, 2001; Pickering, 2004; Corr, 2006).

11.1 General aims of the thesis

The initial three chapters set the premises upon which the experiments reported in the thesis were designed and developed. The main assumption made was that biologically-based personality traits would mediate learning. More specifically, BAS-related traits were expected to mediate performance on CL tasks as a function of reward manipulations. Hence, the first aim of the thesis was to further explore reward-mediated learning. The second aim was to establish the relationship between reward-related learning and personality. In particular, the aim was to identify the underlying personality traits of the BAS system, owing to the dopaminergic nature of both the BAS and reward-dependent learning (Ashby et al., 1998; Schultz, 1998; Gray, 1987; Corr, 2006). Presently, there is contrasting evidence that indicates that either extraversion or impulsivity may be good candidates (e.g. Gray, 1987; Depue and Collins, 1999; Franken and Muris, 2006). Hence, the empirical data reported in the thesis bring further empirical evidence into the ongoing debate to help the field advance towards a solution.

One key aim of the thesis, which arose during the process, was to develop a neural model that could capture dopaminergic activity during reward-related learning and it was able to simulate the behavioural data obtained in the empirical studies. Moreover, the implementation of a neural model allows one to test initial predictions prior to the actual data collection and, therefore, to construct sound empirical tests that are able to explore the interrelation between personality and behaviour.

Another key theme of the thesis was to use formal models of the task to clarify the strategy and processes employed by participants in order to extract the appropriate dependent variables (DVs) from task performance (hence, excluding potential confounding measures). In order to do so, formal signal detection theory 'decision-bound' models were fitted to the performance data using maximum likelihood estimation.

In the initial studies the results were contrasting. In particular, the 'Gabor' tasks (chapter 6) failed to observe a direct behaviour-personality relationship. This may be a result of the fact that the relationship between behaviour (endophenotype) and personality traits[22] (phenotypes) are quite weak (Blum et al., 2000; Munafo et al., 2003) and, therefore, it may be hard to detect this relationship in noisy empirical data (Lee et al., 2007; Eisenberg et al., 2007).

As discussed in chapter 8, the use of a neural model to simulate the relationship between the biological mechanism and behaviour may represent a good indicator of the trait-behaviour relationship, driven by their common genetic nature. Nonetheless, genetic factors only account for 30-60% of the variance in individual differences (Blum et al., 2000; Reuter et al., 2006). Any single biological parameter is likely to be accounting for a much smaller proportion of the trait variance. Moreover, the Gabor tasks seemed to have been too hard for many participants to learn to apply the appropriate dimensional strategy and maximise performance. Indeed, the formal modelling showed that only a third of the sample used the relevant dimensional strategy in the rule-based task whereas none of the participants used the optimal rule during the information integration task, which was even more difficult. Finally, the possibility that motivational manipulations may affect performance differently depending on

---

[22] Personality measures, indexed by questionnaires, are referred to as the phenotype explained by BAS variation (i.e. biological substrate); whereas behavioural measures, indexed by performance on a CL task, are referred to as endophenotypes which are assumed to have a closer relationship with the underlying biology.

the task structure cannot be discounted. Indeed, there are several sub-types of categorisation tasks and within each type the category membership structure can vary (probabilistic vs. deterministic), as can the nature of the stimuli dimensions (continuous vs. discrete). The use of a neural model could also help test these hypotheses prior to data collection. Hence, simulating the data via a neural model represents an effective tool to explore the efficacy of the task design and, if necessary, improve it.

Following these observations, a neural model was developed which could corroborate whether the empirical CL task was a useful tool. It allowed us to explore the relationship between motivation, cognition and personality during an RB category learning task prior to actual data collection. Additionally, the neural model also helped gain further understanding of the human data.

This final section offers a synthesis of the work reported in the thesis and it describes both the implications and limitations of the findings obtained in the different studies. The main findings are discussed in relation to the existing literature, previous research and the rationale and aims that underlie the thesis. Finally, ideas and suggestions for future research are also presented, using the neural model where appropriate to anticipate possible findings.

11.2 Main findings

11.2.1 Impulsivity and decision-making

In the task reported in chapter 5, participants were presented with a decision-making task that required them to choose between an infrequent but immediate reward and a more frequent but delayed (10 seconds) reward. Results suggested that two different processes were active during performance on the task and that each process was mediated by one subtype of impulsivity. Hence, these results are in line with the literature that indicates that impulsivity is a multi-faceted factor and that each aspect of impulsivity mediates different impulsive behaviours (Evenden, 1999; Franken and Muris, 2005 & 2006; Smillie et al., 2007; Vigil-Colet, 2007). In particular, the results were consistent with the idea that rash impulsiveness and reward sensitivity may be distinguished (Dawe and Loxton, 2004). Rash impulsiveness measures [i.e. delay discounting (DD) performance, dysfunctional impulsivity and ImpAss trait scores] were found to be good predictors of 'response disinhibition'. Specifically, participants who scored high on these impulsivity measures tended to respond more frequently to the immediate but infrequent rewarding (i.e. low frequency reward; LR) stimulus than to the more frequently rewarding but delayed (i.e. high frequency reward; HR) stimulus. These results are in line with the literature that indicates that this type of responding is typically impulsive (Bickel et al., 1999; Wade et al., 2000). Additionally, the relationship between disinhibited behaviour and impulsivity scores was moderated by smoking status. Consistent with the literature, smokers scored higher on the impulsivity measures and showed greater disinhibited responding (e.g. DD scores and LR responses; Mitchell, 1999; Bickel, et al., 1999). Moreover, the relationship between disinhibited responding and impulsivity measures (i.e. ImpAss component) was dependent on smoking status.

On the other hand, the results suggested that 'reward sensitivity' (i.e. functional impulsivity; FI; Smillie and Jackson, 2006) mediated the response latencies in response to the stimuli that offered more frequent but delayed rewards (HR). Indeed, individuals who scored high on the FI sub-scale showed shorter reaction times in response to the frequently rewarded stimuli (i.e. HR; approach behaviour). Hence, the latter form of impulsivity seems to be a BAS-related process and its self-report measure may represent the BAS-related trait. Future studies exploring the impact of individual differences on response disinhibition should also aim to disentangle the potential effect of these two processes.

The results also offer insights concerning one of the main criticisms generally directed to the DD task, as a behavioural measure of impulsivity: namely that it generally represents imaginary reward scenarios with long delays (i.e. weeks and years; Kalenscher and Pennartz, 2008). The present results show that delay discounting (DD) scores are a good predictor of impulsive responding in the decision-making task with shortened delays and real rewards. Therefore, they support the validity of the delay discounting task as a measure of impulsivity. Indeed, the DD scores were a good measure of disinhibited responding and, therefore, impulsivity, in the study.

Owing to the strong relationship between the disinhibited responding on the DD task and on the decision-making task, it seems plausible to assume that they rely on similar processes and brain functions. There is growing evidence that indicates that delay discounting is a function of learning guided by the dopaminergic reward prediction error (RPE) signals (Daw and Doya, 2006; Kable and Glimcher, 2007; Kalenscher and Pennartz, 2008). These considerations are supported by the findings reported in chapter 10, where DD scores were found to be positively correlated with reward-related learning on a CL task, which is mediated by RPE signals, at least as captured by the neural simulations. It is possible that RPE signals may mediate decision-making on the task implemented in chapter 5 owing to its relationship with performance on the DD task. Therefore, it may well be possible to develop a neural sub-model capable of simulating decision-making on the 10-second delay task implemented in chapter 5, after implementing some changes to the existing neural model. The main change would require applying a hyperbolic temporal discounting function to formulate the reward prediction signal to delayed rewards. Since the decision-making task does not require setting a step-like boundary between stimuli in different categories (cf. a uni-dimensional CL task), decision-making may purely rely on the implicit system so that the explicit module may be less involved.

Moreover, reward magnitude should be coupled with reward probability in order to capture the impact of temporal discounting. Indeed, it has been suggested that humans may perceive temporal delays as uncertainty (i.e. probability; Kalenscher and Pennartz, 2008), which would also be easier to transform into a parameter and associate with a numerical value. There is neuroimaging and patient data which indicate that firing in the ventral striatum varies linearly as a function of reward prediction error under a monetary incentive delay (MID) task with varying reward probability (Abler et al., 2006; Juckel et al., 2006a; 2006b). Abler and colleagues (2006) defined the RPE as the difference between the expected probability of occurrence and the actual reward occurrence. This evidence indicates that RPE, as a function of either reward magnitude or probability, activate the same brain areas. Thus, it follows that the neural model developed in the thesis (chapter 7) should be able to simulate performance on a probabilistic paradigm such as the decision-making task in chapter 6. Indeed, as described above, relatively few changes may need to be applied to the model in order to capture the task requirement in the simulations. This future work would offer a test of the model as it would allow us to assess whether the same personality variations in the model explain the correlation between simulated performance on two distinct tasks.

11.2.2 Reward manipulation and performance on a CL task

The behavioural results obtained in chapters 8 and 9 showed that asymmetric payoffs in a CL task lead to a response bias in favour of the high payoff category. Participants, performing on an RB CL task, were found to over-classify probabilistic stimuli as members of the category that offered the greatest gains (i.e. they showed a response bias). The response bias was not observed during performance on the same RB task under symmetric payoffs. Thus, these results show that performance on an RB CL task is mediated by manipulations of reward magnitude. These results are somewhat counterintuitive under the COVIS model which postulates that performance on RB tasks should not be affected by the reinforcing and motivating properties of reward manipulations (Ashby et al., 1998; Maddox and Ashby, 2004). Indeed, existing evidence showed that learning during an RB task could occur with no trial-by-trial feedback (Ashby et al., 1999), following pure observational learning (Ashby et al., 2002) and regardless of the timing of the feedback (5-to-10-second delay; Maddox, Ashby & Bohil, 2003). However, in chapter 10, feedback may have enhanced accuracy during an RB task by facilitating the correct criterion shift, away from the perceptual and, towards the optimal accuracy bounds.

Reward manipulations were not found to affect learning *per se* since participants were able to develop a uni-dimensional strategy to classify probabilistic stimuli as members of different categories. However, reward manipulations were responsible for the response bias observed under asymmetric payoffs. The neural model simulations helped understanding the behavioural data in relation to the COVIS model. The RPE-based model was able to simulate the response bias under asymmetric payoffs but it failed to capture the step-wise category boundary function that characterised effective learning (i.e. development of a uni-dimensional rule). The RPE model is a purely DAergic model that explains implicit learning via the 3-factor learning rule (i.e. LTP and LTD), where DA projections from the SNpc into visual associative striatal cells strengthen the appropriate SR associations in corticostriatal synapses (Schultz, 1998; Joel et al., 2002; Seger, 2008). The present model closely resembles the implicit system described by COVIS which describes the basal ganglia as the main substrate of the implicit system and procedural learning as a function of the 3-factor rule and (Ashby et al., 1998; Ashby & Ennis, 2006; Ashby, Ennis & Spiering, 2007). The two models also indicate that the DA learning signal is derived from the difference between the expected and the actual reward, which has been found to guide learning in the basal ganglia (i.e. RPE; Schultz, 1998; Ashby et al., 2007).

The step-like category boundary function was captured by the revised neural model, which includes the explicit system that is responsible for implementing the optimal dimensional rule. According to signal detection theory (SDT), the explicit system solved the categorisation problem by splitting the perceptual space into regions associated to the different categories (Bohil & Maddox, 2001). This feature was added to the neural model in an abstract and formal way, without attempting to specify the neural mechanisms in any detail. The success of the simulations with the revised model suggests that both the explicit and the implicit system are involved during learning on a nominally ‘rule-based’ CL task. In particular, the simulation results suggest that the explicit system mediates the categorical nature of learning whereas the implicit system, activated by the asymmetric payoffs, is responsible for the fine-tuning of criterion placement.

Data from the intertwined, 2-category task indicated that participants developed a modest response bias even following extended practice under symmetric accuracy feedback. These results, therefore, are consistent with the idea that the two systems mediate different aspects of learning and may become active at different stages of the task. More specifically, the explicit system may become active during the first trials of the task in order to mediate RB learning whereas the implicit system may become active over trials as participants experience asymmetric payoffs.

Following these considerations, these results can be reconciled with the COVIS model. In fact, the COVIS postulates that the two systems learn independently and they compete for response control throughout the task (Ashby et al., 1998). Moreover, it states that the explicit system is able to learn faster than the implicit system which, by contrast, learns in an incremental fashion over trials following trial-by-trial feedback (Maddox & Ashby, 2004; Ashby & Valentin, 2005; Shohamy et al., 2008). During the RB CL task, the explicit system may have learnt to solve the categorisation problem over the initial trials (and the details of this process were not explicitly modelled in the revised neural model); whereas in later trials the implicit system may have taken over and mediated the response bias. This is supported by results from the intertwined task where the implicit system became active during the experimental task to maximise performance, even though the explicit system had probably established task success fairly early during practice. The basal ganglia have been found to be responsible for a shift in strategy implemented by participants over trials. Indeed, during performance on a weather prediction task participants were found to implement a simple sub-optimal rule in the first stages of the task whereas they were found to be using the optimal, complex rule in the last block of the task (Gluck, et al., 2002; Shohamy et al., 2008). Moreover, there is computational, pharmacological and patient evidence to indicate that gradual, procedural learning is mediated by SR learning driven by RPE-DA signals in the basal ganglia (Shohamy et al., 2008). It is possible that a similar shifting process took place during performance on the tasks reported in the last chapters where a more optimal strategy was characterised by a criterion shift. In the studies reported in the chapter 8 and 9, the shift increased performance by increasing winnings relative to the optimum accuracy strategy. By contrast, in the study reported in chapter 10, the shift increased performance by maximising accuracy. However, it is not possible to tell whether biases beyond maximum accuracy (towards further increased winnings) would be established with further training. Indeed, the neural model, which was able to capture the human data, showed that response bias in the three studies could be explained by SR learning driven by RPE signals in the striatum (see below).

In conclusion, even though explicit learning may not be highly reliant on feedback since learning has been found to occur under observational and unsupervised training (Ashby et al., 1999; Maddox & Ashby, 2004), subtle manipulations of reward (e.g. payoffs and base-rate) may affect categorisation, as suggested by signal detection theory (SDT; Maddox & Dodd, 2003). The present work strongly suggests that some of these subtle effects may be due to activation and engagement of the implicit system.

11.2.3 Impulsivity and response bias

Results obtained in the blocked, 2-category task (chapter 8) showed that positive schizotypy was the personality component that was associated with the criterion shift; participants who scored high on this component showed a greater response bias than their low scoring counterparts. However, results from this study were inconclusive and hard to interpret due to the small sample size and observed order effect. Therefore, the relationship between criterion and personality is also considered inconclusive.

Impulsivity was found to be the personality dimension that significantly mediated the response bias observed in the RB CL tasks when rewards were presented under an asymmetric matrix. In fact, in the intertwined 2-category task (chapter 9), the ImpAss component was found to be negatively correlated with the criterion shift. Similarly, in the 4-category task (chapter 10), criterion shift was found to be negatively correlated with the behavioural measure of impulsivity, indexed by the delay discounting (DD) task. These results indicated that individuals who scored low on impulsivity showed a greater criterion shift (i.e. a smaller distance between criterion 1 and criterion 3) than their high scoring counterparts.

Criterion shift in both tasks increased the amount of points earned over trials. Indeed, the criterion shift observed in the intertwined task meant that participants shifted their criterion away from the optimal accuracy bounds closer to the optimal reward bound. In other words, participants tended to over-categorise probabilistic stimuli as members of the high payoff category and, therefore, increased their winnings by increasing their chance to respond correctly to the high payoff category. In the 4-category task, participants shifted their criteria away from the initial perceptually determined category bounds towards the optimal accuracy bounds. In this way, they increased their winnings by moving in the direction of maximising accuracy.

These observations support the assumption, made in chapter 8, that criterion shift represents approach behaviour towards reward and, therefore, that it represents a good measure of BAS-mediated behaviour (the criterion shift can also be conceptualised as an endophenotype of impulsivity, which is in turn the phenotype of the BAS). The criterion shifts observed in the studies increase the amount of reward gained. RST has suggested that the BAS represents the bio-motivational system responsive to reward and reward-related stimuli and, subsequently, it directs approach behaviour towards appetitive goals (McNaughton and Corr, 2004; Smillie, Pickering & Jackson, 2006). Therefore, inter-individual differences observed across the task should reflect individual differences in BAS sensitivity. The present results support one of the main RST assumptions, which stipulates that impulsivity is the underlying personality trait of the BAS (Gray, 1987; Corr, 2006). However, the present findings are counterintuitive to the original theory since they indicate that low (cf. high) scores on impulsivity are associated with high BAS sensitivity and approach behaviour. The simulations, run using the neural model, further support that low impulsivity scores correspond to high BAS activation (see below). Future replication may be necessary to confirm these preliminary results.

11.2.4 Delay discounting and response bias

In the 4-category task, the delay discounting scores were the only impulsivity measure to be significantly correlated with criterion shift. Nonetheless, scores on the dysfunctional impulsivity (DI) subscale were found to be positively correlated with DD scores. This association supports that the two measures index a similar construct of impulsivity (i.e. rash impulsiveness; Evenden, 1999; Dawe et al., 2004; Smillie and Jackson, 2006), even though DD was the only significant performance predictor.

A possible reason why DD, but not DI, scores were found to mediate criterion shift in the CL task may be due to the fact that, as an endophenotype, DD is argued to be a more valid index of biobehavioural individual differences than the self-report psychometric impulsivity scores (Lee et al., 2007; Eisenberg et al., 2007). Moreover, there is evidence that indicates that performance on a decision-making task with delayed/probabilistic feedback (e.g. DD task) is mediated by DA-RPE signals in the striatum (Abler et al., 2006; Kalenscher and Pennartz, 2008). These same DAergic processes have been postulated to underlie the implicit system and, therefore, mediate the response biases in the CL tasks. Finally, according to RST, the BAS has a dopaminergic nature and, more specifically, RST suggests that the DA projections to the striatum represent its biological basis (Pickering and Gray, 2001; Corr, 2006). Following these considerations, it can be assumed that performance on the DD task, the CL task and BAS-related individual differences share some common biological mechanism. The link between the biology and questionnaire measures of BAS-related individual differences are assumed to be the weakest and, thus, the associations involving the questionnaire measures are likewise expected to be the weakest in the set.

Recent studies have suggested that personality traits may represent endophenotypes for psychiatric disorders (Benjamin et al., 2001; Rommelse et al., 2008). For example, extreme scores on impulsivity measures (e.g. Novelty Seeking) have been found to be associated with addiction, pathological gambling and ADHD. Additionally, variance of impulsivity measures has been found to be explained by genetic factors (e.g. the A1 allele of the D2 dopamine receptor gene) that mediate DA activity (Benjamin et al., 2001; Eisenberg et al., 2007). In the kind of work presented in this thesis, however, impulsive personality traits are taken to represent the phenotype of the BAS and are the "end point" of the investigations. The work is made more important and relevant by the evidence indicating that extreme scores on BAS-trait measures are associated to pathological behaviours (e.g. addiction and gambling; Ham and Hope, 2003; Franken and Muris, 2006; Vigil-Colet, 2007).

Studies on addiction have indicated that the relationship between endophenotype and phenotype is generally quite weak despite the fact that they are both related to the same genetic variance (Blum et al., 2000; Franken and Muris, 2006; Pardo et al. 2007). These weak relationships are due to the fact that, even though the genotype is a good predictor of the endophenotype (i.e. task behaviour), genes only explain 30-60% of the variance in personality variance (the phenotype; Noble, 1998; Munafo et al., 2003; Reuters et al., 2006). Any single biological component is likely to explain much less personality variance. Therefore, psychometric impulsivity traits may be expected to be only weakly associated with task performance despite the fact that they rely on the same biological mechanism. In contrast, DD scores may have been more strongly associated with the performance on the CL task especially, since DD is a behavioural measure of impulsivity that relies on the same biological mechanisms that was found to mediate biases in category learning (Murray et al., 2007; Kalenscher and Pennartz, 2008). Hence, the relationship between DD scores and criterion shift may be due to their common biological mechanism (i.e. DA activity in the striatum, i.e. BAS; figure 11.1). In Figure 11.1 the criterion shifts in the CL task, and the DD scores may both be regarded as endophenotypes; as both are presumed to be related to the underlying biology more strongly than the link between biology and phenotype. It is reasonable to suppose that the two endophenotypes would correlate together more strongly than either endophenotype would correlate with the phenotype.

