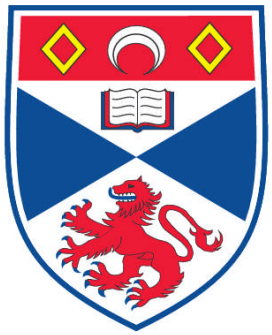
FORMS OF FLEXIBILITY
ASSOCIATIONS BETWEEN EXECUTIVE FUNCTIONS IN THE RAT

E. Alexander Chase

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



2013

Forms of Flexibility

Associations Between Executive Functions in the Rat

E. Alexander Chase

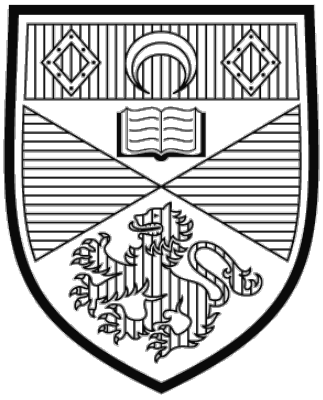
This thesis is submitted in partial fulfilment for the degree of

Doctor of Philosophy

at the

School of Psychology and Neuroscience

University of St Andrews



2013

1. Candidate's declarations:

I, E. Alexander Chase, hereby certify that this thesis, which is approximately 32,500 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in April 2008, and as a candidate for the degree of Doctorate of Philosophy in September 2008; the higher study for which this is a record was carried out in the University of St Andrews between 2008 and 2012.

Date: Signature of candidate:

2. Supervisor's declaration:

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

Date: Signature of supervisor:

3. Permission for electronic publication:

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

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis: access to printed copy and electronic publication of thesis through the University of St Andrews.

Date Signature of candidate:

Date Signature of supervisor:

Table of Contents

Acknowledgments

Although I owe a great deal to a great many of people, foremost I must acknowledge those who have supervised me to a successful ending to this four-year journey. To Prof Verity Brown I will be forever indebted for the gifts of her knowledge, patience, support and humour; and for taking a lost Senior Honours student and starting his career. I am also especially grateful to Dr David Tait for his adroit substitution for "The Disappearing Supervisor", and for ensuring that I would not be lost from the very beginning. To Drs Hugh Marston, Eric Bowman and James Ainge, I am especially grateful for every ounce of wisdom and encouragement that you have provided.

These last eight years in the School of Psychology would not have been nearly as rewarding without lecturers like Paul Gardner and Peter Foldiak; technicians like Mary Latimer, Steven Laing and Michael Kinnear; NACWOs like Wendy Taylor and Jerico Guzman-Mejor; and colleagues like David Wilson, Duncan MacLaren, Nadine Gut, Shuang Xia, Jenny Daggett, Ana Garcia-Aguirre and many, many others. I am also very grateful for the help of several undergraduate interns: Albert Aspinall, Emma Duckworth, Alexandra Hellyer, Alexandra Watson, Laura Graham, Ga Han Hung and Ellen Bowman.

Thanks to the writers – Richard and Annie – for editing this manuscript; and for not appearing afraid to aggressively alter all accidentally alliterative phrases, and my excessively pleonastic verbiage. Thanks to all the musicians – Catriona, Jamie, David, Cameron, Martin, Saif, Alex, Douglas, Letitia, Jassel, Dan, Ben and TAB – for the escape and reprieve from the stress of work (and for replacing it with a whole new breed of stress). Thanks to my family – sisters, parents and cousins – for their unwavering support. And thanks, of course, to Laura; to Chris, Patton, and Alex; to Ashley, Katharina, Fergus, Jessica, Stephane and Katie; to Claire, Freya, Peter, Isabel, Graeme, Stuart and Ryan; to Jen, Pete, Kathleen and Claire.

And my final thanks to Lois St. Pierre, from whom I learned everything I know about learning.

Abstract

Executive control is a vital cognitive function that facilitates the focussing and shifting of attention, planning and working towards a goal, ignoring distractions, and flexibly responding to novel situations. Disruptions to executive control are in many psychiatric and neurodegenerative disorders, as well as healthy aging, which can be profoundly detrimental. Despite having many effective and well-validated methodologies for detecting and quantifying these deficits, there are very few treatments – pharmacological or otherwise – for ameliorating executive dysfunction. This lack of progress can partly be blamed on difficulties associated with identifying drugs that enhance cognition in preclinical research. The work in this thesis aimed to expand our understanding of executive dysfunction – as well as the tasks that measure it – in rats. In results presented in chapter three, middle-aged rats demonstrated impaired reversal learning on the standard attentional set-shifting task, but this was treatable with a novel drug targeting the N-methyl-D-aspartate receptor. The age impairments seen in this experiment were similar to those previously found in young rats with orbital prefrontal cortex (OFC) lesions. The results of chapter four expanded on this similarity to show that, along with reversal deficits, young OFC-lesioned rats are impaired at forming attentional sets when tested on a modified task. In chapter five, another modified set-shifting task revealed that middle-aged rats also suffer from impaired set-formation, but their reversal learning impairments only manifest before attentional set has been formed – not after. Finally, in chapter six, the putative cognitive enhancer modafinil was found to exacerbate middle-aged rats' reversal learning deficit, but it also enhanced their subsequent ability to form attentional set. These experiments reveal that modifying the rat attentional set-shifting task can sometimes make it a more effective tool for testing cognitive enhancers in preclinical settings.

Chapter 1

General Introduction

Executive control is widely involved in our day to day lives: from recalling information stored in long-term memory, to directing behaviour towards a goal, to maintaining focus in the face of distractions. "Higher-order" cognitive functions like these have been associated with the prefrontal cortex since at least the early twentieth century, but quantifying disruptions to these functions is not straightforward. In this chapter, I will present a brief history of the behavioural tasks that led to the development of modern attentional set-shifting tasks, which have proven to be highly sensitive and reliable tests of cognitive function and dysfunction.

In the early autumn of 2009, I travelled to Boston, Massachusetts to visit some school friends: it was the end of my first year as a PhD student, and the start of my "summer" vacation. After I arrived at Logan International Airport, I made my way to the subway station to catch a "T" into town where I was staying. I had brought with me the "CharlieCard" I'd acquired on an earlier trip: a reusable, pay-as-you-go, smart card for use on the city's public transportation. I had never used the card before, but – having grown up in London – I was quite comfortable with the general principle of using smart cards on public transport. I suspected that I wouldn't have enough credit to cover the fare of my journey, and I would therefore need to top-up my card before getting on a train. It seemed logical that recharging a CharlieCard would be a similar process to that of my familiar "Oyster Card" back home, so – imagining I was in my local Tube station – I hazarded my way through the process. First I located a bank of modern-looking ticket machines, eschewing the older ones without touchscreen interfaces. Then, I located the card-reader just below the screen, and held my card to its surface until the screen displayed an acknowledgement of my actions. After a brief "Please Wait...", the machine displayed my balance (the princely sum of $0.00), then guided me through the process of topping-up my card. I took my debit card out of my wallet, swiped it through a different card-reader on the machine, and finally touched my CharlieCard to the first reader once more to finalise my transaction. I then left the machine, touched my card to the reader on the turnstile, and followed the signs for the "INBOUND" platform.

The educated reader will probably assume – quite correctly – that this story is the most mundane and boring that I could tell of my 2009 vacation. However, looking at this ordinary sequence of events in slightly greater detail reveals it to have been a feat of monumental complexity. To successfully navigate "the T" I had to:

1. remember that I possessed a CharlieCard, and that I should pack it,
2. recognise the similarity between Boston's CharlieCard and London's Oyster Card,
3. estimate the likely balance of my unused CharlieCard, and compare that assumption to the signposted fares at the station,
4. react to this calculation by deciding to top-up my card,
5. remember how to top-up an Oyster Card,
6. extrapolate the relevant memories and apply them to my new surroundings, and
7. respond flexibly to any unpredictable occurrences;
8. all the while ensuring I don't get distracted by the myriad of activities going on around me in the busy train station, and allowing this sequence of thoughts and actions to be irreparably disrupted.

Of course, even closer examination reveals that each of these steps can themselves be dissected to further complexity, thus producing a prime example of what computer scientists refer to as "the curse of dimensionality" (Bellman, 1966). In the first step, for example, I had to "mentally time-travel" (Emery & Clayton, 2004) from when I was packing to an imagined future visit to the subway station at Logan Airport, which itself would be based on the assumption that this future visit would be like my past visits to subway stations, which in turn would be based on knowing what a typical past subway station visit involved...

There are, however, common threads that emerge from the increasingly complex solutions that problems like this require. Identifying goals and sub-goals, and then sequencing them to form a plan (as in the list of steps above), is critical to success. Attaining many of these goals then relies on retrieving information from long-term memory, and integrating the relevant aspects of those memories with current information from the environment. Finally, ignoring distracting information that might interfere with

attaining the end goal will also be crucial. These examples often collectively fall under the term "executive control" (Barch *et al.*, 2009), an important aspect of cognitive function. From deciding what to wear in the morning, to taking an alternate route to the office because of traffic, to tuning out your coworkers' conversations so you can focus on your work, to stopping at the supermarket on the way home and navigating the aisles to follow your grocery list: executive control is intimately linked to almost everything we do in our daily lives.

Generally speaking, the neurological substrate of executive control lies within the prefrontal cortex. Given the widespread importance of this aspect of cognition, it may come as no surprise that physical damage to this area – whether through cerebrovascular occlusion, tumour or acquired injury (including as a consequence of neurosurgery) – can lead to profoundly detrimental changes in behaviour. Difficulty recalling long-term memories is frequently reported in this patient group (Damasio *et al.*, 1985), as is a general slowing of movement and thought (Hécaen, 1964). Personality changes (Damasio *et al.*, 1994), a reduced ability to make decisions (Eslinger & Damasio, 1985), and problems taking care of oneself and others (Shallice & Burgess, 1991) are also commonly reported. In extreme examples, problems suppressing typical or 'prepotent' responses (Milner, 1964) can lead to "utilisation behaviour" (Lhermitte, 1983), where patients uncontrollably "use" objects they are presented with: for example, a patient was once described as being unable to pick up a comb without beginning to comb her hair (Goldstein & Scheerer, 1941).

While behavioural descriptions such as these paint a clear picture of severe disruptions to the way the brain guides our behaviour, they lack the specificity necessary to tell us exactly what has gone wrong. Memory loss, for example, may indicate impaired executive control, but it is also reported in patients suffering damage to the temporal lobe or the hippocampus (Damasio *et al.*, 1985; Kapur *et al.*, 1992): are these "amensias" a unified

phenomenon, or is it more likely that they represent a common manifestation of distinct cognitive deficits; that they look the same without actually being the same? Behavioural deficits – like memory loss and poor decision making – are also not unique to patients with physical brain injury: are those suffering from mental illnesses such as schizophrenia demonstrating precisely the same executive control impairments as patients with prefrontal injury? To answer these questions, psychologists in the early twentieth century recognised the need to shift away from descriptive case study reports based on clinical interviews, and in so doing they started on a path toward a range of well-validated, quantitative measures for behavioural disruptions; a path that would forever change the field of psychology.

Early Behavioural Tasks: Abstraction, Concepts and Sorting

The experiments reported in this thesis describe several behavioural tasks designed to measure executive control in rats. While these tasks all represent subtle modifications to a protocol first reported 12 years ago (Birrell & Brown, 2000), their earliest ancestry can be traced back to the end of the First World War. In the years that followed the Armistice, hospitals and hospices were confronted with an unprecedented number of patients with traumatic brain injury. Some of these patients demonstrated curiously specific behavioural abnormalities: they could easily read a clock and report the time, but struggled to set the hands of that clock to a specified hour; they could drink from a glass of water, but were unable to describe the purpose of an identical empty glass; and though they were able to recite the months of the year in order, they were unable to name the months that preceded or followed any given example (Goldstein & Scheerer, 1941). In this respect, these brain-damaged veterans were deemed similar to those suffering from mental illnesses, such as the "de-

mentia praecox" associated with schizophrenia (*ibid.*). The conventional wisdom regarding these "dementias" and "aphasias" was that they were characterised by impairments in forming and assuming "abstract" attitudes or thoughts (*ibid.*; Vigotsky & Kasanin, 1934). The "concrete" action of drinking from a glass persevered in these patients, but the "abstract" representation of glassware's inherent function had been lost.

The first test devised to probe these abstraction deficits is frequently attributed to Gelb and Goldstein (1920; reviewed in Goldstein & Scheerer, 1941). In their experiment, subjects were presented with small pieces of woollen cloth dyed several different colours that varied in intensity – for example, darkest or deepest blue to lightest and least saturated blue; darkest red to lightest red; darkest green to lightest green, *etc.* The authors described several tasks using these materials, but in general they all focused on the subject's ability to sort the fabric swatches in different ways. For example, a subject could be presented with a pile of mixed swatches and asked to sort them by colour, ignoring brightness and saturation, so that all the blue swatches were in one pile, all the reds in another, *etc.* The brain damaged patients – despite showing normal colour vision – demonstrated considerable difficulty on this seemingly simple task. One patient picked out dark and light green swatches, but endlessly vacillated between which of them was "green" and which was some other, unidentifiable colour (Gelb and Goldstein, 1920; cited in Weigl, 1941); others would only select a small number of the most similar swatches (*e.g.* the three darkest reds), and declare there to be no other members of the set (Goldstein & Scheerer, 1941).

Gelb and Goldstein's colour sorting task was later modified by Weigl (1927; cited as Weigl, 1941) such that, instead of fabric swatches that varied in intensity, subjects were presented with coloured pieces of cardboard cut into various shapes. Weigl asked brain-injured patients and healthy controls to sort the shapes however they wished, and then, after

each sort was completed, he asked them to sort the cards in a different way. For example, the subject might begin by putting all the red, green, blue and yellow shapes in separate piles by colour; then he would be asked to pick a new strategy, and he might put all the triangles in one pile, all the circles in another, *etc.*; then he might arrange them into sets where no features (colour or shape) are repeated; and so on until the subject can think of no more ways to sort the shapes. Behaviour in the brain-damaged patient was characterised by poor performance after being asked to switch strategies – they would continue with one method of sorting instead of generating a new one – while the controls found the task to be relatively straightforward (Weigl, 1941).

These and other early sorting tasks represented a significant advance for the field of psychology: by defining a behaviour that hypothetically reflected an aspect of cognition, these tasks enabled the measurement of covert processes through overt events. Unfortunately, these early tasks were marred by significant problems. At their heart, the experiments of the 1920s and 30s intended to produce qualitative analyses that would permit further exploration into the nature of "aphasia", "amnesia" and "dementia": quantifying the impairments' severity was less a priority than describing what the impairments actually were. This is perhaps best exemplified in the fourth and fifth experiments described in Weigl's report (*ibid.*), where the shapes' relationships to one another were manipulated to be less obvious (*e.g.* varied quadrilaterals instead of identical squares), thereby permitting a greater number of – supposedly more difficult – sorting strategies. In these experiments, the purpose was not too see how well the brain-damaged patients could generate these sorts, but whether they were able to perform them at all; thereby revealing precisely which aspects of "abstract thought" differentiated the patient from the control.

It was this distinction between qualitative exploration and quantitative measurement that motivated the development of what is perhaps the

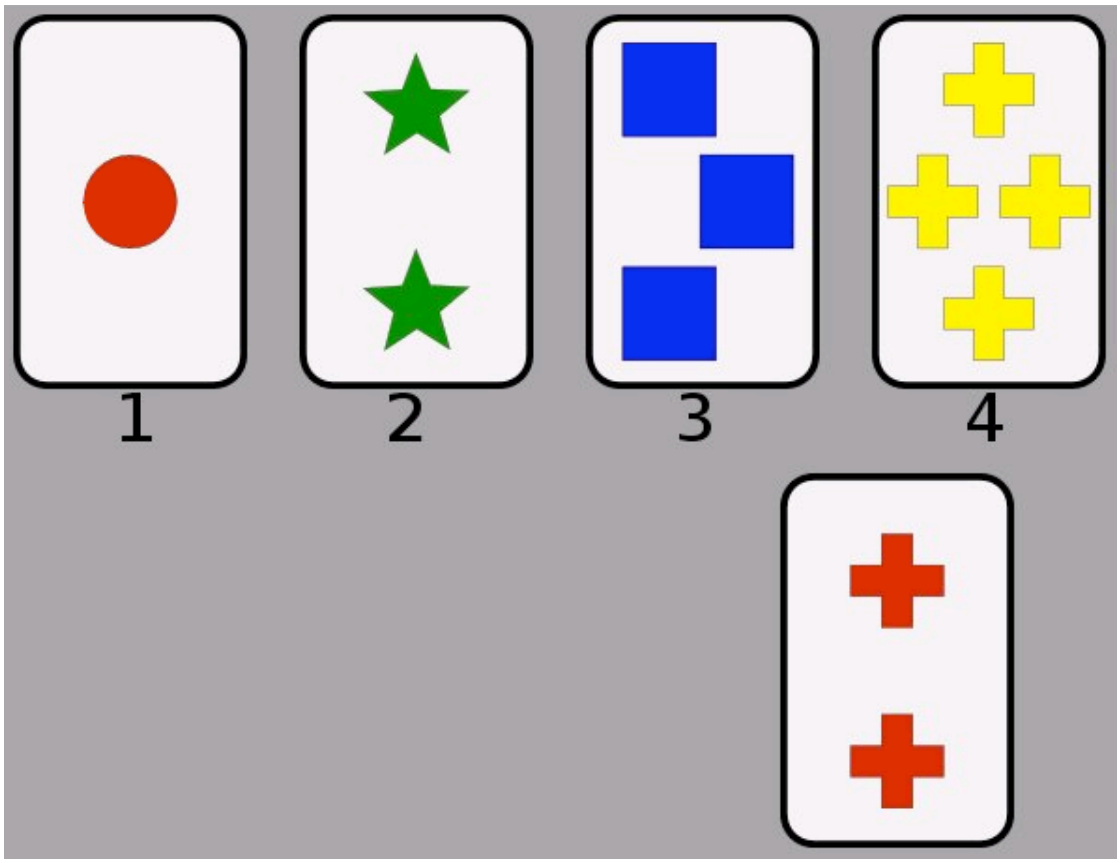


Figure 1.1: A screenshot from a computerised version of the Wisconsin Card-Sorting Task. Four "stimulus" cards (top) form the possible sorted piles onto which the "response" card (bottom) can be placed. In this example, the response card could be placed on piles 1, 2 or 4 depending on whether the correct strategy was by colour, number or form respectively. Image reference included as PEBL Developers (2008).

most famous and successful task in the sorting paradigm, the Wisconsin Card-Sorting Task (Berg, 1948). In this task – still widely used – subjects are given a deck of 60 cards marked with different numbers of coloured shapes, which can thus be sorted in three ways: by the number of shapes on the card, by the colour of those shapes or the form of the shapes themselves. Subjects are given four "stimulus" cards to act as the seeds for the soon to be sorted piles, and are then asked to place cards from the shuffled deck into the pile of their choice (Figure 1.1). By giving trial-by-trial feedback, the experimenter teaches the subject the correct sorting strategy, and – after the subject makes five consecutive correct choices – the experimenter "shifts" to a new sorting strategy to measure subjects' flexibility

(Berg, 1948). In the example in Figure 1.1, the subject might be correct in placing the response card on pile number two – a number-based sort – but in a later stage the card would need to be placed on pile number one; representing a shift to a colour-based sort.

Unlike the earlier tasks, which emphasised descriptions of patients' failures to learn particular sorting strategies, the Wisconsin Card-Sorting Task generated a standard index of performance through the number of trials or errors it took the subject to discover the new method of sorting. In two versions of the task, subjects were either required to shift first between the three novel strategies, and then among these three in a pseudorandom sequence (Berg, 1948), or in a predetermined sequence of novel shifts and shifts back to a previously used strategy (Grant & Berg, 1948). In both experiments, control subjects performed the task reasonably well, but consistently demonstrated behavioural costs – that is, increases in errors – after the experimenter shifted the correct strategy. Future experiments would use the Wisconsin Card-Sorting Task to show that these "shift-costs" were significantly greater in patients with surgical excisions of dorsolateral prefrontal cortex (Milner, 1964), schizophrenia (Goldberg *et al.*, 1987) and age-related dementia (*e.g.* Lees & Smith, 1983).

Early Behavioural Tasks: Discriminations, Reversals and Shifts

Simultaneously to the development of these sorting tasks, a separate school of researchers was making significant advances in understanding the nature of discrimination learning: how – and how well – animals learn to distinguish between stimuli where one signals reward and another does not. A significant strength of this paradigm was its translational nature: discrimination learning presented psychologists with the means for testing learning in humans and other animals in very similar ways. Mazes

(*e.g.* Lashley, 1929) and so-called "escape boxes" (*e.g.* Thorndike, 1898; Jacobsen, 1936) had revealed much about the nature of animal intelligence, but with discrimination learning it was now possible to run formally equivalent tasks in humans as in other animals. This would, in turn, lead to a highly fruitful avenue of research: the measurement of the effects of discrete lesions of cortical and subcortical areas in animals to infer the functional organisation of the human brain.

Though it would not be accurate to credit him with the birth of the movement, many of the successful tasks in the discrimination learning paradigm owe a debt to one of Lashley's (1930) more eccentric apparatuses. In his original design, rats were placed on a small, circular table or pedestal that was only slightly larger than the rat's body length, and raised a few feet from the ground. Approximately 25 cm front of this table was a large piece of plywood – like a wall – with two adjacent square holes permitting access to a larger platform and a food well; with a net hanging below. Each square hole was covered by a piece of heavy cardboard printed with visual patterns: black versus white, horizontal stripes versus vertical, *etc.* One of these cards could be knocked down easily, permitting access to the more comfortable platform and the reward; the other was held into place and would not move. Rats therefore performed the experiment by jumping off the small platform and hurtling into their chosen card; presumably hoping – or experiencing whatever the rodent analogue of "hope" is – that they would not immediately bounce off the incorrect card into the net below. Perhaps unsurprisingly, Lashley's rats learned relatively quickly to discriminate between the moveable and immoveable cards based on their visual features (Lashley, 1930).

Six years later, Jacobsen published his results on the behaviour of monkeys with prefrontal cortex lesions performing a task modified from Lashley's "jumping technique" (Jacobsen, 1936). In his apparatus, Jacobsen presented monkeys with a panel where two food wells could be ac-

cessed by displacing the stimulus cards, and monkeys responded by pushing them over and reaching a hand into the well. This design, in turn, inspired the development of the Wisconsin General Testing Apparatus ("WGTA"; *e.g.* Harlow, 1949), where humans and monkeys alike could be taught discriminations between objects that sat on top of two food wells; where displacing the object to uncover a well counted as a behavioural response (*e.g.* Settlage *et al.*, 1948; Shepp & Turrisi, 1969).

Although a great number of tasks were designed to use the jumping stand or WGTA, two genres of task were particularly influential to the development of the tasks used in this thesis. The first included some of the earliest tasks published using the WGTA: discrimination reversal learning. In these experiments, monkeys were trained on an initial discrimination where displacing one object, a cone for example, would reveal a food reward; while displacing a second object – say, a cube – would result in the monkey's access to the objects, and thus the unobtained reward, being blocked. After the monkeys had learned this discrimination, these contingencies – cone being correct, and cube incorrect – would be "reversed", so that the reward was now underneath the cube, and not the cone. Monkeys with lesions to the prefrontal cortex were shown to be generally worse than controls at performing these tasks (Harlow & Dagnon, 1943; Settlage *et al.*, 1948; 1956).

Later experiments would expand on this result by demonstrating that lesions of the orbital aspect of prefrontal cortex (the "OFC") specifically disrupted the reversal stages, but not the initial discriminations. In 1964, Mishkin described a functional dissociation between lateral and orbital prefrontal lesions in monkeys, with OFC lesions being particularly disruptive to the formation of "learning set" – a kind of practice effect over consecutive discrimination tasks (see Harlow, 1949) – and object reversal learning (Mishkin, 1964). Later experiments would reveal that OFC lesions disrupted reversal learning by increasing habitual or "perseverative" re-

sponding to the previously correct object (McEnaney & Butter, 1969; Jones & Mishkin, 1972).

Following the early experiments on reversal learning in monkeys, Reid (1953) sought to probe the precise mechanism by which a discrimination's reversal was acquired. His task design was relatively straightforward: 45 rats were taught to discriminate between black and white cards, and then they were tested on the reversal of that discrimination. The apparatus was a modified Y-maze, where the entrances to the two goal arms were blocked with hinged doors, upon which the black and white cards were fastened – essentially a Lashley jumping stand without the element of jumping. The correct door could be pushed open to reveal a food reward, while the incorrect door would not open. All the rats were trained to a criterion of nine correct trials in a rolling block of ten, and for 15 of the rats, the reversal came straight after the rats attained criterion. The remaining 30 rats were "overtrained" on the original discrimination: 15 were given 50 extra trials of training, and 15 were given 150 extra trials, before starting the reversal.

The most basic theoretical explanation of reversal learning is that a previously established rule or habit must be extinguished and replaced with its opposite. Thus, in a discrimination where black is rewarded and white is not, the reversal will reward responses to white and not black. If this explanation were accurate, the speed of reversal learning would be inversely proportional to the initial rule's resilience to extinction, which would in turn be determined by the number of times that rule had been reinforced (see Lovejoy, 1966). In other words, the more a rule has been trained, the more difficult that rule's inverse will be to learn. This is what Reid sought to investigate: would overtraining rats on the initial black-white discrimination impair their ability to learn the discrimination's reversal? Perhaps counterintuitively, the data suggested the exact opposite effect: the rats given 150 overtraining trials on the original discrimination learned the

reversal stage in significantly fewer trials than the other groups (Reid, 1953).

This seemingly paradoxical finding, termed the "overtraining reversal effect", generated no small amount of controversy. In their still relevant review, Sutherland and Mackintosh (1971) describe approximately 20 follow-up experiments that replicated Reid's finding, but nearly twice as many that did not. For example, D'Amato and Jagoda (1962) trained rats on a left-right discrimination in a Y-maze, giving some rats up to 160 overtraining trials, but this had no effect on reversal learning performance. In a follow-up experiment, these authors tried 200, 400 and 800 trials of overtraining and were still unable to find an effect (*ibid.*). Erlebacher (1963) trained rats to discriminate between two arms of a T-maze, one painted entirely black and one entirely white. Rats learned this discrimination relatively quickly – at least half as many trials to criterion than in Reid's experiment – and they did not demonstrate an overtraining reversal effect after up to 280 trials. While some interpreted these negative findings as refuting the existence of the overtraining reversal effect (Gardner, 1966), others argued that the effect simply could not be generated in tasks where the discriminations were learned relatively quickly, or for relatively small rewards, or relied on spatial location (Lovejoy, 1966; Mackintosh, 1969; Sutherland & Mackintosh, 1971).

Lovejoy (1966) proposed that the acquisition of a reversal stage, rather than representing the exchange of one habit for another, is a process mediated by two probabilities: the probability of responding to the correct stimulus, as well as the probability of attending to the relevant stimuli. Overtraining therefore facilitates reversal learning by increasing the probability that the subject attends not only to what cue positively predicts reward, but also to what cue negatively predicts reward, while perhaps reducing attention to those cues that are not predictive. In rats, this usually manifests as a difference in the number of responses that are based on side-bias.

Reid observed that overtrained rats spent more trials at the beginning of the reversal performing below chance accuracy than controls. After this phase, though, the overtrained rats spent fewer trials than controls responding at chance accuracy, possibly because they committed fewer errors based on the spatial location of the discriminanda, which enabled them to learn the reversal in fewer trials overall (Reid, 1953).

Reductions in errors based on cue position has been repeated multiple times (Mandler & Hooper, 1967; Siegel, 1967; Mackintosh, 1969; Sutherland & Mackintosh, 1971). Furthermore, if rats are presented with compound stimuli – for example, black and white rectangles of different orientations – the magnitude of the overtraining reversal effect was increased (Mackintosh, 1963), perhaps due to the fact that there are more non-predictive cues to which rats can attend. These results suggest that reversal performance is partly, but crucially, reliant on the establishment of what is and is not relevant to the discrimination – and not simply based on "unlearning" one rule and replacing it with another.

Concurrent to research into this aspect of reversal learning, Buss (1953) set out to better measure what he referred to as the general "flexibility" of behaviour by integrating the results of impaired reversal learning with the impairments seen in the "concept-shifting" tasks described by Weigl (1941) and Berg (1948). By comparing a subject's ability to learn discriminations that differed in some way to what had been previously learned, it was possible to measure how the prior learning "transferred" to the new problem. Buss argued that it was not clear what task manipulations would elicit behaviour that was more or less flexible than others, and in doing so he launched[1] a series of influential experiments of a similar format. In one condition, subjects would be presented with a reversal of the

[1]Unfortunately for Buss, his own experiment was so poorly designed – for example, he tested subjects' abilities to shift before they had demonstrated successful acquisition of the discrimination they were supposed to shift away from – that his hypothesis was far more influential than his results.

previous stage, and in another they would be given a "non-reversal shift discrimination", where the reward was now predicted by a stimulus that was not relevant to the initial discrimination.

Typical "reversal/non-reversal shift" experiments therefore presented subjects with discriminations between stimuli that varied in multiple ways. For example, Kendler and Kendler (1959) asked kindergarten children to find a marble that was hidden under cups that were either large or short, and black or white; Tighe (1965) trained monkeys to discriminate between cards marked with solid or interrupted lines that where printed in red or green; and Tighe *et al.* (1965) trained rats to discriminate between two- or three-dimensional pictures of squares marked with horizontal or vertical lines. In most experiments, subjects were first trained on an initial discrimination (*e.g.* red and not green), and then either presented with the reversal of that discrimination (*i.e.* green and not red) in the reversal condition, or with a discrimination based on an aspect of the stimuli that was not previously relevant (*e.g.* squares and not circles) in the "non-reversal shift" condition. Many referred to these two types of stage respectively as being "intradimensional", within the same stimulus dimension as the previous discrimination, and "extradimensional", in a different dimension to before.

This experimental paradigm therefore proposed that the brain established certain "mediational strategies" to facilitate learning (Kendler & Kendler, 1967; Slamecka, 1967; Sutherland & Mackintosh, 1971), for example the representation of similar stimuli – red and green, for example – as existing in a common dimension, like "colour". Thus, intradimensional stages, being consistent with prior stages, would be facilitated by the prior learning ("positive transfer"), while extradimensional stages, being inconsistent, would be learned less well ("negative transfer").

Unfortunately, the results from reversal/non-reversal shift experiments were largely equivocal. Several groups presented evidence suggesting that reversal shifts are learned faster than non-reversal shifts (*e.g.* Kendler

& D'Amato, 1955; Harrow & Friedman, 1958), while others reported the opposite relationship (Kelleher, 1956; Tighe *et al.*, 1965), and still others reported that the two types of shift do not differ in difficulty (*e.g.* Kendler & Kendler, 1959). This lack of consensus can probably be explained by the common, yet significant, methodological issues seen in these early shifting tasks.

Primary among these was the conflation of the reversal and the intradimensional stages that was inherent to the reversal/non-reversal shift paradigm. While it is true that the reversal stage uses the same dimension of stimuli to the discrimination that preceded it, the hypothesised facilitation of this intradimensional stage over the extradimensional stage will be obscured by the cost of reversing the previous discrimination (Slamecka, 1968). Moreover, because the reversal stage necessarily uses the same stimuli as the initial discrimination, so too must the extradimensional or non-reversal shift in these tasks. This often leads to scenarios where purported impairments on the extradimensional stage were larger than they might otherwise have been, because a previously rewarded stimulus is still rewarded 50% of the time: if a subject first learned to pick the green, but not the red shapes, and then had to pick the square and not the circle, the "green not red" strategy would still be partially reinforced when the subject chose the green square (Kendler & D'Amato, 1955; Buss, 1956; Slamecka, 1968; Sutherland & Mackintosh, 1971). Therefore, impaired performance might be attributed to difficulty shifting attention between the cues, but it might also be attributed to difficulty recognising that the previous strategy is no longer adaptive. This prevents any firm conclusions from being drawn about behavioural flexibility from most of these experiments, as the nature of the reported impairments is often unclear.

Issues of Terminology

The behavioural sciences – like all fields of academia that enjoy wide and diverse membership – suffer from variable and redundant terminology, and this is as true in 2012 as it was in 1932. Where Vigotsky and Kasanin (1934) spoke of abstract concepts and attitudes, Grant and Berg (1948) spoke of mental sets; where Harlow (1949) spoke of "learning how to learn", Sutherland and Mackintosh (1971) spoke of "switching in analysers". Many of the terms used in the papers reviewed above almost certainly refer to the same behavioural and cognitive phenomena, and so before our discussions can get any more complex we must establish a working vernacular for the rest of this thesis.

