Using consummatory behaviour to measure the affective state and welfare of laboratory mice

Jasmine Maria Clarkson

A thesis submitted for the degree of

Doctor of Philosophy

Institute of Neuroscience

Newcastle University

September 2018

Abstract

Mice are the most widely used model in biomedical research, making it important to know how life in the laboratory impacts on their welfare. Whilst it is difficult to evaluate how a mouse might 'feel' because of their inability to self-report, behavioural and physiological measures can give insights into their current welfare state.

One approach is to measure their 'affective state' through how they respond to reward. Humans with depression often report a lack of enjoyment from reward (known as anhedonia), which can also be measured in laboratory rodents by measuring sucrose consumption: the less they consume, the more anhedonic they are. However, consumption is confounded by other factors, particularly motivational state. Therefore, an alternative method assesses the microstructure of the animals' licking patterns, which may better reflect an animal's hedonic response towards reward, i.e. how much it 'likes' it.

The aim of my thesis was to determine how stress influences the hedonic responses of laboratory mice, and determine whether assessing changes in consumption or licking microstructure could be used to infer a mouse's affective state, in order to make evidence-based improvements to their welfare.

Experiments using standard depressogenic methods (i.e. chronic mild stress and chronic corticosterone administration) were ineffective at altering affective state, and sucrose consumption and licking microstructure were unchanged. However, I found that current methods used to handle laboratory mice were sufficient in inducing changes in the animal's affective state. I found that the standard practice of handling mice by their tails causes alterations in reward perception, revealing a depressive-like state. These experiments provide more support for refinements to be made with regards to the existing handling practices of laboratory mice. I discuss my findings in relation to implications for animal welfare and scientific data collection across a number of *in-vivo* models.

Acknowledgements

First I would like to take the opportunity to thank my supervisory team Candy Rowe, Matt Leach and Paul Flecknell, for their continued support, guidance and advice throughout my PhD. I would also like to thank Dominic Dwyer from Cardiff University, for his help and guidance with the set-up and running of the drinking apparatus. A special thanks go to Candy for all her help with the many drafts of this thesis, even working on her holiday, so for that, and everything else across these past four years, a huge thank you!

I would also like to thank everyone in CBE and ION, for creating a research environment that inspires good science and has been invaluable to my own research. Thank you to all CBC staff at Newcastle, in particular a special thanks to Michelle Waddle for all her help with the care of my mice, and for generally making my time in the lab much easier. I would also like to thank the BBSRC for funding my research and the Fibrosis research group for their help with sampling the physiological data.

My final acknowledgments will be dedicated to the people closest to me. First to Grace, I am so grateful to have had your friendship for the past four years, but for this last year especially, I would like to say a huge thank you. You truly have kept me sane knowing that I wasn't the only one going through this. Finally, I would like to take the opportunity to thank both my parents for their continued support throughout my eight years at University, despite not really understanding what, or why I am doing this! Thank you to my grandma and grandad, for their constant support, and whom I am very lucky to have. My granny, for her supportive messages and food parcels, thank you! Finally, a special thanks go to my sister, Bryony, I am so glad that you decided to come to Newcastle!

Table of Contents

List of Figures

List of Tables

Chapter 1: Licking microstructure and its potential to measure the affective state of laboratory mice

1.1 Stress and its effect on animal welfare

The world we live in today can be a stressful place, with an increasing number of us feeling the pressures of everyday life. It is common for individuals to worry about a whole host of everyday problems, whether it be worrying about money or making sure we have an appropriate work-life balance, or even spending enough time with family and friends. Although stress is a normal part of everyday life, it can often become too much for some individuals, affecting their ability to cope and ultimately resulting in poor mental health and wellbeing. Mental health refers to a person's emotional, psychological and social well-being, where the inability to manage and cope with daily stressors can result in the diagnosis of a mental health disorder (MentalHealth.gov 2017). Depression is one of the most common mental health disorders worldwide (Vos et al. 2015). It is characterised by a number of symptoms, ranging from low mood, loss of appetite, insomnia, suicidal thoughts, as well as feelings of guilt and worthlessness (Fava & Kendler 2000). Currently in the UK, it is estimated that approximately 3% of people are suffering from depression, and treatments are becoming a huge economic burden (Gustavsson et al. 2011; McManus et al. 2016). A large body of research now exists dedicated to understanding the aetiology, symptomatology and treatment of depression in order to improve the everyday lives of patients, and reduce the economic cost to society.

Experiencing stress is not unique to us humans: animals are also subject to a wide array of stressors in their natural environments. The term stress is often defined differently by researchers, but is generally considered to be a state that is induced by a stressor that challenges homeostasis, affecting an organism's ability to survive and reproduce (Ramos & Mormède 1998; Moberg 2000; Paul et al. 2005). Stressors might include not finding enough of the right food to eat, fear of predation or even conflicts between conspecifics (Reeder & Kramer 2005; Holmstrup et al. 2010). However, manmade environments, such as zoos, farms and laboratories, whilst reducing the risks of predation and starvation, produce their own stressors to which animals may not be adapted (Morgan & Tromborg 2006). These stressors include: artificial lighting, limited space and resources, living in unnaturally large groups and

overcrowding, increased human interaction, and the restriction of species-specific behaviours (Morgan & Tromborg 2006). Of course, the exposure to stressors is generally considered to be aversive, potentially leading to negative affective states and poor animal welfare (Dawkins 1990; Paul et al. 2005; Dawkins 2008). However, due to the inability to self-report, it is difficult to ascertain exactly how an animal might perceive a given stressor, and how this might affect its state and wellbeing. This means that we need to develop alternative measures if we are to make evidence-based improvements to their environment, experiences and welfare.

My work focuses on the ability to effectively measure and make evidence-based improvements to the welfare of laboratory mice (*Mus musculus*). Laboratories are stressful environments where improving animal welfare remains a priority (Carstens & Moberg 2000). Laboratories not only subject animals to unnatural housing conditions, but also expose them to experimental procedures. However, animal use in scientific experiments often remains necessary (Malakoff 2000). In biomedical research, mice are the most widely used model species. To give an idea of the scale in which they are used, the most recent figures on animal usage show that mice were used in approximately 60% of all experimental procedures for 2016 in the UK (Home Office 2017), and also accounted for approximately 60% of the total animal use across Europe (European Commission 2013). If these figures were an accurate reflection of mouse use on a global scale, this would equate to approximately 70 million mice being used for research each year (Taylor et al. 2008). Their widespread use is largely due to the development of Genetically Modified (GM) mice, which model a number of human diseases, underpinned by advances in molecular genetics to generate knockout and transgenic mice. They also have a short gestational period, quick sexual maturation, and can be kept in large numbers at relatively low cost, meaning that the breeding and maintenance of mouse colonies can be easily achieved and make them conducive to life in the laboratory (Malakoff 2000; Bućan & Abel 2002; Perlman 2016). However, there are significant welfare concerns around their husbandry conditions and the presence of unnecessary pain and suffering inflicted through experimental procedures. Biomedical scientists, animal care staff, regulators and funders all have a duty of care to ensure the best standards of care are provided to animals used in scientific procedures and ensure the appropriate refinements are made. This is reflected in a number of legislative

requirements and frameworks, such the Animals used in Scientific Procedures Act (ASPA) 1986 and the EU Directive (2010/63/EU).

One significant framework which is now embedded in the legislation, regulation, funding and design of *in-vivo* studies, is the framework of the 3Rs. This framework was developed over 50 years ago for conducting humane animal research (Russell & Burch 1959). The 3Rs are: Reduction, which involves methods to minimise the numbers of animals used in scientific research; Replacement, which describes methods to avoid or replace the use of animals for scientific research; and Refinement, which describes methods to minimise pain, suffering, distress or lasting harm, and subsequently make improvements to animal welfare. Refinement does not only apply to making improvements to experimental procedures, but can also include making improvements to an animal's environment and husbandry practices. Examples of refinement include using appropriate analgesia and anaesthesia to minimise pain and unnecessary suffering, or making improvements to housing conditions that allow the expression of species-specific behaviours (Russell & Burch 1959; Tannenbaum & Bennett 2015). An important part of being able to make appropriate refinements, involves the ability to effectively measure the welfare of laboratory animals in order to make evidence-based improvements.

1.2 Physiological and behavioural measures of welfare in animals

Measuring changes in a number of relatively simple physiological and behavioural responses have provided a starting point for making inferences about an animal's welfare. Stressors produce well defined physiological responses, which can be measured scientifically and objectively. For example, it is well known that stressors can immediately increase heart rate and blood pressure due to the release of adrenaline via the Sympathomedullary (SAM) pathway, and longer-term elevations in the glucocorticoid cortisol can occur via the activation of the Hypothalamic-Pituitary-Adrenal (HPA) system (Tsigos et al. 2000). However, although these physiological changes can be quantified, they aren't without their limitations in terms of being able to measure animal welfare. One major limitation is that both positive and negative stimuli can yield the same physiological responses (Rushen 1986; Rushen 1991). For example, exposure to a predator or to a reward can increase blood pressure, heart rate, and circulating glucocorticoid levels (Chabot et al. 1996;

Paul et al. 2005; Zupana et al. 2015; Starcevic et al. 2016). Therefore, it is impossible to know whether an animal responds to a given stimulus or situation positively or negatively using only these measures, making it impossible to fully conclude the impact it has on their welfare. There are also a number of methodological issues around taking physiological measures. For example, the timing at which the sample is taken can significantly affect the results since steroid hormones, and in particular glucocorticoids, show circadian variation and pulsatile secretion patterns (Halberg et al. 1959; Axelrod & Reisine 1984). In addition, the sampling method often uses invasive techniques to try and collect accurate data; for example, in mice the most common way to obtain a blood sample is to make an incision in the tail vein (Morton et al. 1993). Therefore, the stress associated with the sampling method can confound data collection by having effects on circulating neuroendocrine levels (Halberg et al. 1959; Hennessy & Levine 1978; Gärtner et al. 1980; Quirce & Maickel 1981; Riley 1981; Armario et al. 1986; Haemisch et al. 1999; Touma et al. 2004), ultimately masking information about their experiences and welfare. Despite their potential as objective measures that can be measured scientifically, assessing changes in physiological responses alone, cannot indicate the welfare of an animal.