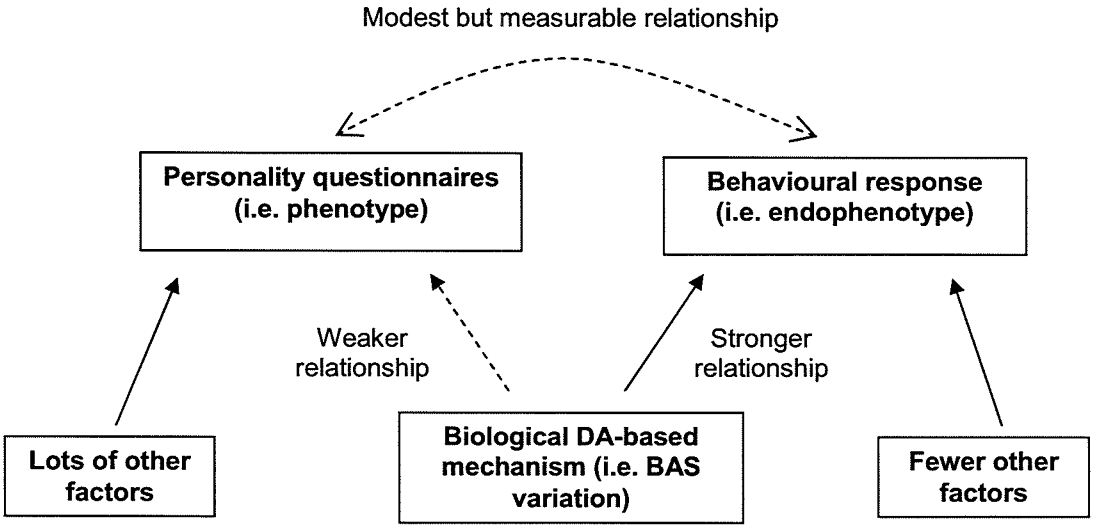


Figure 11.1. Relationship between personality, biology and behaviour (other factors indicate confounding factors)

As previously discussed, a study by Eisenberg and colleagues (2007) supports the above arguments. Indeed, these researchers found that individuals who carried an A1[23] allele showed greater temporal discounting during the DD task than those individuals who did not carry the allele. Genetic variance was not found to be associated with any of the self-report psychometric measures of impulsivity, despite the fact that they were associated with DD scores. These authors concluded that DD scores (as an endophenotype) may be a more reliable measure of the phenotype (i.e. impulsivity trait) than self-report measures, muddled by biases and subjective interpretations (i.e. confounding factors labelled as other factors in figure 11.1).

11.3 Model simulations

The use of the neural model to simulate the human data has proved to be a useful tool to further understand the empirical data as well as to develop future studies.

---

[23] The presence of the A1 allele on the D2 receptor gene has been fount to be associated with a 30-40% reduction in D2 receptors density (Pohjalainen et al., 1998; Blum et al., 1999; Reuter, et al., 2006)

Indeed, the use of several sub-models has allowed us to assess how BAS variation might mediate response bias. Results have shown that BAS variation may lie in responsivity to reward prediction error (RPE) rather than to reward *per se*. This is a shift from the position originally stated in RST (Gray, 1987; Pickering and Gray, 2001; Smillie, Pickering & Jackson, 2006). Indeed, when simulating data from the two CL tasks that offered reliable data (i.e. intertwined and 4-category tasks), the sub-model with BAS acting as a multiplier on the RPE signal which projected to the dorsal striatal (actor SR) cell was able to capture the human data better than the model with BAS variation affecting sensitivity of the reward cell (i.e. on projections to VTA/SNc). In the former model, individual differences (i.e. BAS variations) modulate the sensitivity to RPEs within the dopaminergic projections to the dorsal striatal cells. These RPEs are responsible for strengthening the SR association. Therefore, individual differences vary the extent of SR learning. The simulations indicate that individuals who are more responsive to RPEs (which logically should be deemed high BAS individuals) learn faster and more efficiently how to shift their criteria during CL tasks in the direction of maximising their winnings. In conclusion, the results suggest that individual differences (BAS variations) do not reflect an individual's sensitivity to rewards but to the DA-RPE signals. Nonetheless, the results confirm the dopaminergic nature of the BAS and its impact on reward-related learning according to the 3-factor rule (Pickering and Gray, 2001).

The simulation results indicated that individual differences in sensitivity of the ventral striatal cell to RPE signals (i.e. the reward prediction cell itself, within the critic sub-network) were not effective at leading to individual differences in learning. This is a strong result as the model learning mechanisms are identical in both the actor and the critic sub-networks. The reason for the difference must therefore lie in the function of the two sub-networks within the overall learning process.

The biological realism of the neural model might be enhanced by considering the different DAergic receptors involved in synaptic plasticity in the two cells. Schotanus and Chergui (2008) have observed that synaptic plasticity in the ventral striatum is linked to LTP mediated by D1 receptors; whereas in the dorsal striatum synaptic plasticity is driven by both LTP, mediated by D1 receptors, and LTD which is mediated by a synergic interaction between D1 and D2 receptors. This is immediately of interest as the reported links between impulsivity and dopaminergic processes have been primarily for D2-like receptor type. More specifically, for example, the A1 polymorphism on the dopamine D2 receptor (DRD2) gene has been found to be quite robustly linked with variations on impulsivity measures (e.g. delay discounting and novelty seeking; Benjamin et al., 2001; Eisenberg et al., 2007). Thus, individual differences that characterise inter-individual differences in D2 receptor sensitivity should affect learning that is partly dependent on these receptors (dorsal striatum) but not learning that relies exclusively on D1 receptors (ventral striatum). Hence, these observations suggest that even though D1 receptors are the probable mechanism that mediates synaptic learning in the ventral striatum, the critic unit is not affected by individual differences in RPE sensitivity. Thus, even if there were individual differences (i.e. in some other trait than impulsivity) that were related to D1 receptors functioning, the neural model suggests that they would not affect performance on the responses biases observed in the current CL task. The current CL task is a poor index of individual differences that might affect learning within the ventral striatal synapses of the critic network.

By contrast, individual variance on D2-related personality traits could affect learning in the actor sub-network since D2 receptors have been found to be involved in LTD in dorsal striatum. It is possible that D2 receptors may also mediate LTP by indirectly inhibiting it (Calabresi et al., 2007).

In order to test the above ideas, BAS variation could be added only on negative RPE signals (i.e. LTD; cf. LTP) in the actor network. Thus, simulations were re-run for the 4-category CL task (chapter 10) using the revised model with BAS variation acting only on negative RPE signals projecting to the SR cell (i.e. actor network). Simulations were run with BAS variation ranging uniformly across 0 and 2 for the two payoff conditions. The simulated criteria and noise levels are close to the values observed in the human sample (table 11.1).

Table 11.1. Human criteria and simulated criteria obtained using two perceptual noise values (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Criterion 1	Criterion 2	Criterion 3	Noise
Sim_PO1 **(N = 300)**	177.47 (15.1)	299.80 (9.8)	435.81 (13.9)	80.57 (25.1)
Sim_PO2 **(N = 300)**	178.16 (14.7)	304.42 (8.8)	420.75 (13.7)	78.84 (23.6)
Hum_PO1 **(N = 18)**	182.97 (26.4)	300.13 (14.0)	420.84 (23.8)	80.27 (31.7)
Hum_PO2 **(N = 20)**	191.77 (27.3)	303.21 (21.0)	413.31 (20.7)	79.62 (32.3)

POmatrix 1

Results showed that BAS variation was significantly correlated with criterion 1 (r = .69, p <.001), criterion 2 (r = .56, p <.001) and criterion 3 (r = -.52, p <.001). These correlations are graphically represented in the scatterplots below (figure 11.2).

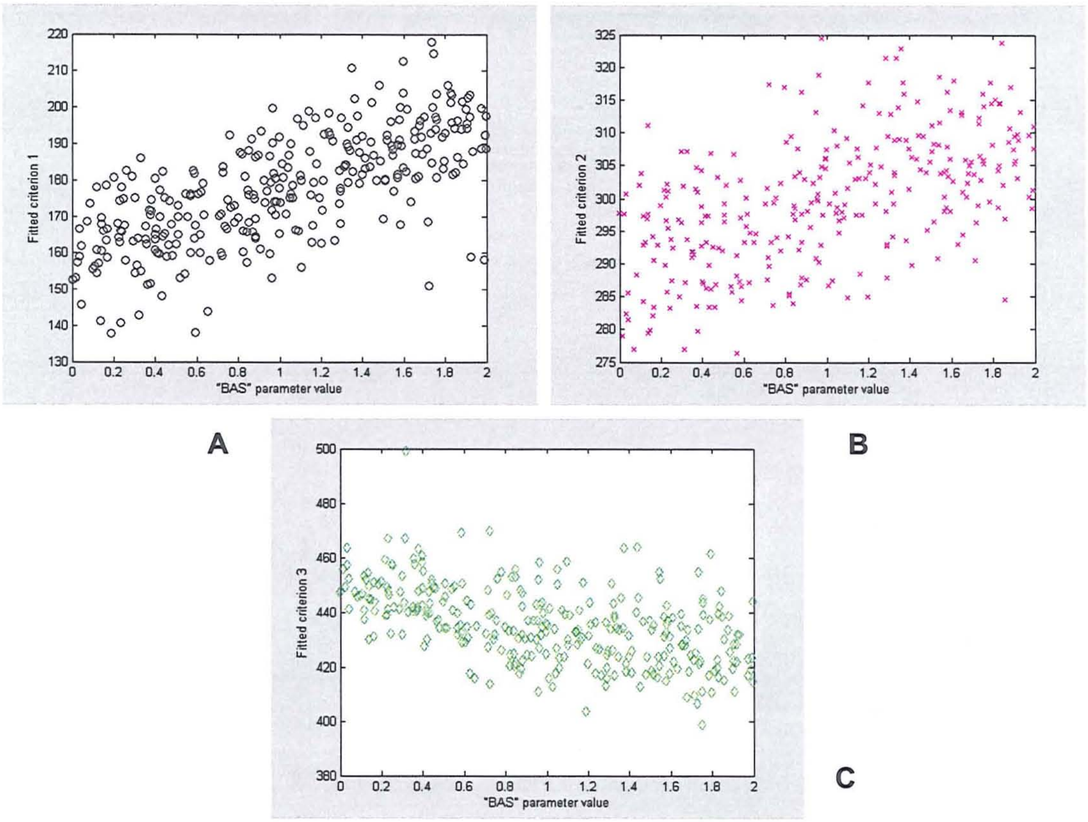


Figure 11.2. Relationship between the BAS variation and the simulated criterion 1 (figure A), criterion 2 (figure B) and criterion 3 (figure C) under PO matrix 1

POmatrix 2

Results showed that BAS variation was strongly significantly correlated with criterion 1 (r = .70, p <.001) and criterion 2 (r = .37, p <.001). The correlation between criterion 3 and BAS variation was significant but weak (r = -.12, p<.001). These relationships are visually represented in figure 11.3.

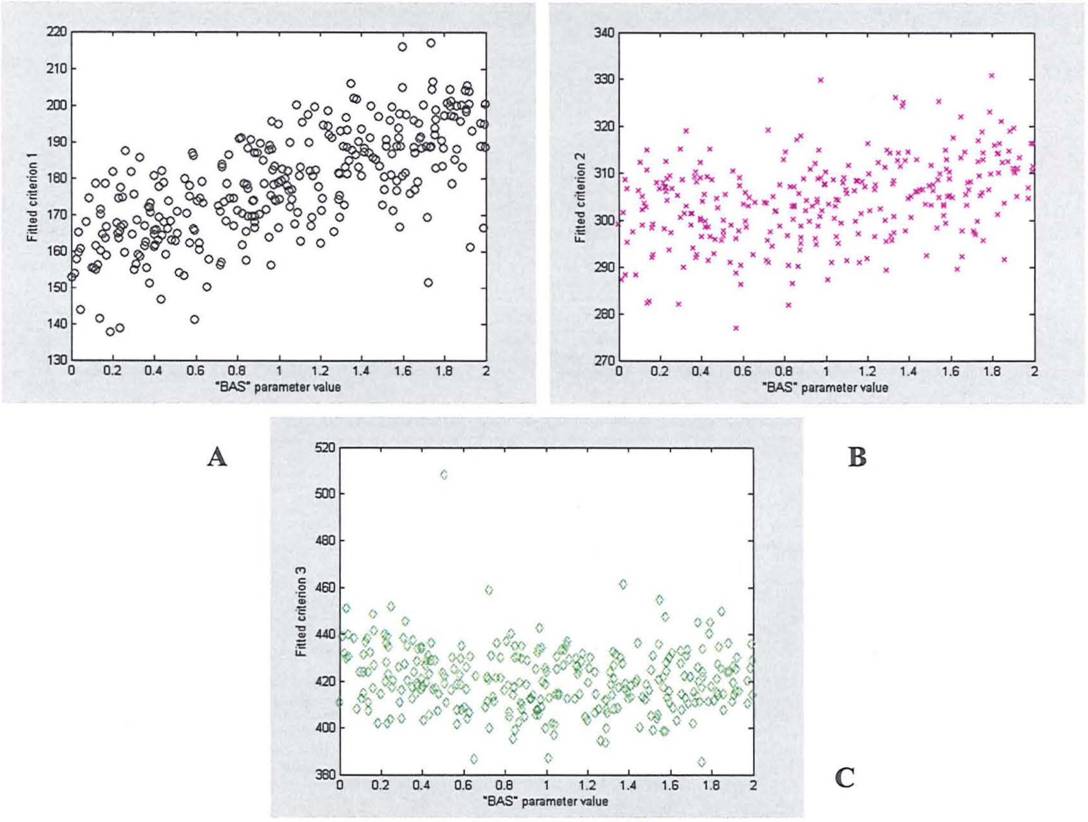


Figure 11.3. Relationship between the BAS variation and criterion 1 (figure A), criterion 2 (figure B) and criterion 3 (figure C) under PO matrix 2

These simulations show that the revised version of the actor, where the BAS variation acts only on LTD, produces results which are not very different from those simulated by the original actor model (BAS variation on both LTP and LTD; table 11.2).

Table 11.2. Correlations between the simulated criteria (C1-C3) and the BAS variation obtained using the original model, with BAS variation acting on both LTP and LTD on the actor network (LTP+LTD), and the model with BAS variation acting only on LTD (i.e. LTD)

	LTP+LTD PO 1	LTP+LTD PO 2	LTD PO1	LTD PO2
C 1	r = .78, p <.001	r = .79, p <.001	r = .69, p <.001	r= .70, p p<.001
C 2	r=.17, p=.002	r = .28, p <.001	r = .52, p<.001	r = .37, p <.001
C 3	r = .58, p <.001	r = -.66, p <.001	r = -.52, p<.001	r = -.12, p <.001

In other words, the observations made by Schotanus and Chergui (2008) were used to constrain the neural model so that BAS variations (linked to impulsivity) were added only to process to have a known D2-receptor involvement. This is appropriate as impulsivity is primarily linked to dopeminergic processes involving the D2-receptor subtype (Benjamin et al., 2001; Eisenberg et al., 2007). Had the data indicated that D2-related learning processes were located mainly in the ventral striatum, then the current neural model would not have been able to explain correlations between impulsivity and the category learning criterion shifts. This is because BAS variations on either LTP or LTD, when added to the ventral striatum (critic) were unable to generate individual differences in simulated CL behaviour.

Overall, the simulations offered a fresh and objective insight into the empirical data. Indeed, if the model had not been implemented the results would have been interpreted in light of the theory and variation on the BAS-measure(s) would have been considered to reflect differences in reward sensitivity. In contrast, the results are distinct from the existing theory and suggest a revision may be necessary. However, further replication of these results is required to support their empirical validity and more conclusively support the revised model of BAS sensitivity.

Additionally, the neural model represents a useful tool to develop future studies and, in particular, CL tasks able to capture the relationship between behaviour and BAS-traits. The section above introduced evidence that show that it is hard to detect the relationship between the endophenotype and the phenotype even though this relationship should be stronger than that between phenotype and the genotype. In order to capture the desired personality-behaviour relationship, it is important to develop and implement a task (endophenotype) whose performance is highly dependent on the biological mechanism. Since personality is conceived as a phenotype, a strong relationship between biological mechanism and task behaviour implies a greater chance of detecting a significant relationship between the task behaviour and personality trait measures. For this reason, tasks should be used for research purposes only if the simulated relationship between behaviour and biological mechanism show strong correlation coefficients. A cut-off would be arbitrary but one might propose the need for simulated correlations equal to or greater than 0.7 (see chapter 8 for detailed mathematical reasoning). The criterion shift measures in the present tasks would just about qualify under this cut-off.

In other words, neural models should be used to test the efficacy of a specific task at assessing the relationship between behaviour and the biological mechanism (e.g. BAS). By doing this, it would be possible to assess the utility of the task as well as identify and rectify potential flaws in the task design prior to data collection. This would prevent spending money and time collecting data for a study with a poor design. The efficacy of the 4-category task was not fully assessed by the neural model prior to data collection (as the neural model was evolving during the thesis and data-collection had to proceed on timetable before the model was fully developed). Had the full model been available, simulations would have shown the task flaws (e.g. independent perceptual and maximum accuracy bounds), which do not allow us to conclude whether participants were shifting their criteria closer to the optimal reward bounds rather than the optimal accuracy bounds (see Appendix 4 for details).

Markman and colleagues (2005) interpreted their results in relation to the regulatory focus theory (RFT) and they concluded that the regulatory fit between the induced regulatory focus and the payoff manipulation facilitated cognitive flexibility and, thereby, allowed a criterion shift. Specifically, it was suggested that regulatory fit enhanced cognitive flexibility thanks to increased midbrain DAergic firing (SNc/VTA), which projects to cortical areas involved in working and episodic memory (Maddox et al., 2006; 2007). In other words, Maddox and colleagues proposed a way for a reward manipulation to affect an RB task other than through its directly reinforcing effects. However, the model simulations indicate that the criterion shift was mediated by an alternative mechanism, which is dependent upon the reinforcing properties of RPE signals. Hence, in the original study, the induced regulatory focus may have affected performance by, indirectly, affecting the impact of RPE on learning.

It was speculated the regulatory focus influenced the activity of the critic sub-network. The critic is considered to play a crucial role in learning to predict future rewards and, by affecting RPEs, it indirectly mediates learning on the actor sub-network (Suri and Schultz, 1998; Joel et al., 2002). It was speculated that the critic would not be active under a prevention focus, where participants receive the ticket prior to performance and their aim is not to lose it (cf. win it). This derived from the assumption that the critic is involved in the mechanism that is sensitive to reward and prone to approach behaviour (i.e. dopaminergic mechanism, BAS). The assumption is supported by simulations run for the asymmetric condition of the blocked 2-category task (chapter 8) using the original parameter values and setting the **rptonic** value equal to 0 across all trials. This turns off the critic circuit's influence on the task so that all rewards remain unpredicted. Indeed, the data showed that the simulated 'participants', in the asymmetric payoff condition, placed their criterion close to the optimal accuracy bounds (mean = 329.52 and sd = 13.1) as in the study by Markman and colleagues. The simulated noise level was low (mean = 29.36 and sd = 10.7)[24]. Setting the **rptonic** value equal to 0 mediates learning by affecting the RPE signals. Indeed, when **rptonic** is set to 0, the critic never develops a reward prediction and, therefore, the RPE signals are always equal to the **rf** values (cf. the normal situation in which they reflect the difference between the actual reward, i.e. **rf,** and the reward prediction). This implies that learning weights would be updated less efficiently (especially following incorrect responses) during the task and, therefore, implicit learning is impaired and performance is at chance. For example, following an incorrect response the RPE signal would be equal to 0.25 and, therefore, the incorrect SR association would be strengthened rather than weakened. When **rptonic** is different from 0 and reward predictions can be learned, the RPE signal following an error is usually negative as the received reward is generally less than the reward expected.