The term "set" has proven very useful in experimental psychology – indeed, at times too useful, as Gibson (1941) called it "at once ambiguous and ubiquitous" – and while its ubiquity has somewhat decreased in recent years, its ambiguity largely remains. For our purposes, the word "set" replaces historical and anthropomorphic terms like "attitude" and "concept", and will be used in similar fashion as in the mathematical discipline, "set theory". Sets are hypothetical stores of finite capacity and specific contents. Although a wide variety of cognitive sets have been proposed in the literature, this thesis will almost exclusively focus on sets which contain perceptual features of stimuli, so-called "attentional sets"; or behavioural patterns, as in "strategy sets"; or representations of how a behavioural task works, as in "learning sets". Sets are "formed" through experience performing a task. For example, over the course of repeated novel object discriminations, monkeys perform better in later discriminations (*i.e.* "positive transfer") than early ones due to the formation of learning set (*e.g.* Harlow, 1949) to contain the rule, "reward can be found by selecting one object and not the other".

The term "correct" will be used to refer to any cue or stimulus that predicts reward significantly above chance (in the following experiments this prediction rate is always 100%), and the term "incorrect" refers to any cue that predicts reward significantly below chance (here that rate is always 0%). Cues that are correct or incorrect – that is, are predictive of reward in either direction – are referred to as "relevant", and cues that only predict reward at chance probability are referred to as "irrelevant". When cues are of one sensory modality – for example colours, shapes, odours, *etc.* – we refer to them as belonging to the same "dimension". Dimensions, like cues, can be relevant or irrelevant depending on the reward prediction of the cues they contain. These terms regarding relevancy and dimensionality therefore only refer to stimulus attributes. To discuss behavioural strategies, the word "adaptive" refers to strategies that enable the subject to find reward (or avoid punishment) significantly above chance, otherwise a strategy is "maladaptive".

Finally, there is the term "shift". This term has meant, and continues to mean, different things in different contexts. Grant and Berg (1948) used the term to describe experimenter actions and task contingencies: in the Wisconsin Card-Sorting Task, the shift refers to the experimenter changing the sorting strategy, or it can refer to the change in sorting rule itself. Buss (1953) used the term to describe subjects' behaviour and responses: in the tasks he developed and inspired, shifts referred to the subjects ceasing to respond to one stimulus and beginning to respond to another. Dias and colleagues (1997) extended Buss' definition to include reference to subjects changing the focus of their attention: in the CANTAB set-shifting task (see the next section), shifts could also refer to the subject transferring behavioural responses from one set of stimuli to another, within or between dimensions.

To maintain a clear focus throughout this thesis, the word "shift" will take on a strict definition, modified slightly from that of the CANTAB

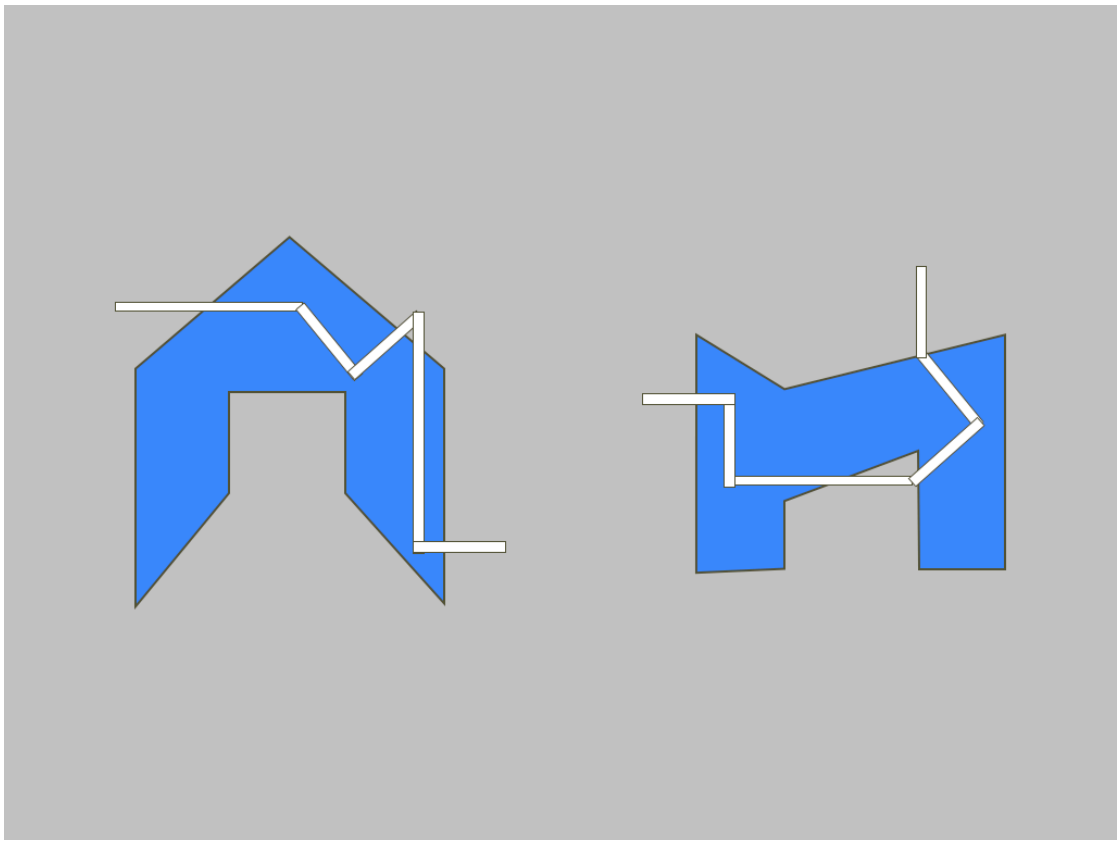


Figure 1.2: An illustrative example of the kind of stimuli used in the Cambridge Neuropsychological Automated Testing Battery (CANTAB). Discriminations are presented on a touchscreen between stimuli that vary among two dimensions: shape and line. The lines can be superimposed on either shape, and all stimuli can appear on the left or right.

set-shifting task. Here, shifting will refer only to the cognitive act of updating the contents of an established set. In attentional set-shifting, therefore, when the perceptual features contained in set, like colour cues, are no longer relevant to the discrimination; that set is shifted onto new features, like shape cues. Therefore, using this narrow definition, only "extradimensional" stages, and not "intradimensional" stages, involve shifting – which is different to many previous experiments, including contemporary examples. In strategy set-shifting, the set may be shifted when the previously adaptive strategy, like "pick the left object", becomes maladaptive (in Settlage *et al.*, 1948, for example, "ignore location and pick the cone").

Like attentional set, only "extramodal[2]" stages can involve a shift in strategy set, while "intramodal" stages do not.

As the shift of set is a cognitive action, it is "covert" and cannot directly be measured (*c.f.* Brown & Tait, 2010). Rather, the shift is inferred when a subject demonstrates attenuated performance (*i.e.* "negative transfer") on a stage that should require a shift of set compared to a stage that should not, as will be discussed further in later sections.

The Contemporary Primate Attentional Set-Shifting Task

The legacy of the early behavioural tasks was complicated, but significant: the establishment of behavioural flexibility as an index to qualify and quantify impairments in executive control was nothing short of revolutionary, but methodological and technological constraints limited the usefulness of the early tasks. The early sorting tasks had revealed impairments associated with adopting new strategies in patients with damage to the dorsolateral prefrontal cortex (Milner, 1964) and schizophrenia (Goldberg *et al.*, 1987). The reversal learning tasks had proven to be particularly sensitive to orbital prefrontal damage (Mishkin, 1964; Jones & Mishkin, 1972), and the overtraining reversal effect demonstrated that reversal learning was a complex cognitive and attentional process (Lovejoy, 1966; Mackintosh, 1969). Unfortunately, though, the best attempts to integrate these two types of flexibility into an omnibus behavioural test – the reversal/non-reversal shift tasks – had produced consistency neither in their methods nor results (Slamecka, 1968; Sutherland & Mackintosh, 1971). The heir to

---

[2]Many contemporary examples refer to intradimensional/extradimensional strategy shifting. This is incorrect, as dimensionality refers to stimulus traits, not behavioural strategies. The terms "intramodal" and "extramodal" (e.g. Ragozzino *et al.*, 1999) are preferable.

the behavioural flexibility throne – a single task that could measure both reversal learning and intradimensional/extradimensional set-shifting – had eluded the psychologists of the 1970s.

In 1988, though, Roberts and colleagues' (1988) demonstrated the value of perseverance. Their new task – like those of Weigl (1941), Berg (1948) and Tighe (1965) – was still based on discriminations between pictorial stimuli, but now used lines superimposed onto abstract shapes (Figure 1.2) displayed on a touchscreen. Primary among the advantages of this new design was the almost infinite supply of shape and line stimuli, thus greatly facilitating not only the use of "total change" designs free from partial reinforcement (Slamecka, 1968). The wider variety of stimuli also permitted the use of within-subjects comparisons of intra- and extradimensional stages. This helps controls against the possibility that performance impairments occur due to innate stimulus discriminability – for example, the lines are harder to tell apart than the shapes – because all subjects learn all types of discrimination; the order of which can be counterbalanced between subjects and groups. In other words, if all subjects perform the extradimensional shift stage worse than the intradimensional stage, even though half the subjects shifted from line cues to shape cues and half shifted *vice versa*, then the impairment is more likely caused by attentional shifting than the inherent characteristics of the cues. These stimuli also do not rely on colour information, thereby permitting their use in subjects whose colour vision is poorer than that of the typical human subject. Furthermore, the computerisation of all aspects of the task enabled more precise behavioural measures – such as trials and errors to criterion, error type, response latency, *etc.*

This new task could also be administered to human and non-human primates[3] alike (Roberts *et al.*, 1988; Owen *et al.*, 1991; 1993; Roberts *et al.*, 1994; Dias *et al.,* 1996a; Dias *et al.*, 1997). In humans, the task has un-

[3]Most frequently the common marmoset (*Callithrix jacchus*).

covered a great deal of information about cognitive deficits in a wide-range of mental and neurological disorders (Owen *et al.*, 1991; Elliott *et al.*, 1995; 1996; Lawrence *et al.*, 1998; Robbins *et al.*, 1998), and it continues to be widely used (e.g. Leeson *et al.*, 2009; Chamberlain *et al.*, 2011; Anderson *et al.*, 2012). In monkeys, the task has greatly increased our understanding of the functional organisation of the prefrontal cortex (Roberts *et al.*, 1992; Dias *et al.,* 1996b; Dias *et al.*, 1997), and the neurochemical basis of behavioural flexibility (Roberts *et al.*, 1994; Clarke *et al.*, 2005).

In the realm of psychopharmacology, however, the monkey task has never been as useful or productive as the human task, which – along with the rest of the CANTAB series (*c.f.* Sahakian & Owen, 1992) – has revealed the effects of numerous cognitively enhancing drugs in healthy and mentally ill populations (Mehta *et al.*, 2001; Turner *et al.*, 2003; Deakin *et al.*, 2004; Turner *et al.*, 2004a; Turner *et al.*, 2004b). This may be due to the one major methodological difference between the human and monkey tasks: while all stages of the human task can be completed within one session – "around seven minutes", according to promotional materials from Cambridge Cognition, Ltd. – the monkey task takes several weeks to complete. Humans can therefore be given a drug, and tested on the task before it wears off; while monkeys would require more complicated dosing schedules. This, along with the ethical and financial implications of working with non-human primates, greatly limits the 'throughput' of the monkey task; and therefore also limits its usefulness in testing novel pharmaceutical agents.

Attentional Set-Shifting In Rats

The Norwegian rat, far more popular in psychopharmacology than any non-human primate, was an obvious choice of subject for an updated set-shifting task. However, the development of such a task was not going to

Figure 1.3: An illustrative example of the discriminations used in the Birrell and Brown "standard 7-stage" set-shifting task. A food reward is buried in the correct bowl, and rats respond by digging to find it. Bowls are discriminable by either their digging media or their odours; and in this example, the odour discrimination is relevant until the ED, when the medium discrimination becomes relevant. Adapted from Chase *et al.* (2012).

be as straightforward as giving the rats tiny touchscreens to perform the primate task. Rats are not particularly visual creatures – indeed, some breeds of rat are almost blind – meaning that shapes and lines are not "species-appropriate" for rats. While rats are capable of learning visual discriminations, they do so reluctantly (*e.g.* Lashley, 1930); often requiring hundreds of trials of training and testing (*e.g.* Bussey *et al.*, 1997). This would mean that the discriminations could not be carried out within one day, and rats would not, therefore, be any more useful than monkeys.

Inspiration for a species-appropriate rat set-shifting task came from Dudchenko and colleagues (2000), who had devised an "odour span task" to test non-spatial memory in rats. When placed in a new environment, rats'

natural tendency is to explore and forage, and this task was designed to exploit this behaviour. In the most basic of terms, rats are rewarded for digging in a cup of sand which is scented with a novel odourant. On the first trial, therefore, rats are presented with one cup filled with sand and odour A, and a food reward is buried in the bottom. On the second trial, rats are presented with two cups – scented with odours A and B – and the reward is in odour B. In the third trial, rats are presented with cups of A, B and C, and the reward is in odour C; and so it continues with a new odour on every trial until the rats make their first mistake (Dudchenko *et al.*, 2000). Although the rats require several days of training to complete this task, they required very little training to dig in the bowls of scented sand.

To create a set-shifting task for the rat, Birrell and Brown (2000) extended the general principle of requiring rats to dig in small bowls for buried rewards such that, instead of using only sand, the rats would be exposed to various "digging media". These media could then be scented with household herbs and spices to present the multidimensional stimuli[4] that lie at the heart of all set-shifting tasks (Figure 1.3). Rats, just like humans and monkeys in the CANTAB task, could therefore be presented with discriminations and reversals, as well as intradimensional and extradimensional stages, to measure behavioural flexibility.

Not only did rats demonstrate attentional set-shifting within a single session on this new task, but lesions to the medial and orbital aspects of the prefrontal cortex appeared to lead to comparable behavioural deficits in rats as had been seen previously in primates with lateral and orbital lesions respectively (Dias *et al.*, 1997; Birrell & Brown, 2000; Brown & Bowman, 2002; McAlonan & Brown, 2003). A new avenue for psychopharmacological research had therefore opened: potential new drugs for treating set-

[4]A third dimension, textured covers for the bowls, was also used; but future experiments – including the work in this thesis – would use only odours and digging media.

shifting or reversal learning deficits in humans could be tested on formally equivalent deficits in rats.

The Psychopharmacologist's Tool

Broadly speaking, current therapies for neuropsychological disorders do not effectively target the impairments in executive control seen in patients with acquired brain injury (*c.f.* Damasio *et al.*, 1985), age-related dementia (*c.f.* Owen *et al.*, 1991), schizophrenia (*c.f.* Goldberg *et al.*, 1987) and others (*c.f.* Elliott *et al.*, 1996). Part of the reason for this lack of progress may be explained by the fact that formally analogous tasks for testing the cognitive effects of drugs in humans (*e.g.* Turner *et al.*, 2004b) have only been available for rats for the last 12 years. Before the set-shifting task, the cognitive effects of drugs were frequently inferred using behavioural tasks like the open-field test of locomotory activity (*e.g.* Walsh & Cummins, 1976), the tail-suspension test (*e.g.* Steru *et al.*, 1985) or the forced-swim test (*e.g.* Porsolt *et al.*, 1977). It is undeniable that drugs affect these behavioural measures (*e.g.* Porsolt *et al.*, 1978; Prut & Belzung, 2003; Belozertseva *et al.*, 2007), but until an enterprising – and ethically dubious – psychologist throws some depressed patients down a well to see if their new antidepressant makes them less likely to stop swimming and accept death by drowning, the results of these tests will not be as easily translated as the results of rat and human set-shifting tasks.

Of course, that is not to say that these tasks are no longer used in drug development (*e.g.* Gould *et al.*, 2004; Arunrut *et al.*, 2009), but some have argued recently that successful development of the next generation of treatments for mental illnesses, such as schizophrenia, will be predicated on the use of rodent behavioural tasks with higher face validity; with the rat attentional set-shifting task (Birrell & Brown, 2000) named as a specific example (Barch *et al.*, 2009; Gilmour *et al.*, 2012). In fact, the use of this

task to screen cognitively enhancing drugs continues to increase since its inception (Goetghebeur & Dias, 2009; Lapiz *et al.*, 2009; Tait *et al.*, 2009; Gastambide *et al.*, 2011; Nikiforuk & Popik, 2011). In at least one case, the United States Food and Drug Administration has even considered data from the rat attentional set-shifting task (*c.f.* Tait *et al.*, 2009) during the approval process of a second-generation antipsychotic (Tait & Brown, personal communication).

Successful use of the rat attentional set-shifting task as a tool for translational psychopharmacology relies on a complete knowledge of what this task actually measures, as well as how these measurements are produced by the task. This, of course, has been a hallmark of the field behavioural testing since the earliest sorting experiments: under what circumstances is behaviour inflexible, and under what circumstances is it flexible? Moreover, when behaviour is inflexible, what cognitive functions (and dysfunctions!) are to blame?

Although our understanding of behavioural flexibility has developed a great deal beyond the nebulous invocations of "abstract attitudes", it unfortunately remains true that we do not fully understand the nature of attentional set, nor its formation and shifting. To fully realise the potential of the rat attentional set-shifting task to help develop a new generation of cognitive enhancers, and perhaps cures for many intractable mental illnesses, it is first necessary to further probe the nature of impaired behavioural flexibility. What populations of rats demonstrate measurable deficits? What do these behavioural impairments tell us about the cognitive processes underlying the task? Is the rat set-shifting task, as described by Birrell and Brown (2000), the most suitable design for psychopharmacological research? These are the questions that guided all that follows.

Chapter 2

General Methods

The following chapters describe several modifications to the intradimensional/extradimensional (ID/ED) test of attentional set-shifting for rats. Although most of the following experiments are based on tasks not described previously in the literature, their similarities far outnumber their differences. This section explains the common methods and materials of the following experiments, as well as the procedure and analysis of "standard 7-stage task". This task, first described by Birrell & Brown (2000), forms the basis for all the modified tasks that follow.

Animals

All animals used in the following experiments were male Lister Hooded rats obtained from breeding establishments registered with the UK Home Office. The aged rats in chapter three were obtained from Harlan UK, while the rats in chapters four, five and six were obtained from Charles River UK. All animals were experimentally naïve before testing. The work described in this thesis complied with the Animals (Scientific Procedures) Act (1986), and licensed procedures were carried out under the authority of Project and Personal Licenses issued by the Home Office. The Project Licenses under which this work was conducted were approved by the University of St Andrews Animal Welfare and Ethics Committee.

Husbandry conditions were virtually identical for all rats, though minor variations may be reported in each chapter. Where possible, rats were socially housed in groups of up to four in cages measuring approximately 50 x 30 x 25 cm (henceforth, "larger cages"). In some cases rats were single- or pair-housed in cages measuring 40 x 23 x 19 cm ("smaller cages"). Chewable toys (e.g. thick cardboard tubes, wooden chew bars, cardboard dome "houses", etc.) were placed in the cage following a pseudorandom, once-weekly sequence to provide "environmental enrichment". The cages also contained at least one clear, plastic tube fastened to the top of the cage, into which rats could climb and sit. The colony room and testing labs were maintained at 21°C ± 2°, with a relative humidity of 55% ± 10%. In the colony room, lighting followed an artificial twelve-hour light cycle (lights on at 7am), and the testing rooms were lit with natural and artificial light.

Not less than two weeks prior to the start of an experiment, but in some cases several months before testing, rats' access to food was moderately restricted to approximately 15–20 g of standard lab diet per rat per day, though water was always freely available. The primary purpose of con-

trolling rats' access to food was not to encourage weight loss – indeed, as can be seen in appendix two, rats continue to grow at a healthy rate under a controlled diet. Rather, controlled access to food was intended first to help prevent rats becoming overweight through "boredom eating"; and second to increase their motivation to work for food rewards.

The rats used in the following experiments were not reused for any other procedures – regulated or otherwise. Therefore, all animals were terminated at the completion of testing by an appropriate method selected from Schedule I of the Animals (Scientific Procedures) Act (1986); with the exception of some rats in chapters four and five, which were – as will be discussed fully in those chapters – intracardially perfused with fixative under terminal anaesthesia.

Procedure

Training and testing took place in a specially modified home-cage, which consisted of a large holding section, with two individually partitionable sections (see Figure 2.1). In each of these smaller sections, ceramic pet food bowls (internal diameter 7 cm; internal depth 4 cm) could be placed, and it was between these two bowls – or more accurately, their contents – that rats discriminated. By filling the bowls with different digging substrates, and then scenting this substrate with a herb or a spice, discriminations along two dimensions could be presented: one between two odours and one between two digging media. To respond, rats dug in the bowls to find a food reward: half a "Honey Loop" (Kellogg Company, UK). The rat set-shifting task is designed to exploit natural foraging behaviour; making it maximally species-appropriate.

Before testing, each rat was given one bowl filled with sawdust and six pieces of food reward. Rats are naturally wary of (and thus slow to consume) foods they have not eaten before, and allowing rats the opportunity

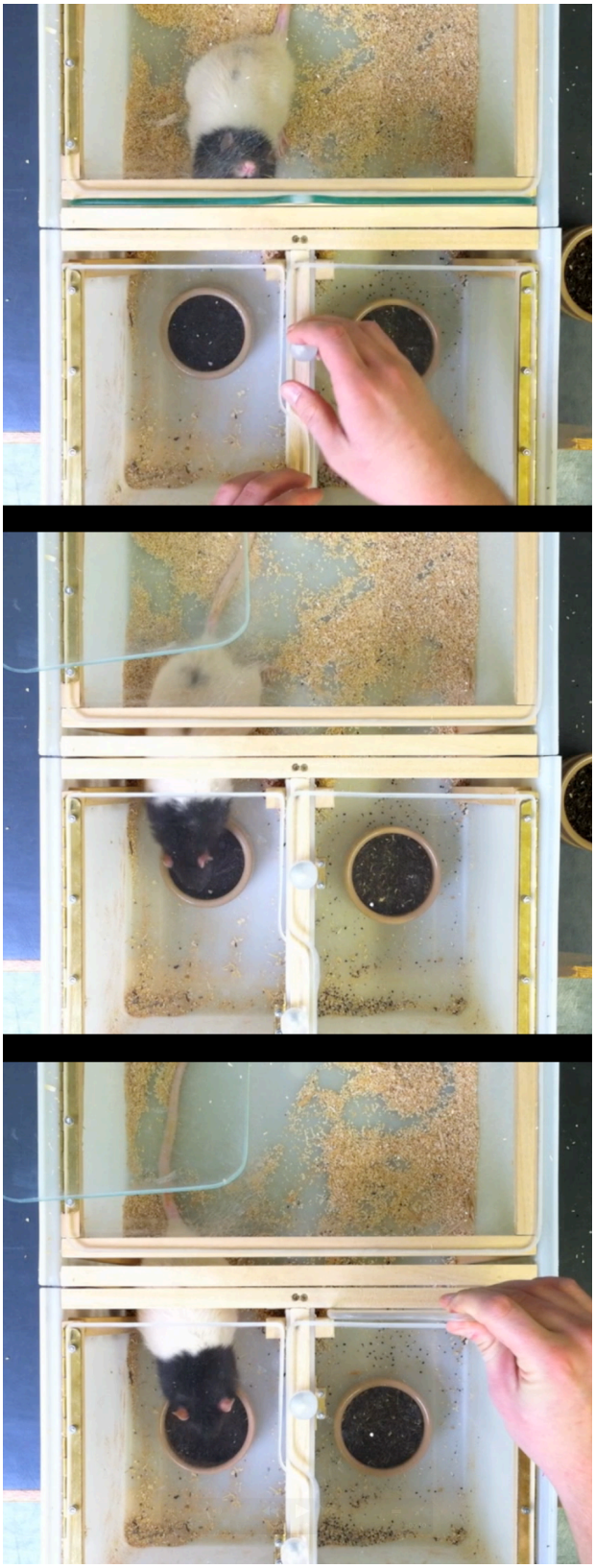


FIGURE 2.1: The set-shifting task in action. Top: the rat waits in the large section for the trial to begin. Middle: the barrier is lifted, allowing the rat to approach a bowl and dig. Bottom: excluding the first four trials, a half-barrier blocks access to the other bowl after the rat responds.

to sample the reward before training reduces this "neophobia". The next day, rats were trained to dig in bowls filled with sawdust to reliably obtain the food reward. This was achieved by first placing the food on top of the sawdust, then burying it progressively further down with each trial. Rats were typically able to reliably find the food reward after six trials in each bowl. Bowls filled with all the media and odours (mixed in sawdust) that would be used in the task were then presented, and rats were allowed to obtain rewards from each stimulus twice. The purpose of this stage was to "pre-expose" the rats to the stimuli so that they would be less likely to refuse to dig in any of the bowls during testing.

In the final stage of training, rats were given two simple discriminations – one between two odours in sawdust, and one between two media with no added odour – using exemplars that would not be used again in any future test. For the first four trials of the discriminations, rats were permitted to "self-correct" following an incorrect dig, and obtain the food reward from the correct bowl; however after these four trials an incorrect dig led to access to the correct bowl being blocked. Rats were deemed to have successfully acquired a discrimination after reaching criterion performance of six consecutive correct responses (including the first four trials; $p = 0.0156$). In a typical experiment, testing took place a day or two after training. However, as it is not necessary to retrain animals between multiple tests (Tait *et al.*, 2009), some of the testing sessions reported in the following chapters may have taken place days, weeks or months after training; as indicated in each chapter.

In the testing phase of the set-shifting task, rats were presented a series of discriminations, the exact order of which can be found in each chapter's methods sections. Common to all tasks were the following stages, which always occurred in this general order, but sometimes with reversal stages interspersed. First rats performed a simple discrimination (SD) between two odour or medium stimuli. After this, the complementary, ir-

relevant pair of stimuli was added to form the compound discrimination (CD). After the CD stage were one or more intradimensional novel acquisition stages (ID), where new compound stimuli were introduced, but the relevancy of the two dimensions remained the same as in previous stages. After the ID stage(s), there followed an extradimensional shift (ED) stage, where rats were again presented with new compound stimuli, but the relevancy of the two dimensions was switched. Thus, if the rats were required, for example, to attend to odour stimuli in the SD, CD and ID stages; they would need to shift their focus to medium stimuli in the ED. Most of the tasks used in this thesis also included reversal stages, where the contingencies of the immediately preceding discrimination were swapped such that the previously incorrect stimulus became correct and vice versa.

The "standard 7-stage task", as described by Birrell & Brown (2000) is the task upon which all the experiments in this thesis are based. The order of stages for the 7-stage task is SD, CD, CD reversal ("Rev1"), ID, ID reversal ("Rev2"), ED and finally ED reversal ("Rev3").

As in training, rats were permitted to self-correct during the first four trials, and the progression through the stages was determined by the rats reaching criterion performance of six consecutive correct digs. A trial began when the barrier blocking access to the bowls was lifted, and a dig was recorded after the rat displaced a significant amount of the contents of a bowl whilst continually investigating the area uncovered. If the rat did not dig within 10 minutes from the start of a trial, a "non-dig" was recorded, the trial was ended, and the rat was left until he appeared engaged in the task again (*e.g.* rearing, investigating the barrier, pacing the length of the holding section, *etc.*). Non-dig trials were not included in the final trials to criterion measurement.

As full counterbalancing cannot be achieved due to the large number of possible exemplar pairings, the same stimulus pairings were always used to reduce the number of possible combinations. Shift type (*i.e.* odour

to medium, or medium to odour), order of exemplar-pair presentation (*e.g.* first pair, second then third; second, third then first, *etc.*), and initially correct stimulus within a pair were balanced as much as possible.

Statistical Analyses

Trials to criterion, errors to criterion, latency to dig, and number of non-digs were recorded on all tests. Trials to criterion and errors to criterion often reveal the same pattern of results, but previous work has suggested that trials to criterion data are usually more powerful (Tait & Brown, 2007). Data were analysed primarily using repeated-measures analyses of variance (ANOVA), with the specific factors and levels reported in each chapter. When significant interactions between the factors were found in the "omnibus" ANOVA tests, simple main-effects analyses were conducted with additional ANOVA tests restricted to the relevant factors and levels. The F-values associated with the simple main-effects or simple interactions were recomputed using the appropriate error term and degrees of freedom from the omnibus ANOVA (Winer, 1971). Where necessary, Huynh-Feldt corrections were applied to data which violated the assumption of sphericity, though uncorrected degrees of freedom are presented throughout this thesis. Graphs were drawn in Prism v5.4 (GraphPad, CA, USA) for Mac OS X, and statistical analyses were computed in SPSS v19.0 (IBM, CA, USA) for Mac OS X.

Chapter 3

Age-Related Cognitive Decline in the Rat: Amelioration Through Modulation of the N-methyl-D-aspartate Receptor

In humans and other animals, cognitive function tends to decline with age. Preventing, slowing or even reversing this trend is of growing importance in the context of our aging society. The attentional set-shifting task appears to be especially sensitive to age-related cognitive decline in humans, but few experiments have investigated whether rats show comparable decline on this task. This experiment demonstrated that middle-aged rats (12–18 months old) are impaired at reversal learning compared to young rats, but they do not demonstrate attentional set-shifting on the standard 7-stage task. An experimental positive allosteric modulator of the N-methyl-D-aspartate receptor (ORG49209) improved reversal learning in these animals, suggesting a potential new avenue for pharmacological cognitive enhancers.

Introduction

An unfortunate consequence of getting older is the gradual decline of our bodies and minds. Cognitive functions – such as planning, ignoring distractions and retrieving information from long-term memory – may start to fade as early as middle-age (Robbins *et al.*, 1998; Cepeda *et al.*, 2001), and this inevitable progression is greatly exacerbated by neurodegenerative diseases (Owen *et al.*, 1993; Lawrence *et al.*, 1998; Perry & Hodges, 1999). The 'real-world' consequences of age-related cognitive decline can be formidable: individuals with greater cognitive dysfunction are more likely to suffer falls and other mobility impairments (Buchman *et al.*, 2011), and difficulty retrieving memories (Gauthier *et al.*, 2006) can be very distressing for older people and their families. In many countries, the proportion of the population aged 60 years or older is growing faster than any other demographic (United Nations, 2011), and many are suggesting that the current generation will need to retire later in life than their parents (*e.g.* Helm, 2012). The need for effective treatments and preventative measures for age-related cognitive decline is therefore greater than ever.

The development of new treatments is principally reliant on accurate diagnosis of the underlying impairment: without full awareness of the problem we cannot begin to generate possible solutions. Furthermore, successful treatment of disorders that get progressively worse – including, in this case, otherwise healthy aging – often depends on early detection and intervention. One tool that satisfies both of these conditions is the attentional set-shifting task, which is capable of detecting subtle impairments in otherwise healthy individuals earlier than many other cognitive tasks (Robbins *et al.*, 1998); and before the development of more serious symptoms, such as chorea in Huntington's disease (Lawrence *et al.*, 1998). This type of task is also available for many different species (Dias *et al.*, 1996a; Birrell & Brown, 2000; Bissonette *et al.*, 2008), making it possible

to quantify a novel drug's effects at every stage of development: from animal model to human patient.

Typically, rats used in attentional set-shifting experiments are aged between six and eight months. Some groups have tested older rats – aged between 18 and 24 months – on modified set-shifting tasks, revealing putative age-related impairments in set-shifting (Barense *et al.*, 2002; Rodefer & Nguyen, 2008). However, it is not clear whether the task modifications made by these groups biased the rats towards performing the task in a certain way. Both groups omitted the first reversal stage of the standard 7-stage task, and neither found a significant age-related impairment in reversal learning. This is contrary to a number of previous operant conditioning experiments showing impaired reversal learning in older rats (*e.g.* Schoenbaum *et al.*, 2002; 2006). It may be that reversal deficits are only detected on tasks requiring extensive training – as in Schoenbaum's odour-based go/no-go tasks (*ibid.*) – and will not be found on the relatively short rat set-shifting task. However, both of the previous set-shifting experiments reported non-significant trends towards reversal learning impairments (Barense *et al.*, 2002; Rodefer & Nguyen, 2008), which raises the possibility that omitting the first reversal of the set-shifting task reduces the likelihood of detecting a reversal impairment. Therefore, the primary purpose of this experiment was to identify the pattern of impairment – if any – in "middle-aged" (approximately 12 months old) rats performing the standard 7-stage task, and then investigate the stability of this performance during senescence.

A secondary purpose of this experiment was to investigate whether potential age-related deficits can be ameliorated through pharmacological intervention. Dysfunction of the glutamatergic system is associated with many psychiatric and neurological disorders, and reduced activity of the N-methyl-D-aspartate (NMDA) receptor appears to be involved in age-related cognitive deficits (Nicolle & Baxter, 2003; Zahr *et al.*, 2008). Reduced

NMDA activity has also been associated with bipolar disorder (Woo *et al.*, 2004; Sanacora *et al.*, 2008) and schizophrenia (Harrison & Weinberger, 2004; Marx *et al.*, 2009). However, despite the clear need, few positive modulators of NMDA are available in the clinic.