An alternative method is to observe an animal's behaviour to evaluate how stressed it is and make inferences about its welfare. Examining relatively simplistic and spontaneous behaviours, such as approach or avoidance, feeding, social behaviour, general activity, or vocalisations have all helped to reveal the welfare state of an animal when exposed to potentially stressful situations (Mason & Latham 2004; Webster & Fletcher 2004; Jeong et al. 2013; da Silva Cordeiro et al. 2013; Vos et al. 2015; Beery & Kaufer 2015; Tarantola et al. 2016). For example, mice that are restrained reduce their exploratory behaviour (Berridge & Dunn 1989), and there are observable changes in bodyweight and composition due to alterations in their feeding behaviour (Jeong et al. 2013). However, just like the physiological measures, assessing spontaneously occurring behaviours in animals aren't without their limitations. One problem is that often interpreting an animal's welfare from these relatively simplistic behaviours relies on the observer's evaluation of the current context in which the animal is in, and can be subjective; for example, a situation we would find aversive and threatening might be very different to what an animal does (Paul et al. 2005). In addition, it is difficult to infer whether a situation is perceived as positive or negative from assessing these behaviours alone because it is not

uncommon for the same behaviours to be elicited from very different situations: approach behaviour can occur towards predators (Humphrey & Keeble 1974; Krams & Krama 2002; Walling et al. 2004) and rewarding stimuli (Tanimoto et al. 2004). Therefore, as with physiological measures, we cannot solely use these relatively simplistic measures to ensure 'good' animal welfare. Instead, this has led to a shift in focus on the development of tests, taking into account more subjective components of an animal's welfare, such as how the animal might 'feel'.

1.3 The importance of understanding animal affect for assessing welfare

An important part of being able to accurately assess an animal's welfare is the ability to understand how they might 'feel'. In particular, concern for animal welfare largely comes from the worry that an animal might be 'suffering' (Dawkins 1990; Dawkins 2008; Weary 2014). When applied to humans, suffering is predominantly considered to reflect negative affect associated with unpleasant affective states such as fear, boredom, pain, hunger or even frustration (Dawkins 1990; Dawkins 2008; Weary 2014). Of course, when we consider suffering in humans, it has a conscious subjective component. For example, we know how we feel when we are suffering from fear, or when we are suffering from boredom, and how these two differ from each other (Dawkins 2008). However, due to their inability for self-report we cannot be certain whether or not, and to what extent, an animal might consciously experience or feel (Dawkins 1990; Dawkins 2008; Mendl, Burman, et al. 2010). Despite this, animals are considered capable of experiencing negative, unpleasant affective states, regardless of their consciousness, and are therefore considered capable of suffering (Dawkins 1990; Dawkins 2008; Mendl, Burman, et al. 2010).

Much of the research studying animal affect has predominantly focused on the ability to assess 'discrete' or 'basic' emotions (Mendl, Burman, et al. 2010). For example, often an animal is exposed to a negative threatening situation and its responses are considered to reflect the discrete emotion of fear (Mendl, Burman, et al. 2010). These discrete emotions are considered to have evolved due to their adaptive value, arising in situations where their primary function is to enable appropriate behavioural decisions to promote survival and reproductive success (Mendl et al. 2009; Mendl, Burman, et al. 2010; Nettle & Bateson 2012). Therefore, the ability to be in either a positive or a negative state can have an adaptive function:

these states have evolved from basic mechanisms that gave animals the ability to avoid harm and punishment, seek reward and resources, and ultimately minimise the presence of negative states associated with suffering (Dawkins 2008; Nettle & Bateson 2012).

More recently, Mendl and colleagues have developed a framework integrating these discrete emotions with 'dimensional' theories of emotion, which can be applied for studying animal affect (Mendl, Burman, et al. 2010). Dimensional theories of emotion have largely been applied to the study of human emotion, where an individual's subjective emotional experience can be represented in two dimensional space (Figure 1.1). This is made up of a 'valence' dimension (i.e. how pleasant or unpleasant something is), and an 'arousal' dimension (i.e. how stimulating something is). Subjective experiences characterised in terms of their valence and arousal are considered an individual's 'core affect' (Russell 2003; Barrett et al. 2007; Mendl, Burman, et al. 2010). The right half of the space represents positive affective states with low or high arousal, and negative affective states with high or low arousal lie on the left hand side of the space (Figure 1.1). Therefore, by conceptualising core affect in this way, it provides a structure for understanding an individual's subjective emotional experience which can be accompanied by behavioural, physiological and cognitive changes (Mendl, Burman, et al. 2010). It has been suggested that this framework can be used to infer an animal's position in core affect space and from this, infer their discrete affective states (i.e. fear, anxiety, etc; Figure 1.1) (Mendl, Burman, et al. 2010). As a result, the development of this framework has provided a structure for generating predictions about the behavioural, physiological and/or cognitive changes that occur with certain affective states, and has facilitated the development of novel measures of these states in animals.

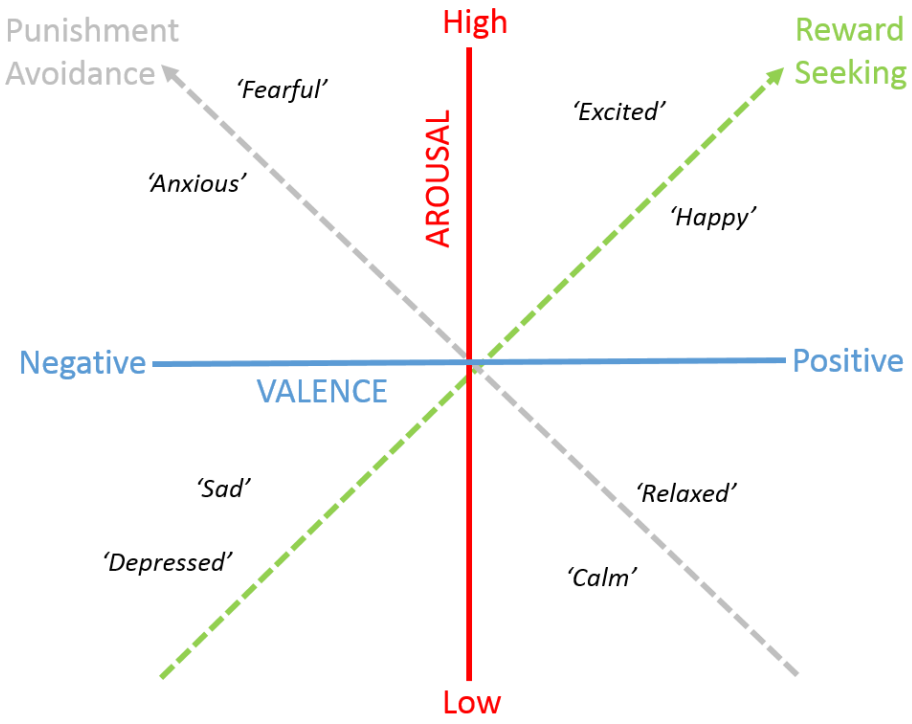


Figure 1.1: Schematic adapted from Mendl et al. (2010) which represents core affect in a two dimensional space with axes of valence and arousal. Words in italics represent basic (discrete) emotions and their place within this framework.

Measuring changes in cognitive processing in animals that are known to be influenced by emotional states in humans has gained significant traction in recent years (Hinde 1985; Forgas 2000; Mathews & Macleod 2002; Harding et al. 2004; Paul et al. 2005; Mendl et al. 2009). This comes from research on human patients that has found that manipulations causing changes in affect (either induced experimentally, or clinical in nature) change the cognitive processing of information (Beck et al. 1979; Bower 1981; Mathews & Macleod 2002). These 'cognitive biases' in how information is processed can be grouped into three main categories: attention biases, memory biases, and judgement biases (Paul et al. 2005). This is because, in humans at least, an individual's affective state can influence their attention towards, perception and memory of, and subsequent judgements and decision-making when exposed to a positive or negative stimulus (Mineka & Sutton 1992; Wright & Bower 1992; Mathews & MacLeod 1994; MacLeod & Byrne 1996; Nygren et al. 1996; Mathews & Macleod 2002; Paul et al. 2005). For example, people in a negative affective state often pay more attention to aversive or threatening stimuli such as angry faces or negative words, and recall negative life events more readily than

people in a more positive affective state (Mineka & Sutton 1992; Mathews & MacLeod 1994). They are also more likely to make negative judgements and be more pessimistic when presented with an ambiguous stimulus compared to people in a more positive affective state (Wright & Bower 1992; MacLeod & Byrne 1996; Nygren et al. 1996). Furthermore, it has now been suggested that the same changes in cognitive processing appear to occur in animals, offering a more direct approach for accessing their underlying affective state (Harding et al. 2004; Paul et al. 2005; Mendl et al. 2009). Assessing these changes in non-human animals offers more insight into how both dimensions of core affect are altered, and therefore considered advantageous over existing behavioural and physiological methods. Cognitive biases can measure the valence of a response towards a stimulus rather than just the degree of arousal it elicits (Figure 1.1). This is important because it enables researchers to make and test hypothesis-driven predictions about an animal's underlying affective state, disentangling arousal from valence.

Cognitive bias studies in animals have focused on assessing judgement biases, i.e. the tendency to show behaviour indicating anticipation of either positive or negative outcomes in response to an ambiguous stimulus (Harding et al. 2004; Bateson & Matheson 2007; Bateson et al. 2011; Brydges et al. 2011). In the first study of its kind looking at judgement biases in rats (Harding et al. 2004), subjects were trained to lever press when presented with a tone that represented a positive event (a reward), and avoid pressing the lever when presented with a tone representing a negative event (no reward and white noise). Once trained, they were presented with intermediate ambiguous stimuli (i.e. sounds lying between the positive and negative stimuli) to see whether they judged them to be more similar to the rewarded or unrewarded stimulus. The study showed that rats housed in 'unpredictable' housing conditions (considered sufficient to induce a negative depressive-like state) were more likely to judge the ambiguous stimulus as being more similar to the unrewarded negative stimulus compared to rats kept under standard housing conditions, i.e. they had more pessimistic judgments (Harding et al. 2004). This study pioneered a new way to measure the valence of affect in animals. To date, these cognitive biases have been studied in a welfare context across a range of species including rats (*Rattus norvegicus*), dogs (*Canis lupus familiaris*), sheep (*Ovis aries*), starlings (*Sturnus vulgaris*) and even honeybees (*Apis mellifera*) (Harding et al. 2004; Bateson & Matheson 2007; Doyle et al. 2010; Brilot et al. 2010;

Bateson et al. 2011; Brydges et al. 2011). However, despite the sensitivity of these tests to measure valence, to date, there is no cognitive bias test validated for use with laboratory mice. There are, however, a wide range of behavioural tests that have been developed to determine the affective state of laboratory mice.