On the other hand, the explicit module can efficiently categorise the stimuli, especially those far from the decision bounds and less efficiently those close to

---

[24] Simulations of the symmetric condition similarly showed that participants placed their criterion close to the optimal accuracy criterion, 320.72 (10.9) and a low noise level (mean = 26.09 and (9.8)

the boundaries. Thus, decision bound(s) would be implemented to solve the categorisation problem but the bound(s) would be determined by the explicit module rather than the implicit module, since learning in the latter is impaired. The decision bounds set by the explicit module would be close to the optimal accuracy bounds since this module is not responsive to payoff manipulations. Overall, these simulations support the assumption that the regulatory focus manipulation may affect performance by acting on the critic network (cf. regulatory fit).

11.4 The utility of the formal model

The formal model used to fit the human data proved to be a useful tool in determining which strategy participants had actually implemented to solve the categorisation problem. In particular, the model was able to establish whether each participant had implemented the appropriate strategy or an alternative one. Indeed, there is evidence that indicates that participants do not necessarily implement the appropriate rule as predicted by the researcher (Gluck et al., 2002; Shohamy et al., 2008). In a study by Gluck and colleagues (2002), 90% of the participants were found to use a simple, singleton strategy to solve an information integration task rather than the optimal, multidimensional strategy. Nonetheless, they were found to perform above chance levels since the simple rule was associated with 75% accuracy. Similarly, participants, who performed above chance during an II task, were found to be using simpler, uni-dimensional strategies following data fitting via a formal model (Ashby and O'Brien, 2007). These authors also indicated that the use of a uni-dimensional rule could ensure 78% accuracy on the task. Overall, these results indicate that accuracy scores are not a good predictor of the type of strategy implemented since alternative, non-optimal rules may sometimes ensure performance well above chance level.

In the II-version of the Gabor task (chapter 6), the formal fitting showed that none of the participants used the optimal II strategy. Nonetheless, 71% of the sample showed accuracy levels above chance despite the fact that they were using simpler and alternative strategies. In fact, a third of the sample used alternative dimensional strategies (i.e. uni-dimensional or conjunctive rules) and the rest used a non-dimensional strategy. Similar results were observed in the simpler RB version of the Gabor task where 27 (out of 64) participants were found to perform above chance although only 19 of them had implemented the appropriate strategy.

The formal model had a double-function in the analysis of the RB CL tasks reported in the last three chapters. The formal model extracted the actual criterion and also evaluated whether this was a 'meaningful' measure of performance for each individual participant. Thus, the fitted criterion is a more accurate measure of decision bound than the standardised criterion (C) that can be calculated using equations of Signal Detection Theory (SDT). Moreover, as for the Gabor tasks, the model could identify those participants who had implemented the optimal uni-dimensional strategy. Since the studies investigated how participants shifted their decision criterion (i.e. showed a response bias) in response to reward manipulations, it was important to measure response bias meaningfully. This was achieved by (conservatively) including in the analysis only those participants whose data was well-fitted by the criterion-based model. In each study, approximately half of the sample implemented the appropriate rule whereas the rest of the sample was found to be either guessing or using an alternative strategy. However, among the latter group accuracy levels were generally at or above chance level. These results further support the view that accuracy scores are not a good index of the strategy implemented by participants since scores above chance do not imply that the optimal rule was implemented. It is, therefore, possible that non-dimensional strategies, driven by alternative processes, may ensure accurate performance and produce a response bias. This claim is supported by the neural model simulations which indicated that the learning pattern observed among those participants who implemented alternative strategies resembled the learning curve simulated by the RPE-only-based model (cf. the revised RPE model incorporating an explicit model). However, there was no good way to measure the response bias shown by these participants, as the formal model could not accurately capture their behaviour. It might have been possible to calculate the SDT location parameter (C) for these participants. The calculated C scores would represent a measure of response bias but they would be meaningless. Indeed, the calculations of the C parameter rest on assumptions about the behaviour of a participant (e.g. the use of a decision bound) that did not apply to these participants; the lack of fit to the assumptions of SDT was indicated by the formal model. Therefore, it was decided to apply a conservative approach and exclude those participants from the data analysis. Even though, the neural model simulated qualitatively similar behaviour, when the implicit procedural module dominated the rule-based module, there is no simple quantitative index (like C) which can capture the behaviour observed.

These observations suggested that different processes may lead to the same apparent behaviour (i.e. response bias); although the study may be interested in exploring only those processes involved in optimal, uni-dimensional performance that can be quantified. The formal model is able to identify the participants implementing the optimal processes and showing the behaviour of interest, and separate them from those using the irrelevant ones. Inclusion of participants who are using alternative strategies and processes would represent a source of noise to the analysis of the phenomenon of interest. This is especially true since the quantitative measure of performance (i.e. decision bound) is an inappropriate measure of performance for the participants that were excluded from the analysis (as discussed above). Indeed, repeating a metaphor used in chapter 8, the inclusion in the analysis of data from those participants who failed to implement the appropriate strategy would correspond to measuring the diameter of tennis balls in a study interested in assessing the impact of fertiliser on orange growth. In other words, the diameter can be calculated but is not meaningful, as the study presumes that the measure (i.e. orange diameter) is something amenable to change. Computational models have been identified as a more reliable tool to exclude noise and irrelevant measures than self-reports and accuracy scores (Gluck et al., 2000; Shohamy et al., 2008).

11.5 Limitations of the studies

Following the formal fitting of the human data, it became evident that some of the tasks may have been too complex and hard for participants to perform optimally. This was indicated by the low retention rates after formal modelling. Indeed, in the rule-based CL tasks with asymmetric payoff matrices approximately half of the sample was found to use the appropriate strategy and, therefore, was retained in the analysis. Several factors may have played a role in rendering the task(s) hard and, subsequently, preventing the majority of the sample to perform optimally.

The main factor that may have hindered learning during the tasks may have been the complexity of the categorisation problem that derived from the great overlap between stimuli from different categories. This may particularly apply to the 2-category tasks that used a d prime equal to 1, which is responsible for the great overlap between the two categories. The choice of this d prime value was driven by the fact that it had been used by Markman and colleagues (2005) following the rationale that motivational factors affect performance only when the task is difficult. Future studies could use a greater d prime value (e.g. 1.5) to ensure learning across the majority of the sample. This d' value will not allow participants to reach optimal (100%) accuracy.

Learning during the 2-category tasks was also affected by the use of the two payoff matrix conditions as two independent tasks which presented stimuli in a slightly different fashion (vertical vs. horizontal). Having two independent tasks raised issues about order of presentation. In the study reported in chapter 8, the two conditions were presented as two separate tasks and the order of presentation was counterbalanced across participants. In this way, half of the sample experienced the asymmetric condition prior to the symmetric condition whereas the opposite order was experienced by the other half of the sample. Results showed that those participants who experienced the symmetric task first showed an aberrant pattern of responding (i.e. response bias only in the symmetric task condition). This pattern could not be fully understood but it is possible it may be an artefact of the task counterbalancing. Moreover, among those participants who experienced the asymmetric task first, a weak response bias was observed during the symmetric task. This bias was indexed by the fact that the criterion was non-significantly shifted away from the optimal accuracy bound. It is possible that this bias was due to a carry-over effect from the preceding asymmetric task since participants may have tended to classify probabilistic stimuli in the symmetric condition as members of the category that offered greater gains in the asymmetric task.

In the study in chapter 9, the above kind of order effect was removed by having participants perform the two tasks simultaneously by presenting them in an intertwined fashion. This task design increased the cognitive and working memory demand of the task since participants had not only to solve the categorisation problems but also to retain simultaneously two separate rules and apply the appropriate one depending on the task. Moreover, two different sets of response-keys were used to respond in the two tasks. Thus, participants had also to learn to associate one of the key-sets to the different tasks and, subsequently, each key to the correct category. Key-mapping may have added extra cognitive and mnemonic effort to the already complex task design. The task complexity in this case may account for the fact that approximately half of the sample failed to learn to perform on the task.

Key-mapping in the intertwined task and the 4-category task was worsened by the fact that the keys for the two tasks were chosen to mismatch the actual category location. This was done to avoid biases due to response position compatibility, although more reflection shows that these concerns were probably not well-founded. For example, in the intertwined task the up and down arrows were used to classify the horizontal stimuli whereas the left and right arrows were used to categorise the vertical stimuli; this almost certainly added complexity to the task and, subsequently, to learning. This complexity adds noise to the categorisation process and, doubtless, requires several trials to be overcome. Therefore, future studies could make the key-mapping more straightforward by eliminating the mismatch element above in order to facilitate learning (e.g. left arrow may correspond to stimuli on the left-side of the screen, category A). On reflection, this seems unlikely to be a problem but will facilitate the rapid establishment of the correct rule-based behaviour (which is not explicitly modelled but is assumed to occur over a small number of early trials). Similarly, in the 4-category task the stimuli corresponded to an arbitrarily assigned arrow-key rather to a more logical key-mapping. In the future, the key-mapping should be simplified to allow learning and a 4-button box may be used to map each key to the actual category location on the screen (e.g. far-left key would correspond to category A vs. far-right key would correspond to category D).

A potential task design, which removed any possible issue raised by task presentation, would require the two conditions to be presented in the same task. Hence, the 2-category task design could be replaced with a 3-category task. Hence, participants would be presented with a task where they have to learn to classify probabilistic stimuli into three different categories. Each category would offer different payoffs for correct responses and equal payoffs for incorrect responses. For example, the high payoff category could offer 400 points and correspond to the category on the left-hand side of the display (category A) whereas the two other low payoff categories could offer 200 points and appear in the middle and on the right-hand side of the display (category B and C).

Several factors may be held responsible for the relatively small response bias obtained in the three studies. The extended practice trials participants experienced prior to performing the experimental session of the intertwined task may be responsible for the small criterion shift observed in the asymmetric task, compared to the symmetric task. It is possible that extensive practice, which offered symmetric accuracy feedback, may ensure efficient category learning by implementing the optimal accuracy criterion. This may imply that during the experimental session, the activation of the implicit system was dampened by the previous activation of the explicit system. Or, indeed, there may have been some "inertia" within the implicit system itself: once it had settled on particular bounds for the symmetric practice trials feedback processing may be less careful and so little further adjustment of the bounds may have taken place. Subsequently, the criterion shift towards the optimal reward bound was constrained by the well-established placement of the bound close to the optimal accuracy one. Therefore, future studies should avoid including extended practice trials or, where necessary, include the minimum number of trials for participants to get acquainted with the task.

In the 4-category task, it became evident that the inclusion of auditory feedback may have been detrimental to performance on the task. Indeed, the task simulations showed that participants probably perceived all 'positive' payoffs (apart from the 400 points) as equivalent. This fitted with the participants' verbal reports. In other words, the asymmetric matrix was treated as a much more symmetric matrix and this may partly account for the fact that reduced response biases (towards optimal reward bounds) were observed in the study. This effect may be exaggerated by the fact that participants relied on the auditory rather than the visual feedback. Several reasons may explain why participants relied on auditory feedback and ignored the visual feedback. One reason may be related to the fact that the task was cognitively demanding due to the task design (e.g. stimuli overlapping, several categories, key-mapping) and therefore, participants attempted to minimise the information to be processed. Therefore, they would rely on just one source of feedback (i.e. auditory) and re-direct the processes that may have been involved in processing the visual feedback to process different information (e.g. update or maintain the rule). It is also possible that visual feedback was mostly discarded since the differentials in the payoffs were quite small and may have not been perceived as very informative. The 400 points, which were identified by the neural model to affect decision bounds, were the only payoff value to be reported as being 'noticed' by participants.

11.6 Future directions

Several potential future studies have emerged that may extend the previous findings and resolve some of the issues discussed earlier that may have hindered the results. Moreover, the section presents potential extensions that could be applied to the model to render it more neurally valid as well as to increase its simulation potential.

11.6.1 Payoff manipulation on category learning

A future study should attempt to replicate the findings obtained in the 4-category task after improving the task design in order to enable participants to display a criterion shift more easily (i.e. towards the optimal reward bounds) in order to increase winnings.

Indeed, the original study failed to observe this type of response bias due to the task design, where the perceptual bounds and the optimal accuracy bounds did not overlap. This has resulted in learning being characterised by participants shifting their bounds away from the perceptual bounds to increase their winnings. The study aimed to explore a criterion shift towards the optimal reward bounds. Such a criterion shift was observed in the study by Markman and colleagues (2005) as well as the studies reported in chapters 8 and 9. Nonetheless, in these studies maximum accuracy and perceptual bounds overlapped exactly. Hence, it may be the case that participants are able to place their criterion closer to the optimal reward bound when the perceptual bound is close (or corresponds) to the optimal accuracy criterion early in the task. Following these considerations, an alternative set of stimuli was generated that had the perceptual bounds overlapping with the accuracy bounds; these stimuli could be used in a follow-up study.

The proposed task consists of 180 trials in total and there would be an equal number of trials for each of the four categories, just as in the original task. Stimuli were generated by sampling from 4 independent but overlapping normal distributions. The stimuli mean values were 100, 200, 400 and 500 for category A, category B, category C and category D, respectively. These values were chosen since they allowed the optimal accuracy bounds to correspond with 150, 300 and 450 pixels, which correspond to the perceptual bounds within the 600*600 display box. The standard deviation value for all categories was equal to 50. Following the choice of mean and standard deviation, the d prime between category A and B was equal to 2 and so was the d' between category C and D. In contrast, d prime was equal to 4 between category B and C, which should enhance categorisation across these two categories. The previous study showed that participants achieved poorer levels of accuracy for stimuli from the central categories (cf. stimuli from the categories near the edges of the display box).

Simulations were run using the same parameter values that were used to simulate the original 4-category task. In particular, the reward prediction tonic value (**rptonic**) was set to 0.5, the **rulebias** value was set equal to 0.4. Owing to the results reported in chapter 10 where participants achieved greater accuracy levels for those categories closer to the external edge of the display box (categories A and D), two different levels of noise were used for stimuli close to the edges of the perceptual box (**pnoise** = 60) and those close to its edges (**pnoise** = 40). These are the same noise values implemented to simulate the data from the 4-category task. Indeed, the task simulated in this section has the same structure as the 4-category task and, therefore, learning is expected to follow a similar pattern (i.e. easier categorisation for the stimuli belonging to the near-the-edge categories). The two simulated asymmetric payoff matrices (i.e. PO1 and PO2) were identical to the one used in chapter 10 (table 11.3).

Table 11.3. Payoff matrix of the four categories across the two payoff (PO) counterbalancing orders

	Correct				**Incorrect**
	Cat A	**Cat B**	**Cat C**	**Cat D**	**All**
PO1	333	200	400	267	100
PO2	200	333	267	400	100

As in the previous studies, the points were recoded into a linear variable (**rf**) that coded 400, 333, 267 and 200 points into 1, 0.83, 0.67 and 0.5, respectively. The study should present only visual feedback to inform participants on the number of points they earned and the potential points they could have earned in each trial. This should be done to prevent participants from relying solely on auditory feedback and, therefore, not focusing on the payoffs.

11.6.1.1 Predictions

Predictions were made following the changes implemented in the task such as the use of visual feedback, only, having the optimal accuracy and perceptual bounds coincide and, finally, the use of larger d prime values (for the middle categories). In particular, participants were expected to place their criterion closer to the optimal reward bounds to maximise their point winning, relative to the optimal accuracy bounds.

11.6.1.2 Preliminary simulations with no individual differences (i.e. no BAS variation)

Initial simulations were run without any inclusion of individual differences and were separately run for the two payoff matrices. Simulations consisted of 300 data sets per payoff condition (PO1 and PO2). The simulated data was then fitted using the formal model to extract the actual criteria.

Results showed that the simulated participants showed the predicted response bias that was characterised by the tendency to over-classify probabilistic stimuli as members of the high payoff categories (table 11.4). The simulated 'participants' in general placed their criteria away from the optimal accuracy bounds in the direction of the optimal reward bounds.

Table11.4. Simulated criteria and the calculated optimal reward criteria across the two asymmetric payoff matrices (sim = simulated; optimal = calculated optimal reward criteria)

	Criterion 1	**Criterion 2**	**Criterion 3**	**Noise**
Sim PO1 (N =300)	172.84 (11.3)	302.58 (10.3)	456.32 (10.2)	86.40 (10.4)
Sim PO2 (N =300)	155.49 (9.9)	304.40 (10.3)	434.44 (10.7)	85.44 (11.3)
Optimal PO1	171.15	286.27	464.64	--
Optimal PO2	128.85	304.16	435.36	--

The simulations suggest that the model would be able to simulate human data, which is characterised by more conservative biases than the optimal classifier (Bohil & Maddox, 2001; Maddox & Dodd, 2001). All the simulated criteria (1 to 3) were found to be significantly different than the optimal accuracy criteria (i.e. 150, 300 and 450, respectively) both under PO matrix 1 ($|ts| > 4.3$, $p <.001$) and under PO matrix 2 ($|ts| > 7.4$, $p <.001$). Thus, even though, the simulated criterion shifts towards the optimal reward bounds seem small, they are significant.

11.6.1.3 Behaviour and BAS variation

Inter-individual differences were simulated by using a uniform distribution of BAS parameter values that ranged between 0 and 2. The simulations were only run for the model that had BAS variation acting as a multiplier on the RPE projections to the synapses of the SR cell (i.e. actor network). This choice was driven by the results, obtained in the previous chapters, which showed that this was the model that best captured the human data. The simulated criteria and noise levels are reported in table 11.5.