The reason for this pharmacological impasse is multifaceted. Primarily, the effects of NMDA receptor activity on cognition are not straightforward. Activity at this receptor is important for the induction of long-term potentiation ("LTP"; Bliss & Collingridge, 1993), an essential mechanism of learning. Negative modulation (or complete blocking) of NMDA receptor activity leads to memory and executive control deficits (e.g. Mathis *et al.*, 1996; Palencia & Ragozzino, 2006), as well as symptoms resembling schizophrenia (*e.g.* Jentsch & Roth, 1999). Superficially, it might therefore seem that increasing NMDA activity will necessarily enhance cognition. However, blocking NMDA activity can inhibit the decay of LTP, which improves spatial memory (Villarreal *et al.*, 2001), and excessive increases in extracellular NMDA levels lead to neuron death (Bonfoco *et al.,* 1995; Liu *et al.*, 2007). Altogether these results show that NMDA levels which are either too low or too high can have formidable adverse consequences.

Positive allosteric modulators of NMDA have received some attention as an avenue for boosting effective NMDA levels without risking adverse events. The endogenous neurosteroid pregnenolone sulphate is one such compound that has shown promise as a cognitive enhancer (Cheney *et al.*, 1995; Mathis *et al.*, 1996; Vallée *et al.*, 1997) without risking neurotoxicity (Paul & Purdy, 1992). However – like many similar compounds – pregnenolone sulphate is not a 'clean' molecule: while positively modulating NMDA, it is an antagonist of the gamma-aminobutyrate-A ($GABA_A$) receptor (Gibbs *et al.*, 2006). Furthermore, pregnenolone sulphate has notably low potency (*ibid.*; Vallée *et al.*, 1997), and can produce convulsions (Paul & Purdy, 1992), which altogether reduces the clinical usefulness of this compound for targeting NMDA hypofunction.

The effects of a novel positive allosteric modulator of NMDA – ORG49209 – with similar structure to pregnenolone sulphate were tested in middle-aged rats. This molecule appears to be uniquely potent and selective (Chase *et al.*, unpublished observations) and so represents an important potential treatment for NMDA-mediated cognitive dysfunction.

Methods

Due to difficulties obtaining and maintaining a sample of aged rats, this experiment was conducted in phases using two cohorts. When testing commenced, the weight-range for the first group (n = 12) was 466–598 g, and for the second group (n = 8) was 433–560 g. Four rats from the first cohort did not survive to complete all phases of the experiment, and their data are not presented here.

Rats were pair-housed in larger cages under standard housing conditions (see chapter 2). When rats were between 10–11 months old they were placed on a moderately restricted diet (15–20 g of standard diet per rat per day) with water freely available.

Habituation and pre-exposure followed the standard protocol, as presented in chapter two, and began when the rats were approximately 12–13 months of age. Previous experiments in our lab have shown that rats can be tested multiple times on the set-shifting task without needing to be trained more than once (Tait *et al.*, 2009). Therefore, although all rats were tested at least twice between the ages of 12 and 18 months, they were only pre-exposed and trained before their first test at 12–13 months of age.

Testing followed the protocol of the standard 7-stage task, as presented in chapter two. A brief summary of the stages and stimuli used in this experiment can be seen in Figure 3.1.

Collecting Baseline Behavioural Data

The first cohort of rats was tested at 12 months of age (n = 12), and again at 18 months of age (n = 8) to assess both the presence of an age-related deficit, as well as its stability over time. To ensure sufficient statistical power for the drug study, the second cohort was tested at 13 months (n = 8), and – given a convincing lack of differences on each cohorts' baseline data (see below) – the data sets were then pooled. To put the older rats' re-

	Discriminanda	*Mixed with...*
Simple Discrimination	Coarse tea, not fine tea	Nothing
Compound Discrimination	Coarse tea, not fine tea	Cinnamon or ginger
First Reversal	Fine tea, not coarse tea	Cinnamon or ginger
Intradimensional Stage	Sand, not grit	Sage or paprika
Second Reversal	Grit, not sand	Sage or paprika
Extradimensional Shift	Turmeric, not clove	Wood chips or sawdust
Third Reversal	Clove, not turmeric	Wood chips or sawdust

Figure 3.1: The stages and stimuli of the standard 7-stage task, first described by Birrell and Brown (2000). Although the stages always occur in this order, the stimuli that are used in each stage are balanced within groups and between tests.

sults in context, data from younger rats (n = 137, typical age approximately seven months) tested in previous experiments are quoted below. Due to the gross inequality of the sample sizes, as well as the *a posteriori* nature of these comparisons, statistical tests between these young and older rats were not appropriate.

Drug Preparation and Administration

Behavioural data on ORG49209 is very limited, and it is not clear if this neurosteroid works best under acute or chronic conditions. To "hedge"

between these two dosing schedules, sixteen rats (eight from each cohort) were given a regimen of once-daily injections (1 ml/kg; intraperitoneal) for 17 days, and then tested approximately 30 minutes after the last injection. On the first two days, rats were given habituation injections of vehicle, which were followed by 15 days of drug or vehicle administration. Assignment to each group was determined pseudorandomly.

Four rats from each of the two cohorts were given 10 mg/kg ORG49209 suspended in 10% hydroxy-propyl-ß-cyclodextrin (HPBCD; Sigma-Aldrich, UK) in sterile water, and the remaining eight rats were given equivalent volumes of the HPBCD vehicle with no active compound.

Statistical Analyses

Trials to criterion data were analysed using multiple repeated-measures analyses of variance (ANOVA) tests. Baseline performance was assessed using Stage and Age as within-subjects factors, and Cohort as a grouping factor. The effects of the drug on performance were assessed using a separate ANOVA test with Stage as a within-subjects factor, and Treatment as a grouping factor. Planned comparisons were performed using simple main-effects analyses, as described in chapter two.

Results

Middle-aged rats show a clear, stable and replicable pattern of impairment on the standard 7-stage task

To first establish an index of baseline performance, rats were tested on the standard 7-stage task when they were 'middle-aged', at 12 (n = 8) or 13 (n = 8) months of age. Half of these rats – the first cohort – were tested again at 18 months of age (n = 8) to assess the stability of the age-related deficit. There was no difference between the two cohorts on performance of any stage of the task (Figure 3.2; main effect of Cohort: $F_{1,98} = 1.44$, $p = 0.233$; all interactions: $F < 1$), nor was there any deterioration seen after six months (main effect of Age and all interactions $F < 1$). Given this convincing lack of difference, data from the test immediately preceding the treatment test were pooled to form a meaningful and statistically powerful "middle-aged baseline" group (n = 16, mean age 15.5 months).

In testing these rats, it was apparent that they performed overall less well than younger animals previously tested in the lab. They explored the box and the bowls less than is typical of younger rats. Although it appeared that all stages of the task were completed more slowly and required more trials, this was particularly true of the discrimination reversals (Figure 3.3). To complete the three reversal stages, young rats require a mean of 12.5 trials (range 8 to 23.67) to reach criterion, while these middle-aged rats require almost twice as many trials, with a mean of 25.5 (range 15.67 to 34.33). This reversal learning impairment did not appear to be due to increased perseveration (Iversen & Mishkin, 1970), as the first time middle-aged rats experience a reversal they make their first correct response after a mean of only 2.25 errors, which represents a mean of 20.7% of their total errors. In the first six trials of the first reversal stage, the rats committed a mean of 3.69 errors, which is less than would be expected if

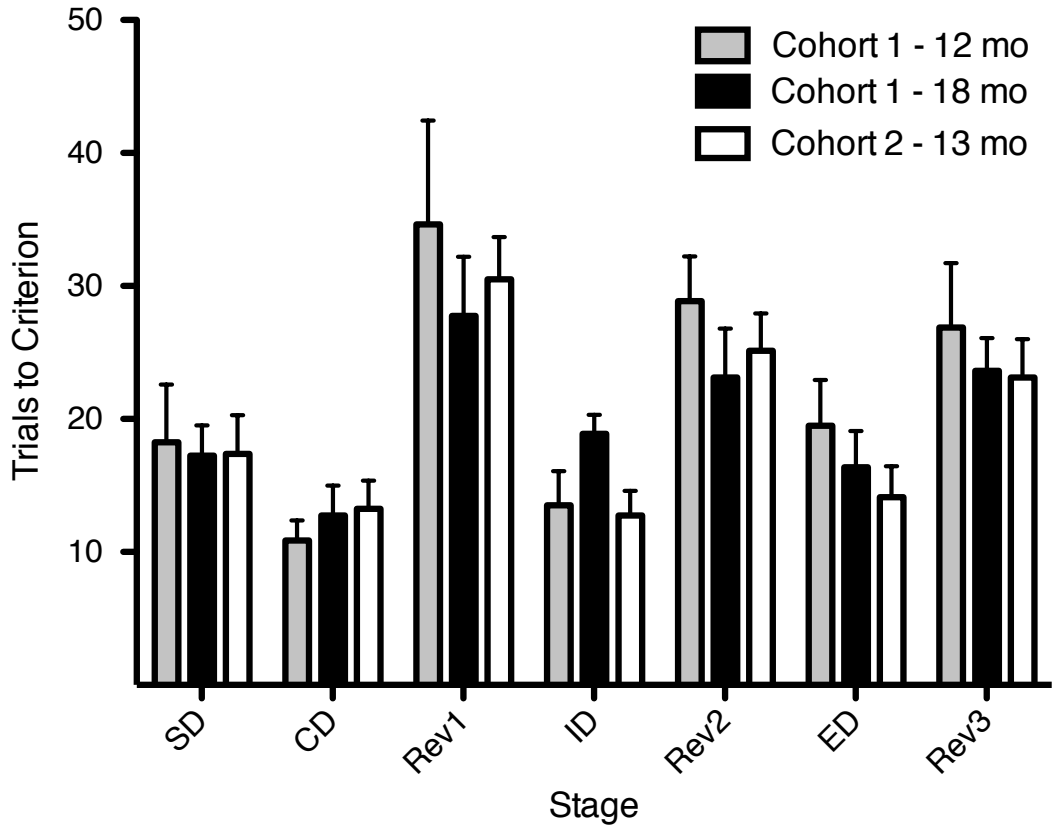


Figure 3.2: Mean trials to criterion (+SEM) on the standard 7-stage task for Cohort 1 at two time-points, and Cohort 2 at baseline. No statistical differences were found within Cohort 1, suggesting that performance impairments on the task are stable for at least six months. Cohort 2 also did not differ from either of the two other baseline tests.

the rat were perseverating and responding preferentially to the incorrect bowl.

The intra- and extradimensional stages of the task are the critical stages for assessing attentional set-shifting, with the greater number of trials to criterion at the extradimensional shift stage compared to the intradimensional stage (the 'shift-cost') indicating the flexibility with which an attentional set can be shifted. In these middle-aged rats, there was no evidence of a shift-cost: the group mean trials to criterion at the ID stage was 15.8, while the group mean for the ED stage was 15.3 trials. This is a strikingly different pattern of performance from young rats, which typically re-

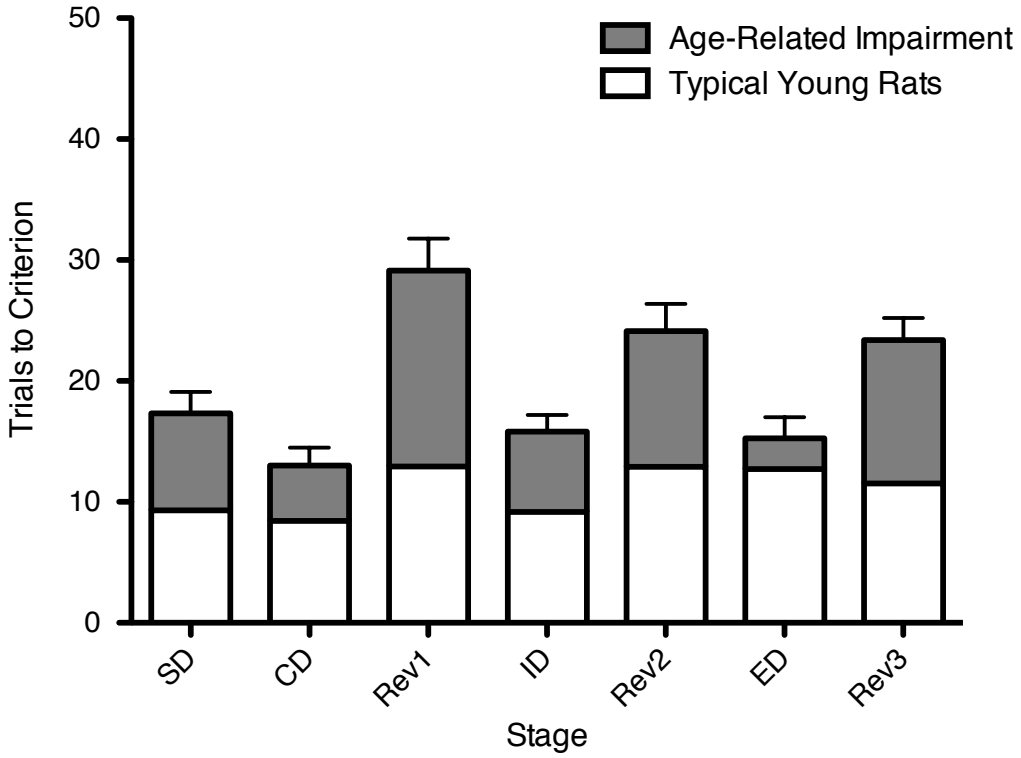


Figure 3.3: Mean trials to criterion (+SEM) on the standard 7-stage task in the pooled-baseline middle-aged group (grey bars), stacked on mean performance of 137 young control rats (white bars) tested in previous experiments. Although there is an apparent "cost" of aging on every stage, its magnitude is particularly large on the reversal stages.

quire 3.6 more trials to learn the ED (mean 12.7 trials to criterion) than the ID (mean 9.12 trials to criterion). The most parsimonious explanation for a lack of shift-cost is that rats have not formed an attentional set, though we cannot say for certain that this absence of evidence is evidence of absence.

Some older rats (6 out of 16) did demonstrate a convincing shift-cost, however (Figure 3.4, left). In these animals the mean performance on the intradimensional stage was 13.17 trials, and on the extradimensional shift stage was 21.17 trials. While the overall group means do not reveal much information about the set-shifting abilities of older rats, those rats

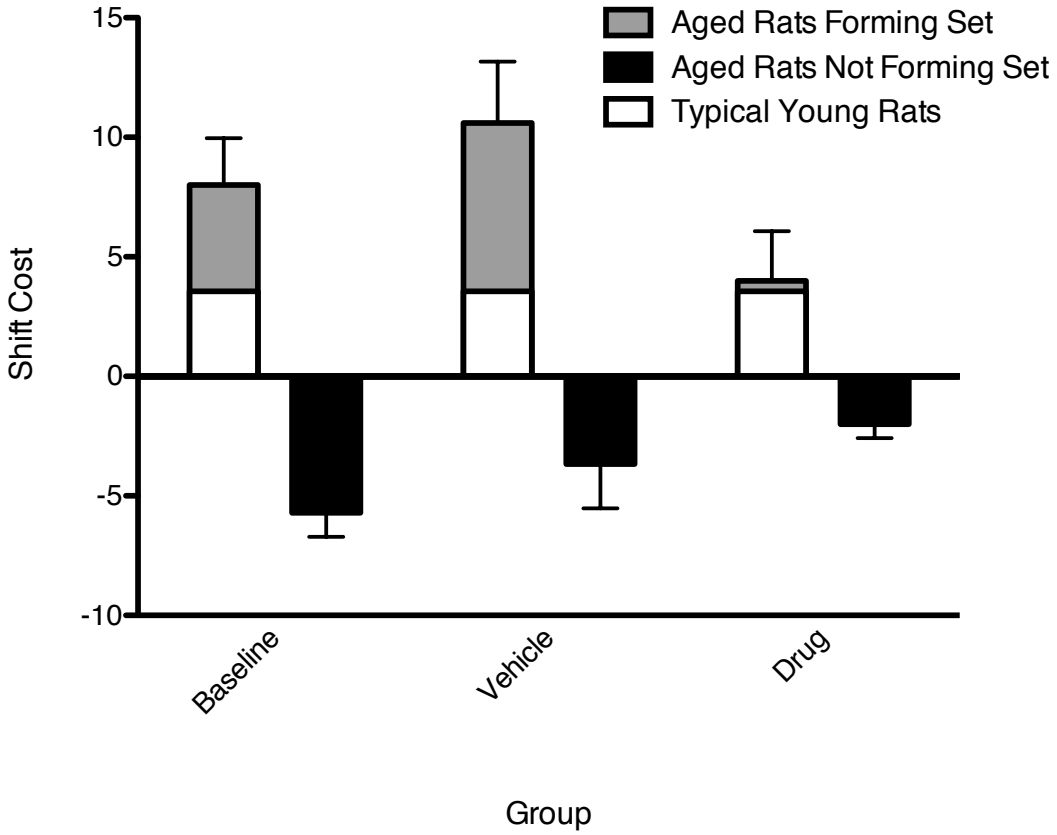


Figure 3.4: Mean shift-costs (+SEM) in numbers of trials to learn the ED less the number of trials to learn the ID. Middle-aged rats demonstrating positive shift-costs (grey bars) are stacked on typical young rat performance (white bars), while middle-aged rats with negative or absent shift-costs are shown in black. Rats in the baseline test (left) and vehicle-treated rats in the drug test (middle) appear to be impaired at shifting relative to young rats, while ORG49209-treated rats (right) appear to have no set-shifting deficit.

that clearly demonstrate an attentional set appear to be impaired at shifting relative to typical young rats.

ORG49209 ameliorates the age-related deficits

Middle-aged (13 months; n = 8) and older (18 months; n = 8) rats were treated with ORG49209 (n = 8, four from each cohort) or vehicle (n = 8). As there was no difference between the middle-aged and the older rats at baseline, the samples were pooled and analysed together.

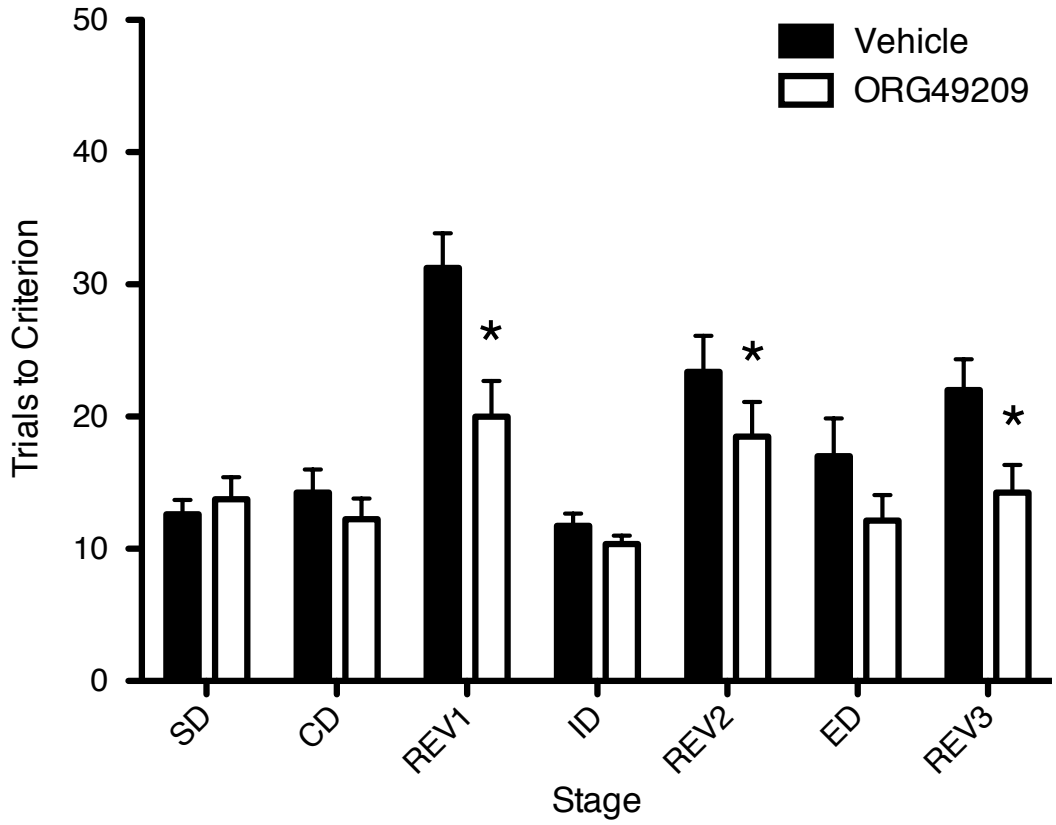


Figure 3.5: Mean trials to criterion (+SEM) in the rats treated with vehicle or ORG49209. Drug-treated rats required significantly fewer trials than controls to learn the three reversal stages, indicating that ORG49209 attenuates the age-related reversal deficit.

ORG49209 significantly attenuated the age-related impairment (Figure 3.5; omnibus main effect of Treatment: $F_{1,14} = 9.4$, $p < 0.05$; and omnibus Treatment by Stage interaction: $F_{6,84} = 2.24$, $p < 0.05$). The improvement in the Org49209 group was particularly obvious at each of the three reversal stages, as revealed by ANOVA restricted to the three reversal stages (main effect of Treatment: corrected-$F_{1,14} = 8.19$, $p < 0.05$; Stage by Treatment interaction: corrected-$F_{6,84} = 1.32$, $p = 0.26$). The improved reversal learning in the group treated with ORG49209 was not due to reduced levels of perseverative responding, as the groups did not differ on number of trials before the first correct dig on the first reversal ($t_{14} = 0.527$, $p = 0.61$), nor on the ratio of these early errors to the total errors before

reaching criterion (t_{14} = 0.176, p = 0.86), nor on the number of errors in the first six trials of the reversal (t_{14} = 1.7, p = 0.11).

ORG49209 had no effect on group performance at the extradimensional shift stage, with the vehicle and drug treated groups showing no significant difference between the ID and ED stages (main effect of Stage: $F_{1,12}$ = 3.48, p = 0.09, interaction: F < 1). However, as in the baseline test, there appeared to be heterogeneity within the groups on performance on the ID and ED stages. Some rats (5 out of 8 in each group) demonstrated positive shift-costs, while the others did not. Comparing performance of only those rats with positive shift-costs in the drug and vehicle groups revealed that rats treated with ORG49209 required fewer trials to learn the ED shift stage (Figure 3.4; t_8 = 2.0, p < 0.05, one-tailed).

Discussion

Cognitive capacity declines with age, which is independent from – although obviously compounded by – age-related dementia. Recent evidence has highlighted the role of the NMDA receptor in this form of cognitive dysfunction (Zahr *et al.*, 2008), linking age-related cognitive decline with a host of other disorders associated with NMDA hypofunction (*c.f.* Harrison & Weinberger, 2004; Woo *et al.*, 2004; Sanacora *et al.*, 2008). Pharmacological interventions for reduced NMDA activity have been marred by issues of excitotoxicity (Bonfoco *et al.,* 1995), low potency (Vallée *et al.*, 1997), and poor receptor selectivity (Gibbs *et al.*, 2006), making extant positive modulators of NMDA suboptimal for clinical use. For example, while the neurosteroid pregnenolone sulphate has received attention as a putative treatment for schizophrenia (Marx *et al.*, 2009), it requires high doses to be effective, and its cognitive effects cannot reliably be assigned to positive modulation of NMDA or negative modulation of $GABA_A$ (Gibbs *et al.*, 2006).

Recently, my collaborators have identified a novel positive allosteric modulator of the NMDA receptor that does not appear to be affected by these issues (Chase *et al.*, unpublished observations). Naturally, this molecule represented an intriguing potential treatment for the cognitive dysfunction associated with NMDA hypofunction, such as that seen in older rats (Nicolle & Baxter, 2003).

From the age of about 13 months, rats show reduced flexibility in the standard 7-stage task, suggesting that the middle-aged rat represents an ecologically valid means for testing age-related cognitive decline in preclinical settings. While middle-aged rats appear to be slower than young rats to learn the discriminations at every stage of the task, the magnitude of this impairment was greatest at the reversal stages. The age-related impairment in reversal learning is consistent with many other reports (Roman *et al.*,

1996; Schoenbaum *et al.*, 2002; 2006; Brushfield *et al.*, 2008; Gilbert *et al.*, 2009).

Previous reports have also suggested that aged rats (18+ months) have set-shifting deficits, as demonstrated by increased shift-costs compared to younger rats (Barense *et al.*, 2002; Rodefer & Nguyen, 2008). This was not the case for the group means in the baseline or drug tests, as not all of the older rats in this experiment demonstrated an attentional set. Over half of our rats (10 out of 16) had absolutely no shift-cost (ED – ID <= 0), but in the remaining rats the shift-cost was convincing. As we did not detect formation of attentional set in more than half of the rats we tested, it is not possible to assess whether they had a set-shifting deficit: a rat does not need to shift a set that was never formed. Nevertheless, the rats that did form an attentional set in our baseline test demonstrated shift-costs that were nearly three times greater than we typically observe in younger rats. This difference is at least as great as those reported by Barense *et al.* (2002) and Rodefer & Nguyen (2008) for their aged rats. In other words, although the baseline mean shift-cost in our experiment appears reduced or absent, those rats that demonstrated an attentional set appeared to be impaired at set-shifting. The finding that a slight majority of our rats did not demonstrate a shift-cost during the baseline test suggests that older rats may be impaired at forming attentional set, which may or may not be related to their impaired reversal learning (Tait & Brown, 2008).

Treatment with ORG49209 significantly improved reversal learning in middle aged rats. It may also be that ORG49209 improves set-shifting, as revealed by a comparison of only those rats which demonstrated positive shift-costs in the two treatment groups. However, this *post-hoc* analysis provides only moderate evidence supporting this possibility, and further investigations are therefore necessary. Nevertheless, these results clearly highlight the potential for positive allosteric modulators of NMDA as a candidate class for next-generation cognitive enhancers.

Chapter 4

Lesions of the Orbital Prefrontal Cortex Impair the Formation of Attentional Set in Rats

More often than not, rats demonstrating impaired performance on the first reversal of the standard 7-stage task do not demonstrate positive shift-costs. This experiment tested the hypothesis that this 'absence of evidence' actually represents a failure to form attentional set by these animals. Young rats with lesions of the OFC – an area that helps signal which cues in the environment predict reward – were tested on a set-shifting task with no reversal stages, and with multiple intradimensional stages before the extradimensional shift stage. This can provide a measures of set-formation as well as set-shifting. Compared to controls, lesioned rats were slower to form attentional set. When they did form a set, they required more trials to complete the extradimensional shift stage.

Introduction

The gradual decline of our cognitive capabilities is mirrored – and almost certainly caused – by physical degradation of the brain itself. Aged rats demonstrate significantly fewer synapses, as well as reduced neurogenesis, in the hippocampus (Kuhn *et al.*, 1996; Rosenzweig & Barnes, 2003), which is likely related to age-related impairments in spatial memory (Fischer *et al.*, 1991; Frick *et al.*, 1995). Older rats, like older humans, show reduced levels of striatal glutamate (Nicolle & Baxter, 2003; Zahr *et al.*, 2008), and this correlates with deficits in behavioural flexibility; while dopaminergic striatal lesions in young rats are associated with reversal learning impairments (O'Neill & Brown, 2007). Also, with age the orbital prefrontal cortex (OFC) becomes less flexible in coding cue-reward relationships, as well as firing less overall in response to cues that predict reward, both of which may explain the reversal impairments seen in older rats (Schoenbaum *et al.*, 2002; 2006).

Degradation of the OFC in older animals is a particularly interesting finding, as the impairments seen in older rats in chapter three were very similar to the results of a previous investigation conducted on young rats with OFC lesions (McAlonan & Brown, 2003). Like the middled-aged rats, OFC-lesioned rats were impaired at all three reversal stages, and demonstrated no significant difference between the intra- and extradimensional stages. This, together with the experiments by Schoenbaum and colleagues (2002, 2006), raises the possibility that the impairments demonstrated by middle-aged rats on the set-shifting task stem from physical changes to the OFC during senescence. However, few experiments have assessed whether this overt similarity reflects a true equivalence of deficits. Put another way, it has generally been under-investigated whether distinct cognitive deficits can produce (misleadingly) similar behavioural deficits. To address this hypothesis, it is first necessary to gain a greater understanding of the

cognitive deficits underlying the behavioural impairments seen in both middle-aged rats, and young rats with OFC lesions.

When a subject forms an attentional set, they become predisposed to attend to those aspects of multidimensional stimuli that have previously predicted reward (Sutherland & Mackintosh, 1971; but see also Esber & Haselgrove, 2011). This predisposition is measured by comparing subjects' performance on two similar discriminations: one that can be solved using the same dimension of cues as prior stages, and one that requires shifting to a new set of cues. The attenuated performance between the latter and former discriminations is often referred to as the "shift-cost", and it is the main behavioural measure of attentional set-shifting tasks.

In several past investigations, including the experiment reported in chapter three, rats that demonstrated reversal learning impairments did not show significant shift-costs. The most notable example of this phenomenon is seen in McAlonan & Brown (2003), which demonstrated that lesions of the OFC result in large reversal learning impairments, but no significant shift-costs. It has previously been argued that the OFC has no role in attentional set (Dias *et al.,* 1996b; Dias *et al.*, 1997; Bissonette *et al.*, 2008), however the failure to observe shift-costs in rats gives cause to doubt this conclusion. As the parsimonious explanation for diminished or absent shift-costs is that attentional set was weaker or never formed – therefore not requiring shifting – it is possible that the OFC lesion impairs rats' abilities to form an attentional set.

More recently, research into the function of the OFC has focussed on this region's involvement in reinforcement learning, but the results appear to support this hypothesis. In grossly oversimplified terms, reinforcement learning theory states that learning is driven by a desire to maximise reward. Learning is therefore mediated by "prediction errors", the difference (positive or negative) between actual obtained rewards and the expectations thereof; and "outcome expectancies", the supposition that a given

cue or action will result in the same reward as it has in the past. On the first few trials of a discrimination on the standard 7-stage task, the discovery of a Honey Loop in the bowl scented with cinnamon would generate a positive prediction error for this cue[5], while a negative prediction error would be generated by ginger[6]. The rat would then begin to respond preferentially to the cinnamon bowl, and over successive trials an outcome expectancy between cinnamon and reward would be established. During a reversal stage, where the rewarded cue in the initial discrimination is now negatively correlated with reward, the prior outcome expectancy – by definition – is violated. To learn the reversal, rats must inhibit their previous response behaviour – that is to say, stop digging in the cinnamon bowl – and identify the new rewarded cue, *i.e.* the ginger, by responding to the new prediction errors and updating their outcome expectancies.

The OFC appears to have two primary roles that facilitate learning in this scenario, which may be dissociated between the area's subregions (*e.g.* Noonan *et al.*, 2010). Principally, neurons in the OFC fire in response to cues that are expected predictors of reward (for review, see Schoenbaum *et al.*, 2009), and the rate of firing is dependent on the predictiveness of these cues, or the probability that their presentation (or a response to their presentation) will be followed by reward (Pennartz *et al.*, 2011). When an expected reward does not materialise, the OFC may use this information to update prediction errors generated elsewhere (Takahashi *et al.*, 2009; Roesch *et al.*, 2010; McDannald *et al.*, 2011; Takahashi *et al.*, 2011; Wallis & Kennerley, 2011); thus helping the subject maintain an accurate representation of what responses and cues are predictive of reward (Takahashi *et al.*, 2011). Animals with lesions of the OFC are

---

[5]Expected: no reward, or an expectation-free state; Actual: half a Honey Loop, the value of which is discounted by the effort of digging in the bowl.

[6]Expected: as above; Actual: no reward, plus time and effort wasted by digging in the bowl.

therefore less able to "assign credit" to the particular response or strategy that led to reward (Walton *et al.*, 2011), and are more likely to respond to cues which are irrelevant to the present discrimination (Kim & Ragozzino, 2005; Walton *et al.*, 2011).

During the standard 7-stage task, rats with lesions of the OFC may be impaired on the reversal stages (*e.g.* McAlonan & Brown, 2003) due to difficulties updating the reward prediction errors established in the initial discriminations, or perhaps because they respond to information that is not relevant – *e.g.* the bowl's spatial location – to the task due to difficulties integrating reward history with choice history. An impaired ability to focus on the relevancy of cues, particularly evident during reversal learning, would explain why OFC damage – whether through lesions in young rats (McAlonan & Brown, 2003), or through age-related dysfunction (Schoenbaum *et al.*, 2006) – appears to disrupt the formation of attentional set.

The standard 7-stage task does not permit the direct measurement of attentional set-formation; it is only capable of detecting set-shifting, from which the prior formation of set can be assumed. Additionally, the presence of reversal stages in this task may interfere with the detection of attentional set, if not with set-formation itself. To measure the formation of attentional set in rats with OFC lesions, it would be necessary to modify the standard set-shifting task to remove the possibly confounding effect of the reversal stages, and instead replace them with multiple intradimensional acquisitions. The present experiment used a task such as this, similar to one previously used in marmosets (Clarke *et al.*, 2005). Four intradimensional stages preceded the extradimensional shift stage, and performance over the course of this series provided an index of set-formation. Rats were also tested on the standard 7-stage task to ensure replication of past results (McAlonan & Brown, 2003), thereby validating the placement of the OFC lesions.