1.4 Measuring affective state in laboratory mice

Although there are no validated cognitive bias tasks for mice, there are a range of behavioural tests which are used to infer the affective state of laboratory mice, particularly aimed at measuring negative states such as fear, anxiety, and depression (Steru et al. 1985; Strekalova et al. 2004; Chiba et al. 2012; Jindal et al. 2013). However, these tests have not been primarily developed for use in a welfare context, but rather in the development and validation of mouse models of depression used to develop novel anti-depressant treatments for clinical use.

Mouse models of depression use a variety of different methods to induce a depressed-like state in laboratory mice. These include pharmacological (e.g. chronic corticosterone, reserpine, tryptophan and psychostimulant withdrawal) and genetic (genetically modified mice) approaches (Barr et al. 2002; Urani et al. 2005; Barr & Markou 2005; Chourbaji et al. 2006; Ardayfio & Kim 2006; Antkiewicz-Michaluk et al. 2014), although the most common method is through the use of stress-inducing manipulations, for example, chronic mild stress, social stress, early life stress and restraint stress (Rygula et al. 2005; Schmidt et al. 2011; Chu et al. 2016; Willner 2017b). These manipulations expose the animal to a variety of different stressors, either acutely or chronically, until they display behavioural and/or physiological changes indicative of a depressed-like phenotype. Due to the co-occurrence of anxiety in humans with depression (Hirschfeld 2001), a depressed-like phenotype in laboratory mice is taken to include increased anxiety, fear and/or despair as well as any other symptomatology associated with depression in humans, such as changes in reward sensitivity (Deussing 2006). Therefore a number of behavioural tests have been developed in order to detect the presence of these negative affective, or depressive-like states in laboratory rodents (Deussing 2006).

One class of behavioural tests use despair-based paradigms. One example is the forced swim test, which involves placing a mouse in a large inescapable

cylinder filled with water and measuring the time taken to exhibit ‘behavioural despair’, i.e. adopting an immobile posture by simply ‘floating’ in the water tank (Porsolt et al. 1977). Another is the tail suspension test, which works on similar assumptions, but this time the mouse is held by its tail until it becomes immobile (Steru et al. 1985). For both of these tests, the idea is that mice in more negative affective states will exhibit behavioural despair quicker than animals in more positive affective states. Although reproducible results have been demonstrated across different laboratories, these tests aren’t without a number of disadvantages including ethical concerns such as a risk of hypothermia for the forced swim test, and that the tail suspension test can only be used for certain strains of mice because some strains have the ability to climb their tails (Nestler & Hyman 2010). Therefore, despite their use for testing the efficacy of novel pharmacological compounds, due to the ethical concerns and strain differences means they are often unsuitable for use in a welfare context.

There are also anxiety-based tests such as the Open Field (OF) and Elevated Plus Maze (EPM), which work on the premise that rodents behave differently to threatening situations depending upon their affective state. Rodents in negative affective states tend to be more fearful of open spaces and therefore avoid large, open and unprotected spaces where risk of predation might be increased (Hall & Ballachey 1932; Hall 1934; Pellow et al. 1985; Gould et al. 2009). The OF consists of an empty arena which the animal is allowed to freely explore; animals in a negative (anxious) affective state will spend significantly less time in the centre region of the arena and much more time close to the walls (i.e. show more thigmotaxis) than a rodent in a more positive (less anxious) affective state (Hall & Ballachey 1932; Hall 1934). The EPM consists of two open, unprotected arms and two closed protected arms elevated from the floor. Once again, mice in more negative affective states will spend more time in the closed protected arms of the EPM and significantly less time on the open, unprotected arms of the EPM than mice that are in a more positive affective state (Pellow et al. 1985; Pellow & File 1986). Although widely used, these tests also have their limitations. Notably, unreliable results are often obtained when antidepressants are administered, and it is not clear whether behavioural responses reflect changes in affect or instead could be explained by differences in locomotion and novelty-seeking behaviours (Deussing 2006). This means that it is often difficult to conclude that these do measure anxiety, and these tests should not be employed

alone, but in combination with others, in order to make conclusions regarding the affective state of laboratory rodents.

Finally, reward-based paradigms are also widely used, due to findings from depressed human patients who report changes in their sensitivity towards reward (American Psychiatric Association 2014). Although these tests were predominantly developed to assess the reinforcing and addictive properties of drugs and identify compounds that are liable to lead to abuse in humans, they have also been used to evaluate the affective properties of drugs in animals (Bardo & Bevins 2000; Spiteri & Le Pape 2000; Panlilio & Goldberg 2007; Simon O'Brien et al. 2011; Roughan et al. 2014). One such test is the Conditioned Place Preference (CPP) test, which works by conditioning the animal to one place which is paired with a drug treatment, and another place that acts as a control by pairing it with saline or nothing at all (Prus et al. 2009). Another is a self-administration paradigm, where animals associate a drug treatment with an instrumental response (e.g. nose-poke, lever press) to obtain the drug (June & Gilpin 2010; Simon O'Brien et al. 2011). Over time, an association forms between the administered drug with either the given context or the instrumental response, and if the drug is positively reinforcing then the animal will spend more time in the place it experienced it or will work progressively harder to obtain the reward (i.e. nose-poke or lever press more) (Bardo & Bevins 2000; Spiteri & Le Pape 2000; Panlilio & Goldberg 2007; Simon O'Brien et al. 2011; Roughan et al. 2014). The degree to which the animals display these behaviours is interpreted as showing how rewarding they find a compound: if a compound improves an animal's affective state, it will spend more time actively trying to obtain it.

However, the most widely documented reward-based paradigm is the sucrose preference or consumption test, which is considered to be indicative of an anhedonic-like symptomatology in rodents (Willner et al. 1987; Papp et al. 1991; Monleon-Paolo et al. 1995; Forbes et al. 1996). One of the core symptoms of human depression is anhedonia, defined as the loss or inability to experience pleasure from a rewarding stimulus (American Psychiatric Association 2014). Anhedonia can present itself in many different forms, including the disengagement with rewarding activities such as: socialising with friends, sexual relationships, or causing changes in a person's appetite and diet. Due to the lack of an animal's ability for self-report, an animal's disinterest in feeding have been assumed to reflect an anhedonic-like state (Willner et al. 1987; Papp et al. 1991; Willner et al. 1992). The sucrose preference or

consumption test builds on the idea that animals in a depressive-like state will be less motivated and get less pleasure from a rewarding tastant such as sucrose; animals are expected to drink less sucrose than 'normal' healthy rodent controls, and show less of a preference for sucrose when given a choice between water and a sucrose solution (Willner et al. 1987; Papp et al. 1991; Monleon-Paolo et al. 1995; Forbes et al. 1996; Nestler & Hyman 2010; Willner 2017b). Although the sucrose preference test has greater face validity (i.e. the ability to exhibit the symptoms to the disease or condition being modelled) to human depression than some of the tests previously described (Deussing 2006; Nestler & Hyman 2010), the possible role of motivation in these tests cannot be overlooked. Assessing total intake is simply an endpoint measure: it is impossible to know whether intake solely reflects how pleasurable the sucrose is and how much they 'like' it, or whether it might be influenced by other factors such as post-ingestive consequences, physiological deficits or motivational differences, affecting how much they 'want' it (Dwyer 2012). Therefore, the ability to dissociate between 'wanting' and 'liking' is important in order to understand the mechanisms underpinning consumption. It has been suggested that behavioural expressions of 'wanting' and 'liking' are driven by different neural systems and that both are required to experience reward (Berridge 1996). How hard an animal might work to obtain a given resource (i.e. how much they want it) does not solely reflect the core positive affect (i.e. how much they like it) when accessing it (Paul et al. 2005). As a result, assessing how much an animal likes something (their hedonic responses) offers the potential to provide more information as to an animal's core affective state, in terms of its valence and not just its arousal. Consequently, alternative measures have been developed to better understand how the hedonic mechanisms, and how much they like a tastant, are related to and drive total consumption.

1.5 Measuring the hedonic responses of laboratory rodents

Behaviours associated with consumption offer alternative methods to provide more valid measures of the hedonic responses of animals rather than just assessing total intake. Grill and Norgren (1978) were the first to show that when rats were exposed to different tastants, they produced stereotypical orofacial 'taste reactivity' responses. When exposed to a palatable tastant, like sucrose or saccharin,

distinctive appetitive responses were observed, such as lip smacking and lateral tongue movements and protrusions. On the other hand, when exposed to unpalatable tastants such as quinine, a rejection pattern of responses are observed including aversive gapes, chin rubbing and grimaces (Berridge 2000; Grill & Norgren 1978b). These responses are considered to be viable measures of an animal's hedonic response for a number of reasons. First, numerous studies have now shown that both palatable and unpalatable tastants yield very different behaviours that can be quantified to provide more detailed behavioural measures compared to just assessing total intake (for review see Dwyer 2012). Second, these responses are not restricted to rats, but are taxonomically widespread and are highly conserved across species: notably, they are also found in humans, including day old babies, but also in other apes and monkeys (Berridge 2000). Finally, these behaviours have also been directly linked to reward pathways in the brain, activating hedonic hotspots in the nucleus accumbens and ventral pallidum (Peciña et al. 2006). However, whilst a valuable measure of hedonic responses to different tastants, orofacial responses are labour-intensive, with each animal's behaviour having to be filmed and then scored retrospectively on a frame-by-frame basis. Therefore, given this limitation there was the requirement for less time consuming and labour intensive methodologies to be developed.