Table 11.5. Simulated criteria and noise mean and sd values for the two payoff matrices

	Criterion 1	**Criterion 2**	**Criterion 3**	**Noise**
PO1 (N =300)	168.50 (12.7)	301.97 (10.6)	453.19 (11.9)	83.22 (26.1)
PO2 (N = 300)	155.03 (10.1)	304.46 (11.2)	438.67 (11.3)	82.34 (24.6)

PO matrix 1

Results showed that BAS variations was significantly correlated with criterion 1 ($r = .45$, $p <.001$) and criterion 3 ($r = .44$, $p <.001$; figure 11.4). The correlation between criterion 2 and BAS variation was significant but weak ($r = .09$, ns)

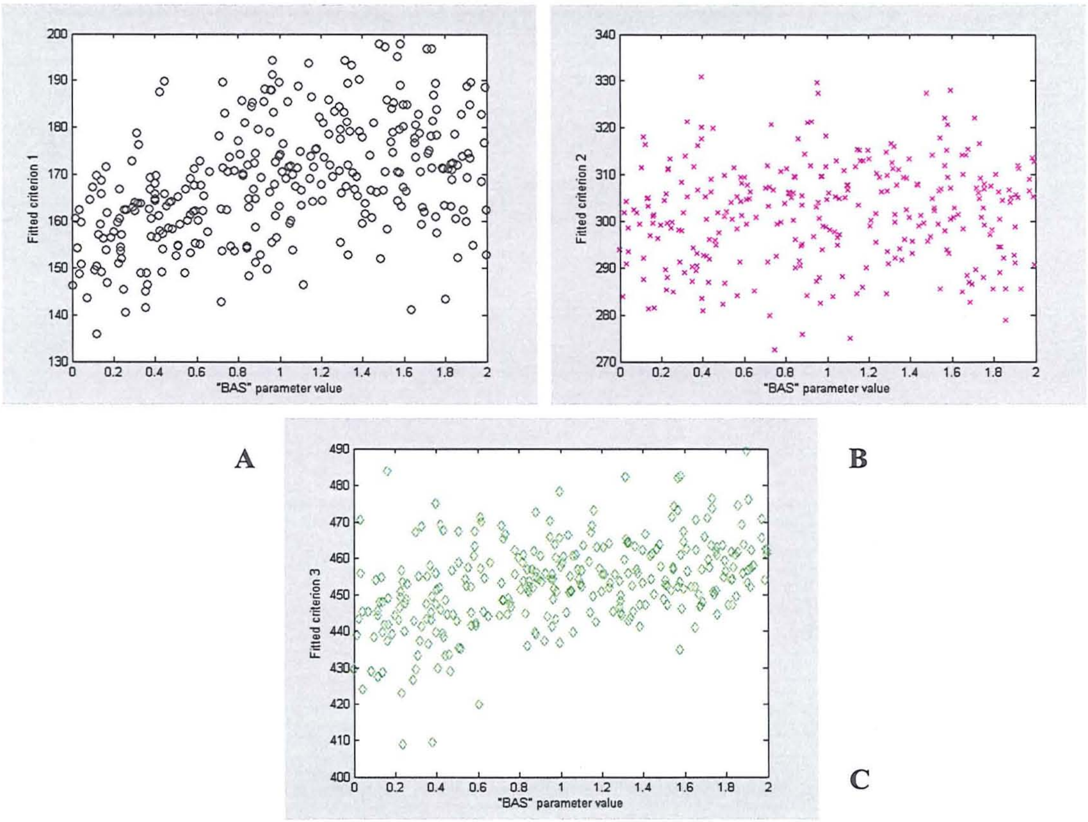


Figure 11.4. Relationship between the BAS variation and criterion 1 (figure A), criterion 2 (figure B) and criterion 3 (figure C) under PO matrix 1

PO matrix 2

Results showed that BAS variation was weakly correlated with criterion 1 ($r = -.22$, $p < .001$) and criterion 2 ($r = .33$, $p < .001$). In contrast, the correlation between criterion 3 and BAS was non-significant ($r = .01$, ns). These relationships are reported in figure 11.5.

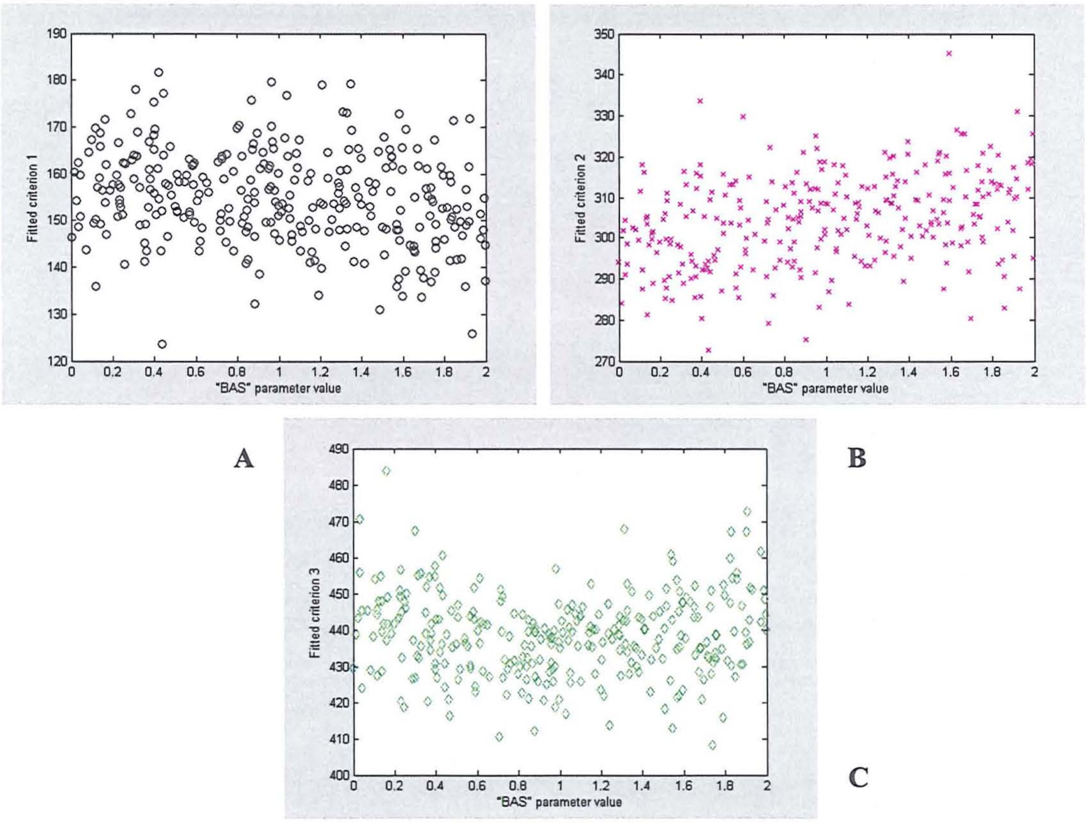


Figure 11.5. Relationship between the BAS variation and criterion 1 (figure A), criterion 2 (figure B) and criterion 3 (figure C) under PO matrix 2

The correlations under this payoff matrix may have been weak due to the fact that the simulated criteria were close to the optimal accuracy bounds (i.e. very small criterion shift towards the optimal reward bounds). Indeed, these criteria were close to the criteria bounds simulated under a symmetric payoff matrix. Under the symmetric matrix correct responses earned 300 points (coded as **rf** = .75) and incorrect responses received 100 points (**rf** = .25). The simulations were run with no individual differences (i.e. no BAS variation) and using the same parameter values reported earlier. Despite the fact that the payoff matrix was symmetrical, the simulated criteria were significantly different from the optimal accuracy criteria (|ts|> 7.5, ps <.001; table 11.6). Results showed that under the symmetric payoff matrix, criteria were placed close to the optimal accuracy/perceptual bounds although they were pushed inwards slightly by the lower noise value chosen for the stimuli near the edges (i.e. stimuli near the edges of the box are expected to be perceived more consistently over time than those stimuli in the middle).

Table 11.6. Simulated criteria under a symmetric payoff matrix

	Criterion 1	Criterion 2	Criterion 3	Noise
Symmetric **(N = 300)**	163.38 (9.8)	304.54 (10.5)	445.34 (10.4)	86.12 (11.1)

This phenomenon was expected to be in action following the simulation results obtained for the 4-category task study (see Appendix 4). This phenomenon may antagonise the shifting force induced by the asymmetric payoffs, when they are in opposition, and it would explain the low shift in criterion 3 under PO matrix 1 and the low shift in criterion 1 under PO matrix 2. In contrast, the criterion shift would be enhanced when the two forces act in the same direction (e.g. criterion 3 under PO 2). Simulations using a noise value equal to 50, regardless of stimulus location, showed criterion placement close to the optimal accuracy bounds [i.e. criterion 1: 157.73 (9.9); criterion 2: 304.52 (10.7) and criterion 3: 450.26 (10.1)] under symmetric payoffs.

Overall, results show that the use of the asymmetric payoff matrix described earlier is effective at producing a weak criterion shift towards the optimal reward bounds. However, the simulated response biases (table 11.5) are quite small and, therefore, it may not be possible to detect them in noisy empirical studies. Subsequently, the BAS-behaviour relationship may also be weak. This small shift may be due both to the near-the-edge force and the small difference across payoffs. The small differential across the payoffs may also be harder to detect under testing conditions. Therefore, using greater payoff differentials may produce greater criterion shifts in both simulated and empirical data. This hypothesis was tested in further simulations with greater payoff differentials.

In order to reduce the complexity of the task, only two payoff values were implemented (i.e. high = 400 points and low = 150 points) for correct responses whereas incorrect responses earned 100 points. The points were recoded into **rf** values equal to 1, 0.375 and 0.25 for 400, 150 and 100 points, respectively.

The use of only two payoff values for correct responses (i.e. 400 vs. 150) allows us to test the impact of asymmetric payoffs in comparison to a symmetric matrix during the same task. Therefore, it eliminates all potential problems encountered in the previous studies such as carry-over effect and over-complicating performance. The payoff matrices are presented in table 11.7.

Table 11.7. Asymmetric payoff matrices

	Correct				Incorrect
	Cat A	**Cat B**	**Cat C**	**Cat D**	**All**
PO 1	400	150	150	400	100
PO 2	150	400	400	150	100

11.6.1.4 Preliminary simulations with no individual differences (i.e. BAS variation)

Two PO matrix structures were used to assess whether the location of the high payoff categories (mid or side of the screen) would have different impact on the criterion shift. This is particularly important since, due to the edge-effect, a response bias seems to be present under symmetric payoffs (table 11.6). The original 4-category study did not find any difference in criterion across the two payoff matrices, nonetheless all payoffs were perceived as equivalent. Hence, it is useful to control for stimulus location even though this may require subdividing the main sample into smaller sub-groups.

The present simulations used the same parameters and parameter values that were implemented in the previous simulations. Results showed that simulated 'participants' showed the predicted response bias as they tended to over-classify probabilistic stimuli as members of the high payoff categories (table 11.8). Nonetheless, the simulated criteria were lower than the optimal reward criteria and this suggests that the model was able to simulate human data which is characterised by more conservative biases than the optimal classifier (Bohil & Maddox, 2001; Maddox & Dodd, 2001).

Table 11.8. Simulated criteria under the two asymmetric payoff matrices (sim = simulated; optimal = calculated optimal criteria)

	Criterion 1	Criterion 2	Criterion 3	Noise
Sim PO1 (N= 300)	180.23 (11.3)	304.06 (10.8)	425.33 (11.4)	93.59 (10.5)
Sim PO 2 (N = 300)	143.58 (10.5)	306.01 (10.3)	468.55 (11.1)	108.17 (11.7)
Optimal PO1	194.80	300.00	405.21	--
Optimal PO2	105.21	300.00	494.79	--

One-sample T-tests showed that the three simulated criteria (1- 3) were significantly different from the optimal accuracy bounds under both payoff matrix conditions (PO1: |ts| > 6.5, ps <.001; PO2: |ts| > 10.1, ps <.001).

11.6.1.5 Behaviour and BAS variation

Inter-individual differences were simulated by using a uniform distribution of BAS parameter values in the range of 0-2. Following early observations, the simulations were only run for the model that had BAS variation acting as a multiplier on the RPE projections to the synapses of the SR cell (i.e. actor network). The simulated criterion shifts under this matrix were a little greater than the shifts observed using the original payoff matrix (table 11.9)

Table 11.9. Simulated criteria when BAS acts on the RPE cell

	Criterion 1	Criterion 2	Criterion 3	Noise
PO1	175.73 (14.8)	303.67 (11.0)	429.30 (12.9)	91.24 (30.1)
PO 2	136.73 (22.5)	303.82 (11.2)	468.37 (21.2)	109.38 (43.7)

PO matrix 1

The simulations under PO matrix 1 showed that there was a positive correlation between criterion 1 and the BAS variation (r = .64, p <.001) and the correlation between BAS and criterion 3 was also significant (r = -.46, p <.001). The correlation between criterion 2 and the BAS, by contrast, was significant but very weak (r = .20, p <.001). The scatterplots below visually represent these relationships (figure 11.6).

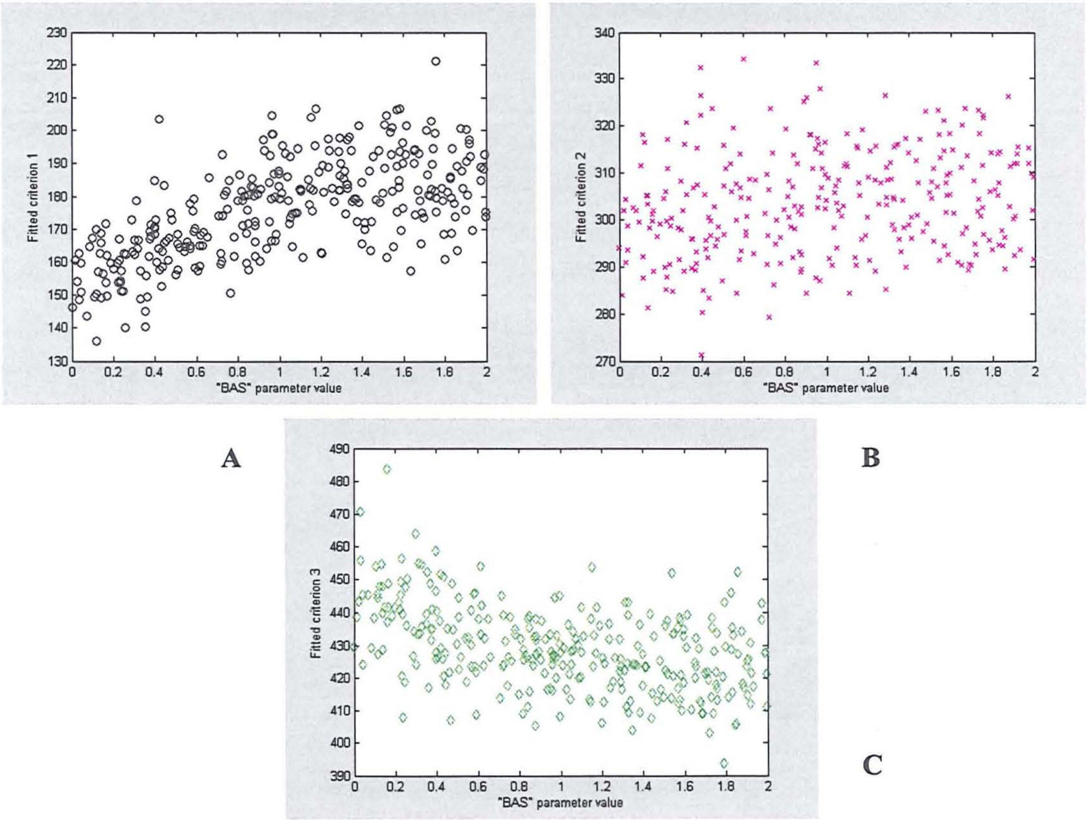


Figure 11. 6. BAS-behaviour relationship under PO matrix 1

PO matrix 2

The simulation results showed strong correlations between the criteria and the BAS variation. In particular, the BAS was negatively correlated with criterion 1 ($r = -.83$, $p < .001$) and positively with criterion 3 ($r = .82$, $p < .001$). In contrast, the correlation between criterion 2 and the BAS was weak ($r = .15$, $p = .01$). Figure 11.7 represents these relationships.

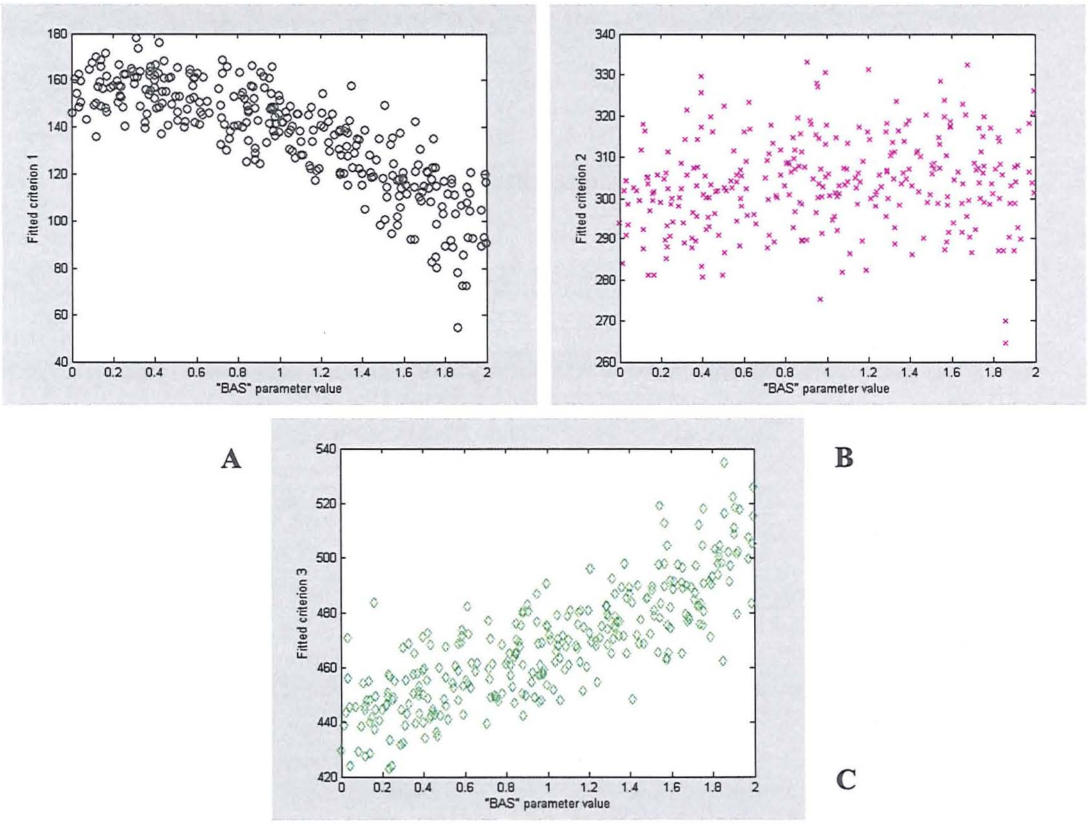


Figure 11.7. BAS-behaviour relationship under PO matrix 2

The lower BAS*criteria correlations simulated under PO matrix 2 (cf. PO matrix 1) may be due to the fact that the criterion shift was enhanced by the 'inward-force' produced by the noise values used for the edge-categories. Indeed, an inward criterion shift was observed under symmetric payoffs (table 11.6). Thus, under PO 1 the criterion shift might have not been purely due to BAS variation and, therefore, the correlations under PO matrix 1 were weaker than under PO matrix 2.

Overall, these simulations indicate that the revised task design could be implemented in a future study to explore the relationship between BAS variations and criterion shift towards the optimal reward bounds in empirical data.

11.6.2 Probabilistic feedback on category learning

As mentioned in the previous sections one of the main limitations of the studies was their high complexity and low d prime values. Evidence from the study in chapter 10, has indicated that increasing the d prime value from 1 to 2 enhanced learning on the RB task. However, in the two Gabor studies a d prime equal to 3 was implemented and the learning rate was still quite low. In fact, only a third of the sample used the optimal rule in the RB task whereas none of the participants implemented the optimal, implicit strategy in the information-integration (II) task. Participants performing the II task were found to use alternative dimensional strategies (e.g. conjunctive and uni-dimensional rules). It was also suggested that participants may have used alternative, non-dimensional strategies that could not be identified by any simple formal model. The reason why participants found it hard to identify the optimal rule amongst the several potential rules may be due to the complexity of the stimuli. Indeed, the stimuli varied along two continuous dimensions and information from both dimensions had to be taken into account to solve the categorisation problem. This is particularly true for the II task since the information from both dimensions had to be combined at a pre-decisional level. Nonetheless, results from the RB task also showed that several participants used the irrelevant dimension to solve the task. Overall, these observations indicate that the task and stimulus complexity, the probabilistic feedback, and possibly the low number of trials, impaired performance during both the RB and the II tasks. Owing to the task complexity, greater training extending over several sessions, should have been offered to participants to develop the optimal strategy. Indeed, published studies of RB and II tasks that required participants to learn to classify Gabor stimuli into different categories have often included hundreds of trials. For example, in a study by Ashby and colleagues (2003), which explored the impact of delayed-feedback on learning, four blocks of 80 trials each constituted both the RB and II tasks. In a recent study that explored the impact of partial feedback during learning on an II task, learning was assessed over five sessions that consisted of 600 trials (Ashby and O'Brien, 2007).