Methods

Animals

Sixteen male, Lister Hooded rats were pair-housed in smaller cages in standard housing conditions. At least seven days before surgery, rats were placed on a moderately restricted diet (15–20 g of standard lab diet per rat per day) with water always freely available. Before surgery the weight range was 327–407 g, before testing the range was 363–463 g, and at the completion of testing the range was 409–546 g. Testing commenced when the rats were approximately eight months of age and was completed over twelve weeks.

Surgery

Rats were anaesthetised by isoflurane and oxygen mix (4% induction, 1.8–2% maintenance), and given a subcutaneous injection of 0.05 ml carprofen (Rimadyl: Pfizer, Kent, UK) to limit pain during recovery; and an intraperitoneal injection of 0.25 ml diazepam (Hameln Pharmaceuticals, Gloucester, UK) to reduce both stress and the risk of post-operative seizure.

Rats were secured in a stereotaxic frame (Kopf, CA, USA) with atraumatic ear bars, and the nose bar was set to –3.3 mm to achieve a level skull. Eight rats then received 0.25 µl of 0.06 M ibotenic acid in each hemisphere from a 30-gauge Hamilton bevelled-tip syringe. A further eight rats were given similar injections with sterile phosphate buffer instead of ibotenic acid. Injections were given bilaterally to form lesions in the OFC. Injection sites were calculated from Paxinos and Watson (Paxinos & Watson, 1998), and, with respect to Bregma, were AP +4.0 mm, ML ±2.0 mm and DV –4.5 mm (from skull surface). The injections were performed manually, taking approximately three minutes each, and after injection the needle was left *in situ* for a further three minutes to allow diffusion of the ibo-

tenic acid. Upon completion of surgery, wounds were sealed with surgical staples.

Rats were single-housed for the 24 hours after surgery. Behavioural and physiological evidence suggested that all rats recovered well, with normal eating and pre-surgery weights returning within 48 hours. Testing began no fewer than twelve days after surgery.

Behavioural Testing

Rats were habituated, pre-exposed and trained following the general protocol in chapter two. The stimuli and order of stages for the two attentional set-shifting tasks used in this chapter are summarised in Figure 4.1.

The 4ID Task

Rats were first presented with a simple discrimination (SD) between either two media with no added scent, or between two odours mixed in sawdust. A second, irrelevant dimension was then added to form a compound discrimination (CD) stage, but the contingencies of the discrimination remained the same as the SD. The next four stages of the task were intradimensional acquisitions (ID1, ID2, ID3 and ID4), where different compound stimuli were presented at each stage, with the relevant dimension remaining constant. Therefore, on the first six stages of the 4ID task rats were required to find rewards by attending to only one stimulus dimension while ignoring the other. This should encourage the formation of attentional set despite the lack of reversal stages, and performance improvements over the course of the four ID stages provide a direct measurement of set-formation.

After the fourth ID, rats were presented with the extradimensional shift (ED), where different compound stimuli were presented, but rats must refocus their attention to the previously irrelevant dimension to solve the

	Discriminanda	*Mixed with...*
Simple Discrimination	Coarse tea, not fine tea	Nothing
Compound Discrimination	Coarse tea, not fine tea	Cinnamon or ginger
First Intradimensional Stage	Sand, not grit	Sage or paprika
Second Intradimensional Stage	Wood chips, not sawdust	Turmeric or clove
Third Intradimensional Stage	Yarn, not cigarette filters	Ylang-ylang or frankincense
Fourth Intradimensional Stage	Coarse cork, not fine cork	Patchouli or lavender
Extradimensional Stage	Bergamot, not rosemary	Long tubes or short tubes
Original Simple Discrimination	Coarse tea, not fine tea	Nothing

Figure 4.1: The stages and stimuli of the 4ID task. A description of the standard 7-stage task can be found in Figure 3.1.

discrimination. The "shift-cost" is calculated by taking the difference in trials required to learn the ED – the stage that requires an attentional shift – from the trials required to learn the final ID, the stage that should require no shifting. Shift-costs therefore provide an index of set-shifting ability that is independent from general learning speed. Finally, the original SD is presented again to control against the possibility that any increase in trials associated with learning the ED were due to issues of fatigue, satiety or memory load, rather than the cost of shifting set.

The Standard 7-Stage Task

The standard task is described extensively elsewhere (Birrell & Brown, 2000; chapters two and three). Briefly, this task begins with a SD and CD, as in the 4ID task. Then, in the first reversal stage (Rev1), the discrimination contingency of the CD is switched such that the previously incorrect stimulus becomes correct and vice versa. There then follows an ID stage as described previously, followed by a second reversal (Rev2). Rats then perform an ED shift stage as above, and finally there is a third reversal (Rev3).

Repeat Testing

Testing followed a within-subjects design (Tait *et al.*, 2009), with all rats performing each task twice, alternating between tests: first they performed the 4ID task, then the standard 7-stage task, and then this two-task sequence was repeated. Typically, when retesting rats, we have found that it is not necessary to repeat the pre-exposure phase outlined above. However, unlike the 7-stage task, some stimuli used in the 4ID task will never be associated with reward, as this task does not contain reversal stages. To control against the possible affect these stimulus-reward associations may have on subsequent task performance, rats were "re-exposed" to the stimuli by following the pre-exposure protocol in between each test.

Histology

All rats were deeply anaesthetised with 0.8 ml Dolethal (intraperitoneal; Univet, Bicester, UK), and then intracardially perfused with 4% paraformaldehyde in 0.1 M phosphate buffer. Brains were refrigerated overnight at 4°C in 20% sucrose solution, then washed with distilled water, dried, and placed in individual wells. Brains were then covered in egg yolk, and left to set in a 40% formaldehyde bath for 72 hours. Set brains were

then cut into 50 μm sections with a freezing microtome (Jung Histoslide 2000, Reichert-Jung, Cambridge Instruments GmbH), which were put into 0.1 M phosphate buffered saline (0.9%). Sections were then double-stained for neuronal nuclei (NeuN) and with cresyl violet to highlight cells in the sections, which were mounted on gel-coated glass slides.

The extent of lesioning of the OFC was determined by light microscopy at 10X and 40X magnifications. Lesions schematics were traced in ImageJ (Abràmoff *et al.*, 2004) to estimate total lesion volume.

Analyses

Trials to criterion data for the 4ID and standard 7-stage tasks were analysed separately using two-way analysis of variance (ANOVA) tests with Stage (8 levels for the 4ID task, and 7 levels for the 7-stage task) and Test (two levels for each task) as repeated-measures factors, and Lesion as a grouping factor. Planned comparisons were conducted as outlined in chapter two.

Results

Histology

All eight rats receiving excitotoxic surgery had lesions centred on the lateral OFC, with damage also being seen in the ventral, medial and dorsolateral OFC; the motor cortex, medial prefrontal cortex and agranular insula (Figure 4.2), which is consistent with previous reports (McAlonan & Brown, 2003; Tait & Brown, 2007). In two rats, the lesions were asymmetrical, with both rats having greater damage on the left hemisphere. However, the behavioural data from these rats were not obviously different from that of the rest of the group. Moreover, when either of these rats is excluded from the statistical analyses, the overall pattern of results and the statistical significance of the effects did not change. Additionally, the deficits we observed did not correlate with overall lesion size (greatest Pearson's $r_8 = 0.63$, $p = 0.097$). Therefore, all rats were included in the analyses presented.

Forming and shifting attentional set is facilitated by the OFC

Rats with lesions of the OFC were impaired relative to controls at several stages of both the 4ID task (Figure 4.3; omnibus Stage by Group interaction after Huynh-Feldt correction: $F_{7,98} = 3.11$, $p < 0.05$) and the 7-stage task (Figure 4.4; omnibus Stage by Group interaction after Huynh-Feldt correction: $F_{6,84} = 6.35$, $p < 0.05$). On the 4ID task, lesioned rats were slower to form attentional set than controls, requiring more trials to complete the second and third IDs (main effects of Group, restricted ANOVA for ID2: corrected-$F_{7,98} = 5.82$, $p < 0.05$; and for ID3: corrected-$F_{7,98} = 2.88$, $p < 0.05$). There was no effect of the lesion on performance of ID4 (main effect of Group: corrected-$F < 1$), indicating that the lesioned rats could acquire a novel ID as rapidly as control rats, once attentional set had been formed.

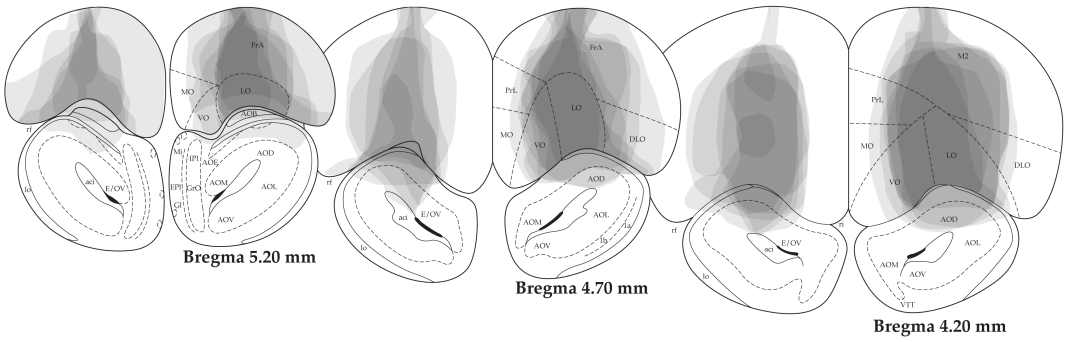


Figure 5.2: Histological assessment of surgical lesions, plotted on schematics of the rat prefrontal cortex using 10% transparency, so that areas of common damage are progressively darker. The overlapping illustrations suggest that the area most consistently damaged was the lateral OFC, with damage also typically seen in the dorsolateral, ventrolateral and medial orbital cortex, as well as surrounding areas. This pattern is consistent with our previous reports (McAlonan & Brown, 2003; Tait & Brown, 2007).

Both lesioned and control rats demonstrated a significant difference between the ID4 and ED shift stages, confirming a shift-cost in both groups on the 4ID task (main effect of Stage: corrected-$F_{7,98}$ = 14.05, $p < 0.05$). This confirms that improvement over the four ID stages reflects the formation of attentional set, and not merely a general discrimination learning practice effect. Rats with lesions of the OFC required more trials to complete the ED stage than control rats, as revealed by ANOVA restricted to only the ED stage (main effect of Group: corrected-$F_{7,98}$ = 4.29, $p < 0.05$). This increase at the ED without a similar increase in the ID reveals that the OFC-lesioned animals demonstrated impaired shift-costs relative to controls (Figure 4.5). This shift-cost impairment did not correlate with damage to the medial prefrontal cortex (in tests 1 and 2, respectively: Pearson's r_8 = 0.36, p = 0.38; and –0.5, p = 0.2).

On the 7-stage task, the lesioned rats required more trials to reach criterion at the ID stage than controls (main effect of Group, corrected-$F_{6,84}$ = 5.56, $p < 0.05$), and showed no cost of shifting. Figure 4.5 shows the shift-costs for each group on the two tasks. Clearly, while control rats

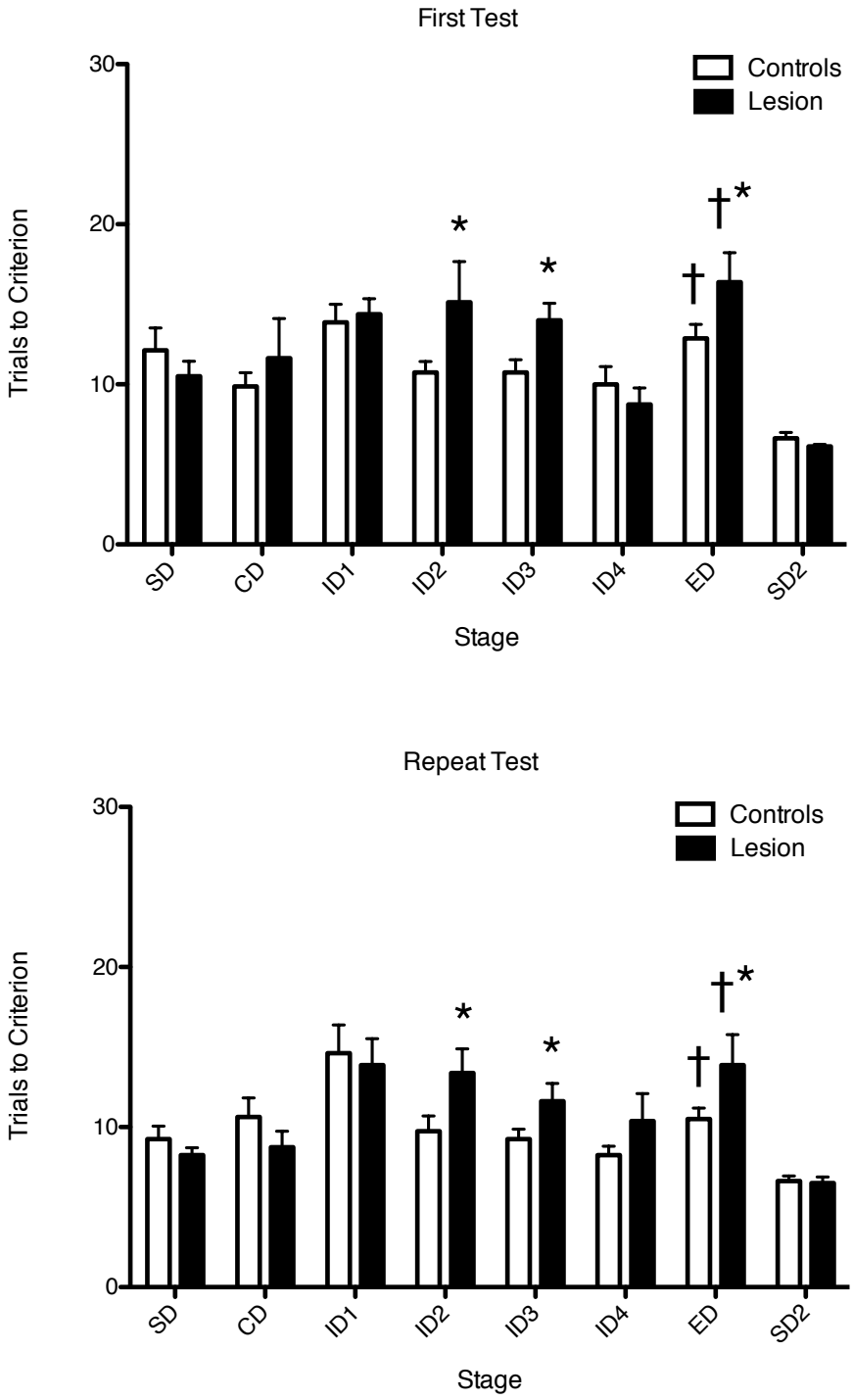


Figure 5.3: Mean trials to criterion (+SEM) on the 4ID task. Rats with lesions of the OFC were significantly slower (*, $p < 0.05$) to learn the second and third intradimensional stages (IDs), indicating an impaired ability to form attentional set. Both groups required more trials to learn the extradimensional shift stage (ED) than the fourth ID (†, $p < 0.05$), and lesioned rats required more trials to shift set than controls. These impairments are robust between tests one and two (top and bottom, respectively).

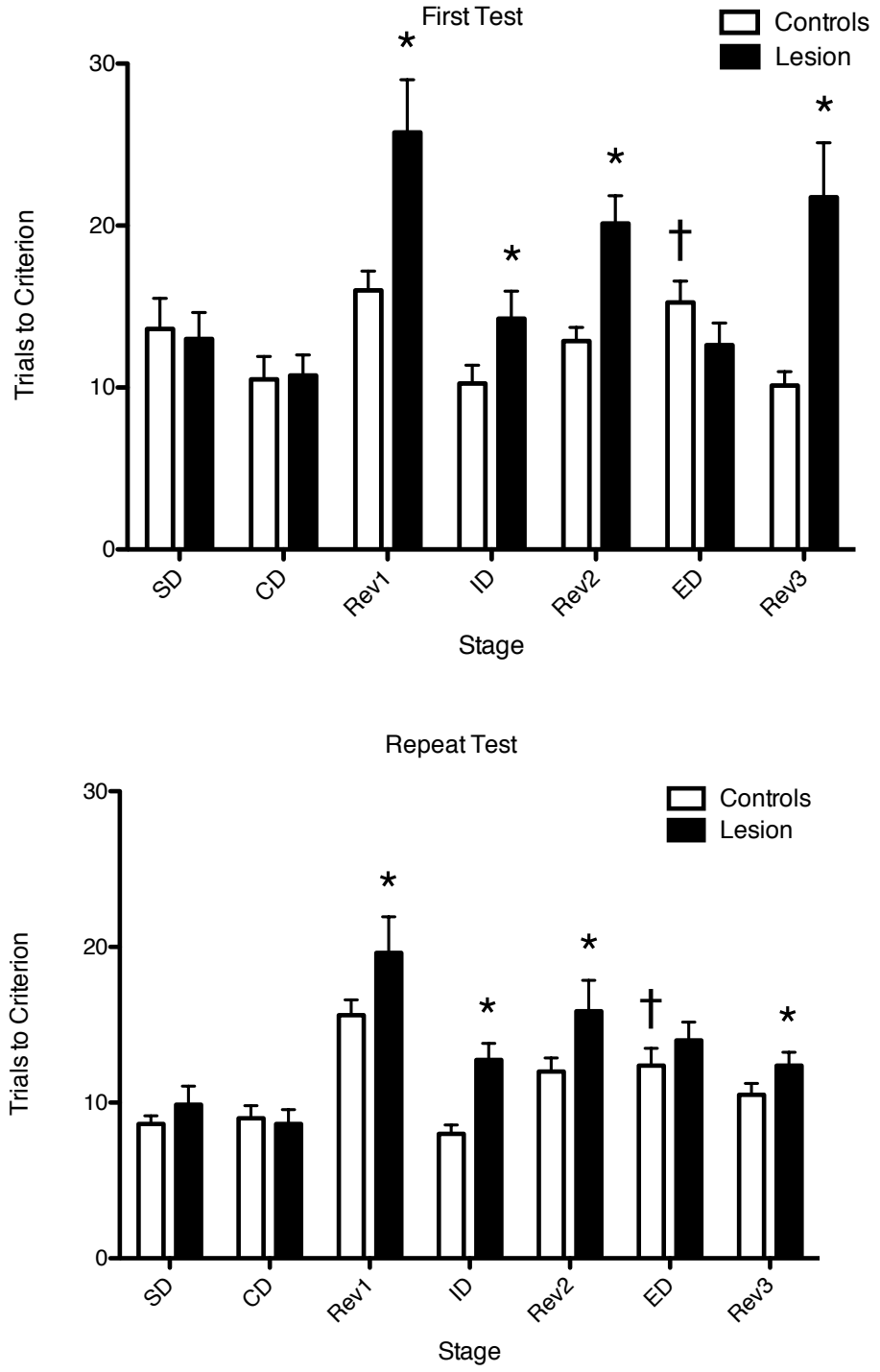


Figure 5.4: Mean trials to criterion (+SEM) on the standard 7-stage task. Rats with lesions of the OFC required significantly more trials (*, $p < 0.05$) to learn all three reversal stages (Rev), as well as the intradimensional acquisition stage (ID). A significant difference between the intra- and extradimensional stages (ED) is only seen in the controls (†, $p < 0.05$). The magnitude of the reversal impairment appears larger in test one than test two (top and bottom, respectively), but it remains statistically significant.

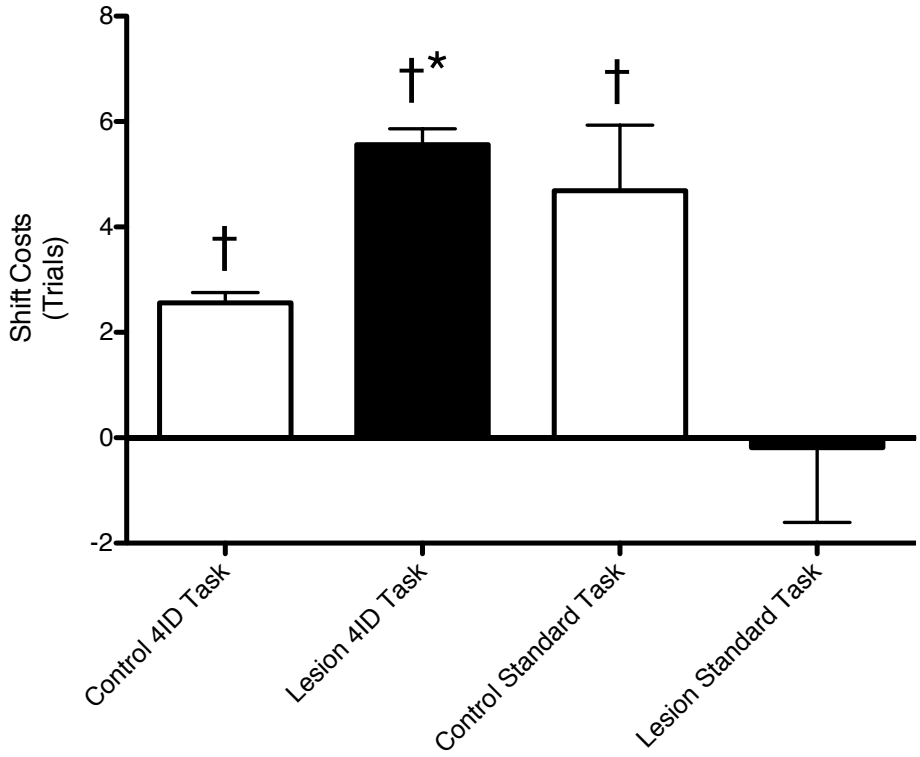


Figure 5.5: Mean shift-costs (+SEM) across the two tests in each task. These are calculated by taking the difference in trials required to learn the ED and the immediately preceding ID. Rats with OFC lesions and control rats both show positive shift-costs on the 4ID task, but on the standard 7-stage task this is only true of controls (†, significantly different from preceding ID; *, significantly higher ED than controls. Both $p < 0.05$). Lesioned rats on the standard task have diminished or absent shift-costs, suggesting that they do not form attentional set on this task. This pattern is similar on tests one and two, and so a mean of both tests is presented for clarity.

demonstrate positive shift-costs on both tasks, rats with lesions of the OFC show positive – and impaired – shift-costs on the 4ID task, but absent shift-costs on the 7-stage task.

Reversal learning is impaired after excitotoxic lesions to the OFC

Rats with lesions of the OFC perform worse than controls on all three reversal stages of the 7-stage task, confirming previous observations (McAlonan & Brown, 2003; ANOVA restricted to each reversal stage: main effects of Group, respectively: corrected-$F_{6,84}$ = 13.73; corrected-$F_{6,84}$

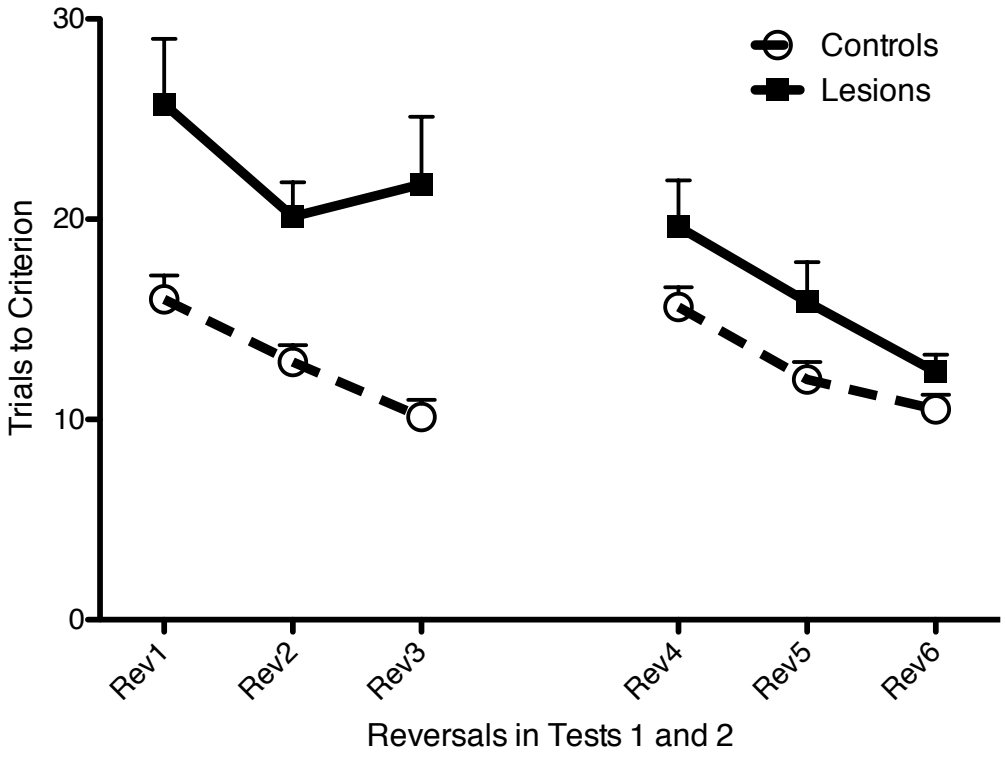


Figure 3.6: Mean trials to criterion (+ SEM) on the three reversal stages (Rev) in both tests of the 7-stage task. In the first test, rats with lesions of the OFC do not show a clear improvement over the course of the three reversals; but on the second test (the fourth, fifth and sixth reversal stages they experience) they show a pattern more similar to the control rats. However, the Stage by Test by Group interaction only approaches significance ($p = 0.057$), while the main effect of group remains significant ($p < 0.05$).

= 8.99; corrected-$F_{6,84}$ = 13.23; all $p < 0.05$). Analyses of the other stages revealed that the groups only differed on the reversal and ID stages (all other corrected-$F < 1$).

Due to the relatively low error-rate across the groups in the experiment, it is difficult to detect patterns of errors on the reversal stages. However, on the first reversal stage the rats experience (Rev1 of the first test), the lesion and control groups take a comparable number of trials to make their first correct response, with control rats requiring a mean of 1.43 trials, and lesioned rats a mean of 2.57 trials ($t_{12} = 1.14$, $p = 0.28$). In the first six trials of this stage, control and lesion rats typically make a similar num-

ber of correct responses with means of 2.29 and 3 respectively ($t_{12} = 0.92$, $p = 0.38$). The number of errors committed before the first correct response relative to total errors also does not differ significantly between groups, with a mean of 27.1% for the control rats, and 26.5% for the lesioned rats ($t_{12} = 0.041$, $p = 0.98$). It seems unlikely, therefore, that the impaired reversal learning seen in the OFC-lesioned animals is due to increased perseveration.

Effects of repeated testing

Rats in both groups generally perform the second 4ID test better than they do the first (main effect of test: $F_{1,14} = 8.76$, $p < 0.05$), however the pattern of performance and, importantly, the impairments, in rats with OFC lesions were unchanged (all interactions $F < 1$).

Performance on the 7-stage task was also improved overall in the second test (omnibus main effect of test $F_{1,14} = 20.15$, $p < 0.05$). The pattern of impairments across the two tests was the same, with the OFC-lesioned rats impaired at all three reversal stages and the ID stage. However, as can be seen in Figure 4.6, there appears to be a tendency for the OFC-lesioned animals to show a diminishing impairment over the course of the six total reversal stages. It was not appropriate to analyse this effect further as the Stage by Test by Group interaction in the omnibus ANOVA was not significant after applying the required Huynh-Feldt correction ($F_{6,84} = 2.53$, $p = 0.057$). Nevertheless, there is clearly an approaching-significance trend.

Discussion

Attentional set-shifting is measured by comparing two acquisitions: one where the reward-predicting aspect of multidimensional stimuli is consistent with previous stages, an "intradimensional" acquisition, and one where the relevant dimension is not consistent, an "extradimensional" acquisition, therefore requiring a shift in attention. While theory predicts that learning an extradimensional shift stage will require more trials than learning an intradimensional stage, this is not because the two stages differ in their inherent difficulty. Rather, such attenuated performance only occurs if the subject had a predisposition to attend to one stimulus dimension over another, or – in other words – had formed an attentional set. After all, a set only requires shifting when it has first been formed, and so these two stages would be formally equivalent from the perspective of a subject that had not formed an attentional set.

Previous experiments using the standard 7-stage task have suggested that when a manipulation impairs the first reversal, the cost of shifting an attentional set diminishes and frequently disappears altogether (McAlonan & Brown, 2003; Tait & Brown, 2008). This "absence of evidence" – the lack of a significant group difference between the trials needed to learn the intra- and extradimensional shift stages – prevents us from commenting on the set-shifting abilities of animals with reversal impairments: when there is no evidence of set-formation in one of the groups, that group's ability to shift set is unknown. A primary goal of the present experiment was to replicate the surgical and behavioural results of McAlonan & Brown (2003): like our previous findings, rats with lesions of the OFC were impaired on all three reversal stages of the 7-stage task, and showed no significant shift-cost. This experiment extends these past findings by showing that learning on the intradimensional stage was significantly impaired in OFC-lesioned rats.

Another major goal of this study was to address the possibility that the lack of difference between the intra- and extradimensional stages associated with impaired reversal learning represented a failure to form attentional set. As it seemed possible that the impairment at the first reversal might causally interfere with the detection of a shift-cost, all rats were tested using the 4ID task before experiencing a reversal stage. By replacing the reversal component of the set-shifting task with multiple intradimensional acquisitions, we found that OFC-lesioned rats could form attentional set, albeit less readily than controls. This reveals that the set-formation impairment is not caused by the impaired first reversal, but rather it is an impairment in its own right. When OFC-lesioned rats did form an attentional set, they were found to be impaired at the extradimensional shift stage relative to controls. This supports the hypothesis that the lack of a significant difference between the intra- and extradimensional stages seen in OFC-lesioned rats on the standard 7-stage set-shifting task indicates a failure to form attentional set, as the increased shift-cost would otherwise have been seen on both tasks.

The question then arises, why are rats with lesions of the OFC able – albeit more slowly – to form attentional set in the 4ID task, but unable to form a set in the standard task? To address this, it is first necessary to consider the nature of the reversal impairments we have observed. It has long been theorised that reversal learning occurs in phases (Jones & Mishkin, 1972): first the original discrimination is "unlearned", then the animal responds at chance accuracy, and finally the new association is acquired. Impaired performance on a reversal stage could therefore stem from increases in errors at one or more of these phases: due to perseveration (Dias *et al.*, 1997) and/or learned non-reward (Tait & Brown, 2007) during the first phase; then due to distraction by irrelevant cues (Sutherland & Mackintosh, 1971; Kim & Ragozzino, 2005) in the second phase; and finally due to learning errors (Clarke *et al.*, 2005) as the animal acquires the new rule.

Because rats make so few errors on bowl-digging tasks, analysis of error type is difficult. Nevertheless, we have observed no evidence in this or other experiments that the OFC-lesioned rats make more early-phase or perseverative errors than controls. It also seems unlikely that the reversal impairment stems from an increase in late-phase or learning errors, as generally poor learning would manifest in the SD and CD acquisitions as well as the reversals.

The remaining possibility is that the reversal impairments we have observed stem from an increase in middle-phase reversal learning, where animals are prone to shift their attention to irrelevant information spontaneously (Sutherland & Mackintosh, 1971). This hypothesis is consistent with the evidence that reversal impairments in OFC-lesioned rats are characterised by responses to cues that do not predict reward (Ghods-Sharifi *et al.*, 2008), and the impairments are exacerbated when the number of irrelevant cues is increased (Kim & Ragozzino, 2005). As to why OFC-lesioned rats were unable to form set on the standard 7-stage task, the reversal stages may not provide sufficient exposure for the OFC-lesioned rats to recognise that the relevancy of the dimensions is constant before the critical intra- and extradimensional stages. In other words, each time the contingencies of a discrimination are reversed in the standard task, the OFC-lesioned rats may temporarily lose sight of which cues are relevant, not only slowing their learning, but preventing them from forming an attentional set on this task.

The presence of deficits in attentional set-formation and set-shifting following OFC lesions is superficially inconsistent with previous reports (Dias *et al.*, 1996b; Dias *et al.*, 1997; Brown & Bowman, 2002; Bissonette *et al.,* 2008). One might suggest that our lesions produced sufficient damage to the medial prefrontal cortex to account for this set-shifting deficit. However, this seems unlikely, as those animals with greater medial damage were not more impaired at the extradimensional shift stage, nor were the

animals with the least damage the least impaired. Furthermore, we have previously found that medial lesions of the size seen here are insufficient to elicit a set-shifting deficit (DS Tait, unpublished observations). It may be more likely that the apparent disparity between past and present behavioural results is due to anatomical and/or task differences between rodents and monkeys. A thorough discussion of these possibilities can be found in Chase *et al.* (2012).