Consequently, an automated methodology was developed to assess the hedonic responses of laboratory rodents when drinking tastants through their licking microstructure. When rodents drink, they produce rapid, rhythmic sets of licks that can be grouped into bouts, which are separated from each other by intervals of varying duration (Davis 1973; Davis & Smith 1992; Smith 2001; Dwyer 2012). The mean number of licks in a bout, referred to as the lick cluster size, can give an indication of an animal's hedonic response to the solution being consumed (Dwyer 2012; Austen et al. 2016). This is because the number of licks in each bout is not random. Instead, they are directly related to the tastant being experienced. Lick cluster size increases with increasing concentration of palatable solutions like sucrose or saccharin (refer to Figure 1.2A), and also decreases with increasing concentration of unpalatable solutions, such as quinine (Davis 1973; Davis 1989; Davis & Smith 1992; Hsiao & Fan 1993; Spector & St. John 1998; Spector et al. 1998; Dwyer 2012). This monotonic relationship is different to the one that exists between total consumption and tastant concentration. This is because the

relationship between consumption and sucrose concentration follows more of a bell-shaped curve (Figure 1.2B), with the highest levels of consumption occurring at intermediate concentrations (Richter & Campbell 1940; Davis 1973; Spector et al. 1984). This bell-shaped relationship is driven by post-ingestive consequences: subjects will drink less of a weak solution because it is not rewarding, but also less of a highly concentrated sucrose solution because of its higher caloric content and greater satiating properties (Richter & Campbell 1940; Davis 1973; Spector et al. 1984). Therefore, subjects can drink the same amount at both low and high concentrations of sucrose, which despite tasting very different and having different nutritional qualities, can still elicit the same quantitative response. This makes licking microstructure a better measure across a range of sucrose concentrations.

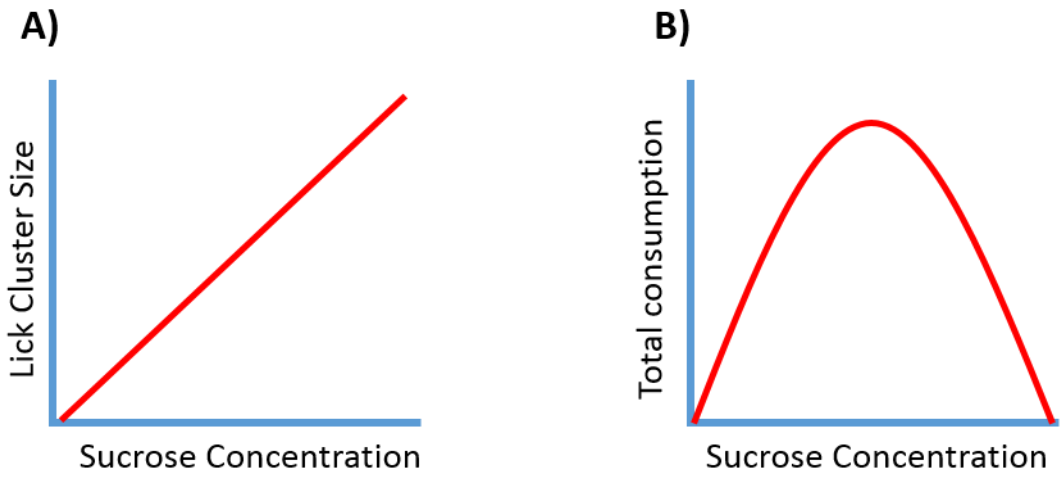


Figure 1.2: Schematic to illustrate the relationship between increasing sucrose concentration with A. Lick cluster size (Davis & Smith 1992; Spector et al. 1998) and B. Total consumption (Davis 1973; Spector et al. 1984; Richter & Campbell 1940).

However, lick cluster size is not solely dependent upon the tastant itself, but is also affected by other factors including learning, memory and prior experience (Dwyer et al. 2009; Dwyer et al. 2011; Dwyer et al. 2013; Austen et al. 2016; Austen & Sanderson 2016). For humans, we know that the palatability of a given food or drink can be influenced by past experience, for example, people often avoid the taste of a given food or drink that has been previously associated with illness (Garb & Stunkard 1974). This is also the case in rats: if a palatable tastant is paired with an aversive stimulus, such as lithium chloride (a nausea-inducing agent), then the lick

cluster sizes obtained will be reduced and become similar to those elicited by an unpalatable tastant such as quinine (Baird et al. 2005; Dwyer 2009). This result parallels that seen when assessing taste reactivity responses, for example, pairing sucrose with lithium chloride also elicits orofacial responses typically seen with quinine (Pelchat et al. 1983). These changes in orofacial behaviour and lick cluster size are considered to be attributable to changes in hedonic mechanisms. This is because studies have shown that when benzodiazepine is administered, which enhances hedonic responses to food in humans (Haney et al. 1997), rats tend to have larger lick cluster sizes and more appetitive orofacial responses (Gray & Cooper 1995; Higgs & Cooper 1998). The fact that manipulations which alter hedonic responses affect both lick cluster size and oro-facial responses in the same way, and in opposing directions, suggests that these two parameters can be considered equivalent in reflecting an animal's hedonic response to a given tastant.

However, despite licking microstructure being a valid method of accessing a rodent's hedonic responses, it has yet to be applied in a welfare context. This is because its use has focused on learning more about the processes underlying learning and memory in the field of experimental psychology (Dwyer et al. 2009; Dwyer et al. 2011; Dwyer et al. 2013; Austen et al. 2016). Therefore, my thesis explores whether or not licking microstructure can be used as a novel method to assess the affective state of laboratory mice, and consequently, be used to help improve their welfare.

1.6 Measuring positive affective states and its importance for animal welfare

Whilst there has been a significant focus on measuring negative affective states in relation to improving animal welfare, it is now becoming increasingly recognised that in order for an animal to be considered in a good welfare state, it needs not only to show the absence of negative events (i.e. suffering) but must also be able to experience positive events (i.e. pleasure) (Dawkins 1990; Boissy et al. 2007; Dawkins 2008; Mendl, Burman, et al. 2010). Simply because an animal is not in pain or does not suffer doesn't necessarily mean it is in a positive affective state analogous to a state of 'happiness' in humans. However, until recently measuring positive affect has been relatively overlooked and understudied (Boissy et al. 2007).

The primary reason for this is that negative experiences tend to be more of a concern from an animal welfare perspective, but also because negative affective states tend to be more intense and cause greater arousal compared to positive affective states, and are consequently often easier to detect (Boissy et al. 2007). However, the ability to successfully measure when animals experience positive affect has obvious implications for understanding and making improvements to their welfare. This has led to a shift in focus in recent years not only in animal welfare science, but also in human psychology, with new disciplines emerging, such as 'positive psychology', to explore the role of positive affect in subjective well-being (Boissy et al. 2007).

A number of studies have now attempted to identify behavioural measures indicative of the presence of positive affect in laboratory rodents (Panksepp & Burgdorf 2000; Burgdorf & Panksepp 2001; Morley-Fletcher et al. 2003; Finlayson et al. 2016). One particular study, that gained significant attention, was the assessment of 'laughter' in rats (Panksepp & Burgdorf 2000). In this study, tickled rats elicited Ultrasonic Vocalisations (USVs) at a frequency of 50kHz, thought to reflect a positive affective state analogous to human joy and laughter (Panksepp & Burgdorf 2000). This comes from behavioural and physiological evidence showing that these USVs at 50-kHz are neurally and functionally homologous to human laughter (Panksepp & Burgdorf 2000; Burgdorf & Panksepp 2001; Rygula et al. 2012).

This discovery of 'laughter' in rats facilitated the development of other novel methods to assess the presence of positive affect in rodents. It is now widely documented that rodents exhibit specific facial expressions when in pain (e.g. the Rat and Mouse Grimace Scale Langford et al. 2010; Sotocina et al. 2011), making it possible that other facial expressions could offer a proxy of positive affect (Finlayson et al. 2016). Positive social interactions, such as the presence of play behaviour, may also signal positive affect in animals (Vanderschuren et al. 1997; Dudink et al. 2006; Boissy et al. 2007; Bateson & Martin 2013). Play behaviour can be described as locomotor activity that has no obvious benefits to the player but is considered to be rewarding and highly important for social and cognitive development, and may be a useful measure of an animal's welfare (Vanderschuren et al. 1997; Dudink et al. 2006; Bateson & Martin 2013). For example, rats exposed to prenatal stress show evidence of impaired social play, which can be reversed through environmental enrichment (Morley-Fletcher et al. 2003). However, although these approaches are

promising, much work is still needed in order to develop and validate more direct measures of positive affect in animals, especially for laboratory mice.

How an animal responds to reward is, of course, measuring how they respond to a positive event. Therefore, measuring how rewarding an animal might find a sucrose solution through assessing changes in their licking microstructure, not only provides more information about their affective state, but also provides insights into their ability to experience pleasure (Dwyer 2012; Austen et al. 2016). It also offers the potential to assess how an animal's responses might change cumulatively over time (Dwyer 2012; Bateson 2016), and explore whether the accumulation of negative experiences might reduce or abolish the typical pleasure response indicative of a state analogous to anhedonia in human patients with depression. It is possible to assess how rewarding an animal finds something given their prior experiences by assessing changes in their licking microstructure following shifts in reward value. This method offers the potential to assess an animal's resilience to negative events (i.e. reward loss) or susceptibility to positive events (i.e. reward gain) (Flaherty 1996; Flaherty et al. 1998; Burman et al. 2008; Neville et al. 2017). Such a paradigm would provide insights into 'how bad' mice might perceive negative events such as a drop in reward value, or 'how good' they perceive positive events such as an increase in reward value, according to their prior experiences and affective state. Therefore, my work has important implications for understanding animal affect and informing animal welfare science. It offers the potential to use a new method to assess how an animal's experiences shape its valuation of reward, and how this is influenced by their affective state. A better understanding of the affective capabilities of laboratory mice would enable evidence-based improvements to their welfare as well as the potential for better animal models of human disease.