Alternatively, the use of a simpler task may have allowed us to detect the impact of probabilistic feedback on category learning. Indeed, probabilistic feedback may be found to affect learning on the RB tasks reported in the last three chapter of the thesis. These tasks are quite simple since the stimuli vary on only one continuous dimension (i.e. location of the stimuli on an imaginary vertical or horizontal line). In those tasks, reward magnitude manipulations (i.e. asymmetric payoffs) have been found to produce a tendency to over-classify probabilistic stimuli as members of the high payoff category (i.e. create a response bias). Moreover, neural model simulations have indicated that reward processing and response bias were mediated by DAergic activity in the striatum. More specifically, higher RPE-DA firing was found to be positively associated with greater reward-related learning and, therefore, bias. Gain-only payoffs offered in a probabilistic fashion should affect learning in a similar fashion since it should be processed by the same brain areas included in the neural model and be involved in the 3-factor learning rule model (Ashby et al., 1998; Pickering and Gray, 2001).

One of the main goals to be achieved in the future is to develop the neural model in such a way as to be able to simulate the effects of probabilistic rewards. This would require substituting reward magnitude with reward frequency and calculate the reward prediction signal as the difference between the actual reward occurrence and the predicted occurrence, which will reflect the past frequency of occurrence. Thus, the neural model implemented to simulate learning under probabilistic reward could be a variation of the delay discounting model (presented above). The main difference relies on the fact that the latter model should simulate the processes that mediate forced-choice decision making whereas the former should include the explicit system processes involved in categorisation solving.

Simulations showed that the implicit model alone was not able to solve the categorisation problem in a human-like fashion because of its inability to generalise responses to similar stimuli over learning. This may have been due to the low overlap between stimuli and the limited number of trials. The model includes a perceptual module which is responsible for applying the appropriate rule and, therefore, solving the categorisation problem. This module is mathematically expressed, although it is not neurally described. Hence, future work should attempt to build into the model the neural processes underlying the perceptual module. Bar-Gad and colleagues (2000) have suggested that the basal ganglia may be responsible for compressing the inputs from the cortex (i.e. dimensionality reduction). This process allows the projection of a great number of inputs to a small number of units while it maintains intact the information. Indeed, it has been proposed that the compression ratio between cortical neurons and striatal neurons is 10:1 (Ashby and Ennis, 2006; Ashby et al., 2007; Seger, 2008). This data compression may lead to the transformation of the cortical inputs into binary-signals (e.g. category A and category B). During each trial, one of the signals would be on whereas the other would be off. For example, if a stimulus were located at pixel location lower than 325, the on-signal would correspond to category A (i.e. on-off, 10 vs. 01 when stimulus belongs to category B). Dimensionality reduction is mathematically obtained by implementing principal component analysis, which summarises the cortical input into smaller components (Bar-Gad et al., 2000; O'Reilly and Munakata, 2000). In relation to category learning tasks, the extracted components should correspond to the different categories. Bar-Gad and colleagues have suggested that dimensionality reduction is mediated by a 3-factor Hebbian learning rule where reinforcement signals from the dopaminergic (SNc/VTA) cell play a key role in the process (Reinforcement Driven Dimensionality Reduction; RDDR). In this way, the system become more efficient at extracting information for reward-related inputs (cf. non-rewarded inputs). Joel and colleagues (2002) suggested that this process may underlie the actor sub-network to facilitate the response selection. Thus, RDDR may be implemented in the model to neurally describe the categorisation processes (i.e. explicit module).

Finally, it would be useful to extend the present (gain) neural model in such a way that it could simulate performance under a loss-only matrix as well as a mixed matrix (as described by Markman and colleagues, 2005). There is evidence that indicates that aversive learning is partly mediated by the striatum. Indeed, during performance on a Pavlovian delay conditioning task, the striatum was found to process positive RPE signals to both appetitive and aversive outcomes (Seymour et al., 2007). However, the appetitive RPE corresponded to activation in the anterior striatum whereas the aversive RPE to the posterior striatum. Moreover, Daw and colleagues (2002) have suggested that the temporal difference (TD) model can be implemented to capture RPE signals in relation to negative rewards/losses. The aversive RPE signal has been proposed to have a serotonergic nature (Daw et al., 2002; Seymour et al., 2007). In other words 5HT is believed to be involved in aversive learning. This is in line with animal studies and RST model since there is evidence that indicates that genes of the serotonin systems are associated with BIS- and FFFS-like traits (e.g. neuroticism and anxiety; Reuter, 2008). Future work should aim to implement a punishment prediction error system, similar to the system, suggested by Daw and colleagues, which could simulate aversive learning under a loss-only and a mixed-matrix.

11.7 Conclusions

In summary, the thesis explored the relationship between motivation, biologically-based personality traits and performance during category learning. In particular, it explored the impact of the BAS on learning novel categories. The use of the CL paradigm was driven by the literature that indicates that effective CL is dependent upon the same DAergic projections that constitute the biological substrate of the BAS (Maddox & Ashby, 2004; Pickering, 2004; Corr, 2006). The main finding showed that impulsivity was the personality trait that mediated categorisation under asymmetric payoffs; more specifically, low impulsivity scores were associated with a criterion shift that worked to maximise rewards. Both self-report and behavioural measures of impulsivity (i.e. delay discounting task) were found to predict the response bias.

Nonetheless, the behavioural measure seemed a more reliable index of impulsivity and, therefore, it may represent a more valid BAS endophenotype than the psychometric indices used as impulsivity phenotypes. The validity of the behavioural measure of impulsivity was further confirmed in a computerised short-delay version of delay discounting: the decision-making task.

The human data together with the simulated data reinforced the COVIS assumption that two systems are required to capture the detailed patterns of performance on CL tasks. The explicit and implicit systems are thought to be independent and compete for control throughout the task (Ashby et al., 1998). Indeed, the explicit system was responsible for implementing the appropriate rule-based strategy whereas the implicit system worked to optimise performance (i.e. winnings) by producing the criterion shifts. The modelling further assumed that explicit learning may be dominant in the early stages of the task whereas the implicit system learns in an incremental fashion following trial-by-trial feedback. The insight offered by the neural simulations indicates the importance of the model to analyse and interpret the results. Moreover, it highlights the utility of a priori simulations to assess the efficacy of the task design at exploring the personality-behavioural-biological relationship of interest.

Therefore, future studies would benefit greatly from testing the validity of the task by using neural simulation, prior to data collection, in order to ensure the use of an optimal task design able to capture the phenomenon under study. The next step should aim to extend the present neural model so that it could simulate performance on different task types (e.g. decision-making task) and CL tasks with alternative types of feedback manipulations. As mentioned above, one of the immediate steps would be to implement the changes necessary for the model to simulate performance on an RB CL task under probabilistic feedback and with differing types of payoff matrices, including losses and mixed incentives.

References

Abler, B., Walter, H., Erk, S., Kammerer, H., & Spitzer, M. (2006). Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *NeuroImage, 31*, 790 - 795.

Apicella, P., Ljungberg, T., Scarnati, E., & Schultz, W. (1991). Responses to reward in monkey dorsal and ventral striatum. *Experimental Brain Research, 85*, 491 - 500.

Arce, E., & Santisteban, C. (2006). Impulsivity: A review. *Psichotema, 18*, 213-220.

Asako, Y., Atsushi, S., Kaory, M., Hiroaki, K., & Tomifusa, K. (2004). Error-related negativity reflexts detection of negative reward prediction error. *NueroReport, 15*, 2561-2565.

Ashby, G. F., Alfonso-Reese, L. A., Turken, U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review, 105*, 442 - 481.

Ashby, G. F., & Ell, S. W. (2001). The neurobiology of human category learning. *TRENDS in Cognitive Sciences, 5*, 204 - 210.

Ashby, G. F., & Ell, S. W. (2002). Single versus multiple systems in learning and memory. In J. W. H. Pashler (Ed.), *Stevens' handbook of experimental psychology: Vol. 4 Methodology in experimental psychology* (3 ed., pp. 655 - 699). New York: Wiley.

Ashby, G. F., Ell, S. W., & Waldron, E. M. (2003). Procedural learning in perceptual categorisation. *Memory and Cognition, 31*, 1114 - 1125.

Ashby, G. F., & Ennis, J. M. (2006). The role of the basal ganglia in category learning. *The Psychology of Learning and Motivation, 46*, 1 - 36.

Ashby, G. F., Ennis, J. M., & Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorisation. *Psychological Review, 114*, 632-656.

Ashby, G. F., & Maddox, T. (2005). Human category learning. *Annual Review of Psychology, 56*, 149 - 178.

Ashby, G. F., Maddox, T., & Bohil, C. J. (2002). Observational versus feedback training in rule-based and information-integration category learning. *Memory and Cognition, 30*, 666 - 677.

Ashby, G. F., Noble, S., Filoteo, J. V., Waldron, E. M., & Ell, S. W. (2003). Category learning deficits in Parkinson's disease. *Neuropschology, 17*, 115 - 124.

Ashby, G. F., & O'Brien, C. (2007). The effects of positive versus negative feedback on information-integration category learning. *Perception and Psychophysics, 69*, 865-878.

Ashby, G. F., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *TRENDS in Cognitive Sciences, 9*, 83 - 89.

Ashby, G. F., Queller, S., & Berrety, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception and Psychophysics, 61*, 1178 - 1199.

Ashby, G. F., & Spiering, B. J. (2004). The neurobiology of category learning. *Behavioural and Cognitive Neuroscience Reviews, 3*, 101 - 113.

Ashby, G. F., & Valentin, V. (2005). Multiple systems of perceptual category learning: Theory and cognitive tests. In H. C. C. Lefebvre (Ed.), *Categorisation in cognitive science*. New york: Elsevier.

Ashby, G. F., & Valentin, V. (2007). Computational cognitive neuroscience: building and testing plausible computational models of neuroscience, neuroimaging, and behavioural data. In M. J. W. C. Schuster (Ed.), *Statistical and process models for cognitive neuroscience and aging* (pp. 15 - 58). Mahwah, NJ: Erlbaum.

Atsushi, S., & Asako, Y. (2004). The nueral basis of error processing: The error-related negativity associated with the feedback reflects response selection based on reward-prediction error. *Japanese Journal of Physiological Psychology and Psychophysiology, 22*, 19-32.

Avila, C. (2001). Distinguishing BIS-mediated and BAS-mediated disinhibition mechanisms: A comparison of disinhibition models of Gray (1981, 1987) and Patterson and Newman (1993). *Journal of Personality and Social Psychology, 80*, 311-324.

Avila, C., & Parcet, M. A. (2000). The role of Gray's impulsivity in anxiety-mediated differences in resistance to extinction. *European Journal of Personality, 14*, 185-198.

Avons, S., Nunn, J., Chan, J., & Armstrong, H. (2003). Exectutive function assessed by memory updating and random generation in schizotypal individuals. *Psychiatry Research, 120*, 145-154.

Ball, S. A., & Zuckerman, M. (1990). Sensation seeking, Eysenck's personality dimensions and reinforcement sensitivity ub concept formation. *Personality and Individual Differences, 11*, 343-353.

Bar-Gad, I., Havazalet, H. G., Goldberg, J. A., Ruppin, E., & Bergman, H. (2000). Reinforcement driven dimensionality reduction - a model for information processing in the basal ganglia. *Journal of Basic and Clinical Physiology and Pharmacology, 11*, 305-320.

Bau, C. H. D., Almeida, S., & Hutz, M. H. (2000). The TaqI A1 allele of the dopamine D2 receptor gene and alcoholism in Brazil: Association and interaction with stress and Harm Avoindance on severity prediction. *American Journal of Medical Genetics (Neuropsychiatric Genetics), 96*, 302-306.

Bau, C. H. D., & Salzano, F. M. (1995). Alcoholism in Brazil: The role of personality and susceptibility to stress. *Addiction, 90*, 693-698.

Beauducel, A., Kersting, M., & Liepmann, D. (2005). A multitrait-multimethod model for the measurement of sensitivity to reward and sensitivity to punishment. *Journal of Individual Differences, 26*, 168-175.

Beninger, R. J., & Miller, R. (1998). Dopamine D1-like receptors and reward-related incentive learning. *Neuroscience and Biobehavioural Reviews, 22*, 335-345.

Benjamin, J., Ebstein, R. P., & Belmaker, R. H. (2001). Genes for human personality traits: 'Endophenotypes' of psychiatric disorders? *World Journal of Biological Psychiatry, 2*, 54-57.

Bentall, R., Claridge, G., & Slade, P. (1989). The multidimensional nature of schizotypal traits: A factor analytic study with normal subjects. *British Journal of Clinical Psychology, 28*, 363-375.

Bickel, W. K., Odum, A. L., & Madden, G. J. (1999). Impulsivity and cigarette smoking: Delay discounting in current, never and ex-smokers. *Psychopharmacology, 146*, 447-454.

Blum, K., Braveman, E., Holder, J., Lubar, J. F., Monastra, V., Miller, D., et al. (2000). Reward deficiency syndrome: A biogenetic model for the diagnosis and treatment of impulsive, addictive, and compulsive behaviour. *Journal of Psychoactive Drugs, 32 supplement.*

Bohil, C. J., & Maddox, T. (2001). Category discriminality, base-rate, and payoff effects in perceptual categorisation. *Perception and Psychophysics, 63*, 361-376.

Bozarth, M. A. (1990). Drug addiction as a psychobiological process. In D. M. Warburton (Ed.), *Addiction controversies* (pp. 112 - 134). London: Harwood Academic Publishers.

Bozarth, M. A. (1994). Pleasure systems in the brain. In W. D.M. (Ed.), *Pleasure: The politics and the reality* (pp. 5 - 14). New York: John Wiley & Sons.

Brown, J., Bullock, D., & Grossberg, S. (1999). How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. *The Journal of Neuroscience, 19*, 10502-10511.

Buchner, A., Erdfelder, E., & Faul, F. (1997). How to use G*Power. http://www.psycho.uniduesseldorf.de/aap/projects/gpower/how_to_use_gpower.html.

Bullock, W. A., & Gilliland, K. (1993). Eysenck's arousal theory of introversion-extraversion: A converging measure investigation. *Journal of Research in Personality, 16*, 139 - 146.

Calabresi, P., Picconi, B., Tozzi, A., & DiFilippo, M. (2007). Dopamine-mediated regulation of corticostriatal synaptic plasticity. *TRENDS in Neurosciences, 30*, 211-219.

Carlsson, N. R. (2002). *Physiology of behaviour* (8 ed.). Boston: Allyn & Bacon.

Carver, C., & Scheier, M. (2003). Biological processes and personality. In C. a. Scheier (Ed.), *Perspective on Personality* (5 ed., pp. 169 - 177). New York: Allyn & Bacon.

Carver, C. S., & White, T. L. (1994). Behavioural inhibition, behavioural activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67*, 319-333.

Childress, A. R., & O'Brien, C. (2000). Dopamine receptors partial agonists could address the duality of cocaine craving. *Trends in Pharmacological Sciences, 21*, 6-9.

Claridge, G., McCreery, C., Mason, O., Bentall, R., Boyle, G. L., Slade, P., et al. (1996). The factor structure of 'schizotypal' traits: A large replication study. *British Journal of Clinical Psychology, 35*, 103-115.

Clark, L., Robbins, T. W., Ersche, K. D., & Sahakian, B. J. (2006). Reflection impulsivity in current and former substance users. *Biological Psychiatry, 60*, 515-522.

Cohen, M. X., Young, J., Baek, J.-M., Kessler, C., & Ranganath, C. (2005). Individual differences in extraversion and dopamine genetics predict neural reward responses. *Cognitive Brain Research, 25*, 851-861.

Comings, D. E., Comings, B. G., Muhleman, D., Dietz, G., Shahbahrami, B., Tast, D., et al. (1991). The dopamine D2 receptor locus as a modifying gene in neuropsychiatric disorders. *Journal of the American Medical Association, 266*, 1793-1800.

Cools, R. (2006). Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. *Neuroscience and Biobehavioural Reviews, 30*, 1-23.

Cools, R., Barker, R. A., Sahakian, B. J., & Robbins, T. W. (2003). L-DOPA medication remediates cognitive inflexibility, but increases impulsivity in patients with Parkinson's disease. *Neuropsychologia, 41*, 1431-1441.

Cools, R., Sheridan, M., Jacobs, E., & D'Esposito, M. (2007). Impulsive personality predicts dopamine-dependent changes in frontostriatal activity during component processes of working memory. *The Journal of Neuroscience, 27*, 5506-5514.

Corr, P. J. (2002). J. A. Gray's reinforcement sensitivity theory and frustrative nonreward: A theoretical note on expectancies in reactions to reward stimuli. *Personality and Individual Differences, 32*, 1247-1253.

Corr, P. J. (2004b). J.A. Gray's reinforcement sensitivity theory: Tests of the joint subsystems hypothesis of anxiety and impulsivity. *Personality and Individual Differences, 33*, 511-532.

Corr, P. J. (2008). Reinforcement Sensitivity Theory (RST): Introduction. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 1-43). Cambridge: Cambridge University Press.

Corr, P. J. (2004a). Reinforcement sensitivity theory and personality. *Neuroscience and Biobehavioural Reviews, 28*, 317-332.

Corr, P. J. (2006). *Understanding biological psychology*. Oxford: Blackwell Publishing.

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

Crowe, E., & Higgins, E. T. (1997). Regulatory focus and strategic inclinations: Promotion and prevention focus in decision-making. *Organizational, Behaviour and Human Decision Processes, 69*, 117-132.

Davis, C., Levitan, R. D., Kaplan, A. S., Carter, J., C., R., Curtis, C., et al. (2008). Reward sensitivity and the D2 dopamine receptor gene: A case-control study of binge eating disorder. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 32*, 620-628.

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology, 16*, 199-204.

Dawe, S., Gullo, M. J., & Loxton, N. J. (2004). Reward drive and rash impulsiveness as dimensions of impulsivity: Implications for substance misuse. *Addictive Behaviour, 29*, 1389 - 1405.

Dawe, S., & Loxton, N. J. (2004). The role of impulsivity in the development of substance use and eating disorders. *Neuroscience and Biobehavioural Reviews, 28*, 343 - 351.

Dayan, P. (1990). Reinforcement comparison. In D. S. Touretzky, J. L. Elman, T. J. Sejnowsky & G. E. Hinton (Eds.), *Proceedings of the 1990 connectionist models summer school* (pp. 45-51). San Matteo, CA: Morgan Kaufman.

Dayton, M. (2003). Model comparisons using information measures. *Journal of Modern Applied Statistical Methods, 2*, 281-292.

de Wit, H., & Wise, R. A. (1977). Blockade of cocaine reinforcement in rats with dopamine receptor blocker pimozide, but not with the noradrenergic blockers phentolamine or phenoxybenzamine. *Canadian Journal of Psychology, 31*, 195 - 203.

DeCaro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than morre. *Cognition, 107*, 284-294.

Delgado, M., Miller, M., Inati, S., & Phelps, E. (2005). An fMRI study of reward-related probability learning. *NeuroImage, 24*, 862-873.

Delgado, M. R., Locke, H. M., Stenger, V. A., & Fiez, J. A. (2003). Dorsal striatum responses to reward and punishment: Effects of valence and magnitude maniupulations. *Cognitive, Affective, & Behavioral Neuroscience, 3*, 27-38.

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses of reward and punishment in the striatum. *Journal of Neurophysiology, 85*, 3072-3077.

Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioural and Brain Sciences, 22*, 491-569.

DiChiara, G. (2002). Nucleus accumbens shell and core dopamine: Differential role in behaviour and addiction. *Behavior and Brain Research, 137*, 75 - 114.

DiChiara, G., & North, A. R. (1992). Neurobiology of opiate abuse. *Trends in Pharmacological Sciences, 13*, 185 - 193.

Dickman, S. J. (1990). Functional and dysfunctional impulsivity: Personality and cognitive correlates. *Journal of Personality and Social Psychology, 58*, 95 - 102.