It is parsimonious to assume that the impairments produced by OFC lesions in reversal learning, attentional set-formation, attentional set-shifting, and perhaps strategy-set formation (*ibid.*) are all due to a single cognitive deficit that is manifesting in multiple ways. Our results conform to the recent hypotheses of OFC function. This region contributes to learning by signalling cue relevancy (Kim & Ragozzino, 2005; Walton *et al.*, 2011), and facilitates the accurate prediction of reward when established outcome expectancies change (Takahashi *et al.*, 2009; Roesch *et al.*, 2010; Kahnt *et al.*, 2011). The importance of this type of learning to discrimination reversal has been discussed clearly in the literature (e.g. Schoenbaum *et al.*, 2009). Here we have shown that this type of learning is also important during set-formation and set-shifting, which both require accurate and flexible representations of cue relevancy, as well as flexible integrations of past choices with their reward outcomes.

Chapter 5

The Complex Relationship Between Deficits in Reversal Learning and Attentional Set-Formation

Given that reversal learning and attentional set-formation impairments are both produced by OFC lesions, it seems possible that a single cognitive deficit in recognising or maintaining cue relevancy may be manifesting as two separate behavioural impairments. This possibility would be highly consequential for drug development, as it suggests a common behavioural impairment – reversal learning – is more complex than is often considered. To further probe the relationship between reversal learning and set-formation, middle-aged rats were tested on one of two set-shifting tasks each with one reversal stage. When the reversal is the third stage of the task, middle-aged rats show impaired reversal learning and set-formation. When the reversal is later in the task, after a series of four ID stages, the middle-aged rats show no significant deficits. This suggests that middle-aged rats are only impaired at reversal learning if they have not first formed attentional set.

Introduction

The results from chapter four demonstrated that lesions of the orbital prefrontal cortex (OFC) impaired not only reversal learning – which has been seen in several prior investigations (Dias *et al.*, 1997; McAlonan & Brown, 2003; Schoenbaum *et al.*, 2003) – but also the formation of attentional set. This finding integrates well with recent evidence for the involvement of the OFC in identifying the cues in the environment that predict reward, as well as updating predictions about the outcomes of behavioural responses (discussed more fully in chapter four). Inferring from these results, it seems possible that the OFC lesion produces a single deficit in recognising relevancy, and this manifests in multiple ways: because rats with OFC lesions are less able to focus on relevant cues, reversal learning and set-formation[7] are impaired.

The importance of relevancy to normal reversal learning has long been established (*e.g.* Reid, 1953). However, relatively few experiments have investigated the contribution of relevancy to *impaired* reversal learning, and no one has investigated the effects of OFC lesions on the overtraining reversal effect (*ibid.*; chapter one). Although it has been repeatedly argued that reversal learning takes place in phases, each characterised by a different type of learning (Sutherland & Mackintosh, 1971; Jones & Mishkin, 1972), it is often difficult to uncover which of these phases elongates to produce an impaired reversal in the rat set-shifting task. Frequently, the explanations for reversal impairments in the literature are based around increases in "perseverative errors": that is, nearly habitual responding to the previously reinforced stimulus. There is, however, no standardised methodology for categorising perseverative and non-perseverative errors. Perseverative errors on a reversal stage can be defined

[7]Including the formation of a replacement set that occurs during set-shifting, which may explain the impaired ED stage (see Chase *et al.*, 2012)

as every error (McAlonan & Brown, 2003), or only the errors before the first correct response (Iversen & Mishkin, 1970; Izquierdo *et al.*, 2009), or the errors that occur before the subject attains chance performance (Jones & Mishkin, 1972; Dias *et al.*, 1997; Schoenbaum *et al.*, 2003; Kim & Ragozzino, 2005; Brigman *et al.*, 2010; Judge *et al.*, 2011). The latter definition is the most popular in the literature, but there is little agreement regarding how to apply this definition to the data themselves.

Although many groups have reported increases in perseveration during reversal learning in animals with OFC lesions (*e.g.* Dias *et al.*, 1997; Meunier *et al.*, 1997), there is evidence to suggest that rats with orbital lesions are impaired at reversal learning for reasons apart from – or at least in addition to – increases in perseveration. Similar to the results of chapter four, Kim and Ragozzino (2005) showed that OFC-lesioned rats were impaired relative to controls on a two-choice odour discrimination reversal, but the magnitude of this deficit increased in a four-choice reversal. The authors interpreted this impairment as stemming from the increase in the number of cues that might predict reward, which confounds the rats with OFC lesions more than controls. This finding is complemented by the work of Tait and Brown (2007) showing that OFC-lesioned rats are influenced more by "learned non-reward" than perseveration – that is to say, difficulty responding to the previously unrewarded cue, rather than difficulty ignoring that which was previously rewarded – when these two components of reversal learning are separated in different behavioural tasks.

These results hint at a possible heterogeneity in the cognitive mechanisms underlying reversal learning deficits, which has been shown directly in relatively few experiments (*e.g.* Mishkin, 1964; Meunier *et al.*, 1997). This has serious potential implications for psychopharmacology. Attentional set-shifting and reversal learning tasks have been used extensively in the quantification of cognitive dysfunction in humans, monkeys,

rats, mice and other animals (Elliott *et al.*, 1995; Dias *et al.*, 1997; Birrell & Brown, 2000; Bissonette *et al.*, 2008; Parker *et al.*, 2012). Recently, the use of these tasks (and tasks which combine elements of both set-shifting and reversal learning) to help identify putative cognitive enhancers has been increasing (Goetghebeur & Dias, 2009; Tait *et al.*, 2009; Gastambide *et al.*, 2011), as using the same task in different species has high face validity. After all, a drug that improves a reversal learning deficit in a rat is also likely to improve the same deficit in humans – but only if the underlying cause of the deficit in both species is the same. If reversal stages can be impaired for a multiple reasons – for example increased perseverative errors in one population, but increased irrelevant errors in another – the precise nature, and therefore the very 'translatability', of these deficits may need to be re-examined. Put another way, one might argue that, as both are characterised by impaired performance on reversal stages (Lange *et al.*, 1995; McAlonan & Brown, 2003), the OFC-lesioned rat could be considered a model of Huntington's disease in humans. However, if the underlying cause of each impaired reversal is different, as would almost certainly be the case in this example, the overtly similar behaviour may be misleading. Successful development of treatments for the cognitive symptoms of such heterogenous disorders as schizophrenia, major depressive disorder, Huntington's disease, *etc.*, will require a precise knowledge of the causes – psychological and physiological – of the symptoms.

It is not just in rat-to-human translations where overt, yet potentially misleading, behavioural deficits occur, however. The middle-aged rats in chapter three were impaired at all three reversal stages of the standard 7-stage task, and demonstrated no significant difference between the number of trials needed to learn the intra- and extradimensional stages. In chapter four, this same pattern of behaviour in young, OFC-lesioned rats coincided with a reduced ability to form attentional set on a set-shifting task without reversal learning stages. As presented in chapter four, some have argued

that the cognitive dysfunction seen in middle-aged rats is due to changes to the OFC during senescence (*e.g.* Schoenbaum *et al.*, 2006). However, age-related dysfunction has also been reported in the striatum (Nicolle & Baxter, 2003; Cruz-Muros *et al.*, 2009), which could explain the reversal learning impairments seen in aged rats (*c.f.* O'Neill & Brown, 2007).

The primary purpose of this experiment was to further elucidate the relationship between reversal learning impairments, such as those we have observed in older rats and young rats with OFC lesions, and attentional set-formation. Sutherland and Mackintosh (1972) explained the overtraining reversal effect as occurring due to an improved representation of what was and was not relevant to the initial discrimination, which facilitated that discrimination's reversal. Extending this hypothesis, a reversal stage that is performed after attentional set has been formed ought to be acquired in fewer trials than a reversal stage occurring before the formation of attentional set. To investigate this possibility, two attentional set-shifting tasks were used in this experiment, each derived from the 4ID task (chapter four), but with one reversal stage added. The first included the reversal after the CD stage, and the second included the reversal after a series of four intradimensional acquisitions.

Middle-aged rats, and young rats with excitotoxic or sham lesions of the OFC, were then tested on these tasks to uncover the potential links between reversal learning, set-formation and set-shifting within each group. Differences between the groups provide a more sensitive analysis of the behavioural deficits produced by each manipulation, thereby potentially revealing differences in the nature of the reversal impairments observed in middle-aged rats and young rats with OFC lesions.

Methods

Animals

The middle-aged group comprised of 16 male, Lister Hooded rats obtained from a registered breeder at the age of eleven months. Rats were tested when they were 18 months old. The weight range at the start of testing was 506–577 g.

For the lesion experiment, 40 male, Lister Hooded rats were obtained from the same breeder at approximately four months of age. At the start of surgery the weight range was 330–377 g, and at the start of testing the weight range was 325–403 g.

Surgery

To prepare rats for surgery, they were first anaesthetised by isoflurane and oxygen mix (4% induction, 1.8–2% maintenance), and then given a subcutaneous injection of 0.05 ml carprofen (Rimadyl: Pfizer, Kent, UK) to limit post-operative pain; and an intraperitoneal injection of 0.2 ml diazepam (Hameln Pharmaceuticals, Gloucester, UK) to reduce the risk of seizure.

Rats were secured in a stereotaxic frame (Kopf, CA, USA) with atraumatic ear bars, and the nose bar was set to –3.3 mm to level the skull surface. Twenty rats received 0.2 μl of 0.06 M ibotenic acid in each injection site, and an additional 20 rats were given similar injections with sterile phosphate buffer. Injection sites were calculated from Paxinos and Watson (1998), with anterior/posterior coordinates given with respect to Bregma, and dorsoventral coordinates with respect to the dura mater. The six sites were: AP +4.0 mm, ML ±0.8 mm and DV –3.4 mm; AP +3.7 mm, ML ±2.0 mm and DV –3.6 mm; and AP +3.2 mm, ML ±2.6 mm and DV –4.4 mm. Injections were made via bolus infusions from pulled pipettes with 30–35 μm tips, and after injection the pipette was left *in situ* for three min-

utes to allow diffusion of the ibotenic acid. Wounds were sealed with surgical staples.

Rats were single-housed for at least 24 hours after surgery. Behavioural and physiological evidence suggested that the majority of rats recovered well, with normal eating and pre-surgery weights returning within 48 hours, though four rats were lost from the group. The first was not permitted to recover from the general anaesthesia after a malfunctioning pipette delivered too much toxin. The second was euthanised after having two post-operative seizures. The last two rats recovered from surgery, but were unable to complete the training phase. Training and testing began no fewer than 16 days after surgery.

Behavioural Testing

All rats were habituated, pre-exposed and trained as laid out in the standard protocol in chapter two. Rats were then assigned pseudorandomly to one of two attentional set-shifting tasks.

The two tasks used in this experiment are summarised in Figure 5.1. In both tasks, rats were first presented with a simple discrimination (SD) between either two media with no added odour, or between two odours mixed in sawdust. In the next stage, the complementary irrelevant dimension was added to the stimuli used in the SD to form a compound discrimination (CD) stage, but the outcome of the discrimination remained the same. The following five stages were intradimensional acquisitions (ID1, ID2, ID3, ID4 and ID5), where different compound stimuli were presented at each stage, but the relevant dimension (*i.e.* odour or medium stimuli) remained constant. Each task contained one reversal stage, where the contingencies of the preceding discrimination were swapped such that the previously correct stimulus became incorrect, and vice versa. In the early reversal task, the reversal stage (CDR) came between the CD and

	Discriminanda	*Mixed with…*
Simple Discrimination	Coarse tea, not fine tea	Nothing
Compound Discrimination	Coarse tea, not fine tea	Cinnamon or ginger
Early Reversal (ER task only)	Fine tea, not coarse tea	Cinnamon or ginger
First Intradimensional Stage	Sand, not grit	Sage or paprika
Second Intradimensional Stage	Wood chips, not sawdust	Turmeric or clove
Third Intradimensional Stage	Yarn, not cigarette filters	Ylang-ylang or frankincense
Fourth Intradimensional Stage	Coarse cork, not fine cork	Patchouli or lavender
Late Reversal (LR task only)	Fine cork, not coarse cork	Patchouli or lavender
Fifth Intradimensional Stage	Long tubes, not short tubes	Bergamot or rosemary
Extradimensional Stage	Coffee, not almond	Cloth or sponge
Sixth Intradimensional Stage	Lemon, not strawberry	BBs or tile-spacers

Figure 5.1: The stages and stimuli of the two set-shifting tasks. The two tasks differ in the placement of their one reversal stage: in the early reversal task, the reversal is the third stage of the task; in the late reversal task, the reversal is the seventh stage.

ID1. In the late reversal task, the reversal (IDR) came between ID4 and ID5.

In both tasks, the fifth ID was followed by an extradimensional shift (ED) stage, where different compound stimuli were presented, but the stimuli which predicted reward were from the previously irrelevant dimension. Comparing the number of trials required to learn the ED, the stage that requires an attentional shift, and the trials required to learn the fifth ID, the stage that does not require shifting of attention, provides a measure of set-shifting ability. The final stage presents new compound stimuli, but the relevant dimension from the ED remains relevant to form a sixth ID (ID6).

Histology

At the end of testing, the rats that had received surgery were deeply anaesthetised with 0.8 ml Dolethal (intraperitoneal; Univet, Bicester, UK), and then intracardially perfused with 4% paraformaldehyde in 0.1 M phosphate buffer. Brains were left overnight in 20% sucrose solution and refrigerated at 4°C. The following day, brains were washed with distilled water, dried, placed in individual wells and covered in egg yolk. These were left to set in a 40% formaldehyde bath for 72 hours. Brains were then cut into 50 μm sections with a freezing microtome (Jung Histoslide 2000, Reichert-Jung, Cambridge Instruments GmbH), and then left overnight in 0.1 M phosphate buffered saline (0.9%). Sections were then double-stained for neuronal nuclei (NeuN) and with cresyl violet to highlight cells in the sections, which were mounted on gel-coated glass slides.

Aged rats were terminated using an appropriate Schedule 1 method, and no histology was performed.

Analyses

As the order of stages differs between the two tasks – for example, the first ID stage comes fourth in the early reversal task, but third in the other task – it is not clear if similar stages will be directly comparable

between tasks. Trials to criterion data for the two tasks were therefore analysed separately using two-way analysis of variance (ANOVA) tests with Stage (10 levels) as repeated-measures factors. The middle-aged and young sham-lesioned rats were analysed together with Age as a grouping factor. Planned comparisons were carried out as described in chapter two, with additional comparisons conducted between the groups on the reversal stages, and each intradimensional stage, using one-way ANOVA tests. The F-ratios of these one-way ANOVA tests were corrected using the relevant error term from the omnibus two-way ANOVA tests, as described in chapter two.

Results

Histology

The surgery protocol produced lesions that were too large to be qualified as OFC lesions, with extensive damage being seen in most areas of prefrontal cortex (see Figure 5.2). Data from the lesioned group were therefore not analysed. Additionally, data from one of the controls were also not analysed after it became clear that the behaviour had not been graded correctly. The remaining sham-lesioned rats (n = 8 in the early reversal task, and n = 9 in the late reversal task) formed the control group for the middle-aged rats.

Middle-aged rats were impaired on the early reversal, but not the late reversal

Middle-aged rats were impaired at several stages of the early reversal task compared to young controls (Figure 5.3; omnibus Stage by Age interaction: $F_{9,117} = 2.61$, $p < 0.05$). The significant interaction in this task was further analysed between the two groups, first on the early reversal stage (CDR). A one-way ANOVA between the two groups revealed that middle-aged rats required significantly more trials to learn the early reversal than young controls (corrected-$F_{1,13} = 11.16$, $p < 0.05$). The older and young rats did not differ in the number of errors committed before the first correct dig (mean of 3.5 in both groups), the ratio of these errors to total errors (means 39.1% and 57.4% respectively, $t_{14} = 1.4$, $p = 0.18$), nor in the number of correct digs in the first six trials (means of 4.0 and 4.5 respectively, $t_{14} = 0.46$, $p = 0.76$). There was no significant difference between the performance of middle-aged and young control rats on any stage of the late reversal task (Figure 5.4; omnibus main effect of Age and Stage by Age interactions both $F < 1$).

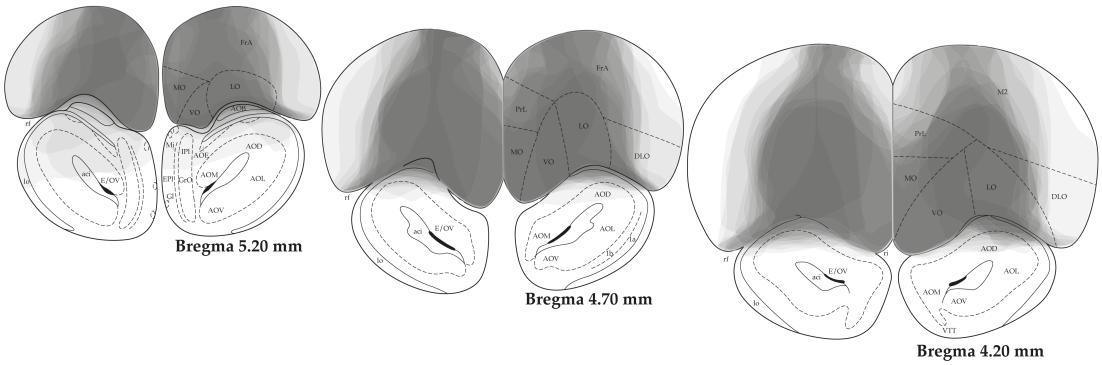


Figure 5.2: Histological assessment of surgical lesions, plotted on schematics of the rat prefrontal cortex using 5% transparency, so that areas of common damage are progressively darker. Extensive damage was seen in almost all areas of prefrontal cortex, thus precluding comment on the involvement of any particular subregion in the performance of this task.

Although not statistically significant, the middle-aged rats appear to take more trials to complete the reversal in the late reversal task than the younger rats (Figure 5.5, top). However, one middle-aged rat took 39 trials to perform the late reversal, which is an outlier in this distribution[8]. Dropping this datum brings the mean difference between young and older rats on the late reversal to just 0.86 trials (Figure 5.5, bottom). Thus, the lack of statistical significance in the omnibus analysis between young and older rats on the late reversal task reflects a true lack of difference in reversal learning between these groups.

Middle aged rats demonstrated impaired attentional set-formation on the early reversal task, but there was no evidence for a similar deficit on the late reversal task

Following the impaired early reversal, middle-aged rats required significantly more trials than young controls to learn the ID stage that followed, as revealed by one-way ANOVA restricted to this stage (Figure 5.6; corrected-$F_{1,13} = 5.57$, $p < 0.05$). Performance did not differ on the other ID

[8]*As defined as any datum greater than or equal to 1.5 times the interquartile range above the third quartile.*

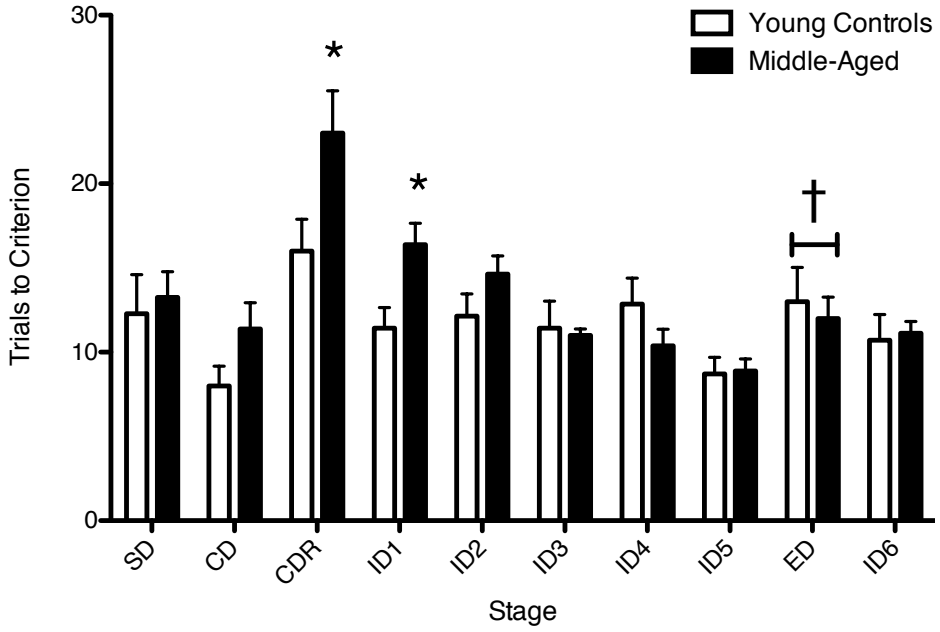


Figure 5.3: Mean trials to criterion (+SEM) for middle-aged rats (n = 8; black bars) and young controls (n = 8; white bars) for the early reversal task. Middle-aged rats were significantly impaired (*, $p < 0.05$) at the early reversal, as well as the intradimensional stage that followed. Both groups required significantly more trials to learn the ED than the preceding ID (†, $p < 0.05$).

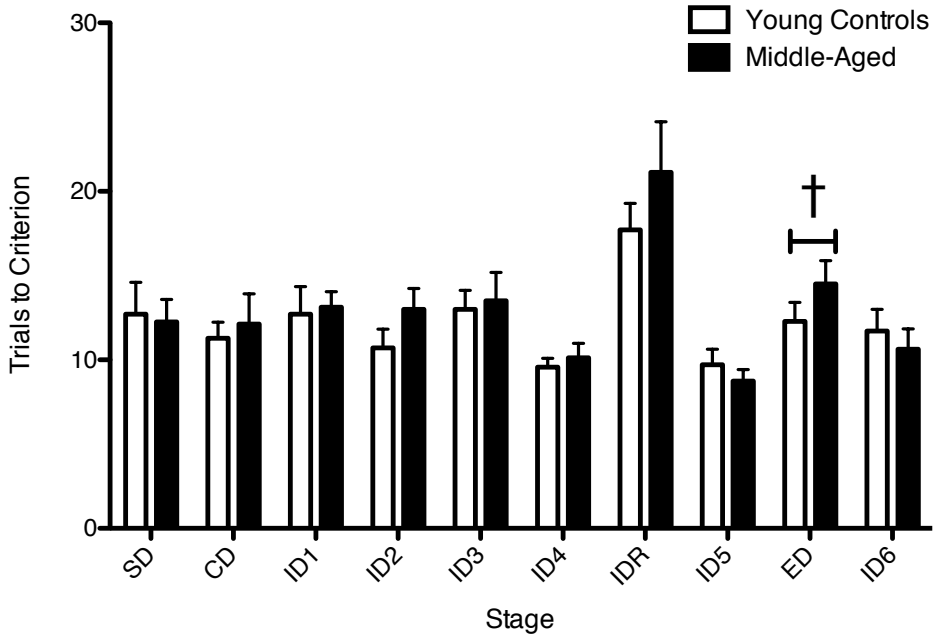


Figure 5.4: Mean trials to criterion (+SEM) for middle-aged rats (n= 8; black bars) and young controls (n = 9; white bars) for the late reversal task. Middle-aged and control rats did not differ significantly on any stage of the task. Both groups required significantly more trials to learn the ED than the preceding ID (†, $p < 0.05$).

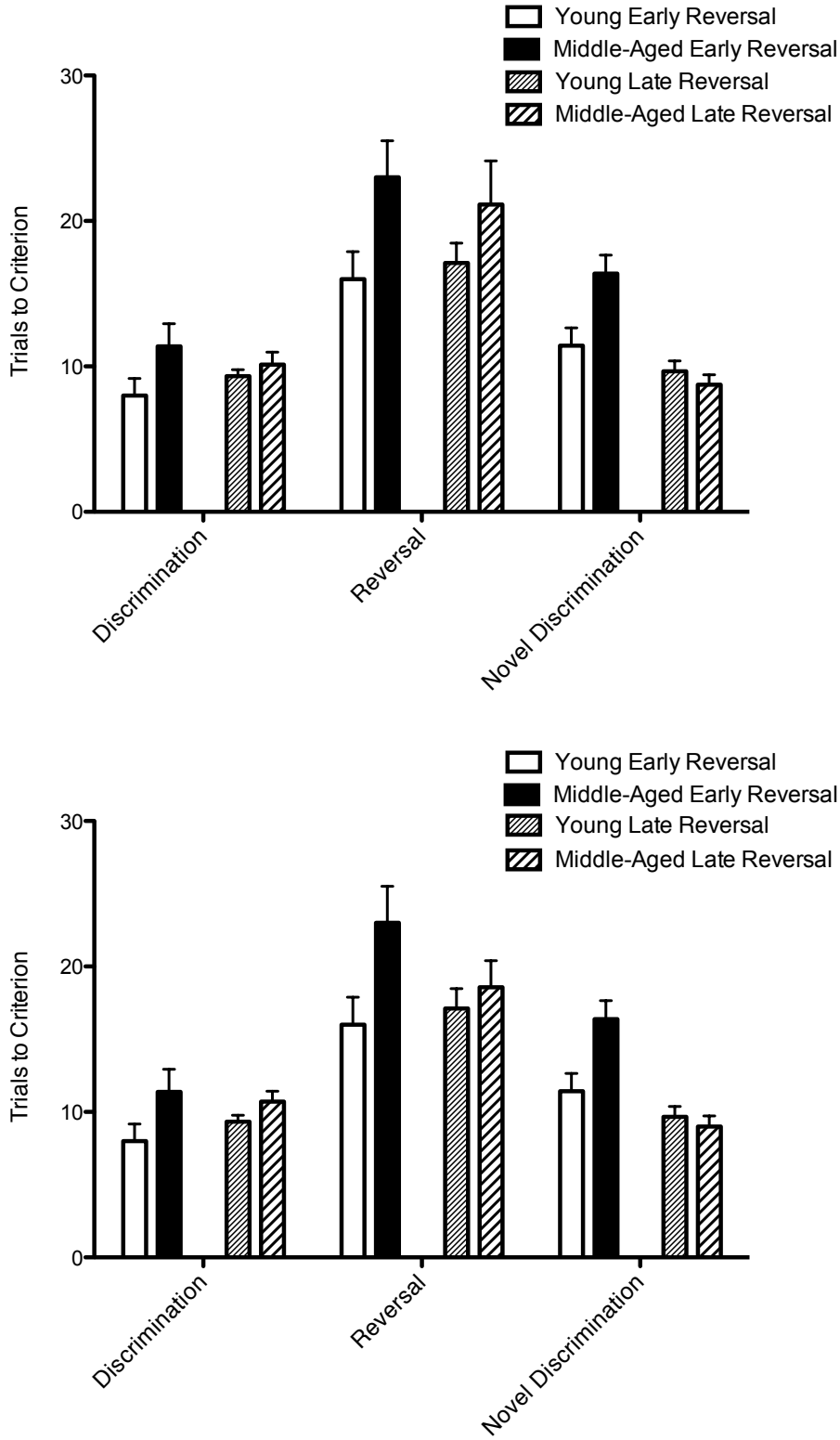


Figure 5.5: To ensure that the lack of difference in reversal learning on the late reversal reflected a true lack of effect, the samples were tested for outliers. Dropping one rat (before: top; after: bottom) from the late reversal task data supports the hypothesis that middle-aged and young rats do not differ on the late reversal.

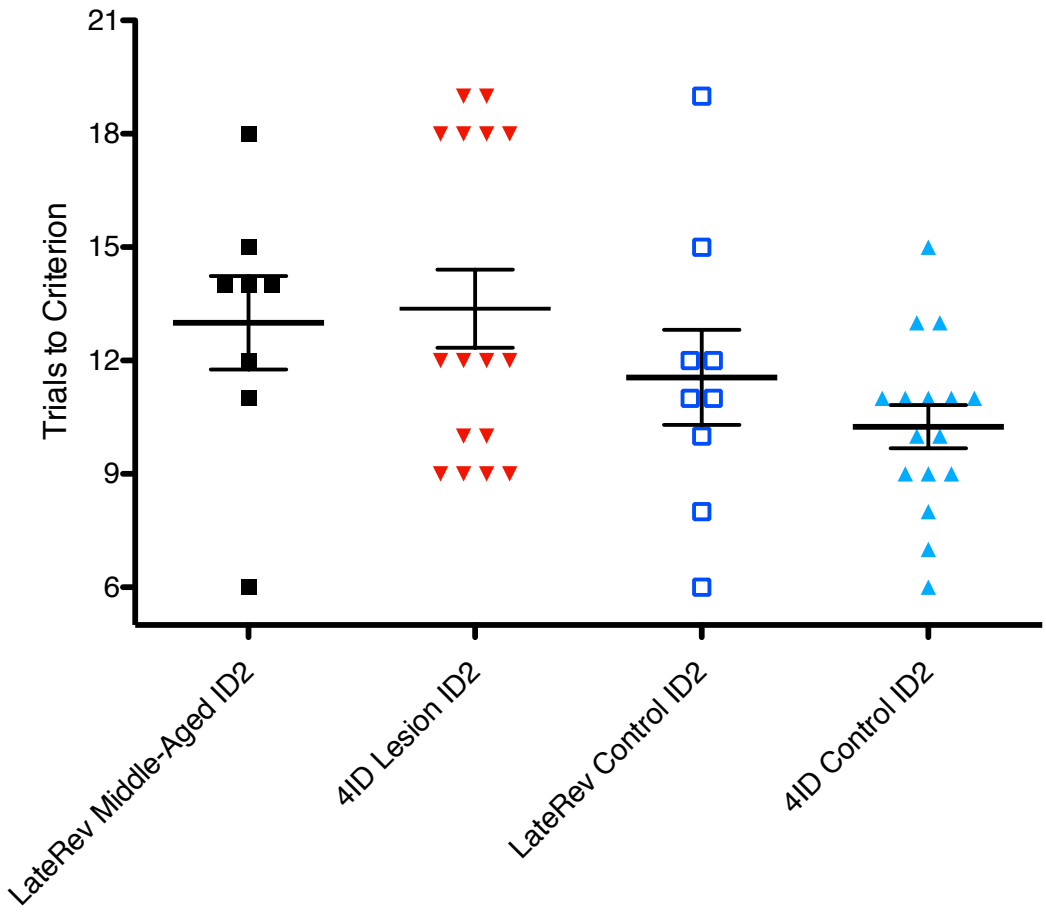


Figure 5.6: Performance on the second ID stage in the late reversal task or the similar 4ID task reported in chapter four. Middle-aged rats (black squares) perform similarly to rats with OFC lesions (red triangles), who were significantly impaired compared to the young controls performing the same task (blue triangles). Young controls performing the late reversal task (blue squares) demonstrated little consistency, and their mean performance appears slightly worse than the 4ID controls.

stages of the early reversal task (largest corrected-$F_{1,13}$ = 1.4, p = 0.28). When the ID stages are not preceded by a reversal stage – the late reversal task – middle-aged rats do not statistically differ from controls; as indicated by the lack of an omnibus interaction reported above.

However, this null result only indicates that the middle-aged rats did not differ from controls, not that they were unimpaired. Comparing the performance on the ID stages of the late reversal task and the 4ID task from chapter four – which are identical until after the fourth intradimen-

sional stage – suggests that this absence of evidence may be due to aberrant performance by the controls in the present experiment. On the 4ID task, OFC lesioned rats required significantly more trials to learn the second ID stage than controls, at means of 14.25 and 10.25 trials to criterion respectively (pooled first and second tests). On the late reversal task, middle-aged rats perform only slightly better than the 4ID lesioned rats on this stage at a mean of 13 trials to criterion, while the late reversal task controls do slightly worse than the 4ID controls at a mean of 11.18 trials to criterion. Analysis of the scatterplots of the performance of these four groups on the second ID stage in each task reveals a high degree of spread on the late reversal task; particularly in the control group (Figure 5.6).

Middle-aged were not impaired at attentional set-shifting

Separate two-way ANOVA tests restricted to the ID5 and ED stages of both tasks revealed that all middle-aged and young groups demonstrated attentional set-shifting (main effects of Stage: corrected-$F_{9,117} = 6.25$, $p < 0.05$; and corrected-$F_{9,117} = 8.43$, $p < 0.05$). The groups did not differ in their abilities to shift set on either task (main effects of Age and Stage by Age interactions all not significant, largest corrected-$F_{9,117} = 1.23$, $p = 0.28$).

Discussion

The overtraining reversal effect is the seemingly paradoxical result that increased training on a discrimination speeds the learning of that discrimination's reversal (Reid, 1953; Lovejoy, 1966; Mackintosh, 1969). The hypothesised mechanism for this effect is that the extra training strengthens the representation of which aspects of the discriminanda predict reward. This enhances reversal learning, because the subject makes fewer errors to irrelevant cues, *e.g.* the position of the stimulus. If reversal learning is improved with a stronger awareness of cue relevancy – as hypothesised by Sutherland and Mackintosh (1971) – then the presence of an attentional set should also improve reversal learning. To test this extension of the overtraining reversal effect, this experiment compared two attentional set-shifting tasks that differed in the placement of their reversal stages. The first task contained an early reversal stage that may have occurred before the formation of attentional set, and the second contained a late reversal stage that may have come after set-formation.