1.7 Aims of my research

There is no standard method for assessing the affective state and no current method for measuring positive affective states of laboratory mice. Consequently, there is an urgent need to develop novel methodologies that can provide more understanding of the affective states of mice used in biomedical research. Using established protocols, I aimed to manipulate the affective state of laboratory mice, and if successful, test if licking microstructure was also altered. My overall aim was to develop and evaluate licking microstructure as a novel measure of affective state (both positive and negative) in laboratory mice. I specifically asked:

1. How does stress affect the hedonic responses of laboratory mice towards reward by assessing changes in licking microstructure and total consumption of sucrose solutions?

2. Is licking microstructure a better and more robust measure compared to existing methodologies for assessing the affective state of laboratory mice?

3. Can we use changes in licking microstructure to measure the impact of standard husbandry practice on mice, and advise current legislation and guidelines to improve the welfare of laboratory mice?

Chapter 2: General Methods and Materials

The methodologies and tests described in this chapter were the main techniques used throughout my thesis. However, where other techniques and methodologies were used these will be described and discussed accordingly in the respective experimental chapter.

2.1 Ethical Statement

All experiments were conducted within the Comparative Biology Centre at Newcastle University following approval from the University's Animal Welfare and Ethical Review Body (AWERB), and approved by the Home Office for regulated work, as appropriate (PPL: PC6981D63, 60/4431, PIL: IBE41DE17). All work was conducted in accordance with the EU Directive (2010/63/EU), ASPA (1986) and the NIH Guidelines for Care and Use of Animals for Experimental Procedures (National Institutes of Health 2011). All mice were free from all recognised pathogens, and the health status of the colony was monitored following the FELASA health monitoring recommendations (Guillen 2012). All reporting abides by the ARRIVE guidelines (Kilkenny et al. 2010).

2.2 C57BL/6 mice as a model species

Mice (*Mus musculus*) are the most widely used model species for biomedical research (Home Office 2017). For my experiments, I used the C57BL/6 strain because they are the most commonly used background strain for genetically modified mice in biomedical research, making my findings translatable across a wide range of scientific disciplines, and with the potential to improve the welfare of a vast number of mice used for scientific research worldwide.

C57BL/6 are also good models because their phenotype is well established because they serve as the reference genome for laboratory mice (Mouse Genome Sequencing Consortium et al. 2002). C57BL/6 mice are well characterised in terms of their behaviour and cognitive abilities (Lepicard et al. 2000; Lepicard et al. 2006;

Sankoorikal et al. 2006; Moy et al. 2007; Heinla et al. 2018), and have a high preference for sweet tastants which made them the ideal strain for use in my experiments (Lepicard et al. 2000; Bachmanov et al. 2001; Pothion et al. 2004; Lewis et al. 2005). They are also considered a good model for use in biomedical research due to their relatively low levels of stress and anxiety, and are considered relatively stress resilient in comparison to other strains such as Balb/c mice (Griffiths et al. 1992; Ducottet & Belzung 2004b; Ducottet & Belzung 2005). Therefore, in order to make my findings translatable and to keep my results consistent, I used the C57BL/6 mouse strain across all experiments.

2.3 Drinking apparatus and lickometers

All my studies used custom-built drinking chambers made from standard mouse Individually Ventilated Cages (IVC) home cages measuring (34 (L) x 19 (W) x 14 (D) cm). Drinking chambers had transparent perspex sides, a metal perforated floor, and a wire cage lid with two modified attachments (approximately 3.5cm apart) to connect the sipper tubes to the right and/or left hand side of the cage (see Figure 2.1A-D). Solutions were presented to the animals on the left hand side of the cage for all studies, using 50ml falcon tubes attached to sipper tube lids with a metal spout.

The drinking chambers were connected to contact sensitive Med Associates dual contact lickometers (ENV-250B, Med Associates Inc., St. Albans, Vermont), which recorded all licks from the sipper tubes. The metal perforated floor and bracket lid connected to the sipper tube created an electrical circuit that was closed each time the animal licked or otherwise contacted the spout. Therefore, a single lick could be detected and recorded to the nearest 0.01 second through an interface and a computer running MED-PC software and custom written programmes (courtesy of Prof Dominic Dwyer).

I had a total of eight drinking chambers (Figure 2.1D), which could collect all data simultaneously. Therefore, in each study, I ran animals in groups of eight. Mice would be allocated to a chamber, counterbalanced with respect to the experimental treatment, and drink in the same chamber at the same time on each test day. Animals were initially habituated to the drinking chamber and the novel taste of sucrose in order to avoid any neophobia and ensure adequate consumption across

testing days. This consisted of a training phase, where the spout was positioned slightly protruding into the drinking cage in order to ensure the animal's engagement with the task (i.e. they were consuming some of the solution presented to them). Once all the animals were consistently drinking, the spout was positioned so that it was 'flush' with the bracket. This was to minimise accidental contact with the spout, and reduce the likelihood of 'false licks'.

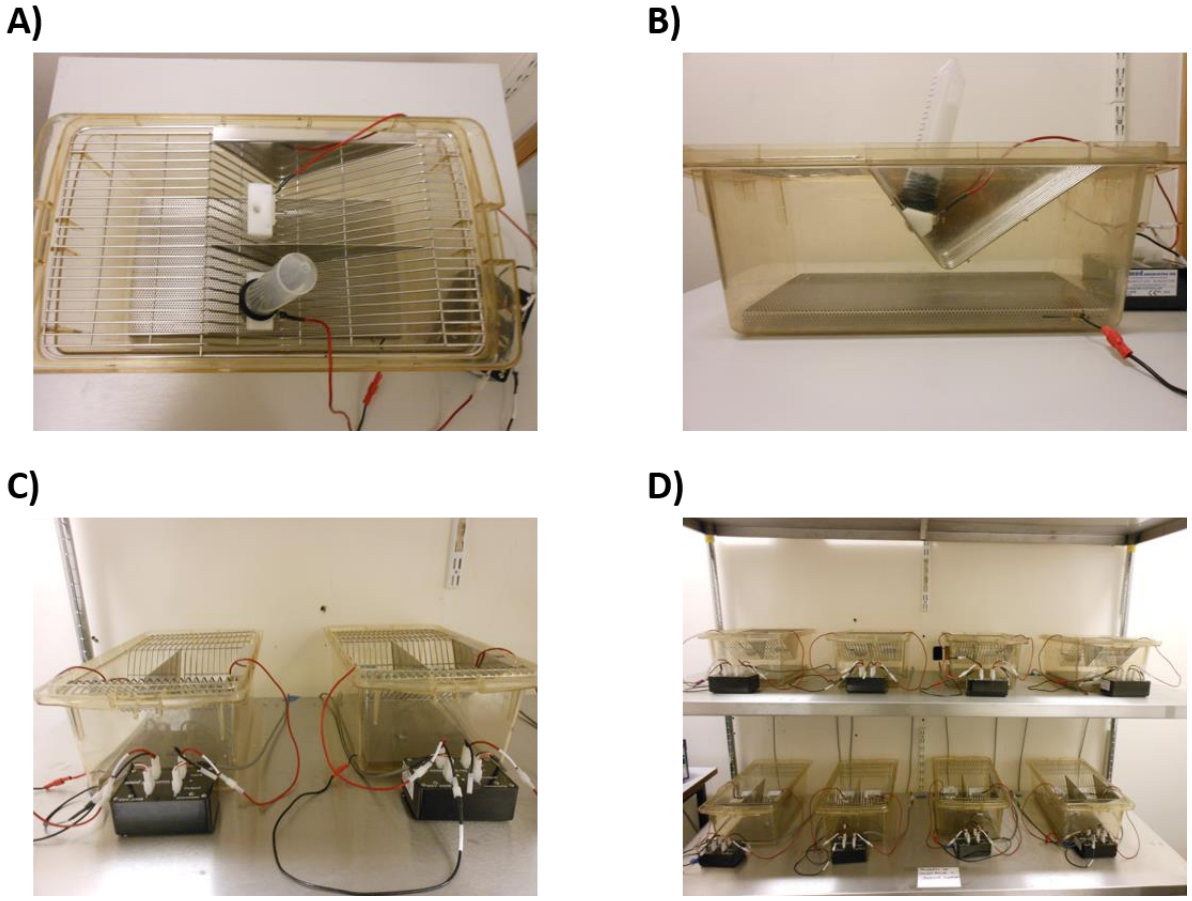


Figure 2.1: Photographs of drinking apparatus used for all experiments presented in this thesis A. View from above a single drinking chamber B. View from the side of a single drinking chamber C. Front view of two drinking chambers with med associates lickometer D. Front view of all eight drinking chambers on shelving.

2.4 Drinking parameter definitions

I collected a range of measures for each mouse in every trial. The total amount of sucrose solution drank was measured by weighing the sipper tube both before and after each trial, using a pair of weighing scales that were accurate to the nearest 0.1g. The Med Associates lickometer system recorded every lick with a time stamp, from which the total number of licks and number of drinking bouts could be

calculated. Bouts were defined using interbout intervals (IBI) of <250ms, <500ms or <1000ms, which refer to the time required between two subsequent licks to classify a single bout, and which have been widely used in published work (Davis & Smith 1992; Davis & Perez 1993; Spector & St. John 1998; Dwyer 2009; Dwyer 2012; Dwyer et al. 2013).

Whilst these bout criteria may seem rather arbitrary, they were originally established in rats by John Davis (Davis & Smith 1992), who identified three inter-lick intervals that defined different drinking and consummatory behaviours: a) <250ms, b) <500ms and c) >500ms (Davis & Smith 1992). The majority of inter-lick intervals reflect continuous licking and are less than 250ms (Davis & Smith 1992). However, longer pauses of licking were identified as being useful in disentangling different components of food intake behaviour. Relatively short pauses in the licking that occur between 251-500ms are thought to reflect behaviours such as swallowing and lateral tongue movements, i.e. the stereotypical oro-facial response associated with palatable solutions outlined in Chapter 1 (Grill & Norgren 1978a; Davis & Smith 1992; Johnson et al. 2010). However, longer pauses of >500ms were suggested to reflect the animal leaving the area or performing other competing behaviours such as grooming (Davis & Smith 1992; Johnson et al. 2010). However, this is not agreed across laboratories; more recently, Spector et al. (1998) suggested using a pause criterion of 1000ms to reflect palatability and hedonic responses and therefore consider it a more conservative pause between licking than using 500ms. Whilst these IBIs were established from the licking behaviour in rats, they have also been used to assess the licking microstructure in mice (Austen et al. 2016; Austen & Sanderson 2016; McNamara et al. 2016), although much less is known whether these are also optimal for use in mice. Therefore, given that the choice of criterion is debated within the literature, means that I classified bouts according to the three pause criteria <250ms, <500ms and <1000ms in order to ensure results were robust across criteria. It is important to highlight that these three pause criteria are not independent, for example the <500ms criterion will also include pauses <250ms.