Eisenberg, D. T. A., MacKillop, J., Modi, M., Beauchemin, J., Dang, D., Lisman, S. A., et al. (2007). Examining impulsivity as an endophenotype using a behavioural approach: a DRD2 TaqI and DRD4 48-bp VNTR association study. *Behavioural and Brain Functions, 3.*

Evenden, J. L. (1999). Varieties of impulsivity. *Psychopharmacology, 146*, 348-361.

Eysenck, H. J. (1997). Personality and experimental psychology: The unification of psychology and the possibility of a paradigm. *Journal of Personality and Social Psychology, 73*, 1224 - 1237.

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire (Junior and Adult)*. London: Hodder and Stoughton.

Eysenck, S. B. G., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences, 6*, 21-29.

Field, M., Rush, M., Cole, J., & Goudie, A. (2007). The smoking Stroop and delay discounting in smokers: effects of environmental smoking cues. *Journal of Psychopharmacology, 21*, 603-610.

Field, M., Santarcangelo, M., Sumnall, H., Goudie, A., & Cole, J. (2006). Delay discounting and the behavioural economics of cigarette purchases in smokers: the effects of nicotine deprivation. *Psychopharmacology, 186*, 255-263.

Filoteo, J. V., Maddox, T., Ing, A. D., & Song, D. D. (2007). Characterizing rule-based category learning deficits in patients with Parkinson's disease. *Neuropsychologia, 45*, 305 - 320.

Filoteo, J. V., Maddox, T., Ing, A. D., Zizak, V., & Song, D. D. (2005b). The impact of irrelevant dimensional variations on rule-based category learning in patients with Parkinson's disease. *Journal of the International Neuropsychological Society, 11*, 503 - 513.

Filoteo, J. V., Maddox, T., Simmons, A. N., Ing, A. D., Cagigas, X. E., Matthews, S., et al. (2005a). Cortical and subcortical brain regions involved in rule-based category learning. *Brain Imaging, 16*, 111 - 115.

Fiorino, D. F., & Phillips, A. G. (1999). Facilitation of sexual behaviour in male rats dollowing d-amphetamine-induced behavioural sensitisation. *Psychopharmacology, 142*, 200-208.

Floresco, S. B., Magyar, O., Ghods-Sharifi, S., Vexelman, C., & Tse, M. T. (2006). Multiple dopamine receptors subtypes in the medial prefrontal cortex of the rat regulate set-shifting. *Neuropsychopharmacology, 31*, 297-309.

Fowles, D. (2006). Jeffrey Gray's Contributions to Theories of Anxiety, Personality, and Psychopathology. In C. T. (Ed.), *Biological bais of personality and individual differences* (pp. 14-34). New York: Guilford Press.

Fox, H. C., McLean, A., Turner, J. J. D., Parrott, A. C., Rogers, R., & Sahakian, B. J. (2002). Neuropsychological evidence of a relatively selective profile of temporal dysfunction in drug-free MDMA ("ecstasy") polydrug users. *Psychopharmacology, 162*, 203-214.

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 16311-16316.

Franken, I. H. A. (2002). Behavioural approach system (BAS) sensitivity predicts alcohol craving. *Personality and Individual Differences, 32*, 349-355.

Franken, I. H. A., Booij, J., & van den Brink, W. (2005a). The role of dopamine in human addiction: From reward to motivated attention. *European Journal of Pharmacology, 526*, 199-206.

Franken, I. H. A., & Muris, P. (2006). BIS/BAS personality characteristics and college students' substance use. *Personality and Individual Differences, 40*, 1497-1503.

Franken, I. H. A., & Muris, P. (2006). Gray's impulsivity dimension: A distinction between Reward Sensitivity versus Rash impulsiveness. *Personality and Individual Differences, 40*, 1337 - 1347.

Franken, I. H. A., & Muris, P. (2005b). Individual differences in decision-making. *Personality and Individual Differences, 39*, 991-998.

Franken, I. H. A., Muris, P., & Georgieva, I. (2006). Gray's model of personality and addiction. *Addictive Behaviour, 31*, 399-403.

Franken, I. H. A., Van Strien, J. W., Franzek, E. J., & van de Wetering, B. J. (2007). Error-processing deficits in patients with cocaine dependence. *Biological Psychology, 75*, 45-51.

Franken, I. H. A., van Strien, J. W., Nijs, I., & Muris, P. (2008). Impulsivity is associated with behavioural decision-making deficits. *Psychiatry Research, 46*, 263-269.

Gibbons, H., & Rammsayer, T. H. (1999). Differential effects of personality traits related to the P-ImpUSS dimension on latent inhibition in healthy female subjects. *Personality and Individual Differences, 27*, 1157-1166.

Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic category learning. *Learning and Memory, 9*, 408 - 418.

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron, 14*, 477 - 485.

Gottesman, I. J., & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry, 106*, 636-645.

Grace, A. A., Floresco, S. B., Goto, Y., & Lodge, D. J. (2007). Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends in Neurosciences, 30*, 220 - 227.

Grant, S., Contoreggi, C., & London, E. D. (2000). Drug abusers show impaired performance in a laboratory test of decision making. *Neuropsychologia, 38*, 1180-1187.

Gray, J. A. (1970). The psychophysiological basis of introversion-extraversion. *Behavioural Research and Therapy, 8*, 249-266.

Gray, J. A. (1987). Perspectives on anxity and impusivity: A commentary. *Journal of Research in Personality, 21*, 493 - 509.

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety*. Oxford: Oxford University Press.

Gray, N. S., Fernandez, M., Williams, J., Ruddle, R. A., & Snowden, R. J. (2002). Which schizotypal dimensions abolish latent inhibition? *British Journal of Clinical Psychology, 41*, 271-284.

Gray, N. S., Pickering, A. D., Hemsley, D. R., Dawling, S., & Gray, J. A. (1992). Abolition of latent inhibition by a single 5 mg dose of d-amphetamine in man. *Psychopharmacology, 107*, 425-430.

Gray, N. S., Pickering, A. D., Snowden, R. J., Hemsley, D. R., & Gray, J. A. (2002). The partial reinforcement extinction effect in humans: Effects of schizophrenia, schizotypy and low doses of amphetamine. *Behavioural Brain Research, 133*, 333-342.

Ham, L. S., & Hope, D. A. (2003). College students and problematic drinking: A review of the literature. *Clinical Psychology Review, 23*, 719-759.

Harnad, S. (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. In H. S. (Ed.), *Categorical Perception: The Groundwork of Cognition* (pp. Chapter 1). New York: Cambridge University Press.

Heath, R. G. (1964). Pleasure response of human subjects to direct stimulation of the brain: Physiologic and psychodynamic considerations. In R. G. Heath (Ed.), *The role of pleasure in human behaviour* (pp. 219 - 243). New York: Hoeber.

Heery, E., Robinson, B., McMahon, R., & Gold, J. (2007). Delay discounting and schizophrenia. *Cognitive Neuropsychiatry, 12*, 213-221.

Heinz, A. (2002). Dopaminergic dysfunction in alcoholism and schizophrenia - psychological and behavioural correlates. *European Psychiatry, 17*, 9-16.

Hellman, S. G., Kern, R. S., Neilson, L. M., & Green, M. F. (1998). Monetary reinforcement and Wisconsin Card Sorting performance in Schizophrenia: Why show me the money? *Schizophrenia Research, 34*, 67-75.

Heym, N., Ferguson, E., & Lawrence, C. (in press). An evaluation of the relationship between Gray's revised RST and Eysenck's PEN: Distinguishing BIS and FFFS in Carver and White's BIS/BAS scales. *Personality and Individual Differences*.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*, 319 - 340.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*, 1280-1300.

Hogan, M. J. (1966). Influence of motivation on reactive inhibition in extraversion-introversion. *Perceptual and Motor Skills, 22*, 187 - 192.

Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature neuroscience, 1*, 304 - 309.

Honkanen, A., Hyytia, P., Korpi, E. R., & Ahtee, L. (1999). Effects of morphine on metabolism of dopamine and serotonin in brains of alcohol preferring AA and alcohol-avoiding ANA rats. *Alcohol, 18*, 3-10.

Jenkins, B. G., Sanchez-Pernaute, R., Brownell, A.-L., Chen, Y.-V. I., & Isacson, O. (2004). Mapping dopamie function in primates using pharmacologic magnetic resonance imaging. *The Journal of Neuroscience, 24*, 9553 - 9560.

Joel, D., Niv, Y., & Ruppin, E. (2002). Actor-critic models of the basal ganglia: New anatomical and computational perspectives. *Neural Networks, 15*, 535-547.

Johnson, M. W., Bickel, W. K., & Baker, F. (2007). Moderate drug use and delay discounting: A comparison of heavy, light and never smokers. *Experimental and Clinical Psychopharmacology, 15*, 187-194.

Jonsson, E. G., Nothen, M., Gustavsson, J. P., Neidt, H., Brene, S., Tylec, A., et al. (1997). Lack of evidence for allelic association between personality traits and the dopamine D4 receptor gene polymorphisms. *American Journal of Psychiatry, 154*, 697-699.

Jouhn, O. P., Donohue, E. M., & Kentle, B. L. (1991). *The big five inventory - version 4a and 54*. Berkley, CA: University of California, Institute of Personality and Social Research.

Juckel, G., Schlagenhauf, F., Koslowski, M., Filonov, D., Wüstenberg, T., Villringer, A., et al. (2006b). Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not typical, neuroleptics. *Psychopharmacology, 187*, 222 - 228.

Juckel, G., Schlagenhauf, F., Koslowski, M., Wüstenberg, T., Villringer, A., Knutson, B., et al. (2006a). Dysfunction of ventral striatal reward prediction in schizophrenia. *NeuroImage, 29*, 409 - 416.

Julien, R. M. (2003). *A primer of drug action* (9 ed.). New York: Worth Publisher.

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience, 10*, 1625 - 1634.

Kalenscher, T., & Pennartz, C. M. A. (2008). Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. *Progress in Neurobiology, 84*, 284-315.

Katsuragy, S., Kiyota, A., Tsutsumi, T., Isogawa, K., Nagayama, H., Arinami, T., et al. (2001). Lack of association between a polymorphism in the promoter region of the dopamine D2 receptor and personality traits. *Psychiatry Research, 105*, 123-127.

Keri, S. (2003). The cognitive neuroscience of category learning. *Brain Research Reviews, 43*, 85 - 109.

Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin addicts discount delayed rewards at higher rates than non-drug using controls. *Journal of Experimental Psychology: General, 128*, 78-87.

Klein, T., Neumann, J., Reuter, M., Hennig, J., von Cramon , D. Y., & Ullsperger, M. (2007). Genetically determined differences in learning from errors. *Science, 318*, 1642-1645.

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

Knutson, B., Bjork, J. M., Fong, G. W., Hommer, D., Mattay, V. S., & Weinberger, D. R. (2004). Amphetamine modulates human incentive processing. *Neuron, 43*, 261 - 269.

Koo, M., Levitt, J., McCarley, R., Seidman, L., Dickey, C., Niznikiewicz, M., et al. (2006). Reduction of caudate nucleus volumes in neuroleptic-naive female subjects with schizotypal personality disorder. *Biological Psychiatry, 60*, 40-48.

Koob, G. F., & LeMoal, M. (1997). Drug abuse: Hedonic homeostatic dysregulation. *Science, 278*, 52-58.

Kruschke, J. K. (2005). Category learning. In K. L. L. Goldstone (Ed.), *The handbook of cognition* (pp. 183 - 201). London: Sage.

Kumari, V., Cotter, P., Mulligan, O., Checkley, S., Gray, N. S., Hemsley, D. R., et al. (1999). Effects of d-amphetamine and haloperidol on latent inhibition in healthy male volunteers. *Journal of Psychopharmacology, 13*, 398-405.

Lee, S. H., Ham, B., Cho, Y., Lee, S., & Shim, S. H. (2007). Association study of dopamine receptor D2 TaqI A polymorphism and reward-related personality traits in healthy Korean young females. *Neuropschobiology, 56*, 146-151.

Ljungberg, T., Apicella, P., & Schultz, W. (1991). Responses of monkey midbrain dopamine neurons during delayed alternation performance. *Brain Research, 567*, 337 - 341.

Ljungberg, T., Apicella, P., & Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *Journal of Neurophysiology, 67*, 145 - 163.

Maddox, T., & Ashby, G. F. (2004). Dissociating explicit and procedural learning based systems of perceptual category learning. *Behavioural Processes, 66*, 309 - 332.

Maddox, T., Ashby, G. F., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology: Learning, Memory & Cognition, 29*, 650-662.

Maddox, T., Ashby, G. F., Ing, A. D., & Pickering, A. D. (2004). Disrupting feedback processing interferes with rule-based but not information-integration category learning. *Memory and Cognition, 32*, 582 - 591.

Maddox, T., Baldwin, G. C., & Markman, A. B. (2006). A test of the regulatory fit hypothesis in perceptual classification learning. *Memory and Cognition, 34*, 1377 - 1397.

Maddox, T., & Bohil, C. J. (1998). Base-rate and payoff effects in multi-dimensional perceptual categorisation. *Journal of Experimental Psychology: learning, Memory and Cognition, 24*, 1459-1482.

Maddox, T., & Bohil, C. J. (2003). A theoretical framework for understanding the effects of simultaneous base-rate and payoff manipulations on decision criterion learning in perceptual categorisation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*, 307 - 320.

Maddox, T., & Bohil, C. J. (2005). Optimal classifier feedback improves cost-benefit but not base-rate decision criterion learning in perceptual categorisation. *Memory and Cognition, 33*, 303-319.

Maddox, T., Bohil, C. J., & Dodd, J. L. (2003). Linear transformation of the payoff matrix and decision criterion learning in perceptual categorisation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*, 1174 - 1193.

Maddox, T., & Dodd, J. L. (2001). On the relation between base-rate and cost-benefit learning in simulated medical diagnosis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27*, 1367 - 1384.

Maddox, T., & Filoteo, J. V. (2001). Striatal contributions to category learning: Quantitative modeling of simple linear and complex nonlinear rule learning in patients with Parkinson's disease. *Journal of the International Neuropsychological society, 7*, 710 - 727.

Maddox, T., Filoteo, J. V., Hejl, K. D., & Ing, A. D. (2004). Category number impacts on rule-based but not information-integration category learning: Further evidence for dissociable category-learning systems. *Journal of Experimental Psychology: Learning, Memory and Cognition, 30*, 227 - 245.

Maddox, T., Markman, A. B., & Baldwin, G. C. (2007). Using classification to understand the motivation-learning interface. *The Psychology of Learning and Motivation, 47*, 213-249.

Markman, A. B., Baldwin, G. C., & Maddox, T. (2005). The interaction of payoff structure and regulatory focus in classification. *Psychological Science, 16*, 852-855.

Martin-Sölch, C., Magyar, S., Künig, G., Missimer, J., Schultz, W., & Leenders, K. L. (2001). Changes in brain activation associated with reward processing in smokers and nonsmokers. *Experimental Brain Research, 139*, 278-286.

Mason, O., Claridge, G., & Jackson, M. (1995). New scales for the assessment of schizotypy. *Personality and Individual Differences, 18*, 7-13.

McNaughton, N., & Corr, P. J. (2004). A two-dimensional neuropsychology of defense: Fear/anxiety and defensive distance. *Neuroscience and Biobehavioural Reviews, 28*, 285-305.

Mendrek, A., Blaha, C. D., & Phillips, A. G. (1998). Pre-exposure of rats to amphetamine sensitizes self-administration of this drug under a progressive ration schedule. *Psychopharmacology, 135*, 416 - 422.

Mirenowicz, J., & Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *Journal of Neurophysiology, 72*, 1024 - 1027.

Mitchell, S. H. (1999). Measures of impulsivity in cigarette smokers and non-smokers. *Psychopharmacology, 146*, 455-464.

Mitchell, S. H. (2004). Effects of short-term nictoine deprivation on decision-making: Delay, uncertainty and effort discounting. *Nicotine and Tobacco Research, 6*, 819-828.

Mitchell, S. H. (2004). Measuring impulsivity and modeling its association with cigarette smoking. *Behavioural and Cognitive Neuroscience Reviews, 3*, 261-275.

Mohanty, A., Herrington, J. D., Koven, N. S., Fisher, J. E., Wenzel, E. A., Webb, A. G., et al. (2005). Neural mechanisms of attentive interference in schizotypy. *Journal of Abnormal Psychology, 114*, 16-27.

Monterosso, J., & Ainslie, G. (1999). Beyond discounting: Possible experimental models of impulsive control. *Psychopharmacology, 146*, 339-347.

Morgan, M. J. (1998). Recreational use of "ecstasy" (MDMA) is associated with elevated impulsivity. *Neuropsychopharmacology, 19*, 252-264.

Motulsky, H., & Christopulos, A. (2004). *Fitting models to biological data using linear and nonlinear regression*. New York: Oxford University Press.

Munafo, M. R., Clark, T. G., Payne, E., Walton, R., & Flint, J. (2003). Genetic polymorphism and personality in healthy adults: A systematic review and meta-analysis. *Molecular Psychiatry, 8*, 471-484.

Murray, G., Corlett, P., Clark, L., Pessiglione, M., Blackwell, A., Honey, G., et al. (2007). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry*, 1-10.

Newman, J. P., Kosson, D. S., & Patterson, C. M. (1992). Delay of gratification in psychopathic and nonpsychpathic offenders. *Journal of Abnormal Psychology, 101*, 630-636.

Newman, J. P., Widom, C. S., & Nathan, S. (1985). Passive avoidance in syndromes of disinhibition: Psychopathy and extraversion. *Journal of Personality and Social Psychology, 48*, 1316-1327.

Nichols, S. L., & Newman, J. P. (1986). Effects of punishment on response latency in extraverts. *Journal of Personality and Social Psychology, 50*, 624-630.

Noble, E. P. (1998). The D2 dopamine receptor gene: A review of association studies in alcoholism and phenotypes. *Alcohol, 16*, 33-45.

Nomura, E. M., Maddox, T., Filoteo, J. V., Ing, A. D., Gitelman, D. R., Parrish, T. B., et al. (2007). Neural correlates of rule-based and information-integration visual category learning. *Cerebral Cortex, 17*, 37 - 43.

Nomura, E. M., & Reber, P. J. (2008). A review of medial temporal lobe and caudate contributions to visual category learning. *Neuroscience and Biobehavioural Reviews, 32*, 279-291.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature neuroscience, 4*, 95 - 102.

O'Doherty, J. P. (2004). Reward representation and reward-related learning in the human brain: Insight from neuroimaging. *Current Opinion in Neurobiology, 14*, 769 - 776.

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.

O'Reilly, R. C. O., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience*. Cambridge: Massachussets Institute of Technology.

Ozkaragoz, T., & Noble, E. P. (2000). Extraversion: Interaction between D2 dopamine receptor polymorphisms and parental alcoholism. *Alcohol, 22*, 139-146.

Pallant, J. (2001). *SPSS survival manual*. Buckingham: Open University.

Pappata, S., Dehaene, S., Poline, J. B., Gregoire, M. C., Jobert, A., Delforge, J., et al. (2002). in vivo detection of striatal dopamine release during reward: A PET study with [11C] raclopride and a single dynamic scan approach. *NeuroImage, 16*, 1015 - 1027.

Pardo, Y., Aguilar, R., Molinuevo, B., & Torrubia, R. (2007). Alcohol use as a behavioural sign of disinhibition: Evidence from J. A. Gray's model of personality. *Addictive Behaviour, 32*, 2398-2403.

Patterson, C. M., Kosson, D. S., & Newman, J. P. (1987). Reaction to punishment, reflectivity, and passive avoidance learning in extraverts. *Journal of Personality and Social Psychology, 52*, 565-575.

Patterson, C. M., & Newman, J. P. (1993). Reflectivity and learning from aversive events: Towards a psychological mechanism for the syndromes of disinhibition. *Psychological Review, 100*, 716-736.