In young control rats, there was no difference between the tasks, offering no support for the hypothesis that reversal learning will be improved post-set-formation. However, during the standard 7-stage task – which is identical to the early reversal learning task until after the (first) intradimensional stage – young control rats can demonstrate attentional set after the CD reversal (Birrell & Brown, 2000). It is therefore premature to reject the hypothetical enhancement of post-set over pre-set reversal learning in normal animals based on the current data, as typical control rats form attentional set before or during the early reversal stage. Thus, as far as the controls rats were concerned, there was no difference between the early and late reversals: they were both likely performed with the benefit of an established attentional set.

In middle-aged rats, however, there was a clear difference between the two types of reversal stage. When middle-aged rats must acquire a reversal before the series of multiple intradimensional acquisition stages, they show a significant impairment compared to younger rats. Additionally, the older rats require significantly more trials than the younger rats to learn the first intradimensional stage that follows the reversal. These results are consistent with the findings of chapter three, where age-related impairments occurred on the first reversal and intradimensional stages of the standard 7-stage task. When the reversal stage follows the series of four intradimensional stages, however, the middle-aged and younger rats perform comparably well. Additionally, the intradimensional stage that follows the post-set reversal is not impaired in middle-aged rats, suggesting that once set is formed, it is not disrupted by the reversal.

The nature of the age-related deficit in attentional set-formation is somewhat less clear than the reversal deficit. After the impaired early reversal, middle-aged rats were significantly impaired at the first intradimensional acquisition, suggesting impaired set-formation. However, no statistical differences were found between older and younger rats on the intradimensional stages of the late reversal task. Qualitative comparisons between the late reversal task and the 4ID task reported in chapter four suggests that middle-aged rats and young rats with OFC lesions learn the second ID in a similar number of trials. This performance by the OFC-lesioned rats in chapter four represented a significant impairment, revealing their deficit in attentional set-formation relative to controls. However, the performance of the control rats on the late reversal task appears to have been more variable than the controls in chapter four, which may have obscured an age-related deficit in set-formation. Further investigations are necessary to address whether middle-aged rats show a pervasive deficit in attentional set-formation, or if this deficit only occurs after an impaired early reversal stage.

Nevertheless, the results from these two tasks altogether support the hypothetical relationship between reversal learning and attentional set-formation. Performance on a reversal stage in middle-aged rats is slower in the absence of an attentional set. Furthermore this performance impairment in turn interferes with the subsequent formation of set. However, when older rats are given a longer opportunity to form attentional set – that is, four intradimensional stages before the reversal – reversal learning occurs as fast as in young control rats, and attentional set remains intact after the reversal stage. This supports the hypothesis that the reversal learning impairment in middle-aged rats is highly influenced by an impaired representation of what cues are relevant during the reversal stage. If a similar impairment is seen in young rats with OFC lesions, as one might expect given the apparent importance of relevancy to the reversal learning deficits seen in these animals (Kim & Ragozzino 2005, Ghods-Sharifi *et al.* 2008; chapter four), this would suggest that organic OFC dysfunction plays an important factor in age-related deficits on the set-shifting task. Unfortunately, the rats lesioned for this experiment cannot provide data to support or refute this hypothesis, and repetition of this condition is therefore necessary.

On both tasks middle-aged rats formed and shifted attentional set, but no rats demonstrated a set-shifting impairment. This result is in contrast to the observation in chapter three – that the positive shift-costs in older rats were approximately three-times larger than typical controls – as well as previous reports in the literature (Barense *et al.*, 2002; Rodefer & Nguyen, 2008). This disparity may be explained by the differing number of stages before the extradimensional shift stage in this task compared to the 7-stage task used in chapter three, and the 5- or 6-stage tasks used in past experiments (*ibid.*).

As increased training on an initial discrimination speeds that discrimination's reversal (*i.e.* the overtraining reversal effect), so too might in-

creasing the number of stages that are consistent with an attentional set speed the shifting of that set. As the set-shifting tasks used throughout this thesis all use "total-change" designs (Slamecka, 1968), the repeated presentation of novel stimuli – through the series of intradimensional stages – may permit the rat to learn about both the relevant and irrelevant stimuli. The aspects of the discriminanda that predict reward will become more salient as attentional set forms (see Esber & Haselgrove, 2011). As new stimuli in the relevant dimension are always accompanied by new irrelevant stimuli, the non-predictive stimuli – because they are novel – will likely also gain salience relative to, for example, spatial location, which is ever present. The overtraining reversal effect occurs when rats are faster to identify the next best alternative after the previously rewarded cue is no longer rewarded. A similar effect may be seen when the relevant – and therefore most salient – stimuli no longer predict reward: rats given a long series of intradimensional stages may be more likely to investigate the predictiveness of the next most salient stimuli (*e.g.* the media cues, following a series of odour discriminations), rather than the hypothetically infinite number of low-salience irrelevant cues. The standard 7-stage task would provide the rats less opportunity to establish this hierarchical salience, and therefore shorter tasks than those used here may be more likely to uncover set-shifting deficits.

Chapter 6

Modafinil Exacerbates Reversal Learning Deficits in Middle-Aged Rats

The results of the previous experiment demonstrated that middle-aged rats are impaired at "pre-set" reversal learning, but not "post-set" reversal learning, and they demonstrate impaired set-formation after an impaired reversal stage. Given this increased understanding of the nature of age-related cognitive decline in the rat, we can now tailor a set-shifting task to specifically quantify the various deficits that a putative cognitive enhancer might affect. The acute effects of modafinil, an atypical stimulant, were tested on middle-aged rats. Modafinil significantly impaired the early reversal stage compared to vehicle treated controls, who learned the reversal similarly poorly to the rats in the previous chapter. The modafinil-treated rats learned the following ID stage better than controls, however, suggesting the cognitive effects of this drug are complex.

Introduction

The primary strength of rat attentional set-shifting tasks is their ability to identify cognitive enhancers relatively early on in the drug development process. This, in turn, helps establish which candidate drugs or drug-targets are most promising for future research and clinical testing. However, to best uncover the cognitive effects of a drug it is necessary to have a precise understanding of the deficit this drug is supposed to ameliorate. For example, the results of chapters four and five suggest that impairments in reversal learning are closely linked to the formation of attentional set, but this aspect of reversal learning may often be overlooked. This raises the possibility that a drug designed to improve behavioural flexibility on reversal stages may not adequately target the underlying dysfunction due to a latent heterogeneity of cognitive deficits between the patient and the animal model.

Our greater awareness of the nature of age-related cognitive decline in the rat now permits the tailoring of a behavioural task to best quantify the deficits we expect to find, and, therefore, their potential amelioration as well. This may provide a more accurate profile of a drug's cognitive effects: instead of measuring only discrimination learning, reversals and set-shifting, we can now compare pre- and post-set reversal learning, as well as attentional set-formation.

Recently, a great deal of research has focused on the effects of the atypical stimulant modafinil ("Provigil"), which is well-tolerated, with few side-effects and very little abuse potential (Schmitt & Reith, 2011). While it is currently approved by the US Food and Drug Administration for treating narcolepsy, shift-work sleep disorder and sleep apnoea/hyponoea; it has also been shown to have potential for treating chronic schizophrenia (Turner *et al.*, 2004b), adult attention-deficit/hyperactivity disorder (Turner

et al., 2004a), major depressive disorder (Abolfazli *et al.*, 2011) and acute mania (Schoenknecht *et al.*, 2010).

To date, little research has been conducted on the cognitive effects of modafinil in healthy, older subjects: human or animal. In middle-aged (18+ months old) rats, modafinil has been shown to ameliorate age-related deficits in the acoustic startle response (McFadden *et al.*, 2010) and sustained attention (Morgan *et al.*, 2007). However, no one has yet investigated whether modafinil affects the deficits in reversal learning (chapters three and five) and attentional set-formation (chapter five) seen in older rats, and so extrapolating possible clinical benefits for older humans, who show a range of deficits on the set-shifting task (Owen *et al.*, 1991; Lange *et al.*, 1995; Lawrence *et al.*, 1998; Robbins *et al.*, 1998), is difficult.

The purpose of this experiment was to investigate the effects of modafinil on the impairments seen in middle-aged rats, using a set-shifting task similar to those used in chapter five. Due to difficulties obtaining a large sample of middle-aged rats, the two tasks from chapter five were combined into one task with both early and late reversal stages. Rats aged 18 months were given an acute challenge of modafinil or vehicle before being tested on this extended set-shifting task.

Methods

Animals

Sixteen male, Lister Hooded rats were obtained from a registered breeder at approximately 11 months of age. Training and testing began when rats were approximately 18 months old. Rats were pair-housed in larger cages under standard housing conditions (chapter two). Approximately one week after their arrival, rats' access to food was controlled, but only mildly restricted (15–20 grams of standard diet per rat per day). Water was always freely available. When testing began, the weight range was 429–562 g. One rat had to be euthanised due to poor health before testing, and three rats (two from the drug group) were unable to complete the test in one session. Only the data from the twelve remaining rats (n = 6 per group) is reported below.

Testing Protocol

Habituation, pre-exposure and training followed the standard protocol laid out in chapter two, but with the following modifications. Two days before testing, rats were given the first of two habituation injections of vehicle solution (intraperitoneal, see below), and then given the habituation bowls. The next day, rats were given their second habituation injection, and then trained to dig in bowls, and pre-exposed to the stimuli.

The day after training, rats were administered drug or vehicle injections (see below) 30 minutes before the start of testing. The order of testing stages is summarised in Figure 6.1. In the first stage, rats were presented with a simple discrimination (SD), and in the next stage, the complementary irrelevant dimension was added to form a compound discrimination (CD). There then followed the first reversal stage (Rev1), where the contingencies of the CD were reversed such that the previously correct stimulus became incorrect and vice versa. The next four stages of the task were in-

	Discriminanda	*Mixed with...*
Simple Discrimination	Coarse tea, not fine tea	Nothing
Compound Discrimination	Coarse tea, not fine tea	Cinnamon or ginger
First Reversal	Fine tea, not coarse tea	Cinnamon or ginger
First Intradimensional Stage	Sand, not grit	Sage or paprika
Second Intradimensional Stage	Wood chips, not sawdust	Turmeric or clove
Third Intradimensional Stage	Yarn, not cigarette filters	Ylang-ylang or frankincense
Fourth Intradimensional Stage	Coarse cork, not fine cork	Patchouli or lavender
Second Reversal	Fine cork, not coarse cork	Patchouli or lavender
Fifth Intradimensional Stage	Long tubes, not short tubes	Bergamot or rosemary
Extradimensional Stage	Coffee, not almond	Cloth or sponge
Sixth Intradimensional Stage	Lemon, not strawberry	BBs or tile-spacers

Figure 6.1: The stages and stimuli of the extended set-shifting task.

tradimensional acquisitions (ID1, ID2, ID3 and ID4), where different compound stimuli were presented at each stage, with the relevant dimen-

sion remaining consistent with the previous stages. After ID4, there was a second reversal (Rev2) stage, and then a fifth intradimensional acquisition (ID5), as described above. The following stage was an extradimensional shift (ED) stage, where different compound stimuli were presented, but rats needed to attend to the previously irrelevant dimension to solve the discrimination. Finally, new compound stimuli are presented, but the relevant dimension from the ED stage remained relevant to form a sixth intradimensional stage (ID6).

Drug Preparation and Administration

Assignment to the drug and vehicle groups was determined pseudorandomly. Six rats were given a single 30 mg/kg injection (1 ml/kg; intraperitoneal) of modafinil (Sequoia Research Products, UK) dissolved in 1% w/v methylcellulose (Sigma-Aldrich, UK) in sterile saline before testing. The remaining six rats were given similar injections of methylcellulose vehicle with no active compound.

Statistical Analyses

Trials to criterion data were analysed using two-way analysis of variance (ANOVA) with Stage (11 levels) as a repeated-measures factor, and Drug as a grouping factor. Planned comparisons were conducted as described in chapters two and five. To put the performance of middle-aged rats in context, the results of typical young rats from the same data set described in chapter three are quoted below. As this young rat data set was tested on a different set-shifting task (the standard 7-stage task), statistical tests between the young and middle-aged data are not appropriate.

Results

Acute modafinil challenge impairs pre-set reversal learning

Treatment with modafinil significantly affected the performance of middle-aged rats on this extended set-shifting task (Figure 6.2; Stage by Drug interaction after Huynh-Feldt correction: $F_{10,100} = 3.16$, $p < 0.05$). One-way ANOVA confirmed that modafinil-treated rats required significantly more trials to complete the first reversal stage than controls (corrected-$F_{1,10} = 22.67$, $p < 0.05$). On the impaired first reversal stage, the modafinil and vehicle treated groups did not differ in the number of errors committed before the first correct dig (means 2.5 and 2.17 respectively, $t_{10} = 0.31$, $p = 0.76$), nor in the number of correct digs in the first six trials (means of 2.17 and 2.5 respectively, $t_{10} = -0.41$, $p = 0.7$), nor in the ratio of errors before the first correct dig to total errors (means 22.6% and 38.7% respectively, $t_{10} = -1.04$, $p = 0.33$). Altogether these results suggest that impairment was not due to perseveration, as the groups did not appear to differ in their tendencies to commit errors at the early trials of the first reversal.

The groups did not differ on the second reversal stage (corrected-$F_{1,10} = 1.37$, $p = 0.27$). However, it should be noted that the half-life of modafinil has been reported to be very short in rats (Waters *et al.*, 2005). The best data available (*ibid.*), which are unfortunately for a different route of administration (gavage versus intraperitoneal) and a slightly different dose (32 mg/kg versus 30 mg/kg), suggest that plasma and brain concentrations of modafinil may reach peak concentration after 60 minutes, but then drop to 20% of peak in the following hour. Mean performance of modafinil-treated rats in the present experiment (Figure 6.3) suggests the drug may not have been effective after the first reversal.

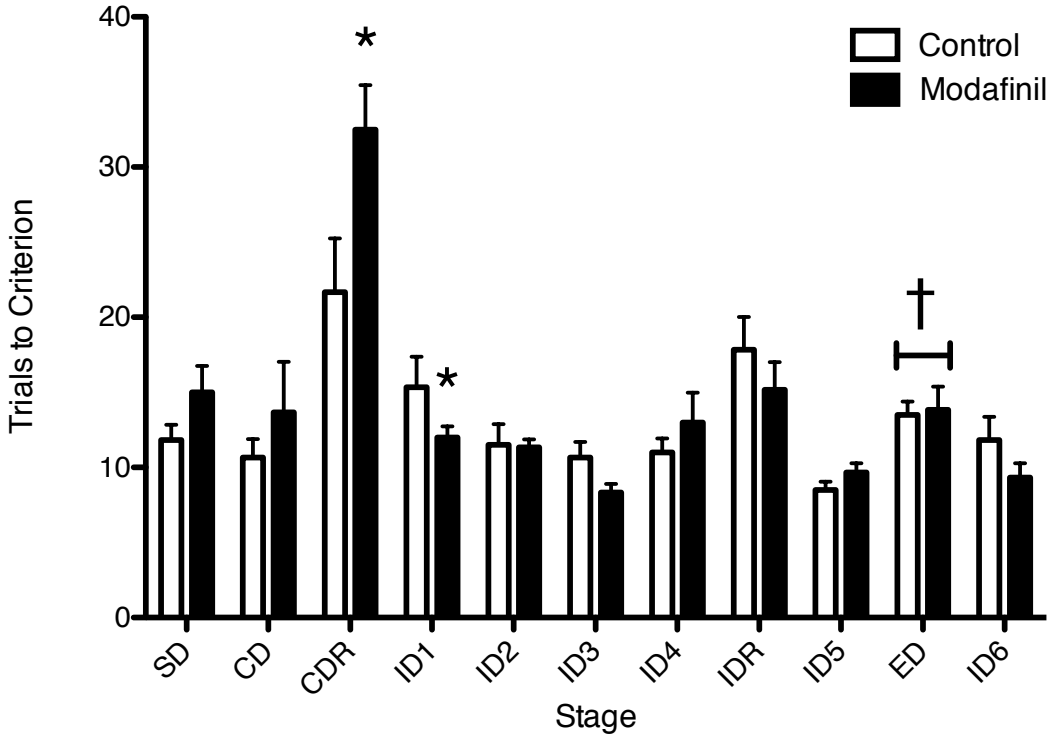


Figure 6.2: Mean trials to criterion (+SEM) on the extended set-shifting task by middle-aged rats treated with modafinil or vehicle. Rats treated with modafinil were significantly impaired at the early reversal, but were significantly faster at learning the first intradimensional stage. The groups did not differ at any other stage.

Attentional set-formation was facilitated by modafinil treatment

Vehicle-treated rats required a mean of 15.33 trials to complete the first ID stage (range 11–24 trials), which is similar to the performance seen on the analogous stage of the early reversal task in chapter five (mean 16.38, range 10–22). Rats treated with modafinil performed significantly better than vehicle-treated controls on the first ID stage (mean 12, range 9–15; corrected-$F_{1,10} = 2.15$, $p < 0.05$), suggesting that acute modafinil attenuates the age-related set-formation deficit.

The groups did not differ in attentional set-shifting

Both groups of middle-aged rats demonstrated a behavioural cost of shifting attentional set, as revealed by ANOVA restricted to the fifth ID

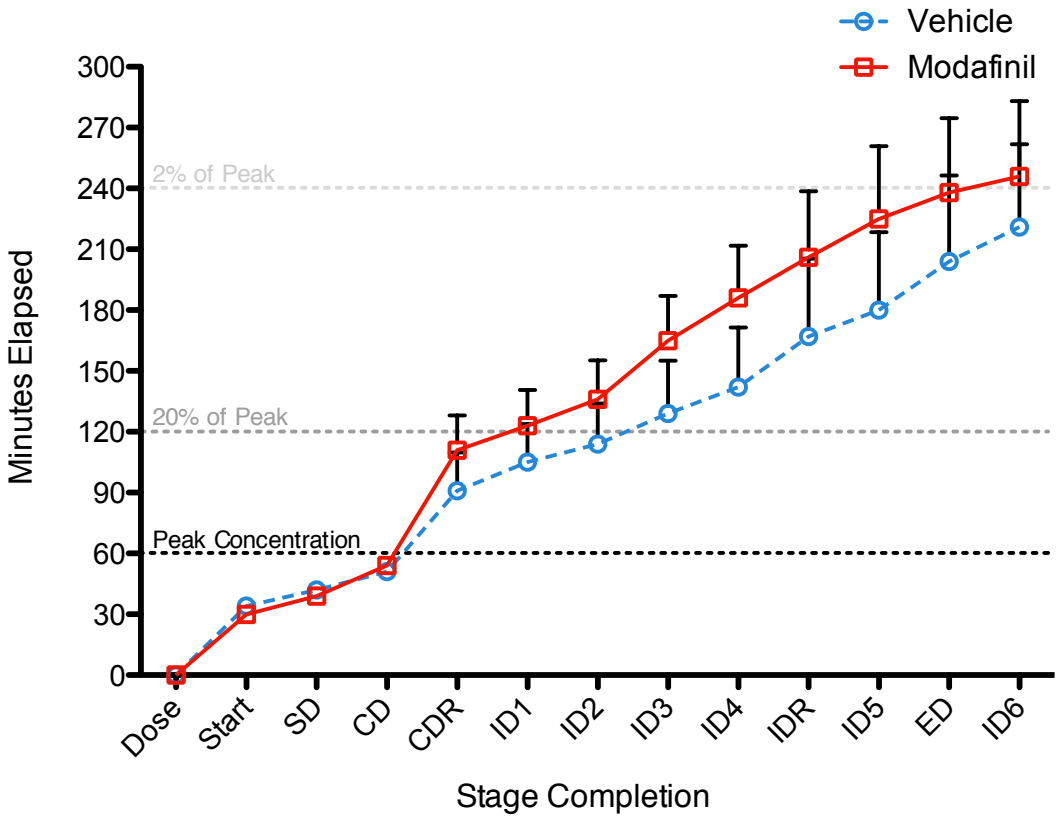


Figure 6.3: Modafinil has a notably short half-life in rats. Time to complete each stage is plotted for the modafinil-treated rats (red line), with horizontal lines marking peak plasma concentration, and the drug's tendency to wear off rapidly (adapted from Waters *et al.*, 2005). To provide context, vehicle-treated rats' performance is plotted in the blue-dashed line. One vehicle-treated rat's duration data were corrected by subtracting a period of 54 minutes spent sleeping during the CDR, which would otherwise skew this group's mean performance.

stage and the ED stage (main effect of Stage, corrected-$F_{1,10}$ = 6.52, $p < 0.05$). Modafinil did not affect set-shifting in middle-aged rats (main effect of Drug and Stage by Drug interactions both corrected-$F < 1$), though the drug may not have been effective at this point of the task (Figure 6.3). The mean shift-cost, calculated by taking the difference between the number of trials to learn the ED with the trials to learn the immediately preceding ID, in the middle aged rats was 4.58 trials. This is only slightly higher than typical young rats performing the standard 7-stage task (mean 3.55 trials), and is similar to the results of chapter five (mean of both tasks 4.44 trials).

Discussion

Attentional set-shifting tasks are an important tool for diagnosing age-related cognitive decline in humans (Robbins *et al.*, 1998), and as we have seen in chapter three, they can also be used to identify potential treatments in rats. The use of set-shifting tasks in drug development for all types of disorders is increasing (*e.g.* Lapiz & Morilak, 2006; Rodefer & Nguyen, 2008; Goetghebeur & Dias, 2009; Tait *et al.*, 2009). It is therefore very important that we have a thorough and complete understanding not only of the cognition being assessed, but the precise effects of putative cognitive enhancers. In chapter four, we were first presented with the possibility that reversal impairments are caused by failures to form attentional set, and this hypothesis was supported in chapter five. Middle-aged rats, which demonstrate reversal learning impairments (Schoenbaum *et al.*, 2006; chapter three), are not impaired at reversal learning when they are first given the opportunity to form set (chapter five). However, they appear less able to form set than controls (*ibid.*), suggesting that their "pre-set" reversal learning impairment is caused by a reduced ability to maintain focus on relevant cues during the reversal.

Here we have replicated previous findings (Tait *et al.*, unpublished observations; chapter three; chapter five) that older rats, like humans, demonstrate reduced behavioural flexibility on an attentional set-shifting task compared to younger controls. The 18-month old rats in this experiment required more trials to learn the first reversal – an early, pre-set reversal – than typical young rats we have tested on a similar set-shifting task (chapter five), and they also appeared to take longer than young rats to form attentional set than young rats (*ibid.*). The shift-costs in these middle-aged rats appeared to be within the normal range of young rats performing the standard 7-stage task. These results also mirror those of chapter five.

The age-related deficit at the first reversal is exacerbated by treatment with modafinil, as drug-treated rats required more trials to learn this stage than the rats given vehicle injections. We saw no evidence that this increase in reversal impairment magnitude was due to an increase in perseveration, as the groups did not differ in their tendency to commit so-called "early phase" reversal errors (Iversen & Mishkin, 1970; Jones & Mishkin, 1972). Additionally, we saw no evidence to suggest that modafinil impaired discrimination learning, as the groups did not differ on the simple or compound discrimination stages, suggesting this reversal deficit was not due to an increase in "learning errors" (*c.f.* Clarke *et al.*, 2005).

In chapter five, the impairment at the pre-set reversal, analogous to the first reversal in the present task, coincided with an impairment at the first intradimensional stage. This result, similar to the findings of chapter four, suggested that the middle-aged rats were impaired at attentional set-formation, and this may therefore have been the mechanism by which their reversal learning was disrupted. Interestingly, rats treated with modafinil in this experiment were significantly better than controls at the first intradimensional stage, and it therefore seems unlikely that the exacerbated reversal deficit was caused by worsening the impairment in attentional set-formation. It could be that modafinil-treated rats were more likely to attend to cues in both dimensions during the first reversal. While this would have hindered their reversal learning due to a relatively increased attentional load, it would also facilitate their representation of what cues were relevant to the discrimination reversal. This in turn would benefit them in the first intradimensional stage.

The groups did not differ on the second reversal stage. This may be because the cognitive requirements of the second reversal stage are different from the first reversal: for example, both groups will perform the second reversal with the benefit of an established attentional set. Additionally, the second reversal may be less "surprising" than the first, and they may

perform it with an increased awareness of the stage's algorithm – that is, with the benefit of a learning set. Alternatively – or perhaps, additionally – it is likely that the drug was no longer present in the brain in large enough concentrations during this stage (*c.f.* Waters *et al.*, 2005). Some experiments have administered a second dose of modafinil during the task, for example before the extradimensional shift stage (*e.g.* Goetghebeur & Dias, 2009; Dawson *et al.*, 2012). Administering a second challenge of modafinil between the first and second hours of testing might be worth considering in future investigations.

These results highlight the value of tailoring a set-shifting task to best quantify the deficits expected from a given population of rats. However, it would be incorrect to infer from these results that the current task – or any modified task that measures discrimination learning and their reversals, attentional set-formation and set-shifting – represents the new paragon of rat set-shifting tasks. Indeed, the very notion of tailoring a set-shifting task to a certain animal model begs the question of what deficits one expects that model to demonstrate. As we had previously established that there would be reversal learning deficits without evidence of set-formation in middle-aged rats performing the standard 7-stage task (chapter three), we had a firm hypothesis regarding the deficits older rats would demonstrate on modified tasks (chapter five), and therefore where we would want to measure a drug's effects. However, for manipulations that have never been tested on any set-shifting task, or manipulations that affect set-shifting but not set-formation (*e.g.* mPFC lesions: Birrell & Brown, 2000), tasks primarily targeting reversal learning and set-formation, such as in the current experiment, would not be most appropriate.

Chapter 7

General Discussion

The rat attentional set-shifting task has revealed a great deal about the cognitive effects of various novel pharmaceutical agents. Over the course of the last four experiments we have seen that certain behavioural impairments, such as deficits in attentional set-formation, can sometimes go undetected in the standard 7-stage task; and modifying the task may therefore be necessary to uncover more information about behavioural flexibility in certain groups of rats. This work has also generated several theoretical questions and testable hypotheses regarding the nature of reversal learning, the function of the OFC, and the best way to probe cognitive impairments through behavioural testing. The answers to these questions, along with the data reported in this thesis, will bring us closer than ever to curing mental illness and cognitive dysfunction.

Findings

Testing the effects of novel cognitive enhancers with the rat attentional set-shifting task generally follows one basic protocol: select the most valid animal model, administer a treatment, then look for changes in task performance relative to controls. While this simple script has produced many informative studies (*e.g.* Tait *et al.*, 2009; Gastambide *et al.*, 2011), the usefulness of the set-shifting task to psychopharmacology can yet be increased. First, more work needs to be done to identify manipulations that better model aspects of various human disorders in rats. Second, a greater understanding of the nature of flexibility impairments is required to help us interpret what measured changes in performance actually reveal about the brain.

The experiment reported in chapter three of this thesis was carried out with the first of these two goals in mind. Older rats potentially represented an ecologically valid way to test age-related cognitive decline in animals. Although the effects of aging on set-shifting had been reported in rats before (Barense *et al.*, 2002; Rodefer & Nguyen, 2008), the changes these groups made to the task rendered their experiments difficult to relate to previous descriptions of young rats' performance (*e.g.* Birrell & Brown, 2000; McAlonan & Brown, 2003). When middle-aged rats were tested on the standard 7-stage task (chapter three) they demonstrated reversal learning impairments, which was not the case in previous aged rat experiments. Also unlike past results, our middle-aged rats showed no evidence of attentional set-shifting – that is, they showed no significant difference between the intra- and extradimensional stages of the standard 7-stage task.

It was this null result that provided the motivation for the experiment reported in chapter four, which pursued the second of the above two goals: to expand our understanding of the behaviour produced by the task. A lack of significant difference between the intra- and extradimensional

stages after an impaired first reversal is a fairly consistent finding in our lab, no matter what the manipulation was: aging (chapter three), lesions of the OFC (McAlonan & Brown, 2003) or lesions of the basal forebrain (Tait & Brown, 2008). This raised the possibility that the impaired first reversal was disrupting the formation – or at least the detection – of attentional set, and so the set-shifting task was modified to involve no reversal stages (chapter four). Young rats with OFC lesions were significantly slower than controls to form attentional set on this modified task, and when they were tested on the standard 7-stage task their data replicated past results (McAlonan & Brown, 2003): impaired reversal learning, and no significant shift-costs.

This finding was broadly consistent with experiments showing the OFC to be intimately involved in calculating the relevancy of stimuli, particularly during reversal learning. It seemed possible, therefore, that the reversal impairments we had observed in chapters three and four were simply alternate manifestations of a single relevancy-based deficit. In chapter five, the 4ID task was modified to have one reversal stage either early in the task, or after the rats had been given a series of intradimensional stages to promote the formation of attentional set. On the early reversal stage, middle-aged rats demonstrated similar reversal learning impairments to those reported in chapter three, and they were also slower to form attentional set. On the task with a late reversal stage, however, the middle-aged rats were not impaired at reversal learning. Though not fully conclusive, these data support the hypothesis of a single deficit: as training the rats in one manifestation, attentional set-formation, suppressed an alternative manifestation, reversal learning.

At this point, our understanding of the nature of age-related cognitive decline in rats was far greater than in the investigation reported in chapter three. Relating this back to the screening of cognitive enhancers, we were then able to tailor the set-shifting task to permit us to measure all

the deficits we expected to find in middle-aged rats. This extended set-shifting task tested early and late reversal learning, as well as set-formation and set-shifting (chapter six). Middle-aged rats given an acute dose of modafinil were significantly impaired at the early reversal stage compared to vehicle-treated controls. While this suggested that modafinil exacerbated the age-related deficits, we also measured a significant improvement in the first intradimensional stage in modafinil-treated rats; signifying improved set-formation in this group. Modafinil may therefore function to increase rats' attention to irrelevant cues during the reversal, which impaired the acquisition of this stage, but subsequently facilitated the formation of attentional set.

Context and Stakes

Deficits in executive control are highly debilitating, affecting almost everything we do in our daily lives; and this can be distressing not only for those suffering from them, but their families and caregivers as well. In older individuals, impaired executive control can lead to difficulties retrieving long-term memories and planning daily chores (Gauthier *et al.*, 2006), but it also predicts mobility impairments and falls (Buchman *et al.*, 2011). Furthermore, in the age-related dementias, widely used drug treatments may confer only mild cognitive benefits (*e.g.* Courtney *et al.*, 2004), and in some cases can even exacerbate cognitive deficits (Cools, 2006). Suboptimal drug therapies for mental illnesses such as schizophrenia and major depressive disorder (Amado-Boccara *et al.*, 1995; Keefe *et al.*, 1999) may increase the risk of disease relapse (*c.f.* Teasdale *et al.*, 2000; DeRubeis *et al.*, 2008). Effective treatment for these diseases must therefore be capable of enhancing cognition as well as targeting what may be more obvious symptoms, such as low mood or hallucinations. Despite this clear need for effective treatments of cognitive dysfunction, few are available.

Detecting these cognitive deficits has primarily been achieved through measuring the flexibility of thought and action; a hallmark feature of executive control (*c.f.* Brown & Tait, 2010). As early as the close of the First World War, psychologists recognised the potential of behavioural flexibility tasks as tools for investigating the cognitive impairments afflicting veterans with damage to the frontal areas of the brain, as well as patients with schizophrenia (Vigotsky & Kasanin, 1934; Goldstein & Scheerer, 1941; Weigl, 1941). These early flexibility tasks would go on to inspire the Wisconsin Card-Sorting Task (Berg, 1948) and the CANTAB intradimensional/extradimensional set-shifting task (*e.g.* Owen *et al.*, 1991), which are still widely used to quantify flexibility impairments in a variety of human populations. More importantly, the set-shifting task – as part of the entire CANTAB package – has proven to be highly effective at screening the effects of putative cognitive enhancers (*e.g.* Mehta *et al.*, 2001; Deakin *et al.*, 2004; Turner *et al.*, 2004b). Unfortunately, improving human cognitive tests is only part of the calculus for getting better treatments from the lab bench to the patient's bedside: the majority of the drug development process occurs long before any human subjects are involved.

Broadly speaking, the prolonged and expensive process of getting a drug into the clinic breaks down into three major phases. The first phase belongs to chemists and electrophysiologists: molecules are synthesised and screened *in vitro* for the selectivity with which they affect various neurotransmitter receptors, as well as the magnitude of this activation. To illustrate the scale of this stage, the drug described in chapter three – ORG49209 – was identified as one of the three most promising candidates out of nearly 47,000 similar molecules (Madau *et al.*, 2009; HM Marston, personal communication).