Once I had established the number of bouts according to the three bout criteria (250ms, 500ms or 1000ms), I then calculated the mean lick cluster size in Microsoft Excel according to the following equation:

$$Mean\ Lick\ Cluster\ Size = \frac{\text{Total licks - single licks}}{\text{Total number of bouts}}$$

It was important to subtract the number of single licks because these often reflect accidental contact with the spout or false licks, and a single lick (or contact) on its own cannot be considered a bout of licking. My calculations resulted in mean lick cluster sizes at each bout criterion (250, 500 and 1000ms) for each mouse for each trial.

2.5 Open field test (OF)

The open field test (OF) is a standard test of anxiety in laboratory rodents (Hall & Ballachey 1932; Hall 1934; Gould et al. 2009). Mice and rats typically tend to avoid large open spaces, and will spend less time in the centre of a large empty arena and more time near the walls, a behaviour known as thigmotaxis (defined as the tendancy to remain close to vertical surfaces). In my studies, I used a white plastic arena measuring (54.5cm (L) x 35.5cm (W) x 17cm (H)) with a transparent Perspex lid (Figure 2.2). Mice were placed in the centre of the open field and left to freely explore for 10 minutes. This was filmed from above and the behaviour was later analysed using automated tracking software (Ethovision XT version 5.1, Noldus, Virginia, USA). This calculated the total duration spent in the centre, crosses to the centre, total distance travelled, the total time spent moving and the velocity of movement for each mouse. Presence of defecation during this test was also noted. The arena and protocol used in my studies were in line with those reported in the literature (for a review see; Gould et al. 2009). Although the size of the arena varies across studies, it is generally recommended that the area required must be relatively large (at least 1600cm^2) in order to produce reliable data for assessing thigmotaxis (Gould et al. 2009). Traditionally the open field test is relatively short, between 2 and 10 minutes in duration (Gould et al. 2009), to capture the animal's behaviour towards novelty and initial exploration. In order to minimise odour cues between animals, I always disinfected the arena between animals with 70% ethanol and the animals were always placed in a separate holding cage after testing to minimise disruption to other animals.

A)

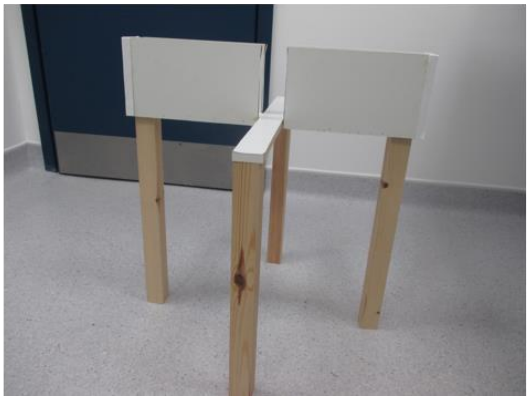
B)

Figure 2.2: Photographs of the open field arena used throughout my thesis A. View from above the arena demonstrating arena size B. View from the side of the arena demonstrating wall height.

2.6 Elevated Plus Maze (EPM)

The elevated plus maze (EPM) is also a standard test of anxiety in laboratory mice (Pellow et al. 1985; Walf & Frye 2007). The maze consists of two open and unprotected arms without walls, and two closed and protected walled arms elevated off the ground (Figure 2.3). Mice showing greater levels of anxiety are predicted to spend significantly less time on the open, unprotected arms of the maze and more time in the closed, protected arms compared to control mice. This is because mice naturally seek dark enclosed spaces and have an unconditioned fear of heights and open spaces (Walf & Frye 2007). My elevated plus maze was made out of white chipboard, with each arm being 30cm (L) x 5cm (W). The side walls of the two closed arms were 15cm (H). The maze was elevated 50cm from the ground (see Figure 2.3). Mice were placed facing an open arm of the maze and left to freely explore for five minutes. Their movement and behaviour was filmed from above and later scored using Observer XT (version 21, Noldus, Virginia, USA). The dimensions used and the duration for which I ran this test were in line with previous work; most avoidance behaviour is often thought to occur in the first five minutes of the test (Montgomery 1955; Walf & Frye 2007). In order to minimise odour cues between individuals, I always disinfected the maze with 70% ethanol in between subjects and mice were placed in a separate holding cage following testing.

A)



B)

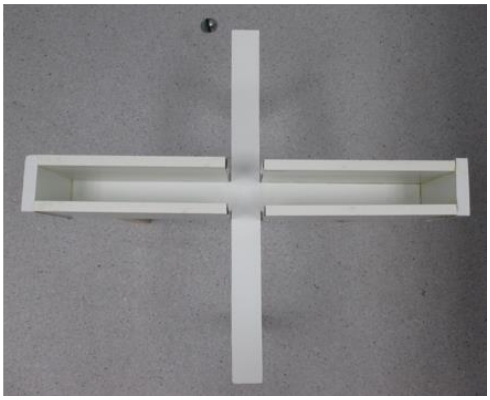


Figure 2.3: Photographs of the elevated plus maze (EPM) used for experiments in my thesis A. View from the side of the EPM demonstrating leg height B. View from above demonstrating the two open and two closed arms of the maze.

Chapter 3: The effects of Chronic Mild Stress on licking microstructure

3.1 Abstract

Mice are the most widely used model organism for biomedical research, making it important to develop methods to determine the effects of laboratory stressors on their welfare. Measuring changes in the way that animals respond to reward offers a way of accessing how they might feel. Anhedonia, or the reduced sensitivity to reward, is a core symptom of human depression, and is currently assessed in laboratory rodents through the amount of sucrose they consume: lower consumption is thought to reflect an anhedonic-like state. However, this is often highly influenced by the animal's motivation. Therefore, an alternative measure suggested to better reflect their hedonic evaluations of sucrose involves assessing changes in the way that rodent's drink, specifically the size of their licking bouts or 'lick cluster size'. The aim of this study was to manipulate the mice's affective state using the well-established Chronic Mild Stress (CMS) regime to determine the effect of chronic stress on licking microstructure.

Despite being well established within the literature, I found no evidence that the CMS protocol induced negative affect in my mice. Mice undergoing CMS did not demonstrate greater anxiety-like behaviour in the open field test and had similar body weights to control mice kept under standard husbandry conditions. Consequently, it was unsurprising that I failed to find an effect of CMS on licking microstructure. Future work would benefit from inducing negative affect using other well-established methodologies to determine the effects of stress on licking microstructure.

3.2 Introduction

Laboratory conditions are stressful for many animals, with subjects often exposed to different experimental procedures, artificial housing conditions and high levels of human interaction. Given that mice are the most widely used model organism for biomedical research, developing methods to assess their welfare has the potential to improve the lives of millions of mice housed in laboratories worldwide.

One way to evaluate the welfare state of an animal is to try and determine their underlying affective state (Dawkins 1990; Mendl, Burman, et al. 2010). People exposed to significant stress in their lives tend to be more likely to develop negative affective states such as depression (Shishkina & Dygalo 2017). However, what causes an animal stress, and what effect stressors have on an animal's underlying affective state, can be challenging to establish (Dawkins 1990; Mendl, Burman, et al. 2010). Because animals can't self-report how they 'feel', appropriate behavioural and physiological methods are required in order to make evidence-based improvements to the housing conditions and experimental procedures they experience (Dawkins 1990; Paul et al. 2005; Mendl, Burman, et al. 2010). As outlined in Chapter 1, one method which has gained significant traction in recent years has been to assess changes in animal's judgement biases as a way of accessing their underlying affective state. There is now good evidence from a number of species to suggest that the way that animals make judgements and decisions are influenced in predictable ways by their underlying affective state (Harding et al. 2004; Bateson & Matheson 2007; Brilot et al. 2010; Mendl, Brooks, et al. 2010). When animals are trained to associate one stimulus with a positive outcome (i.e. a reward) and another stimulus with a negative outcome (i.e. no reward), those animals in negative affective states are more likely to judge a novel intermediate stimulus more negatively, than animals that are in more positive affective states (Paul et al. 2005; Mendl, Burman, et al. 2010). However, these judgment bias tasks can be time consuming, since animals need to be trained to discriminate well between a positive and negative stimulus before being tested with intermediate stimuli (Harding et al. 2004; Paul et al. 2005) This, combined with lengthy training periods, means that there is yet to be a cognitive bias task specifically developed for use in laboratory mice (Harding et al. 2004; Bateson & Matheson 2007; Brilot et al. 2010; Mendl, Brooks, et al. 2010; Bateson et al. 2011). Therefore, alternative measures of affective state are required in order to make evidence-based improvements to the welfare of laboratory mice.

One feasible and relatively simple approach might be to measure changes in the ways that animals respond to reward. This is because humans suffering from depression often report lowered sensitivity to reward and experience less pleasure, a symptom known as anhedonia (American Psychiatric Association 2014). Anhedonia has traditionally been measured in rodents by assessing their voluntary consumption of a rewarding sucrose solution, where a lower consumption or preference for

sucrose is taken to reflect an anhedonic-like state (Willner et al. 1987; Papp et al. 1991; Monleon-Paolo et al. 1995). However, as previously discussed in Chapter 1, assessing how much an animal drinks of something does not solely reflect how much they like it. This is because the amount that an animal eats or drinks of something, is also highly influenced by the animal's motivational state (Berridge 1996; Dwyer 2012). Studies suggest that measuring mice's specific licking pattern, i.e. their lick cluster sizes, could offer a more sensitive measure of their hedonic response (Davis 1973; Davis 1989; Hsiao & Fan 1993; Dwyer 2012; Austen et al. 2016). When animals drink they do so by a series of licks. These licks can be grouped into bouts or clusters according to pre-defined interbout interval criteria, whereby the average number of licks that occur in a single bout (lick cluster size) is considered to reflect the hedonic 'liking' towards that given tastant (Chapter 1: Section1.5, p12) (Davis 1973; Davis 1989; Hsiao & Fan 1993; Dwyer 2012; Austen et al. 2016). Although this methodology has been proven effective at measuring an animals hedonic 'liking' towards a given tastant in relation to memory and learning (Dwyer et al. 2009; Dwyer et al. 2013; Austen et al. 2016), it has yet to be applied in a welfare context.