Pearce-McCall, D., & Newman, J. P. (1986). Expectation of success following noncontingent punishment in introverts and extraverts. *Journal of Personality and Social Psychology, 50*, 439-446.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature, 332*, 1042 - 1045.

Pickering, A. D. (2004). The neuropsychology of Impulsive Antisocial Sensation Seeking personality traits: From dopamine to hippocampal function? In R. M. Stelmack (Ed.), *Psychology of Personality*: Elsevier Ltd.

Pickering, A. D. (2008). Formal and computational models of Reinforcement Sensitivity Theory. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 453-481). Cambridge: Cambridge University Press.

Pickering, A. D., & Gray, J. A. (2001). Dopamine, appetitive reinforcement and the neuropsychology of human learning: An individual differences approach. In E. Angleitner (Ed.), *Advances in Research on Temperament* (pp. 113-149). Lengerich, Germany: PABST Science Publisher.

Pinel, P. J. (2006). *Biopsychology* (6 ed.). London: Allyn & Bacon.

Pohjalainen, T., Rinne, J. O., Nagren, K., Lehikoinen, P., Anttila, K., Syvalahti, E., et al. (1998). The A1 allele of the human D2 dopamine receptor gene predicts low D2 receptor availability in healthy volunteers. *Molecular Psychiatry, 3*, 256-260.

Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. E. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropschology, 13*, 564 - 574.

Ponce, G., Jimenez-Arriero, M. A., Rubio, G., Hoenicka, J., Ampuero, I., Ramos, J. A., et al. (2003). The A1 allele of the DRD2 gene (TaqIA polymorphism) is associated with antisocial personality in a sample of alcohol-dependent patients. *European Psychiatry, 18*, 356-360.

Powell, J., Dawkins, L., & Davis, R. E. (2002). Smoking, reward responsiveness, and response inhibition: Tests: of an incentive motivational model. *Biological Psychiatry, 51*, 151-163.

Price, A. (2006). Explicit category learning in Parkinson's disease: deficits related to impaired rule generation and selection processes. *Neuropschology, 20*, 249 - 257.

Raine, A. (1991). The SPQ: A scale for the assessment of schizotypal personality based on DSM-III-R criteria. *Schizophrenia Bulletin, 17*, 555-564.

Rao, S. M., Bobholz, J. A., Hammeke, T. A., Rosen, A. C., Woodley, S. J., Cunningham, J. M., et al. (1997). Functional MRI evidence for subcortical participation in conceptual reasoning skills. *Neuroreport, 8,* 1987 - 1993.

Reuter, M. (2008). Neuro-imaging and genetics. In P. J. Corr (Ed.), *The Reinforcement Sensitivity Theory of personality* (pp. 317-344). Cambridge: Cambridge University Press.

Reuter, M., Schmitz, A., Corr, P. J., & Hennig, J. (2006). Molecular genetics support Gray's personality theory: The interaction of COMT and DRD2 polymorphisms predicts the behavioural approach system. *International Journal of Neuropsychopharmacology, 9,* 155-166.

Richards, J. B., Zhang, L., Mitchell, S. H., & de Witt, H. (1999). Delay of probability discounting in a model of impulsive behaviour: Effect of alcohol. *Journal of the Experimental Analysis of Behaviour, 71,* 121-143.

Roberts, A. C., De Salvia, M. A., Wilkinson, L. S., Collins, P. F., Muir, J. L., & Everitt, B., J. (1994). 6-hydroxydopamine lesions of the prefrontal cortex in monkeys enhance performance on an analog of the Wisconsin Card Sort test: Possible interactions with subcortical dopamine. *Journal of Neuroscience, 14,* 2531 - 2544.

Robinson, T. E., & Berridge, K. C. (2001). Incentive-sensitization and addiction. *Addiction, 96,* 103 - 114.

Robinson, T. E., & Berridge, K. C. (2003). Addiction. *Annual Review of Psychology, 54,* 25 - 53.

Rommelse, N. N. J., Altink, M. E., Martin, N. C., Buschgens, C. J. M., Faraone, S. V., Buitelaar, J. K., et al. (2008). Relationship between endophenotype and phenotype in ADHD. *Behavioral and Brain Functions, 4,* 1-14.

Routtenberg, A., & Lindy, J. (1965). Effects of the availability of rewarding septal and hypothalamic stimulation on bar pressing for food under conditions of deprivation. *Journal of Comparative and Physiological Psychology, 60*, 158 - 161.

Ruchsow, M., Spitzer, M., Gron, G., Grothe, J., & Kiefer, M. (2005). Error-processing and impulsiveness in normals: Evidence from event-related potentials. *Cognitive Brain Research, 24*, 317-325.

Schotanus, S., & Chergui, K. (2008). Dopamine D1 receptors and group I metabotropic glutamate receptors contribute to the induction of long-term potentiation in the nucleus accumbens. *Neuropharmacology, 54*, 837-844.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology, 80*, 1-27.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron, 36*, 241-263.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology, 57*, 87 - 115.

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*, 900-913.

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

Schultz, W., Tremblay, L., & Hollerman, J. (2003). Changes in behaviour-related neuronal activity in the striatum during learning. *Trends in Neurosciences, 26*, 321-328.

Seger, C. A. (2008). How do basal ganglia contribute to categorisation? Their roles in generalisation, response selection, and learning via feedback. *Neuroscience and Biobehavioural Reviews, 32*, 265-278.

Seymour, B., Daw, N. D., Dayan, P., Singer, T., & Dolan, R. J. (2007). Differential encoding of losses and gains in the human striatum. *The Journal of Neuroscience, 27*, 4826-4831.

Shah, J., Higgins, E. T., & Friedman, R. S. (1998). Performance incentives and means: How regulatory focus influences goal attainment. *Journal of Personality and Social Psychology, 74*, 285-293.

Shohamy, D., Myers, C., & Gluck, M. A. (2008). Basal ganglia and dopamine contribution to probabilistic category learning *Neuroscience and Biobehavioural Reviews, 32*, 219-236.

Smillie, L. D., Dalgleish, L. I., & Jackson, C. J. (2007). Distinguishing between learning and motivation in behavioural tests of the reinforcement sensitivity theory of personality. *Personality and Social Psychology Bulletin, 33*, 476 - 489.

Smillie, L. D., & Jackson, C. J. (2005). The appetitive motivation scale and other BAS measures in the prediction of Approach and Active Avoidance. *Personality and Individual Differences, 38*, 981-994.

Smillie, L. D., & Jackson, C. J. (2006). Funtional impulsivity and reinforcement sensitivity theory. *Journal of Personality, 74*, 47 - 84.

Smillie, L. D., Jackson, C. J., & Dalgleish, L. I. (2006). Conceptual distinctions among Carver and White's (1994) BAS scales: A reward-reactivity versus trait impulsivity perspective. *Personality and Individual Differences, 40*, 1039-1050.

Smillie, L. D., Pickering, A. D., & Jackson, C. J. (2006). The new reinforcement sensitivity theory: Implications for personality measurement. *Personality and Social Psychology Review, 10*, 320 - 335.

Smith, J. D., & Kemler-Nelson, D. G. (1984). Overall similarity in adults' classification: The child in all of us. *journal of Experimental Psychology: General, 113*, 137 - 159.

Stainslaw, H., & Todorov, N. (1999). Calculations of signal detection theory measures. *Behaviour Research methods, Instruments and Computers, 31*, 137 - 149.

Stanfield, K. H., & Kirstein, C. L. (2005). Neurochemical effects of cocaine in adolescence compared to adulthood. *Developmental Brain Research, 159*, 119-125.

Stewart, J., de Wit, H., & Eikelboom, R. (1984). Role of unconditioned and conditioned drug effects in self-administration of opiates and stimulants. *Psychological Review, 91*, 251 - 268.

Stuettgen, M. C., Hennig, J., Reuter, M., & Netter, P. (2005). Novelty Seeking but not BAS is associated with high dopamine as indicated by a neurotransmitter challenge test using mazindol as a challenge substances. *Personality and Individual Differences, 38*, 1597-1608.

Sullivan, P. F., Fifield, W. J., Kennedy, M. A., Mulder, R. T., Sellman, J. D., & Joyce, P. R. (1998). No association between Novelty Seeing and the type 4 dopamine receptor gene (DRD4) in two New Zealand samples. *American Journal of Psychiatry, 155*, 98-101.

Suri, R. E., & Schultz, W. (1998). Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Experimental Brain Research, 121*, 350-354.

Suri, R. E., & Schultz, W. (1999). A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience, 91*, 871-890.

Surmeier, D. J., Ding, J., Day, M., Wang, Z., & Shen, W. (2007). D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *TRENDS in Neurosciences, 30*, 228-235.

Sutton, R. S., & Barto, A. G. (1981). Towards a modern theory of adaptive networks: Expectation and prediction. *Psychological Review, 88*, 135 - 170.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4 ed.). Needham Heights: Allyn & Bacon.

Tharp, I. (2007). *The effect of personality on attentional strategy in category learning*. Unpublished PhD thesis, Goldsmiths (University of London), London.

Tharp, I., & Pickering, A. (under review). A discussion of DeCaro, Thomas and Beilock (2008): Further data demonstrates complexities in the assessment of category learning performance. *Cognition*.

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005a). Adaptive coding of reward value by dopamine neurons. *Science, 307*, 1642 - 1645.

Tobler, P. N., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2005b). Human neural learning depends on reward predictions errors in the blocking paradigm. *Journal of Neurophysiology, 95*, 301 - 310.

Torrubia, R., Avila, C., Molto, J., & Caseras, X. (2001). The sensitivity to punishment and sensitivity to reward questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Personality and Individual Differences, 31*, 837 - 862.

Vigil-Colet, A. (2007). Impulsivity and decision-making in the balloon analogues risk-taking task. *Personality and Individual Differences, 37*, 43-45.

Volkow, N. D., Fowler, J. S., & Wang, G.-J. (2003). The addicted human brain: Insights from imaging studies. *The Journal of Clinical Investigation, 111*, 1444 -1451.

Volkow, N. D., Wang, G.-J., Fowler, J. S., Logan, J., Gatley, S. J., Gifford, A., et al. (1999). Prediction of reinforcing responses to psychostimulants in humans by brain dopamine D2 receptor levels. *American Journal of Psychiatry, 156*, 1440-1443.

Volkow, N. D., Wang, G.-J., Fowler, J. S., Logan, J., Gatley, S. J., Hitzeman, R., et al. (1997). Deecreased striatal dopaminergic responsiveness in detoxified cocaine-dependent subjects. *Nature, 386*, 830-833.

Volkow, N. D., Wang, G.-J., Fowler, J. S., Logan, J., Gatley, S. J., Wong, C., et al. (1999). Reinforcng effects of psychostimulants in humans are associated with increases in brain dopamine and occupancy of D2 receptors. *The Journal of Pharmacology and Experimental Therapeutics, 291*, 409-415.

Vollema, M. G., & van den Bosch, R. (1995). The multidimensionality of schizotypy. *Schizophrenia Bulletin, 21*, 19-31.

Wacker, J., Chavanon, M.-L., & Stemmler, G. (2006). Investigating the dopaminergic basis of extraversion in humans: A multilevel approach. *Journal of Personality and Social Psychology, 91*, 171-187.

Wade, T. R., De Witt, H., & Richards, J. B. (2000). Effects of dopaminergic drugs on delayed reward as a measure of impulsive behaviour in rats. *Psychopharmacology, 150*, 90-101.

Waldek, T. L., & Miller, L. S. (1997). Gender and impulsivity differences in licit substance use. *Journal of Substance Abuse, 9*, 269-275.

Waldron, E. M., & Ashby, G. F. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin & Reviews, 8*, 168 - 176.

Weiner, I., Schiller, D., & Gaisler-Salomon, I. (2003). Disruption of potentiation of latent inhibition by Risperidone: The latent inhibition model of atypical antipsychotic action. *Neuropschophamacology, 28*, 499-509.

Williams, R. J. (1992). Simple statistical gradient - Following algorithms for connectionist reinforcement learning. *Machine Learning, 8*, 229-256.

Wrase, J., Kahnt, T., Schlagenhauf, F., Beck, A., Cohen, M. X., Knutson, B., et al. (2007). Different neural systems adjust motor behaviour in response to reward and punishment. *NeuroImage, 36*, 1253-1262.

Wyvell, C. L., & Berridge, K. C. (2000). Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: Enhancement of reward "wanting" without enhanced "liking" or response reinforcement. *The Journal of Neuroscience, 20*, 8122 - 8130.

Zeithamova, D., & Maddox, T. (2006). Dual-task interference in perceptual category learning. *Memory and Cognition, 34*, 387 - 398.

Zinbarg, R. E., & Mohlman, J. (1998). Individual differences in the acquisition of affectively valenced associations. *Journal of Personality and Social Psychology, 74*, 1024-1040.

Zinbarg, R. E., & Revelle, W. (1989). Personality and conditioning: A test of four models. *Journal of Personality and Social Psychology, 57*, 310-314.

Zuckerman, M. (1979). *Sensation seeking: Beyond the optimal level of arousal.* Hillsdale, NJ: Erlbaum.

Zuckerman, M. (1993). P-Impulsive Sensation Seeking and its behavioural, psychophysiological and biochemical correlates. *Neuropsychobiology, 28*, 30-36.

Appendix 1

Correlations

		b5 - extraversion	b5 - neuroticism	bis	bassum	epq_p	epq_e	epq_n	spq-cognitive/perceptual	spq-disorganised	olife-unex	olife-cogdis	olife-intan	olife-impnon	olife-ext	sss_tot
b5 - extraversion	Pearson Correlation	1	-.213**	-.093	.354**	.097	.778**	-.173**	.114	-.066	.069	-.218**	-.471**	.300**	.776**	.272**
	Sig. (2-tailed)		.001	.160	.000	.145	.000	.009	.087	.326	.298	.001	.000	.000	.000	.000
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
b5 - neuroticism	Pearson Correlation	-.213**	1	.638**	-.031	-.062	-.338**	.699**	.113	.268**	.192**	.553**	.296**	.062	-.299**	-.158*
	Sig. (2-tailed)	.001		.000	.643	.350	.000	.000	.090	.000	.004	.000	.000	.356	.000	.017
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
bis	Pearson Correlation	-.093	.638**	1	.198**	-.160*	-.220**	.559**	.105	.197**	.116	.493**	.097	.062	-.159*	-.142*
	Sig. (2-tailed)	.160	.000		.003	.016	.001	.000	.114	.003	.081	.000	.147	.356	.017	.033
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
bassum	Pearson Correlation	.354**	-.031	.198**	1	.098	.395**	.039	.121	.143*	.156*	.051	-.297**	.292**	.416**	.176**
	Sig. (2-tailed)	.000	.643	.003		.143	.000	.556	.070	.031	.019	.442	.000	.000	.000	.008
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
epq_p	Pearson Correlation	.097	-.062	-.160*	.098	1	.126	-.020	.250**	.327**	.251**	.097	.058	.490**	.075	.331**
	Sig. (2-tailed)	.145	.350	.016	.143		.058	.760	.000	.000	.000	.147	.382	.000	.260	.000
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
epq_e	Pearson Correlation	.778**	-.338**	-.220**	.395**	.126	1	-.264**	.060	-.079	.058	-.301**	-.658**	.320**	.931**	.345**
	Sig. (2-tailed)	.000	.000	.001	.000	.058		.000	.366	.238	.386	.000	.000	.000	.000	.000
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
epq_n	Pearson Correlation	-.173**	.699**	.559**	.039	-.020	-.264**	1	.259**	.410**	.357**	.769**	.306**	.252**	-.216**	-.087
	Sig. (2-tailed)	.009	.000	.000	.556	.760	.000		.000	.000	.000	.000	.000	.000	.001	.191
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
spq-cognitive/perceptual	Pearson Correlation	.114	.113	.105	.121	.250**	.060	.259**	1	.588**	.793**	.442**	.177**	.434**	.076	.078
	Sig. (2-tailed)	.087	.090	.114	.070	.000	.366	.000		.000	.000	.000	.007	.000	.256	.240
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
spq-disorganised	Pearson Correlation	-.066	.268**	.197**	.143*	.327**	-.079	.410**	.588**	1	.611**	.567**	.253**	.535**	-.075	.177**
	Sig. (2-tailed)	.326	.000	.003	.031	.000	.238	.000	.000		.000	.000	.000	.000	.261	.008
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
olife-unex	Pearson Correlation	.069	.192**	.116	.156*	.251**	.058	.357**	.793**	.611**	1	.532**	.192**	.504**	.058	.048
	Sig. (2-tailed)	.298	.004	.081	.019	.000	.386	.000	.000	.000		.000	.004	.000	.387	.473
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
olife-cogdis	Pearson Correlation	-.218**	.553**	.493**	.051	.097	-.301**	.769**	.242**	.267**	.332**	1	.378**	.353**	-.278**	-.044
	Sig. (2-tailed)	.001	.000	.000	.442	.147	.000	.000	.000	.000	.000		.000	.000	.000	.508
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
olife-intan	Pearson Correlation	-.471**	.296**	.097	-.297**	.058	-.658**	.306**	.177**	.253**	.192**	.378**	1	-.029	-.676**	-.284**
	Sig. (2-tailed)	.000	.000	.147	.000	.382	.000	.000	.007	.000	.004	.000		.666	.000	.000
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
olife-impnon	Pearson Correlation	.300**	.062	.062	.292**	.490**	.320**	.252**	.234**	.335**	.304**	.353**	-.029	1	.302**	.343**
	Sig. (2-tailed)	.000	.356	.356	.000	.000	.000	.000	.000	.000	.000	.000	.666		.000	.000
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
olife-ext	Pearson Correlation	.776**	-.299**	-.159*	.416**	.075	.931**	-.216**	.076	-.075	.058	-.278**	-.676**	.302**	1	.307**
	Sig. (2-tailed)	.000	.000	.017	.000	.260	.000	.001	.256	.261	.387	.000	.000	.000		.000
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
sss_tot	Pearson Correlation	.272**	-.158*	-.142*	.176**	.331**	.345**	-.087	.078	.177**	.048	-.044	-.284**	.343**	.307**	1
	Sig. (2-tailed)	.000	.017	.033	.008	.000	.000	.191	.240	.008	.473	.508	.000	.000	.000	
	N	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227

Appendix 2

TableA2.1. The table reports the main parameter of the neural model and the parameter values implemented to simulate the human data from the various studies

Symbol	Description	Markman	2-cat blocked task	2-cat intertwined task	4-category
θr	Threshold for the raw RPE signal	0.05			
W_{mean}	Mean to generate SR weights from cortical inputs	0.3			
W_{sd}	Std. to generate SR weights from cortical inputs	0.02			
Scaler1	Learning rate on the RP unit	2			
Scaler2	Learning rate on the SR unit	2			
b_1 and b_2	Learning rates of RP unit	0.05* scaler 1			
b_3 and b4	Learning rates of SR unit	0.05* scaler 2			
θ	Category threshold (bound)	325			200-300-400
nx	Number of cortical units	650			600
Pnoise	Perceptual and criterial noise	50	50	50	40 & 60
rulebias	Strength of explicit module	0	0.6	0.65	0.4
m	Reinforcement (rf) values multiplier	.5	.8	.5	1
Rptonic	Reward prediction tonic value	.5	.5	.5	.4

Appendix 3

Pilot studies for the 4-category task (chapter 10)

Introduction

In the two previous studies, a d' equal to 1 had been selected to render the task difficult enough to allow motivational factors to come into play (Markman et al., 2005). However, in the present study the number of categories is doubled (from 2 to 4) which renders the task more complex and, subsequently, requires extra working memory effort. Hence, it was decided to increase the d prime value in order to facilitate and, therefore, ensure sufficient learning (i.e. accuracy above chance level). Nonetheless, the d prime value could not be too high otherwise the task would be too easy and no emotional factors would come into play and mediate performance. A pilot study was conducted in order to establish whether a d' equal to 1.5 was sufficient to ensure sufficient performance during a probabilistic rule-based CL task.