The second phase belongs to preclinical neuroscientists and experimental psychologists: having received the molecules most likely to have an effect from the chemists, the drugs are given to mice and rats for a se-

ries of *in vivo* experiments. In the case of ORG49209, this work focused on measuring glutamate release in different brain regions after the drug was administered (Connick *et al.*, 2009), testing the drug's effects on locomotor activity (*ibid.*), and – of course – investigating its effects on behavioural flexibility (chapter three).

Finally, the drug makes its way into the hands of psychiatrists and physicians for clinical testing: first it is given to a small group of (three or four) healthy people to test its tolerability; then to successively larger groups, including patient populations, to eventually test its effects in the double-blind, randomised-control, multisite trial – the gold standard of clinical trials. If successful, the drug can then be sent to an external regulatory body for approval; before being marketed to patients, pharmacists and doctors.

The cost of this entire process is difficult to ascertain, but recent industrywide estimates have ranged from USD$92 million to USD$883.6 million (Morgan *et al.*, 2011), or – in 2012 sterling[9] – GBP£73.7 million to GBP£708.6 million. One estimate of total research and development costs for drugs targeting the central nervous system, like antipsychotics and cognitive enhancers, put the total bill for a typical compound at GBP£102.8 million, with the lion's share – GBP£64.6 million – accruing in the third phase of drug development: human clinical trials (DiMasi *et al.*, 2004).

Of course, there's no guarantee of success after all this investment. As noted above, in the case of ORG49209, only three compounds stood out from a group of 47,000 molecules. Of these three, only one was extensively tested in animals (*e.g.* chapter three), and it is unlikely that this particular drug will ever be tested in humans due to the highly problematic way that it is absorbed and metabolised by mammals. This is common: of

[9]Converting from historical rates provided by http://www.xe.com/, and mean annual inflation estimates courtesy of http://www.bankofengland.co.uk

the subset of molecules that show some potential for further testing – like ORG49209, but excluding the 46,997 others – approximately 30% will never leave the first phase of testing, and 70% of those will not progress beyond the second phase (DiMasi *et al.*, 2004). This low success rate is not a weakness of the development process, but a strength: there is a great deal of money to be saved by separating the wheat from the chaff before too much has been invested in any particular compound, and less money spent on bad drugs means more money available to develop the good ones.

The economics of drug development reveal the true value of the rat attentional set-shifting task, as well as the necessity for continuing its development. Given the effectiveness of the CANTAB set-shifting task in diagnosing and quantifying the cognitive impairments associated with aging (*e.g.* Robbins *et al.*, 1998) and mental illness (*e.g.* Elliott *et al.*, 1995; 1996), this task represented a logical choice for assessing the cognitive effects of putative pharmacological enhancers in humans (*e.g.* Turner *et al.*, 2004b). This, in turn, provided the next logical step in drug development, which was to assess drugs' cognitive effects in rats – long before the costly human testing phase – using a fundamentally analogous task. After all, the drugs that are most likely to ameliorate executive dysfunction in rats will also likely ameliorate executive dysfunction in humans, thus identifying the most promising candidate drugs quickly and economically. It should surprise no one that this has been the most popular use for the task in recent years (*e.g.* Lapiz & Morilak, 2006; Goetghebeur & Dias, 2009; Tait *et al.*, 2009; Gastambide *et al.*, 2011).

Conclusions

The standard 7-stage task of attentional set-shifting (Birrell & Brown, 2000) possibly represents the most successful attempt to date for measuring behavioural flexibility in a formally analogous way in rats as in

humans. This success is reflected in the recent recommendations for greater use of this task during the development of cognitively enhancing drugs, as well as treatments for diseases like schizophrenia (Barch *et al.*, 2009; Gilmour *et al.*, 2012). Over the course of this thesis, however, we have seen that a behavioural impairment commonly reported on this task – reversal learning – is perhaps more complex than is often considered; being intimately linked to disruptions of attentional set-formation that may frequently go undetected (*e.g.* McAlonan & Brown, 2003; chapter four).

The possible consequences of this are well illustrated by the results from chapter six. It is not clear if the complicated effects modafinil had on middle-aged rats – exacerbating their reversal deficit while attenuating their set-formation deficit – would have been revealed had the rats been tested on the standard 7-stage task. The results of chapter three suggests that – as is the case in OFC-lesioned rats (chapter four) – the 7-stage task does not provide most middle-aged rats enough opportunity to form set. Using this shorter task might have prevented us from measuring potential effects of the drug on attentional set-formation, which in turn may have led us to summarily dismiss this drug as worsening the age-related deficit. Rather, the complicated pattern of behaviour that was revealed by the modified task suggests that the effects of this drug in older subjects is a topic well worth further investigation.

Though it may be tempting to conclude that the extended set-shifting task reported in chapter six should replace the standard task described by (Birrell & Brown, 2000), this would be incorrect. The tasks used in this thesis were modified with the intention of better quantifying the behavioural deficits expected from rats that had previously demonstrated null results on the standard 7-stage task. Thus a more prudent interpretation of the work reported in this thesis is simply that the order of stages prescribed by Birrell & Brown (2000) need not be taken as dogma; and that prospective and hypothesis-driven modifications to the task can

yield more informative data. Increasing our awareness of the psychology behind the set-shifting task will better enable us to use these tasks to quantify novel drugs' effects on the brain, and this will be key to the search for the next-generation of cognitive enhancers.

Future Questions and Testable Hypotheses

While the experiments reported in this thesis have clearly demonstrated the value of rat set-shifting tasks for testing potential cognitive enhancers, the quest to improve these tasks is ongoing. The work in this thesis raises a number of testable hypotheses to be addressed in future investigations.

Is it possible to produce dissociable reversal impairments through different manipulations?

One of the central themes of these experiments was the possibility that reversal stages can be impaired for multiple reasons. This hypothesis was partly based on data suggesting that reversal learning occurs in phases (Mishkin, 1964; Jones & Mishkin, 1972; Meunier *et al.*, 1997), and therefore an impairment could manifest at one or more of these phases (for full discussion, see chapter four). One of the possible mechanisms by which a reversal stage might be impaired is through increased responding or attention to irrelevant cues. This appears to be why middle-aged rats are impaired at reversal learning, as those rats which were first permitted to form attentional set – or, in other words, were given more opportunity to establish what cues are relevant – did not show the expected age-related reversal impairment (chapter five). Importantly, the intradimensional stage that followed this post-set reversal was not disrupted – as the ID that followed the pre-set reversal was – suggesting that the relevancy impairment takes primacy over the reversal impairment, and not vice versa.

Reversal impairments have also been reported in rats performing the standard 7-stage task after lesions of the OFC (McAlonan & Brown, 2003), lesions of the basal forebrain (Tait & Brown, 2008), prefrontal serotonin depletion (Lapiz *et al.*, 2009), striatal dopamine depletion (O'Neill & Brown, 2007), prenatal administration of methylazoxymethanol acetate (Gastambide *et al.*, 2011), acute challenge of Δ9-tetrahydrocannabinol (Egerton *et al.*, 2005), and others. Unfortunately, we were not able to directly test the hypothesis that different manipulations would lead to dissociable reversal impairments on the set-shifting task, and relatively few experiments have tested this hypothesis in other tasks (*e.g.* Mishkin, 1964). If this hypothesis is supported, the consequences for psychopharmacology would be significant. The use of reversal learning as an index of behavioural flexibility would be less reliable if impairments were found to occur due to subtly different cognitive deficits – for example, a perseverative deficit compared to a set-formation deficit. To test this possibility, dissociable patterns of behaviour should be investigated on the 4ID or early/late reversal tasks between the various manipulations that produce reversal impairments on the standard 7-stage task.

Why haven't we observed perseveration during a reversal stage?

The reversal impairments seen in young rats with lesions of the OFC in chapter four did not appear to be due to increased perseverative responding, which is contrary to previous reports in OFC-lesioned marmosets (Dias *et al.,* 1996b; Dias *et al.*, 1997). Perhaps the most parsimonious explanation for this disparity is that the functional homology between rodent and primate OFC, which was originally assumed because of the similar behavioural effects produced by surgical lesions in both orders of animal (*c.f.* Brown & Bowman, 2002), is not as valid as it originally seemed.

However, two alternative possibilities may also explain this failure to detect perseveration.

As argued more fully in Chase and colleagues (2012), there are considerable differences between the set-shifting tasks used in rats and monkeys. Primary among these are the nature of the stimuli and responses used in the task, and the vastly increased number of trials and errors that monkeys require to perform the task (Dias *et al.*, 1997) compared to rats. On all the stages in the tasks used in this thesis, rats were permitted to "self-correct" by retrieving the food from the correct bowl if they make an error in the first four trials. The original intent of this aspect of the protocol was to permit the rats to experience both discriminanda, thus making the rat task more like the primate task, where the stimuli – images on a touchscreen – can be perceived in instantaneous succession. A consequence of this correction procedure may be that rats are less likely to make long chains of errors before their first correct response on a reversal stage, as many rats will experience the updated cue-reward outcome on the first trial of the reversal. Additionally, given that the discriminations are acquired very quickly in rats, the old cue-reward outcomes may be relatively easy to abandon, making it less likely that perseverative responses will be produced during reversal learning.

The other possible explanation for our failure to detect perseveration is simply that our methods for quantifying it were ineffective. Perseveration during reversal learning represents a subject's failure to disengage the previously rewarded cue-outcome association (or response strategy) in favour of the newly rewarded cue. Given the compound nature of the stimuli used in the set-shifting task, it is difficult to ascertain why a rat chose a particular bowl: was he responding to a cue in the relevant dimension, the irrelevant dimension, their combination, the bowl's spatial location...? In this sense, every error has the potential to be a perseverative error because it is necessarily a response to the previously rewarded stimulus

(McAlonan & Brown, 2003). To overcome this, we used three simple metrics that we reasoned might reveal early-phase, or perseverative (Iversen & Mishkin, 1970; Jones & Mishkin, 1972), errors: the number of errors committed before the first correct dig, the ratio of this number to the total number of errors, and the number of errors committed in the first six trials. In every chapter, these indices revealed an unremarkable number of early errors on the first reversal stage the rats experienced.

Many other experiments have attempted to detect perseverative behaviour by analysing subjects' performance in subsets, or "blocks", of trials: if a subject is perseverating – that is, preferentially responding to the incorrect stimulus – that subject will perform below chance accuracy in a given block (*e.g.* Jones & Mishkin, 1972; Kim & Ragozzino, 2005; Brushfield *et al.*, 2008). In some experiments, including those conducted in monkeys (*e.g.* Dias *et al.*, 1997), the animals are trained in daily blocks of trials of predetermined length, so analysing performance in this way is easily justified.

However, in many experiments with rats – including all of those reported in this thesis – stages are completed when rats attain a criterion of consecutive correct responses, and the resulting trials to criterion scores usually do not divide evenly into blocks of uniform length. Blocks of trials applied to the data *a posteriori* therefore tend not to produce consistent or informative results. For example, while some of the OFC-lesioned rats in chapter four demonstrated one block of four trials – the block size used in Kim & Ragozzino (2005) – below chance accuracy on the first four trials of the first reversal, others attained chance accuracy within this block (data not shown). Little can therefore be inferred about perseveration at the level of the group. Furthermore, this type of analysis is often impossible to apply to control data. The large sample of typical young control rats (reported in chapter three) perform the three reversal stages in a mean of 12.5 trials; permitting only three blocks of four trials to be fit to the data. Given the

performance criterion of six consecutive correct responses, the last of these blocks is necessarily 100% correct, and the middle block can only be 50% or 75% correct. Thus, even if this type of analysis produced a significant result in an impaired group, it is unlikely that any meaningful comparisons could be made to controls.

If perseveration can be detected in a bowl-digging task, then it may be more likely in rats not permitted to self-correct than in rats following the standard protocol. It may also be easier to detect by changing the criterion from six consecutive correct responses to a measure of performance based on blocks of trials: for example, two consecutive four-trial blocks with only one or two errors between them.

Is it possible to glean even more information from the behaviour produced by discrimination learning tasks?

As mentioned above, it may be impossible to know why a rat picked a particular bowl on any one trial. However, it may be possible to detect patterns in the rat's choices over blocks of trials. In each stage of the tasks used in this thesis, the stimuli are presented to the rats in a predetermined, pseudorandom sequence that generally follows a Gellermann order (Gellermann, 1933); meaning that stimuli can appear on the same side for no more than three consecutive trials. These orders were then modified slightly before beginning the experiment reported in chapter five such that in a rolling block of six trials (*i.e.* the length of the performance criterion) each stimulus appears equally with each irrelevant stimulus, and on each side[10], as far as possible. For example, in a block of six trials, the coarse tea will appear with the cinnamon and ginger three times each. The order of

[10] To prevent the spatial pattern from becoming too predictable, in some of the 36 possible six-trial blocks – not all of which will be experienced by any particular rat – the stimuli appear on each side of the box in a 4:2 ratio, not 3:3. The ratio of these left-biased to right-biased blocks is approximately 1:1.

these presentations is balanced to ensure that the only adaptive strategy is to follow the correct cue, and that strategies based on alternation behaviour are maladaptive.

This arrangement may enable us to identify when rats use alternative strategies through patterns in their responding. For example, in a block of six trials a rat could dig in the correct and incorrect bowls an equal number of times, but closer analysis of his responses might reveal that – rather than randomly alternating – the rat made six consecutive digs in bowls containing one of the irrelevant stimuli. Analyses of this type were piloted during the experiment reported in chapter four (data not shown), but a lack of consistency in the behaviour in the lesion group – which represented the largest set of data to analyse – yielded only non-significant trends. Mainly due to lack of time, these analyses were not pursued, and so never perfected.

It is difficult to overstate the value of exploring analyses of this type. The work in this thesis primarily sought to uncover the cognitive underpinnings of impairments in behavioural flexibility by modifying the set-shifting task and closely examining the resulting behaviour. In chapter four, for example, the 4ID task and the standard 7-stage task respectively revealed that OFC-lesioned rats were impaired at attentional set-formation as well as reversal learning. This enabled us to infer that a single deficit in recognising relevancy might cause both impairments, thereby proposing a mechanism by which the OFC lesion impairs reversal learning. However, with a system for analysing runs of responses, or otherwise estimating to which cues a rat is attending, testing on multiple modified tasks in this way would become unnecessary. Rather, the rat would simply perform a reversal stage, and the analysis would reveal – perhaps even in real time, were it integrated into the data collection process – how the rat was solving the stage. This analysis could also potentially reveal whether the deficit on the extradimensional stage seen in OFC lesioned rats (chapter four) represented an impaired

ability to shift set, or an impaired ability to form a replacement set. It might also help explain why those older rats that are capable of forming set are impaired at shifting in the standard task (chapter three), but no age-related set-shifting deficits were seen in longer tasks (chapters five and six).

This type of analysis might represent a convergent line of research to recent advances in the computational modelling of human and animal behaviour (*e.g.* Yu & Dayan, 2005; Takahashi *et al.*, 2009; Walton *et al.*, 2010; Wilson & Niv, 2012). Integrating more sophisticated behavioural analyses such as these may be crucial for the continued development of tasks for translational psychopharmacology. As the development of so-called "computational phenotypes" (Montague *et al.*, 2012) of mental illnesses become more common, more precise animal models will be needed for testing potential drug treatments. The rat set-shifting task owes its value as a test of cognitive enhancers to the success of the CANTAB set-shifting task. This celebrated position may therefore be somewhat precarious: if our understanding of human cognitive impairments advances too far beyond our understanding of rat behaviour, the usefulness of the rat set-shifting task – indeed, any rat behavioural task – will be significantly reduced. This would seriously hinder the pursuit of more effective treatments for mental illness and cognitive dysfunction, and avoiding this possibility is therefore of utmost importance.

Can additional rat behavioural tasks reveal more about impaired cognition?

Attentional set-shifting and reversal learning tasks are the only successful bowl-digging tasks for testing putative cognitive enhancers in rats. This is in stark contrast to the panoply of tests available for quantifying executive control in humans; indeed, the CANTAB set-shifting task is only one in a battery of up to 25 tests. Executive control is a multifaceted aspect of cognition, and the effects of drugs across the various domains of execu-

tive control can differ between subject populations. An especially relevant example of this phenomenon can been seen in the effects of modafinil on individuals suffering from schizophrenia (Turner *et al.*, 2004b) compared to those with adult attention deficit/hyperactivity disorder (Turner *et al.*, 2004a): pro-cognitive effects have been measured in both patient groups, but the set-shifting task was only informative in those suffering from schizophrenia.

Unfortunately, developing alternative bowl-digging tasks to probe executive control in rats is not as straightforward as selecting another task from the CANTAB and translating it for use in rats. As discussed more fully in chapter one, the visual nature of the stimuli used in touchscreen tasks makes them quite difficult for rats to discriminate between. In developing the rat attentional set-shifting task, this was not an insurmountable problem: the visual dimensions of line and shape were replaced with the somatosensory dimensions of odour and medium, thus forming a species-appropriate task. In some other CANTAB tasks, however – such as the visual pattern recognition task, the Tower of London task, and the rapid visual information processing task (*e.g.* Turner *et al.*, 2004b) – the visual component is more difficult to replace.

Another factor that makes many CANTAB tasks difficult to transfer to the rat is the element of spatial memory; as is seen in the spatial span task, the spatial working memory task and the Tower of London task (*ibid.*). Rats are very good at learning spatial information, and may preferentially attempt to solve problems spatially before engaging higher-order cognitive processes. Evidence of this can be seen in the overtraining reversal effect: increasing the amount of training increases the probability that the animal will respond to the relevant cues (Lovejoy, 1966), and decreases the probability that the animal will respond to the cues' spatial locations (Reid, 1953; Mackintosh, 1969; Sutherland & Mackintosh, 1971). This supposed preference for spatial tasks would also explain the failures to

generate an overtraining reversal effect on spatial reversal learning problems (*e.g.* D'Amato & Jagoda, 1962). Purported indices of executive control tasks are therefore difficult to interpret when the tasks use spatial variables, as the extent to which the rat frontal cortex is involved in this type of learning remains unclear.

A promising avenue for task development may be found in the recent advances in testing episodic-like memory in rats (for review, see Eichenbaum *et al.*, 2012), as memory recall is an important facet of executive control. Ainge and colleagues (2010) have recently described a variant of the novel object recognition task (Ennaceur & Delacour, 1988). Specifically, rats in this experiment were presented with two objects in one testing box, and two different objects in a second testing box. Rats were then placed back in the first box and presented with one object from the second box, and one that had always been in the first. Though the rats had experienced both objects for a similar amount of time, they spent more time investigating the object that had not previously been associated with that location (Ainge *et al.,* 2010). This result suggests that the rats recognised that there was something new about this old object; indicating that they attended not only to the object's identity, but also the context in which the object was first experienced.

Another method of testing context-dependent learning can be seen in the acquired equivalence and distinctiveness tasks (Coutureau *et al.*, 2002; Iordanova *et al.*, 2007). Generally speaking, the theory behind this task states that the inherent differences between contexts or locations become less salient – that is, the contexts become "equivalent" – when stimuli presented in those contexts predict reward in similar ways. For example, if the outcome of a discrimination between two stimuli is the same in two testing boxes with checkerboard and polka-dot patterns on the walls, but different in two testing boxes that are heated or cooled beyond room temperature; then the patterned boxes will be perceived as equivalent, and dis-

tinct from the temperature boxes. This effect can be measured by presenting the rats with free food in one of the boxes, and then seeing if they appear to also expect free food in the hypothetically equivalent box. Indeed, normal rats clearly demonstrate this expectation – an effect of acquired equivalence (Coutureau *et al.*, 2002) – while rats with medial prefrontal cortex lesions do not (Iordanova *et al.*, 2007). This raises the possibility that this operant conditioning task engages similar executive control processes as attentional set-shifting (*c.f.* Birrell & Brown, 2000), which one might expect if context-dependent learning really is related to episodic memory.

The extensive training involved in these past experiments – which take 20 days from start to finish (Iordanova *et al.*, 2007) – renders them of little use to psychopharmacology. Simultaneously to conducting the experiment reported in chapter three, I made several attempts to devise an acquired equivalence and distinctiveness bowl-digging task. Unfortunately, none of these attempts were successful, and they are therefore not described elsewhere in this thesis. This may have been because the training phase of the pilot tasks used a similar performance criterion as in the set-shifting task, which may not have been sufficiently lengthy to produce the effect. Alternatively, the behavioural measures taken to quantify expectation – time spent digging in an empty bowl, or in some cases the number of trials the rats will dig in empty bowls before losing interest – may not have been optimal.

Impairments in retrieving memories are common features of the executive control deficits reported in age-related cognitive decline (Gauthier *et al.*, 2006) and major depressive disorder (Watkins & Teasdale, 2001). A bowl-digging task that can measure memory-retrieval, similar to the episodic-like memory or context-dependent learning tasks (Iordanova *et al.*, 2007; Ainge *et al.*, 2010), may represent an important goal for future task development. A bowl-digging battery of executive control tasks may

bring us closer than ever to curing mental illness and age-related cognitive decline.

Appendix I: References

Abolfazli, R., Hosseini, M., Ghanizadeh, A., Ghaleiha, A., Tabrizi, M., Raznahan, M., Golalizadeh, M., & Akhondzadeh, S. (2011) Double-blind randomized parallel-group clinical trial of efficacy of the combination fluoxetine plus modafinil versus fluoxetine plus placebo in the treatment of major depression. *Depression and Anxiety*, **28**, 297–302.

Abràmoff, M.D., Magalhães, P.J., & Ram, S.J. (2004) Image Processing with ImageJ. *Biophotonics International*, **11**, 36–42.

Ainge, J.A., Schlesiger, M.I., & Langston, R.F. (2010) Lateral entorhinal cortex is activated by memory for object-context associations. Presented at the Society for Neuroscience Annual Meeting 2010.

Amado-Boccara, I., Gougoulis, N., Poirier Littré, M.F., Galinowski, A., & Lôo, H. (1995) Effects of antidepressants on cognitive functions: a review. *Neuroscience & Biobehavioral Reviews*, **19**, 479–493.

Anderson, A.L., Li, S.-H., Biswas, K., McSherry, F., Holmes, T., Iturriaga, E., Kahn, R., Chiang, N., Beresford, T., Campbell, J., Haning, W., Mawhinney, J., McCann, M., Rawson, R., Stock, C., Weis, D., Yu, E., & Elkashef, A.M. (2012) Modafinil for the treatment of methamphetamine dependence. *Drug and Alcohol Dependence*, **120**, 135–141.

Arunrut, T., Alejandre, H., Chen, M., Cha, J., & Russo-Neustadt, A. (2009) Differential behavioral and neurochemical effects of exercise, reboxetine and citalopram with the forced swim test. *Life Sciences*, **84**, 584–589.

Barch, D.M., Braver, T.S., Carter, C.S., Poldrack, R.A., & Robbins, T.W. (2009) CNTRICS final task selection: executive control. *Schizophrenia Bulletin*, **35**, 115–135.

Barense, M.D., Fox, M.T., & Baxter, M.G. (2002) Aged rats are impaired on an attentional set-shifting task sensitive to medial frontal cortex damage in young rats. *Learning & Memory*, **9**, 191–201.

Bellman, R. (1966) *Adaptive Control Processes: a Guided Tour*. Princeton University Press.

Belozertseva, I., Kos, T., Popik, P., Danysz, W., & Bespalov, A. (2007) Antidepressant-like effects of mGluR1 and mGluR5 antagonists in the rat forced swim and the mouse tail suspension tests. *European Neuropsychopharmacology*, **17**, 172–179.

Berg, E.A. (1948) A Simple Objective Technique for Measuring Flexibility in Thinking. *Journal of General Psychology*, **39**, 15–22.

Birrell, J.M. & Brown, V.J. (2000) Medial frontal cortex mediates perceptual attentional set shifting in the rat. *Journal of Neuroscience*, **20**, 4320–4324.

Bissonette, G., Martins, G., Franz, T.M., Harper, E., Schoenbaum, G., & Powell, E.M. (2008) Double Dissociation of the Effects of Medial and Orbital Prefrontal Cortical Lesions on Attentional and Affective Shifts in Mice. *Journal of Neuroscience*, **28**, 11124–11130.

Bliss, T.V. & Collingridge, G.L. (1993) A synaptic model of memory: long-term potentiation in the hippocampus. *Nature*, **361**, 31–39.

Bonfoco, E., Krainc, D., Ankarcrona, M., Nicotera, P., Lipton, S.A. (1995) Apoptosis and Necrosis: Two Distinct Events Induced, Respectively, by Mild and Intense Insults with N-Methyl-D-Aspartate or Nitric Oxide/Superoxide in Cortical Cell Cultures. *Proceedings of the National Academy of Sciences of the United States of America,* **92**, 7162-7166.

Brigman, J.L., Mathur, P., Harvey-White, J., Izquierdo, A., Saksida, L.M., Bussey, T.J., Fox, S., Deneris, E., Murphy, D.L., & Holmes, A. (2010) Pharmacological or Genetic Inactivation of the Serotonin Transporter Improves Reversal Learning in Mice. *Cerebral Cortex (New York, NY)*, **20**, 1955–1963.

Brown, V.J. & Bowman, E.M. (2002) Rodent models of prefrontal cortical function. *Trends in Neuroscience,* **25**, 340–343.

Brown, V.J. & Tait, D.S. (2010) Behavioral Flexibility: Attentional Shifting, Rule Switching and Response Reversal. *Encyclopedia of Psychopharmacology*, 1–5.

Brushfield, A.M., Luu, T.T., Callahan, B.D., & Gilbert, P.E. (2008) A comparison of discrimination and reversal learning for olfactory and visual stimuli in aged rats. *Behavioral Neuroscience,* **122**, 54–62.

Buchman, A.S., Boyle, P.A., Leurgans, S.E., Barnes, L.L., & Bennett, D.A. (2011) Cognitive function is associated with the development of mobility impairments in community-dwelling elders. *American Journal of Geriatric Psychiatry*, **19**, 571–580.

Buss, A.H. (1953) Rigidity as a function of reversal and non-reversal shifts in the learning of successive discriminations. *Journal of Experimental Psychology*, **45**, 75–81.

Buss, A.H. (1956) Reversal and nonreversal shifts in concept formation with partial reinforcement eliminated. *Journal of Experimental Psychology*, **52**, 162–166.

Bussey, T.J., Muir, J.L., Everitt, B.J., & Robbins, T.W. (1997) Triple dissociation of anterior cingulate, posterior cingulate, and medial frontal cortices on visual discrimination tasks using a touchscreen testing procedure for the rat. *Behavioral Neuroscience*, **111**, 920–936.

Cepeda, N.J., Kramer, A.F., & Gonzalez de Sather, J.C. (2001) Changes in executive control across the life span: examination of task-switching performance. *Developmental Psychology*, **37**, 715–730.

Chamberlain, S.R., Robbins, T.W., Winder-Rhodes, S., Müller, U., Sahakian, B.J., Blackwell, A.D., & Barnett, J.H. (2011) Translational Approaches to Frontostriatal Dysfunction in Attention-Deficit/Hyperactivity Disorder Using a Computerized Neuropsychological Battery. *Biological Psychiatry*, **69**, 1192–1203.

Chase, E.A., Tait, D.S., & Brown, V.J. (2012) Lesions of the orbital prefrontal cortex impair the formation of attentional set in rats. *European Journal of Neuroscience,* **36**, 2368–2375.

Cheney, D.L., Uzunov, D., & Guidotti, A. (1995) Pregnenolone sulfate antagonizes dizocilpine amnesia: role for allopregnanolone. *Neuroreport*, **6**, 1697–1700.

Clarke, H.F., Walker, S.C., Crofts, H.S., Dalley, J.W., Robbins, T.W., & Roberts, A.C. (2005) Prefrontal serotonin depletion affects reversal learning but not attentional set shifting. *Journal of Neuroscience*, **25**, 532–538.

Connick, J.H., Cadzow, L., Jeggo, R., Ge, J., Madau, P., Erdemli, G., & Thomson, F. (2009) Org1 **[the reference for ORG49209 while it was under embargo]** potentiates NMDA-receptor mediated responses *in vivo*. Presented at the Society for Neuroscience Annual Meeting 2009.

Cools, R. (2006) Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. *Neuroscience & Biobehavioral Reviews*, **30**, 1–23.

Courtney, C., Farrell, D., Gray, R., Hills, R., Lynch, L., Sellwood, E., Edwards, S., Hardyman, W., Raftery, J., Crome, P., Lendon, C., Shaw, H., Bentham, P., AD2000 Collaborative Group (2004) Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): randomised double-blind trial. *Lancet*, **363**, 2105–2115.

Coutureau, E., Killcross, A.S., Good, M., Marshall, V.J., Ward-Robinson, J., & Honey, R.C. (2002) Acquired equivalence and distinctiveness of cues: II. Neural manipulations and their implications. *Journal of Experimental Psychology: Animal Behavior Processes*, **28**, 388–396.

Cruz-Muros, I., Afonso-Oramas, D., Abreu, P., Pérez-Delgado, M.M., Rodríguez, M., & González-Hernández, T. (2009) Aging effects on the dopamine transporter expression and compensatory mechanisms. *Neurobiology of Aging*, **30**, 973–986.

D'Amato, M.R. & Jagoda, H. (1962) Overlearning and position reversal. *Journal of Experimental Psychology*, **64**, 117–122.

Damasio, A.R., Graff-Radford, N.R., Eslinger, P.J., Damasio, H., & Kassell, N. (1985) Amnesia following basal forebrain lesions. *Archives of Neurology*, **42**, 263–271.

Damasio, H., Grabowski, T., Frank, R., Galaburda, A., & Damasio, A. (1994) The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*, **264**, 1102–1105.

Dawson, N., Thompson, R.J., McVie, A., Thomson, D.M., Morris, B.J., & Pratt, J.A. (2012) Modafinil Reverses Phencyclidine-Induced Deficits in Cognitive Flexibility, Cerebral Metabolism, and Functional Brain Connectivity. *Schizophrenia Bulletin,* **38**, 457–474.

Deakin, J.B., Aitken, M.R.F., Dowson, J.H., Robbins, T.W., & Sahakian, B.J. (2004) Diazepam produces disinhibitory cognitive effects in male volunteers. *Psychopharmacology*, **173**, 88–97.

DeRubeis, R.J., Siegle, G.J., & Hollon, S.D. (2008) Cognitive therapy versus medication for depression: treatment outcomes and neural mechanisms. *Nature Reviews Neuroscience*, **9**, 788–796.

Dias, R., Robbins, T.W., & Roberts, A.C. (1996a) Primate analogue of the Wisconsin Card Sorting Test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behavioral Neuroscience*, **110**, 872–886.

Dias, R., Robbins, T.W., & Roberts, A.C. (1996b) Dissociation in prefrontal cortex of affective and attentional shifts. *Nature,* **380**, 69–72.

Dias, R., Robbins, T.W., & Roberts, A.C. (1997) Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. *Journal of Neuroscience*, **17**, 9285–9297.

DiMasi, J.A., Grabowski, H.G., & Vernon, J. (2004) R&D Costs and Returns by Therapeutic Category. *Drug Information Journal*, **38**, 211–223.

Dudchenko, P.A., Wood, E.R., & Eichenbaum, H. (2000) Neurotoxic hippocampal lesions have no effect on odor span and little effect on odor recognition memory but produce significant impairments on spatial span, recognition, and alternation. *Journal of Neuroscience*, **20**, 2964–2977.

Egerton, A., Brett, R.R., & Pratt, J.A. (2005) Acute Δ9-Tetrahydrocannabinol-Induced Deficits in Reversal Learning: Neural Correlates of Affective Inflexibility. *Neuropsychopharmacology*, **30**, 1895–1905.

Eichenbaum, H., Sauvage, M., Fortin, N., Komorowski, R., & Lipton, P. (2012) Towards a functional organization of episodic memory in the medial temporal lobe. *Neuroscience and Biobehavioral Reviews*, **36**, 1597–1608.

Elliott, R., McKenna, P.J., Robbins, T.W., & Sahakian, B.J. (1995) Neuropsychological evidence for frontostriatal dysfunction in schizophrenia. *Psychological Medicine*, **25**, 619–630.

Elliott, R., Sahakian, B.J., Mckay, A., Herrod, J., Robbins, T.W., & Paykel, E. (1996) Neuropsychological impairments in unipolar depression: the influence of perceived failure on subsequent performance. *Psychological Medicine*, **26**, 975–989.

Emery, N.J. & Clayton, N.S. (2004) The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science*, **306**, 1903–1907.

Ennaceur, A. & Delacour, J. (1988) A new one-trial test for neurobiological studies of memory in rats. 1: Behavioral data. *Behavioral Brain Research,* **31**, 47–59.

Erlebacher, A. (1963) Reversal learning in rats as a function of percentage of reinforcement and degree of learning. *Journal of Experimental Psychology*, **66**, 84–90.

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

Eslinger, P.J. & Damasio, A.R. (1985) Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology*, **35**, 1731–1741.