Therefore, in order to determine how licking microstructure was affected by laboratory mice's affective state, I aimed to manipulate the affective state of laboratory mice using the well-established Chronic Mild Stress (CMS) paradigm, and compare this to the traditional method of assessing total consumption alone. The choice to employ the CMS paradigm was because it is the most commonly used method to induce a negative depressed-like state in laboratory rodents. It has been estimated to have been used in over 1300 published studies, with around 200 studies reporting its use per year (Willner 2017b). This methodology was first developed by Katz and colleagues as the 'chronic stress model' (Katz 1982a) and was later refined to the Chronic Mild Stress (CMS) model by Willner (Willner et al. 1987). It consists of randomly applying a variety of mild stressors which are unpredictable in nature for a number of weeks or months, until a negative affective state is observed. Although there are no strict set of stressors, typical stressors often applied include, but are not limited to, food and water deprivation, overnight illumination, cage tilt, soiled cage and exposure to white noise (Muscat & Willner 1992; Willner 1997). In Willner's original work, rats exposed to CMS showed a negative affective state through drinking less sucrose compared to their respective controls, an effect which could be reversed by the administration of antidepressant compounds (Willner et al. 1987).

CMS is now considered to be a valid method of inducing a negative affective state both rats and mice, and has been used to assess the effect on a number of behavioural and physiological measures (Willner 1997; Ducottet et al. 2003; Pothion et al. 2004; Grippo et al. 2005; Yalcin et al. 2005; Goshen et al. 2008; Schweizer et al. 2009; Malki et al. 2015; Willner 2017b).

However, despite its effect on sucrose consumption being well validated (Willner 1997), its effect on licking microstructure remains unknown. Therefore, in this study, I aimed to manipulate laboratory mice's affective state using the Chronic Mild Stress (CMS) paradigm and assess whether it was sufficient to cause a negative affective state in laboratory mice using the Open Field (OF) test, body mass changes and sucrose consumption. If successful, I then aimed to determine whether assessing changes in licking microstructure might offer more information about hedonic changes, and establish whether measuring changes in lick cluster size might be more useful to infer the affective state of laboratory mice. I predicted that mice undergoing the CMS regime would not only drink less sucrose but the size of their drinking bouts (i.e. their lick cluster size) would also be significantly smaller than mice that were maintained under standard 'control' conditions.

3.3 Methods and Materials

3.3.1 Ethical statement

All experiments were conducted at Newcastle University following approval from the University's Animal Welfare and Ethical Review Body (AWERB Project ID: 307) and were completed in full compliance with the UK Home Office (PPL: 60/4431, PIL: IBE41DE17). All work was conducted in accordance with the EU Directive (2010/63/EU), ASPA (1986) and the NIH Guidelines for care and use of animals for experimental procedures (National Institutes of Health 2011). All animals were checked daily and no adverse effects were reported. At the end of the experiment, animals were humanely killed via intraperitoneal (i.p.) overdose of anaesthetic (Pentobarbital) in accordance to Schedule 1 guidance. One animal from the Chronic Mild Stress (CMS) group was euthanized after one week of CMS due to health concerns (pre-putial abscess).

3.3.2 Animals, housing and husbandry

Forty male C57BL/6J mice (*Mus musculus*) were purchased from Charles River Laboratories, UK and were approximately 10 weeks of age (Mean±SEM mass: 26.8±0.17g) on arrival and were allowed a one week acclimation period. They were tail marked using permanent marker for identification and were handled according to standard practice (tail handled). Mice were free from all recognised pathogens, and the health status of the colony was monitored following the FELASA health monitoring recommendations (Guillen 2012). Mice were housed in groups of four in standard MB1 cages (45cm (L) x 28cm (W) x 13cm (H), North Kent Plastics), with sawdust bedding, nesting material and cardboard tubes for environmental enrichment (NestPak, 4HK Aspen chips and Sizzlepet nesting, Datesand Ltd, Manchester). All cages were subject to two full cage cleans per week. All animals had access to food (Special Diet Services, RM3E diet) and water *ad libitum,* except prior to training and testing and CMS food and water manipulations (see below). Animals were housed in two separate climate control chambers, in order to have strict control over their environmental conditions, and were randomly assigned to one of two treatment groups. Animals in one chamber underwent the CMS regime referred to as the Chronic Mild Stress (CMS) group (details below), and those in the other climate control chamber acted as controls, the Control group. Animals were maintained on a reverse 12:12 hour light/dark cycle (lights off 10:30 until 22:30), and experiments were conducted under red light illumination. Mice were kept under relatively constant temperature 21±4°C and relative humidity 55±10% (Home Office 2014). All mice were weighed twice a week to monitor any weight changes occurring due to treatment effects; typically, mice undergoing CMS would have a lower body weight or reduced weight gain compared to animals housed under standard conditions (Schweizer et al. 2009; Willner 1997).

3.3.3 Study Design and timeline

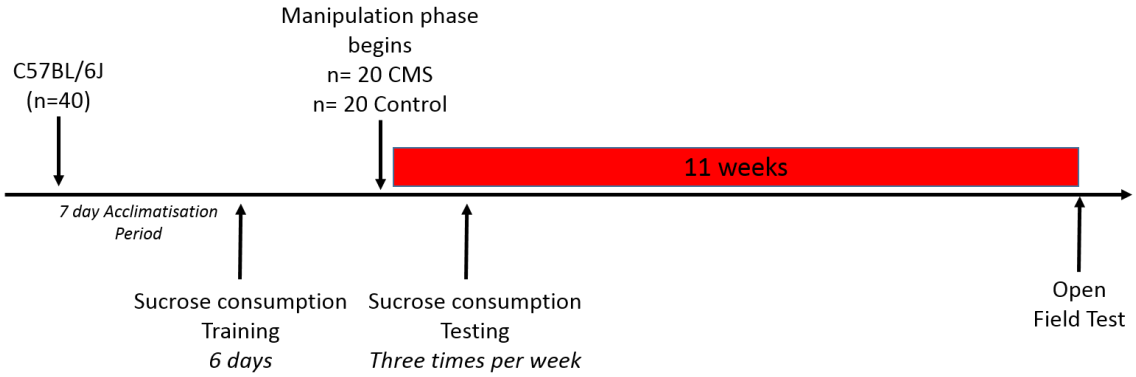


Figure 3.1: Schematic illustrating the study timeline showing when each behavioural test were conducted relative to the state manipulation phase.

3.3.4 Sucrose consumption tests

Water bottles on the home cages were removed 3hrs before training and testing (08:00) to ensure no water was consumed immediately prior to the test. All drinking sessions were conducted once the lights had gone off at 10:30 under red light illumination, meaning that animals were tested towards the start of their active period. Mice were trained and tested in the eight custom-made drinking chambers (see Chapter 2 for full details). Briefly, the drinking chambers were connected to contact sensitive Med Associates dual contact lickometers (Med Associates Inc., St. Albans, Vermont), which transmit the time of each lick to the nearest 0.01 second to a computer using MED-PC software. Custom-written software (courtesy of Prof Dominic Dwyer) calculated the lick cluster sizes according to a range of interbout intervals (IBI), which is the length of time used to determine when licks can be considered to be in a single bout (Davis & Perez 1993; Davis & Smith 1992; Dwyer 2012; Dwyer et al. 2013). The data presented here use interbout intervals of 250ms, 500ms or 1000ms in order to ensure data were robust across criteria. This means that, for example, when an interbout interval of 250ms was applied, any duration of 250ms or longer between two licks defined the end of one bout and the start of the next.

Mice were randomly sub-divided into five testing groups, where each testing group consisted of eight mice (4 mice from each treatment group), which were tested in the same order at the same time each day. The first group typically entered the

boxes 30mins after lights off (11:00), and the final group started 3hrs after lights off (13:30). Each mouse was placed in the same box across drinking sessions, where the position was balanced across the five groups with respect to treatment. The sucrose solution (8% w/w) was made available to the animals through drinking spouts presented on the left hand side of the chamber. The amount of fluid consumed for each mouse was calculated by weighing the drinking bottle before and after each test session using a set of scales accurate to the nearest 0.1g.

Training involved familiarising the animals to the drinking chamber and the sucrose solution. Mice were put into their chamber for 15 minutes across six consecutive days (Figure 3.1), at which point all animals were consistently consuming the sucrose solution (defined as making at least 100 licks). For the first three training sessions, the metal spout protruded into the chamber to ensure engagement with the task. After this, the spout was positioned 'flush' to the wall to reduce the possibility of accidental contact with the spout. Following six days of training, the state manipulation phase began (Figure 3.1).

3.3.5 State manipulation phase

In order to alter the affective state of some animals, I employed the well-established Chronic Mild Stress (CMS) paradigm (Willner et al. 1987; Papp et al. 1991; Muscat & Willner 1992; Willner 2017b). The mice in the CMS group underwent the Chronic Mild Stress regime for 11 weeks, whilst mice in the Control group were kept under standard laboratory conditions. The CMS regime involved the semi-random administration of any one of the six following manipulations throughout the day or night to make their occurrence unpredictable:

1. Overnight Illumination: Due to the animals being housed on a reversed light dark cycle this consisted of 'overday' illumination where animals were exposed to 36hr of light or intermittent periods of light during their dark phase.

2. Soiled Cage: Animals were transferred to a dirty cage that had previously held another cage of mice from the CMS treatment group.