Task

The task used in the pilot study was the same 4-category RB task implemented in the learning phase task during the experimental testing session. Hence, the pilot consisted of 180 trials in total. The task also closely resembled the structure and format of the studies reported in the previous chapters. During the pilot participants had to learn to classify stimuli varying in location across an imaginary, 600-pixel long horizontal line into four categories. They received visual and auditory feedback for each response they made. The feedback was given in points according to the payoff matrices described in the main chapter. The aim of the task was to earn as many points as possible to win an entry into a £25 lottery. As in the experimental session, participants had to reach a performance criterion in order to earn the ticket. The performance criterion was set at 40800 points, which corresponds to 80% points earned by the optimal classifier.

Pilot with a d prime value equal to 1.5

Only three participants were tested using a d' equal to 1.5 since it was clear straight away that this value was far too small to ensure learning on the task.

Participants

An opportunity sample was tested and it consisted of 3 female psychology postgraduate whose mean age was equal to 28.3 (s.d. = 4.0).

Results

The proportion of correct responses for each participant was calculated, in order to assess the level of accuracy achieved with a d' equal to 1.5 (table A3.1).

Table A3.1. Accuracy levels with a d prime equal to 1.5

Participants	# correct	Accuracy (%)	Points won
1	51	28.3	28032
2	100	55.6	37401
3	48	26.7	28467

Looking at table A3.1 it can easily be observed that 2 of the 3 participants were performing at chance level (0.25). Overall, those three participants reached a very low proportion of correct (pc) responses over the 180 trials constituting the main, pre-payoff switch phase (pc: mean = 32.25, s.d. = 21.9). The proportion of correct responses reached by these three participants is just above chance (25% for a four-category task). Similarly, the total amount of points earned was way below the bonus criterion which (mean = 31233, s.d. = 5342). These results suggest that a d prime equal to 1.5 in a category learning tasks with a high working memory demand may be too small.

Pilot with a d prime value equal to 1.5

A second pilot study was subsequently carried out in order to assess whether a d prime equal to 2 could enhance participants' learning above chance levels. Performance with a d prime equal to 2 showed to be more promising both by looking at the data and from participants' verbal reports. Eight participants were tested under this condition.

Participants

The 8 participants were sampled from the student population at Goldsmiths. The sample included 6 postgraduate and 2 undergraduate psychology students. The mean age was 29.38 (s.d. 10.11).

Results

Table A3.2 shows that under this condition, six of the eight participants performed above chance levels. However, their accuracy level was below 100%, which suggests that the task was not. Hence, these results indicate that a d prime equal to 2 offers a good balance in difficulty. Indeed, results show that learning occurs above chance level but it does not reach optimal accuracy (100%). Therefore, the task is still hard enough to let emotional come into play and mediate performance (i.e. response bias) according to the observations made by Markman and colleagues (2005).

Table A3.2. Accuracy levels with a d' equal to 2

Participants	# correct	Accuracy (%)	Points won
1	109	60.6	41032
2	45	25	27661
3	96	53.3	37065
4	108	60	39800
5	43	23.9	25565
6	70	38.9	32192
7	93	51.7	37292
8	97	53.9	38.097

The variation in performance across the pilot sample during this pilot resembles the pattern observed in the studies reported in chapters 8 and 9. In these studies, some participants did not learn to perform on the task; others were quite good (i.e. accuracy levels of 60%) and the majority of participants lied somewhere between the two extremes.

Overall, these results indicate that it is necessary to implement a dprime equal to 2 for the 4-category task, in order to ensure learning above chance levels. However, the data also showed that among those participants who performed above chance, only one of them earned enough points to win an entry into the £25 lottery. Hence, according to these results, one out of eight participants could reach the bonus criterion and, subsequently, it was decided decrease the number of points necessary to win the lottery ticket. This was done in order to prevent participants from experiencing frustration as the task progressed. Indeed, participants may have felt like their performance was poor and would not allow them to earn the ticket and, subsequently, would have given up half-way through the task.

The average score of the other five participants who performed above chance was approximately 37000 points, which corresponds to an optimal level equal to 67% accuracy level (cf. 80%). It was decided to use this accuracy level and hence to reduce the performance criterion from the initial 40807 to 37100. In this way, it was ensured that participants would maintain a promotion focus rather than experiencing frustration as they monitored their point-earning over trials.

Appendix 4

Inter-block instructions presented in the 4-category task

At the end of the first block, the task stopped and participants received the same inter-block message received in the study in chapter 8. The message represented a lottery ticket that was struck through by a red cross and had a message saying:

> 'If that had been the end of this section of the experiment, you would have not earned an entry into the lottery'.

Participants also had the chance to take a short break if they needed a rest. At the end of the learning phase, the task paused and presented the same ticket as the one in the inter-block message. The ticket was either crossed out or not depending on whether participants had reached the performance criterion or not. The window dialog also reported a message saying:

> 'That is the end of this section of the experiment and you earned enough points to get an entry into the £25 lottery'

if participants had reached the performance criterion. Otherwise, the message would read:

> 'That is the end of this section of the experiment and you did not earn enough points to get an entry into the £25 lottery'.

The subsequent window offered participants a summary on their performance. In fact the message said:

> 'That is the end of this section of the experiment.
> [Sorry, but] you earned (number of) points, which means
> that you did [did not] earn an entry into the lottery for £25.'

At the end of the first task participants also received a message that informed them:

> 'You can take a short break now if you would like. Press any key when you are ready to continue on to the next section. You need to earn xxxxx points in order to win an entry.'

The breaks were not compulsory and participants could chose to carry on with the task but if they decided to take a short break its length was up to them. Once participants pressed any key to proceed, the second task started. It consisted of only 60 trials with no breaks in between. At the end of the second task, participants received the same message they had received at the end of the first one to notify them on whether they had won the lottery ticket or not.

At the end of the two tasks (i.e. learning phase), participants were debriefed by a message which said:

> 'This is the end of this experiment. Thanks for your participation!
> You earned x entries into the £25 lottery,'
> Please tell the experimenter you have finished, and
> that you earned x entries.'

Depending on their performance over the two tasks, x could be 0, 1 or 2.

Appendix 5

Preliminary simulations for the 4-category task

A parameter set that was able to simulate the behavioural data was identified by adopting values from the previous simulations and making adjustment by trial-and-error to fit those parameters to the requirements of the present task. In order to successfully simulate the behavioural data the **scaler** parameter was set equal to 2 and the noise value to 50 (as for previous studies); whereas **rulebias** was set equal to 0.4, the reward prediction signal (**rptonic**) was set equal to 0.5. Following the logic implemented in the previous studies' simulations, it was decided to code the points earned into a linear variable (rf). Since the maximum winning was equal to 400 points, a winning of 400 points was coded as 1 and, therefore, 200 points as 0.5. It follows that 333 and 267 points were re-coded as 0.83 and 0.67, respectively, while the 100 points earned for incorrect responding was equal to 0.25.

Owing to the categories' mean locations, the optimal accuracy criteria were equal to 200, 300 and 400. Initial simulations were run using these values as the initial boundaries, which represent the criteria applied by the perceptual rule module. However, the simulations indicated that the use of these values offered a bad fit to the human data (i.e. criteria were close to the optimal accuracy criteria and the noise levels were low; table A5.1).

Table A5. 1. Human and simulated criteria (s.d.) with initial boundaries set at 200, 300 & 400 across the two payoff conditions (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Criterion 1	**Criterion 2**	**Criterion 3**	**Noise**
Sim_PO1 (N=300)	211.54 (7.6)	293.80 (7.7)	400.73 (8.2)	63.56 (7.0)
Sim_PO2 (N=300)	197.40 (8.9)	303.47 (7.4)	387.73 (8.3)	62.69 (7.0)
Hum_PO1 (N = 18)	182.97 (26.4)	300.13 (14.0)	420.84 (23.8)	80.27 (31.7)
Hum_PO2 (N = 20)	191.77 (27.3)	303.21 (21.0)	413.31 (20.7)	79.62 (32.3)

There is no reason why a participant would set the boundaries at the maximum accuracy criteria, as these can be determined only once the properties of the distribution of the category stimuli are reasonably well-established. It is much more plausible that the initial boundaries are set using simple perceptual principles.

Since stimuli were displayed along a 600-pixel long horizontal line, it is possible that on starting the task, participants placed their criteria at symmetric distances across the line (i.e. 150, 300 & 450 pixel) in order to split the box into four equal (vs. unequal: large-small-small-large) spatial zones (i.e. perceptual criteria; figure A5.1). Thus, it is possible that participants might have initially placed their decision criteria close to the perceptual, rather than the optimal accuracy, criteria as the two types of bounds did not coincide.

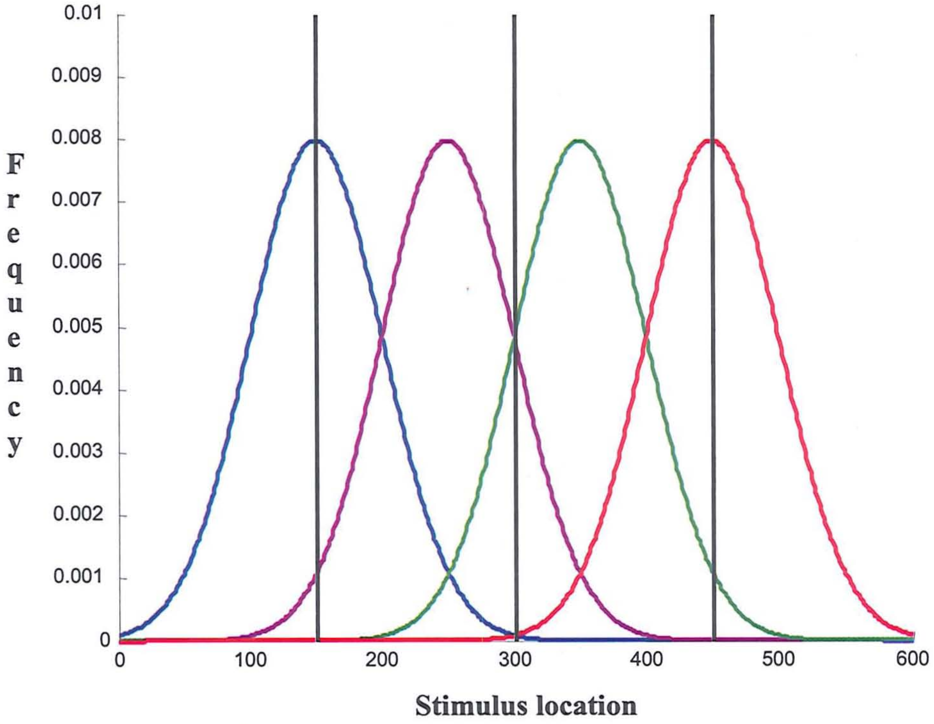


Figure A5. 1. Stimulus distribution and symmetrical perceptual decision criteria

Therefore, the simulations were run again with these perceptual criteria values being used as starting points to identify the actual perceptual boundaries participants implemented over trials to categorise the stimuli. Simulations with these values better captured the human data than the optimal accuracy bounds (i.e. 200, 300 and 400 pixels; table A5.2). However, the simulated criteria were still quite far from the criteria set by human participants (especially criterion 3 for PO matrix 1 and criterion 1 for PO matrix 2).

Table A5. 2. Human and simulated criteria (s.d.) using initial boundaries set at 150, 300 & 450 across the two payoff conditions (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Criterion 1	Criterion 2	Criterion 3	Noise
Sim_PO1 (N = 300)	180.25 (11.3)	296.54 (8.1)	442.83 (12.3)	88.22 (9.8)
Sim_PO2 (N = 300)	158.42 (11.4)	303.79 (7.7)	422.92 (12.6)	87.80 (10.2)
Hum_PO1 (N = 18)	182.97 (26.4)	300.13 (14.0)	420.84 (23.8)	80.27 (31.7)
Hum_PO2 (N = 20)	191.77 (27.3)	303.21 (21.0)	413.31 (20.7)	79.62 (32.3)

Moreover, the simulations did not reproduce the accuracy levels obtained in the empirical data. Indeed, the simulated 'participants' showed greater accuracy level for the categories associated to greater payoffs whereas human participants showed greater accuracy for the categories near the edges of the display box (i.e. categories A and D; table A5.3). Participants verbally reported having found it easier to learn to identify category membership for those stimuli.

Table A5.3. Simulated and human accuracy level under the two asymmetric matrices (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Cat A	Cat B	Cat C	Cat D
Sim_PO1	0.69	0.55	0.63	0.54
Sim_PO2	0.56	0.69	0.52	0.64
Hum_PO1	0.68	0.53	0.54	0.66
Hum_PO2	0.73	0.48	0.50	0.68

Owing to these preliminary simulations that indicate that the perceptual boundaries better simulate the human data than the actual optimal accuracy boundaries, all simulations reported below were run using the perceptual bounds (150, 300 and 450 pixels) as the initial boundaries.

Moreover, it was decided to adopt two values of perceptual noise (cf. original noise level equal to 50) in order to capture this accuracy differences in the simulations. Thus, the new perceptual noise values were equal to 40 and 60 for the stimuli of the edge and middle of the display, respectively. Simulations were run with the new noise levels under the two payoff matrices (table A5.4). The simulated criteria were still not close enough to the human criteria (especially criterion 1 for PO matrix 2 and criterion 3 for PO matrix 1).

Table A5.4. Human criteria and simulated criteria obtained using two perceptual noise values (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Criterion 1	Criterion 2	Criterion 3	Noise
Sim_PO1 (N = 300)	185.92 (11.0)	295.93 (8.6)	437.21 (11.8)	86.12 (9.3)
Sim_PO2 (N = 300)	164.36 (11.1)	304.50 (8.1)	417.54 (12.2)	85.56 (10.0)
Hum_PO1 (N = 18)	182.97 (26.4)	300.13 (14.0)	420.84 (23.8)	80.27 (31.7)
Hum_PO2 (N = 20)	191.77 (27.3)	303.21 (21.0)	413.31 (20.7)	79.62 (32.3)

These results suggest that there may be two simultaneously active forces that affect the criterion location during performance on the task. Specifically, one force may be pushing the criteria close to the optimal accuracy criteria (i.e. driven by learning) and the other force push the criteria close to the optimal winning criteria (i.e. driven by the asymmetric payoffs). In order to disentangle the effect of the two forces, further simulations were conducted with a symmetrical payoff matrix, where correct responses earned 300 point and incorrect responses 100 points. Following the previous studies, 300 and 100 points were re-coded as **rf** values equal to 0.75 and 0.25, respectively, for simulation purposes. These values were also chosen since they ensure a point-symmetry across the asymmetric and symmetric matrix. Table A5.5 reports the simulated criteria and noise levels under a symmetric payoff matrix. The simulated criteria and noise levels resemble those observed in the human data.

Table A5.5. Simulated criteria (s.d.) under a symmetric payoff matrix

	Criterion 1	Criterion 2	Criterion 3	Noise
Symmetric matrix	179.42 (10.5)	305.21 (8.1)	429.42 (12.1)	83.11 (9.6)

Thus, these results suggest that, in the present task, learning was characterised by the fact that over trials participants placed their criteria away from the perceptual bounds (i.e. 150, 300 and 450 pixels) and closer to the criterion optimal accuracy bounds i.e. 200 300 and 400 (figure A5.2). Thus, learning was mainly driven by the force which pushed the bounds closer to optimal accuracy bounds. These observations suggest that different payoffs might have been perceived as equivalent which, subsequently, suggests that in the human data the asymmetric payoffs were perceived as symmetric and guided learning.

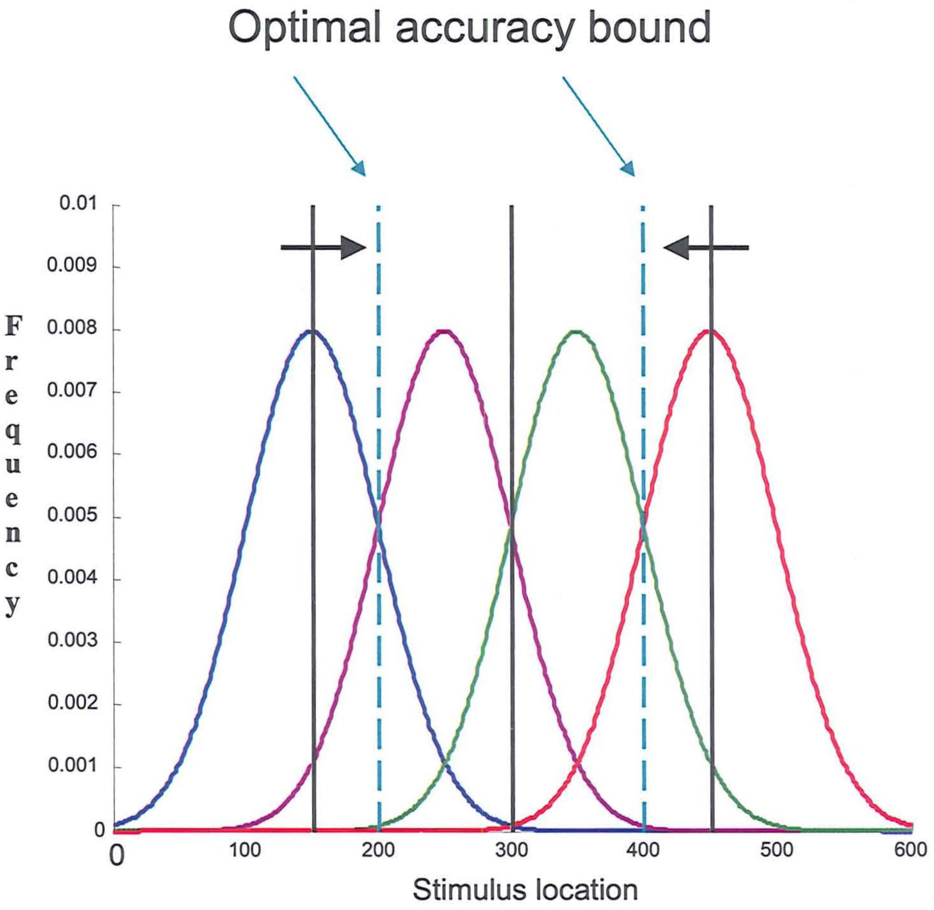


Figure A5. 2. Criterion shift from the perceptual bounds (black lines) towards the optimal accuracy bounds (cyan lines)

Participants verbally reported having noticed the 400 points more frequently than the other payoffs so it was assumed that the 400 points may have affected learning and criterion placement. Therefore, simulations were re-run using the symmetric matrix implemented earlier but an **rf** value of 1 was substituted for the category that offered 400 points for correct responses [i.e. partially asymmetric payoff matrices; PO matrix 1: 0.75 - 0.75 1 - 0.75; PO matrix 2: 0.75 - 0.75 - 0.75 – 1][25].

---

[25] These **rf** values code winnings received following correct responses for categories A to D, whereas an **rf** value equal to 0.25 was used to code winnings for incorrect responses.

Results show that these 'asymmetric' payoff matrices offered a better fit than the symmetric matrix for the criteria and noise level (table A5.6).

Table A5.6. Human and simulated criteria (s.d.) under the revised asymmetric matrices across the two payoff matrices (sim = simulated; hum = human; PO1 = POmatrix 1 and PO2 = POmatrix 2)

	Criterion 1	**Criterion 2**	**Criterion 3**	**Noise**
Sim_PO1 (N = 300)	178.98 (10.7)	300.02 (8.2)	432.71 (11.7)	82.79 (9.5)
Sim_PO2 (N=300)	180.33 (10.3)	303.35 (7.7)	418.02 (12.0)	79.38 (8.9)
Hum_PO1 (N = 18)	182.97 (26.4)	300.13 (14.0)	420.84 (23.8)	80.27 (31.7)
Hum_PO2 (N = 20)	191.77 (27.3)	303.21 (21.0)	413.31 (20.7)	79.62 (32.3)