Fischer, W., Nilsson, O.G., & Björklund, A. (1991) In vivo acetylcholine release as measured by microdialysis is unaltered in the hippocampus of cognitively impaired aged rats with degenerative changes in the basal forebrain. *Brain Research*, **556**, 44–52.

Frick, K.M., Baxter, M.G., Markowska, A.L., Olton, D.S., & Price, D.L. (1995) Age-related spatial reference and working memory deficits assessed in the water maze. *Neurobiology of Aging*, **16**, 149–160.

Gardner, R.A. (1966) On box score methodology as illustrated by three reviews of overtraining reversal effects. *Psychological Bulletin*, **66**, 416–418.

Gastambide, F., Cotel, M.C., Gilmour, G., O'Neill, M.J., Robbins, T.W., & Tricklebank, M.D. (2011) Selective Remediation of Reversal Learning Deficits in the Neurodevelopmental MAM Model of Schizophrenia by a Novel mGlu5 Positive Allosteric Modulator. *Neuropsychopharmacology*, **37**, 1057–1066.

Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R.C., Ritchie, K., Broich, K., Belleville, S., Brodaty, H., Bennett, D.A., Chertkow, H., Cummings, J.L., de Leon, M., Feldman, H., Ganguli, M., Hampel, H., Scheltens, P., Tierney, M.C., Whitehouse, P., Winblad, B., & International Psychogeriatric Association Expert Conference on mild

cognitive impairment (2006) Mild cognitive impairment. *Lancet*, **367**, 1262–1270.

Gellermann, L.W. (1933) Chance Orders of Alternating Stimuli in Visual Discrimination Experiments. *The Pedagogical Seminary and Journal of Genetic Psychology*, **42**, 206–208.

Ghods-Sharifi, S., Haluk, D.M., & Floresco, S.B. (2008) Differential effects of inactivation of the orbitofrontal cortex on strategy set-shifting and reversal learning. *Neurobiology of Learning and Memory*, **89**, 567–573.

Gibbs, T.T., Russek, S.J., & Farb, D.H. (2006) Sulfated steroids as endogenous neuromodulators. *Pharmacology Biochemistry and Behavior*, **84**, 555–567.

Gibson, J.J. (1941) A critical review of the concept of set in contemporary experimental psychology. *Psychological Bulletin*, **38**, 781–817.

Gilbert, P.E., Pirogovsky, E., Brushfield, A.M., Luu, T.T., Tolentino, J.C., & Renteria, A.F. (2009) Age-related Changes in Associative Learning for Olfactory and Visual Stimuli in Rodents. *Annals of the New York Academy of Sciences*, **1170**, 718–724.

Gilmour, G., Arguello, A., Bari, A., Brown, V.J., Carter, C., Floresco, S.B., Jentsch, J.D., Tait, D.S., Young, J.W., & Robbins, T.W. (2012) Measuring the construct of executive control in schizophrenia: Defining and validating translational animal paradigms for discovery research. *Neuroscience & Biobehavioral Reviews*.

Goetghebeur, P.J.D. & Dias, R. (2009) Comparison of haloperidol, risperidone, sertindole, and modafinil to reverse an attentional set-shifting impairment following subchronic PCP administration in the rat–a back translational study. *Psychopharmacology*, **202**, 287–293.

Goldberg, T.E., Weinberger, D.R., Berman, K.F., Pliskin, N.H., & Podd, M.H. (1987) Further Evidence for Dementia of the Prefrontal Type in

Schizophrenia? A Controlled Study of Teaching the Wisconsin Card Sorting Test. *Archives of General Psychiatry*, **44**, 1008–1014.

Goldstein, K. & Scheerer, M. (1941) Abstract and concrete behavior an experimental study with special tests. *Psychological Monographs,* **53**, 1– 151.

Gould, T.D., Einat, H., Bhat, R., & Manji, H.K. (2004) AR-A014418, a selective GSK-3 inhibitor, produces antidepressant-like effects in the forced swim test. *The International Journal of Neuropsychopharmacology*, **7**, 387–390.

Grant, D.A. & Berg, E.A. (1948) A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, **38**, 404–411.

Harlow, H.F. (1949) The formation of learning sets. *Psychological Review,* **56**, 51–65.

Harlow, H.F. & Dagnon, J. (1943) Problem solution by monkeys following bilateral removal of the prefrontal areas. *Journal of Experimental Psychology*, **32**, 351–356.

Harrison, P.J. & Weinberger, D.R. (2004) Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. *Molecular Psychiatry*, **10**, 40–68.

Harrow, M. & Friedman, G.B. (1958) Comparing reversal and nonreversal shifts in concept formation with partial reinforcement controlled. *Journal of Experimental Psychology*, **55**, 592–598.

Helm, T. (2012) New graduates will have to work until 71 before qualifying for state pension. *The Guardian*.

Hécaen, H. (1964) Mental Symptoms Associated with Tumors on the Frontal Lobe. In Warren, J.M. & Akert, K. (eds), *The Frontal Granular Cortex and Behavior*. McGraw-Hill, pp. 335–352.

Iordanova, M., Killcross, A.S., & Honey, R.C. (2007) Role of the medial prefrontal cortex in acquired distinctiveness and equivalence of cues. *Behavioral Neuroscience*, **121**, 1431–1436.

Iversen, S.D. & Mishkin, M. (1970) Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Experimental Brain Research*, **11**, 376–386.

Izquierdo, A., Belcher, A.M., Scott, L., Cazares, V.A., Chen, J., O'Dell, S.J., Malvaez, M., Wu, T., & Marshall, J.F. (2009) Reversal-Specific Learning Impairments After a Binge Regimen of Methamphetamine in Rats: Possible Involvement of Striatal Dopamine. *Neuropsychopharmacology*, **35**, 505–514.

Jacobsen, C.F. (1936) Studies of cerebral function in primates. *Compartive Psychology Monographs*, **13**, 1–60.

Jentsch, J.D. & Roth, R.H. (1999) The neuropsychopharmacology of phencyclidine: from NMDA receptor hypofunction to the dopamine hypothesis of schizophrenia. *Neuropsychopharmacology*, **20**, 201–225.

Jones, B.C. & Mishkin, M. (1972) Limbic lesions and the problem of stimulus–Reinforcement associations. *Experimental Neurology*, **36**, 362–377.

Judge, P.G., Evans, D.W., Schroepfer, K.K., & Gross, A.C. (2011) Perseveration on a reversal-learning task correlates with rates of self-directed behavior in nonhuman primates. *Behavioral Brain Research*, **222**, 57–65.

Kahnt, T., Heinzle, J., Park, S.Q., & Haynes, J.D. (2011) Decoding the Formation of Reward Predictions across Learning. *Journal of Neuroscience*, **31**, 14624–14630.

Kapur, N., Ellison, D., Smith, M.P., Mclellan, D.L., & Burrows, E.H. (1992) Focal Retrograde Amnesia Following Bilateral Temporal Lobe Pathology. *Brain*, **115**, 73–85.

Keefe, R.S.E., Silva, S., Perkins, D., & Lieberman, J. (1999) The effects of atypical antipsychotic drugs on neurocognitive impairment in schizophrenia: a review and meta-analysis. *Schizophrernia Bulletin*, **25**, 201–222.

Kelleher, R.T. (1956) Discrimination learning as a function of reversal and nonreversal shifts. *Journal of Experimental Psychology*, **51**, 379–384.

Kendler, H.H. & D'Amato, M.F. (1955) A comparison of reversal shifts and nonreversal shifts in human concept formation behavior. *Journal of Experimental Psychology*, **49**, 165–174.

Kendler, H.H. & Kendler, T.S. (1967) Vertical and Horizontal Processes in Problem Solving. In *Human Learning and Memory*. Oxford University Press, New York, pp. 445–459.

Kendler, T.S. & Kendler, H.H. (1959) Reversal and nonreversal shifts in kindergarten children. *Journal of Experimental Psychology*, **58**, 56–60.

Kim, J. & Ragozzino, M.E. (2005) The involvement of the orbitofrontal cortex in learning under changing task contingencies. *Neurobiology of Learning and Memory*, **83**, 125–133.

Kuhn, H.G., Dickinson-Anson, H., & Gage, F.H. (1996) Neurogenesis in the dentate gyrus of the adult rat: age-related decrease of neuronal progenitor proliferation. *Journal of Neuroscience,* **16**, 2027–2033.

Lange, K.W., Sahakian, B.J., Quinn, N.P., Marsden, C.D., & Robbins, T.W. (1995) Comparison of executive and visuospatial memory function in Huntington's disease and dementia of Alzheimer type matched for degree of dementia. *Journal of Neurology, Neurosurgery and Psychiatry*, **58**, 598–606.

Lapiz, M.D.S. & Morilak, D.A. (2006) Noradrenergic modulation of cognitive function in rat medial prefrontal cortex as measured by attentional set shifting capability. *Neuroscience,* **137**, 1039–1049.

Lapiz, M.D.S., Soto-Piña, A.E., Hensler, J.G., & Morilak, D.A. (2009) Chronic intermittent cold stress and serotonin depletion induce

deficits of reversal learning in an attentional set-shifting test in rats. *Psychopharmacology*, **202**, 329–341.

Lashley, K.S. (1929) *Brain Mechanisms and Intelligence*, 1963 edition. Dover, New York.

Lashley, K.S. (1930) The Mechanism of Vision: I. A Method for Rapid Analysis of Pattern-Vision in the Rat. *The Pedagogical Seminary and Journal of Genetic Psychology*, **37**, 453–460.

Lawrence, A.D., Sahakian, B.J., & Robbins, T.W. (1998) Cognitive functions and corticostriatal circuits: insights from Huntington's disease. *Trends in Cognitive Sciences*, **2**, 379–388.

Lees, A.J. & Smith, E. (1983) Cognitive Deficits In The Early Stages Of Parkinson's Disease. *Brain*, 257–270.

Leeson, V.C., Robbins, T.W., Matheson, E., Hutton, S.B., Ron, M.A., Barnes, T.R.E., & Joyce, E.M. (2009) Discrimination learning, reversal, and set-shifting in first-episode schizophrenia: stability over six years and specific associations with medication type and disorganization syndrome. *Biological Psychiatry*, **66**, 586–593.

Lhermitte, F. (1983) "Utilization behaviour" and its relation to lesions of the frontal lobes. *Brain*, **106**, 237–255.

Liu, Y., Wong, T.P., Aarts, M., Rooyakkers, A., Liu, L., Lai, T.W., Wu, D.C., Lu, J., Tymianski, M., Craig, A.M., & Wang, Y.T. (2007) NMDA Receptor Subunits Have Differential Roles in Mediating Excitotoxic Neuronal Death Both In Vitro and In Vivo. *Journal of Neuroscience*, **27**, 2846–2857.

Lovejoy, E. (1966) Analysis of the overlearning reversal effect. *Psychological Review*, **73**, 87–103.

Mackintosh, N.J. (1963) The Effect Of Irrelevant Cues On Reversal Learning In The Rat. *British Journal of Psychology*, **54**, 127–134.

Mackintosh, N.J. (1969) Further analysis of the overtraining reversal effect. *Journal of Comparative Physiological Psychology*, **67**, Suppl:1–Suppl: 18.

Madau, P., Clark, A.G., Neale, S.A., Smith, L.H., Hamilton, N., Thompson, F., Connick, J.H., Belelli, D., & Lambert, J.J. (2009) Org (1), Org (2) and Org (3) potentiate NMDA-receptor mediated responses in vitro. Presented at the Society for Neuroscience Annual Meeting 2009.

Mandler, J.M. & Hooper, W.R. (1967) Overtraining and goal approach strategies in discrimination reversal. *Quarterly Journal of Experimental Psychology*, **19**, 142–149.

Marx, C.E., Keefe, R.S.E., Buchanan, R.W., Hamer, R.M., Kilts, J.D., Bradford, D.W., Strauss, J.L., Naylor, J.C., Payne, V.M., Lieberman, J.A., Savitz, A.J., Leimone, L.A., Dunn, L., Porcu, P., Morrow, A.L., & Shampine, L.J. (2009) Proof-of-Concept Trial with the Neurosteroid Pregnenolone Targeting Cognitive and Negative Symptoms in Schizophrenia. *Neuropsychopharmacology,* **34**, 1885–1903.

Mathis, C., Vogel, E., Cagniard, B., & Criscuolo, F. (1996) The neurosteroid pregnenolone sulfate blocks deficits induced by a competitive NMDA .... *Neuropharmacology,* **35**, 1057–1064.

McAlonan, K. & Brown, V.J. (2003) Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behavioral Brain Research*, **146**, 97–103.

McDannald, M.A., Lucantonio, F., Burke, K.A., Niv, Y., & Schoenbaum, G. (2011) Ventral Striatum and Orbitofrontal Cortex Are Both Required for Model-Based, But Not Model-Free, Reinforcement Learning. *Journal of Neuroscience*, **31**, 2700–2705.

McEnaney, K.W. & Butter, C.M. (1969) Perseveration of responding and nonresponding in monkeys with orbital frontal ablations. *Journal of Comparative Physiological Psychology*, **68**, 558–561.

McFadden, S.L., Zulas, A.L., & Morgan, R.E. (2010) Age-dependent effects of modafinil on acoustic startle and prepulse inhibition in rats. *Behavioral Brain Research*, **208**, 118–123.

Mehta, M., Swainson, R., Ogilvie, A., Sahakian, B.J., & Robbins, T.W. (2001) Improved short-term spatial memory but impaired reversal learning following the dopamine D 2 agonist bromocriptine in human volunteers. *Psychopharmacology*, **159**, 10–20.

Meunier, M., Bachevalier, J., & Mishkin, M. (1997) Effects of orbital frontal and anterior cingulate lesions on object and spatial memory in rhesus monkeys. *Neuropsychologia*, **35**, 999–1015.

Milner, B. (1964) Some Effects of Frontal Lobectomy in Man. In Warren, J.M. & Akert, K. (eds), *The Frontal Granular Cortex and Behavior*. McGraw-Hill, pp. 313–334.

Mishkin, M. (1964) Perseveration of Central Sets after Frontal Lesions in Monkeys. In Warren, J.M. & Akert, K. (eds), *The Frontal Granular Cortex and Behavior*. McGraw-Hill, pp. 219–241.

Montague, P.R., Dolan, R.J., Friston, K.J., & Dayan, P. (2012) Computational psychiatry. *Trends in Cognitive Sciences*, **16**, 72–80.

Morgan, R.E., Crowley, J.M., Smith, R.H., LaRoche, R.B., & Dopheide, M.M. (2007) Modafinil improves attention, inhibitory control, and reaction time in healthy, middle-aged rats. *Pharmacology Biochemistry and Behavior*, **86**, 531–541.

Morgan, S., Grootendorst, P., Lexchin, J., Cunningham, C., & Greyson, D. (2011) The cost of drug development: A systematic review. *Health Policy*, **100**, 4–17.

United Nations (2011) *World Population Prospects*. United Nations Publications, New York.

Nicolle, M. & Baxter, M.G. (2003) Glutamate receptor binding in the frontal cortex and dorsal striatum of aged rats with impaired attentional set-shifting. *European Journal of Neuroscience*, **18**, 3335–3342.

Nikiforuk, A. & Popik, P. (2011) Effects of quetiapine and sertindole on subchronic ketamine-induced deficits in attentional set-shifting in rats. *Psychopharmacology*, **220**, 65–74.

Noonan, M.P., Walton, M.E., Behrens, T.E.J., Sallet, J., Buckley, M.J., & Rushworth, M.F.S. (2010) Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proceedings of the National Academy of Sciences*, **107**, 20547–20552.

O'Neill, M. & Brown, V.J. (2007) The effect of striatal dopamine depletion and the adenosine A2A antagonist KW-6002 on reversal learning in rats. *Neurobiology of Learning and Memory*, **88**, 75–81.

Owen, A.M., Roberts, A.C., Hodges, J.R., Summers, B.A., Polkey, C.E., & Robbins, T.W. (1993) Contrasting mechanisms of impaired attentional set-shifting in patients with frontal lobe damage or Parkinson's disease. *Brain*, **116**, 1159–1175.

Owen, A.M., Roberts, A.C., Polkey, C.E., Sahakian, B.J., & Robbins, T.W. (1991) Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia*, **29**, 993–1006.

Palencia, C.A. & Ragozzino, M.E. (2006) The effect of N-methyl-D-aspartate receptor blockade on acetylcholine efflux in the dorsomedial striatum during response reversal learning. *Neuroscience*, **143**, 671–678.

Parker, M.O., Gaviria, J., Haigh, A., Millington, M.E., Brown, V.J., Combe, F.J., & Brennan, C.H. (2012) Discrimination reversal and attentional sets in zebrafish (Danio rerio). *Behavioral Brain Research*, **232**, 264–268.

Paul, S. & Purdy, R. (1992) Neuroactive steroids. *The FASEB Journal*, **6**, 2311-2322.

Paxinos, G. & Watson, C. (1998) *The Rat Brain in Stereotaxic Coordinates*. Academic Press.

PEBL Developers (2008) A Public Domain Screenshot From a Computerised Version of the Wisconsin Card Sorting Task. *Wikimedia*: http://en.wikipedia.org/wiki/File:WisconsinCardSort.png

Pennartz, C.M.A., van Wingerden, M., & Vinck, M. (2011) Population coding and neural rhythmicity in the orbitofrontal cortex. *Annals of the New York Academy of Sciences*, **1239**, 149–161.

Perry, R.J. & Hodges, J.R. (1999) Attention and executive deficits in Alzheimer's disease. A critical review. *Brain*, **122**, 383–404.

Porsolt, R.D., Bertin, A., & Jalfre, M. (1978) "Behavioural despair" in rats and mice: Strain differences and the effects of imipramine. *European Journal of Pharmacology*, **51**, 291–294.

Porsolt, R.D., Le Pichon, M., & Jalfre, M. (1977) Depression: a new animal model sensitive to antidepressant treatments. *Nature*, **266**, 730–732.

Prut, L. & Belzung, C. (2003) The open field as a paradigm to measure the effects of drugs on anxiety-like behaviors: a review. *European Journal of Pharmacology*, **463**, 3–33.

Ragozzino, M.E., Detrick, S., & Kesner, R.P. (1999) Involvement of the prelimbic-infralimbic areas of the rodent prefrontal cortex in behavioral flexibility for place and response learning. Journal of Neuroscience, **19**, 4585–4594.

Reid, L.S. (1953) The development of noncontinuity behavior through continuity learning. *Journal of Experimental Psychology*, **46**, 107–112.

Robbins, T.W., James, M., Owen, A.M., Sahakian, B.J., Lawrence, A.D., McInnes, L., & Rabbitt, P. (1998) A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: Implications for theories of executive functioning and cognitive aging. *Journal of the International Neuropsychological Society*, **4**, 474–490.

Roberts, A.C., de Salvia, M.A., Wilkinson, L.S., Collins, P., Muir, J.L., Everitt, B.J., & Robbins, T.W. (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.

Roberts, A.C., Robbins, T.W., & Everitt, B.J. (1988) The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *Quarterly Journal of Experimental Psychology*, **40**, 321–341.

Roberts, A.C., Robbins, T.W., Everitt, B.J., & Muir, J.L. (1992) A specific form of cognitive rigidity following excitotoxic lesions of the basal forebrain in marmosets. *Neuroscience*, **47**, 251–264.

Rodefer, J.S. & Nguyen, T.N. (2008) Naltrexone reverses age-induced cognitive deficits in rats. *Neurobiology of Aging*, **29**, 309–313.

Roesch, M.R., Calu, D.J., Esber, G.R., & Schoenbaum, G. (2010) All That Glitters ... Dissociating Attention and Outcome Expectancy From Prediction Errors Signals. *Journal of Neurophysiology*, **104**, 587–595.

Roman, F.S., Alescio-Lautier, B., & Soumireu-Mourat, B. (1996) Age-related learning and memory deficits in odor-reward association in rats. *Neurobiology Aging*, **17**, 31–40.

Rosenzweig, E.S. & Barnes, C.A. (2003) Impact of aging on hippocampal function: plasticity, network dynamics, and cognition. *Progress in Neurobiology*, **69**, 143–179.

Sahakian, B.J. & Owen, A.M. (1992) Computerized assessment in neuropsychiatry using CANTAB: discussion paper. *Journal of the Royal Society of Medicine*, **85**, 399–402.

Sanacora, G., Zarate, C.A., Krystal, J.H., & Manji, H.K. (2008) Targeting the glutamatergic system to develop novel, improved therapeutics for mood disorders. *Nature Reviews Drug Discovery*, **7**, 426–437.

Schmitt, K.C. & Reith, M.E.A. (2011) The atypical stimulant and nootropic modafinil interacts with the dopamine transporter in a different manner than classical cocaine-like inhibitors. *PLoS ONE*, **6**, e25790.

Schoenbaum, G., Nugent, S.L., Saddoris, M.P., & Gallagher, M. (2002) Teaching old rats new tricks: age-related impairments in olfactory reversal learning. *Neurobiology of Aging*, **23**, 555–564.

Schoenbaum, G., Roesch, M.R., Stalnaker, T.A., & Takahashi, Y.K. (2009) A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature Reviews Neuroscience*, **10**, 885–892.

Schoenbaum, G., Setlow, B., Nugent, S.L., Saddoris, M.P., & Gallagher, M. (2003) Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. *Learning & Memory*, **10**, 129–140.

Schoenbaum, G., Setlow, B., Saddoris, M.P., & Gallagher, M. (2006) Encoding changes in orbitofrontal cortex in reversal-impaired aged rats. *Journal of Neurophysiology*, **95**, 1509–1517.

Schoenknecht, P., Olbrich, S., Sander, C., Spindler, P., & Hegerl, U. (2010) Treatment of acute mania with modafinil monotherapy. *Biological Psychiatry*, **67**, e55–e57.

Settlage, P., Butler, R., & Odoi, H. (1956) Perseverative interference in monkeys following bilateral removal of the prefrontal areas. *Journal of General Psychology*, 255–262.

Settlage, P., Zable, M., & Harlow, H.F. (1948) Problem solution by monkeys following bilateral removal of the prefrontal areas; performance on tests requiring contradictory reactions to similar and to identical stimuli. *Journal of Experimental Psychology*, **38**, 50–65.

Shallice, T. & Burgess, P.W. (1991) Deficits in Strategy Application Following Frontal Lobe Damage in Man. *Brain*, **114**, 727–741.

Shepp, B. & Turrisi, F. (1969) Effects of overtraining on the acquisition of intradimensional and extradimensional shifts. *Journal of Experimental Psychology*, **82**, 46–51.

Siegel, S. (1967) Overtraining And Transfer Processes. *Journal of Comparative Physiological Psychology*, **64**, 471–477.

Slamecka, N.J. (1967) *Human Learning and Memory*. Oxford University Press, New York.

Slamecka, N.J. (1968) A methodological analysis of shift paradigms in human discrimination learning. *Psychological Bulletin*, **69**, 423–438.

Steru, L., Chermat, R., Thierry, B., & Simon, P. (1985) The tail suspension test: A new method for screening antidepressants in mice. *Psychopharmacology*, **85**, 367–370.

Sutherland, N.S. & Mackintosh, N.J. (1971) *Mechanisms of Animal Discrimination Learning*, Second edition. Academic Press.

Tait, D.S. & Brown, V.J. (2007) Difficulty overcoming learned non-reward during reversal learning in rats with ibotenic acid lesions of orbital prefrontal cortex. *Annals of the New York Academy of Sciences*, **1121**, 407–420.

Tait, D.S. & Brown, V.J. (2008) Lesions of the basal forebrain impair reversal learning but not shifting of attentional set in rats. *Behavioral Brain Research*, **187**, 100–108.

Tait, D.S., Marston, H., Shahid, M., & Brown, V.J. (2009) Asenapine restores cognitive flexibility in rats with medial prefrontal cortex lesions. *Psychopharmacology*, **202**, 295–306.

Takahashi, Y.K., Roesch, M.R., Stalnaker, T.A., Haney, R.Z., Calu, D.J., Taylor, A.R., Burke, K.A., & Schoenbaum, G. (2009) The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron*, **62**, 269–280.

Takahashi, Y.K., Roesch, M.R., Wilson, R.C., Toreson, K., O'Donnell, P., Niv, Y., & Schoenbaum, G. (2011) Expectancy-related changes in fir-

ing of dopamine neurons depend on orbitofrontal cortex. *Nature Neuroscience*, **14**, 1590–1597.

Teasdale, J.D., Segal, Z.V., Williams, J.M.G., Ridgeway, V.A., Soulsby, J.M., & Lau, M.A. (2000) Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology*, **68**, 615–623.

Thorndike, E.L. (1898) Animal intelligence: An experimental study of the associate processes in animals. *Psychological Review*, **2**, 1–8.

Tighe, T.J. (1965) Effect Of Overtraining On Reversal And Extradimensional Shifts. *Journal of Experimental Psychology*, **70**, 13–17.

Tighe, T.J., Brown, P.L., & Youngs, E.A. (1965) The effect of overtraining on the shift behavior of albino rats. *Psychonomic Science*, **2**, 141–142.

Turner, D.C., Clark, L., Dowson, J., Robbins, T.W., & Sahakian, B.J. (2004a) Modafinil improves cognition and response inhibition in adult attention-deficit/hyperactivity disorder. *Biological Psychiatry*, **55**, 1031–1040.

Turner, D.C., Clark, L., Pomarol-Clotet, E., McKenna, P.J., Robbins, T.W., & Sahakian, B.J. (2004b) Modafinil Improves Cognition and Attentional Set Shifting in Patients with Chronic Schizophrenia. *Neuropsychopharmacology*, **29**, 1363–1373.

Turner, D.C., Robbins, T.W., Clark, L., Aron, A.R., Dowson, J., & Sahakian, B.J. (2003) Cognitive enhancing effects of modafinil in healthy volunteers. *Psychopharmacology*, **165**, 260–269.

Vallée, M., Mayo, W., Darnaudéry, M., Corpechot, C., Young, J., Koehl, M., Le Moal, M., Baulieu, E.E., Robel, P., & Simon, H. (1997) Neurosteroids: deficient cognitive performance in aged rats depends on low pregnenolone sulfate levels in the hippocampus. *Proceedings of the National Academy of Sciences*, **94**, 14865–14870.

Vigotsky, L.S. & Kasanin, J. (1934) Thought In Schizophrenia. *Archives of Neurology and Psychiatry*, **31**, 1063–1077.

Villarreal, D.M., Do, V., Haddad, E., & Derrick, B.E. (2001) NMDA receptor antagonists sustain LTP and spatial memory: active processes mediate LTP decay. *Nature Neuroscience*, **5**, 48–52.

Wallis, J.D. & Kennerley, S.W. (2011) Contrasting reward signals in the orbitofrontal cortex and anterior cingulate cortex. *Annals of the New York Academy of Sciences*, **1239**, 33–42.

Walsh, R.N. & Cummins, R.A. (1976) The open-field test: A critical review. *Psychological Bulletin*, **83**, 482–504.

Walton, M.E., Behrens, T.E.J., Buckley, M.J., Rudebeck, P.H., & Rushworth, M.F.S. (2010) Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron*, **65**, 927–939.

Walton, M.E., Behrens, T.E.J., Noonan, M.P., & Rushworth, M.F.S. (2011) Giving credit where credit is due: orbitofrontal cortex and valuation in an uncertain world. *Annals of the New York Academy of Sciences*, **1239**, 14–24.

Waters, K.A., Burnham, K.E., O'connor, D., Dawson, G.R., & Dias, R. (2005) Assessment of modafinil on attentional processes in a five-choice serial reaction time test in the rat. *Journal of Psychopharmacology (Oxford)*, **19**, 149–158.

Watkins, E. & Teasdale, J.D. (2001) Rumination and overgeneral memory in depression: effects of self-focus and analytic thinking. *Journal of Abnormal Psychology*, **110**, 353–357.

Weigl, E. (1941) On the psychology of so-called processes of abstraction. *Journal of Abnormal and Social Psychology*, **36**, 3–33.

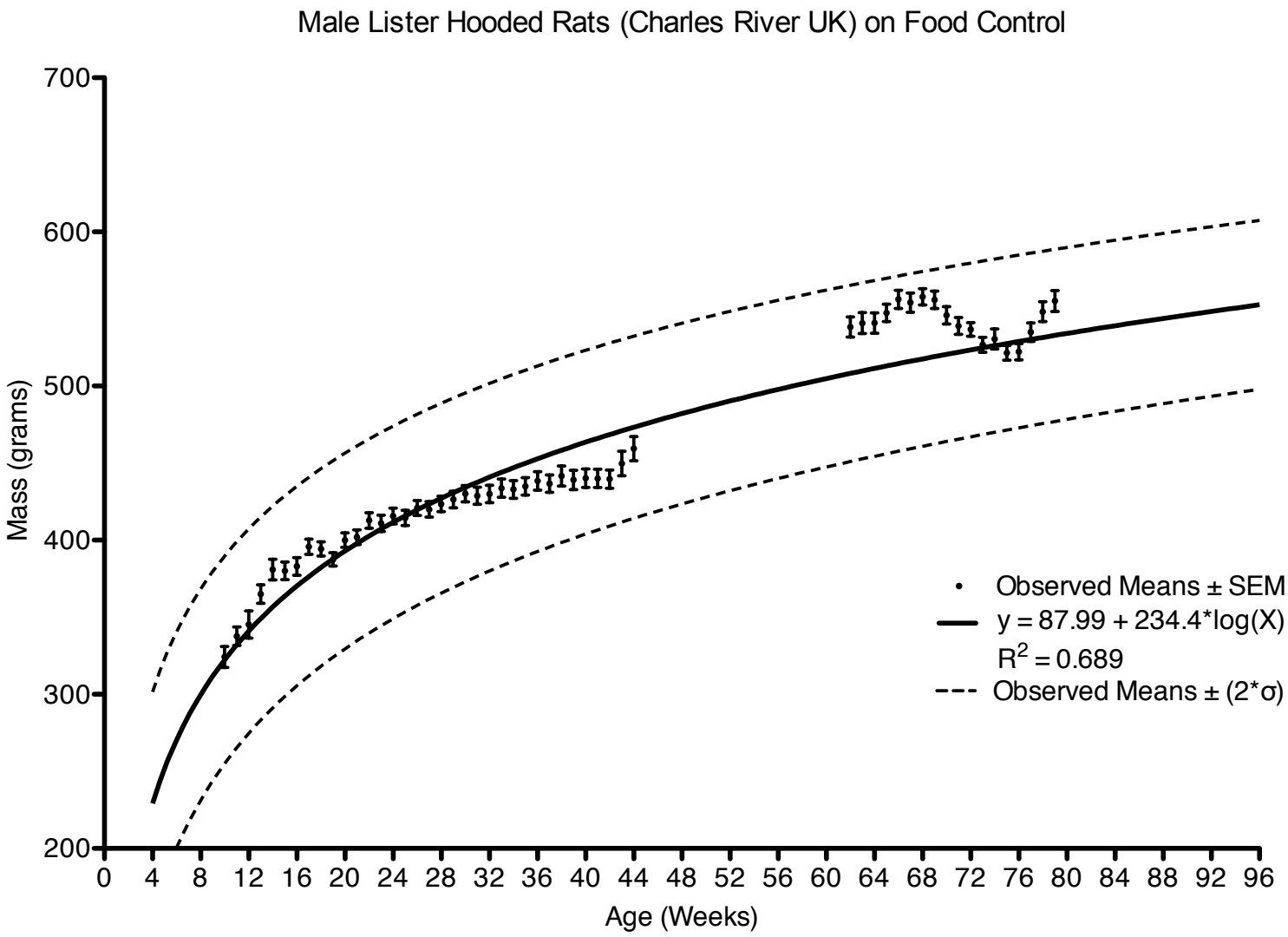
Wilson, R.C. & Niv, Y. (2012) Inferring relevance in a changing world. *Frontiers in Human Neuroscience*, **5**, 189.

Winer, B.J. (1971) *Statistical Principles in Experimental Design*, Second. edition. McGraw-Hill Book Company, New York.

Woo, T.-U.W., Walsh, J.P., & Benes, F.M. (2004) Density of glutamic acid decarboxylase 67 messenger RNA-containing neurons that express the N-methyl-D-aspartate receptor subunit NR2A in the anterior cingulate cortex in schizophrenia and bipolar disorder. *Archives of General Psychiatry*, **61**, 649–657.

Yu, A.J. & Dayan, P. (2005) Uncertainty, neuromodulation, and attention. *Neuron*, **46**, 681–692.

Zahr, N.M., Mayer, D., Pfefferbaum, A., & Sullivan, E.V. (2008) Low Striatal Glutamate Levels Underlie Cognitive Decline in the Elderly: Evidence from In Vivo Molecular Spectroscopy. *Cerebral Cortex*, **18**, 2241–2250.


Male Lister Hooded Rats (Charles River UK) on Food Control
Mass (grams)
700
600
500
400
300
200
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60 64 68 72 76 80 84 88 92 96
Age (Weeks)
Observed Means ± SEM
y = 87.99 + 234.4*log(X)
R2 = 0.689
Observed Means ± (2*σ)