3. Damp bedding: 100ml of water was added to the sawdust of the cage.

4. Food and water deprivation: Both food and water were removed from the cages for up to 3hrs.

5. Exposure to white noise: Animals were exposed to an untuned radio (<90db) for up to 6hrs.

6. Cage shaking: Researchers shook the cage rack for a total of 30 seconds every 5-minutes for a total session duration of up to 30 minutes.

For the first 4 weeks, stressors were applied in the dark phase (10:30-22:30), but after that, some were carried out during the light phase (22:30-10:30). This was due to concerns regarding the effectiveness of only carrying out these manipulations during the animal's dark phase when they were already active. Since some of the stressors required human presence, this meant that I could only carry out exposure to white noise, overnight illumination, soiled cage or damp bedding overnight.

Throughout the manipulation phase, the drinking behaviour of both the CMS group and Control group continued to be monitored. The protocol for measuring sucrose consumption and licking microstructure was the same as that described for training. Drinking trials were conducted three times per week (Monday, Wednesday and Friday) for 11 consecutive weeks and therefore coincided with the 11 week state manipulation phase.

3.3.6 Open Field Test

At the end of the manipulation phase, each mouse was individually placed in a corner of a rectangular arena (54.5cm (L) x 35.5cm (W) x 17cm (H)) made of white plastic with a transparent perspex lid (see Chapter 2 for full details). Individuals were allowed to freely explore for 10 minutes, whilst being filmed from above (Sony Handycam HDR-CX220). The order in which CMS and Control mice were tested was counterbalanced across the testing day. The total duration spent in the centre, crosses to the centre, total distance travelled, the total time spent moving and the velocity of movement for each mouse were measured using Ethovision XT (v5.1, Noldus, Virginia, USA). Occurrence of defecation during testing was also noted.

3.3.7 Statistical analyses

All statistical analyses were conducted using IBM Corp. SPSS (v23, SPSS Inc, Chicago, USA). Datasets were tested for normality using the Shapiro-Wilk test and homogeneity of variance using and the Levene statistic test and where assumptions were not met, non-parametric statistical methods were used. Where significant main effects were found, Bonferroni post hoc tests were performed to investigate pairwise comparisons (Table 3.1). In all statistical tests differences were considered significant using a p value <0.05.

For the sucrose drinking tests, data were averaged across the three weekly sessions in order to obtain a mean value per mouse per week. I calculated averages for the total consumption and also for lick cluster size according to the three IBI criteria for the last three sessions of the sucrose consumption training phase to be used as a baseline (week 0), and for the 11 subsequent weeks of the state manipulation phase.

Table 3.1: Statistical tests carried out on each data set, including factors included in the model, experimental unit and sample size.

Data	Dependent Variable	Statistical Test	Factor(s)	Unit	Sample Size
Sucrose Drinking	Consumption (g)	Repeated measures ANOVA	Between subject factor: Treatment (2 levels: CMS or Control) Within subject factor: Week (12 levels: weeks 0-11)	Mouse	n=19 CMS n=20 Control
	Lick cluster size (at 250ms, 500ms or 1000ms)	Repeated measures ANOVA	Between subject factor: Treatment (2 levels: CMS or Control) Within subject factor: Week (12 levels: weeks 0-11)	Mouse	n=19 CMS n=20 Control
Open Field Test	Distance travelled; Duration in centre; Duration of movement; Frequency in centre; Velocity	Independent t-tests	Between subject factor: Treatment (2 levels: CMS or Control)	Mouse	n=19 CMS n=20 Control
	Defecation	Binary logistic regression	Between subject factor: Treatment (2 levels: CMS or Control)	Mouse	n=19 CMS n=20 Control
Body Weight	Weekly body weight (g)	Repeated measures ANOVA	Between subject factor: Treatment (2 levels: CMS or Control) Within subject factor: Week (12 levels: weeks 0-11)	Mouse	n=19 CMS n=20 Control
	Percentage weight gain	Independent t-test	Between subject factor: Treatment (2 levels: CMS or Control)	Mouse	n=19 CMS n=20 Control

3.4 Results

3.4.1 Sucrose consumption tests

There was no evidence that the CMS treatment affected the amount of sucrose consumed during training (week 0) and across the 11 weeks of the state manipulation phase, with mice undergoing the CMS regime consuming a similar amount of sucrose as control mice (ANOVA: $F_{1,37}=0.63$, $p=0.434$; Figure 3.2). There

was a significant main effect of week, with consumption changing over time irrespective of treatment (ANOVA: $F_{5.795,214.405}$=6.95, p<0.001), and there was no significant interaction between these factors (ANOVA: $F_{5.795,214.405}$=1.99, p=0.07).

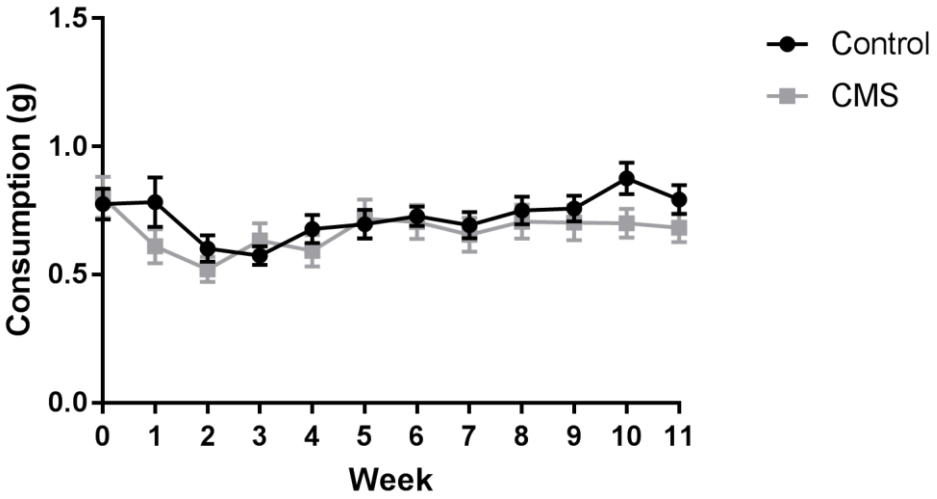


Figure 3.2: Mean (±SEM) consumption (g) for CMS and Control groups in the training phase (week 0) and the 11 weeks of the manipulation phase.

Similar to the results presented for consumption, I found no evidence to suggest an effect of the CMS regime on the licking microstructure of the mice (ANOVA: 250ms: $F_{1,37}$=0.17, p=0.687; Figure 3.3A; 500ms: $F_{1,37}$=0.01, p=0.927; Figure 3.3B; 1000ms: $F_{1,37}$=0.02, p=0.878; Figure 3.3C). Again, there were significant changes across the weeks (ANOVA: 250ms: $F_{3.93,145.51}$=3.22, p=0.015; Figure 3.3A; 500ms: $F_{3.93,145.36}$=3.93, p=0.005; Figure 3.3B; 1000ms: $F_{4.03,149.03}$=4.49, p=0.002; Figure 3.3C), and no significant interaction between treatment and week (ANOVA: 250ms: $F_{3.93,145.51}$=0.91, p=0.455; Figure 3.3A; 500ms: $F_{3.93,145.36}$=1.01, p=0.406; Figure 3.3B; 1000ms: $F_{4.03,149.03}$=0.94, p=0.443; Figure 3.3C). Therefore, I found no evidence to suggest that the CMS regime had any effect on the licking microstructure of mice compared to their respective controls.

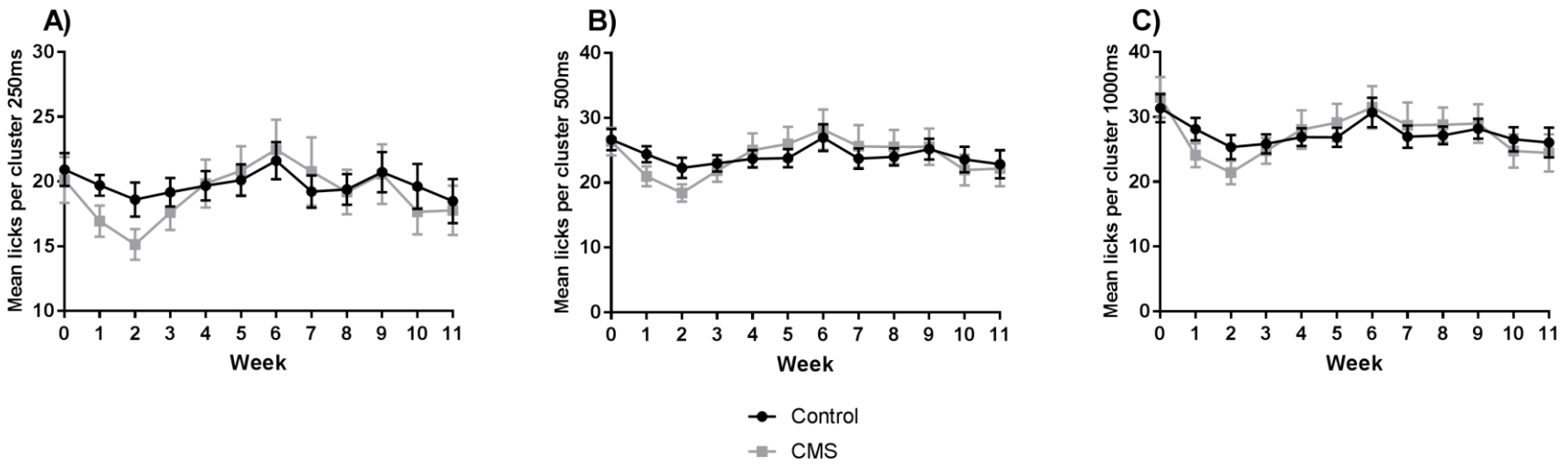


Figure 3.3: Mean (±SEM) Lick cluster size for CMS and Control mice according to the three different interbout (IBI) interval criteria A. 250ms IBI interval criterion B. 500ms IBI interval criterion C. 1000ms IBI interval criterion.

3.4.2 Open Field Test

The only evidence to suggest that the CMS regime increased stress and anxiety in CMS mice relative to their respective controls was that more mice from the CMS group than the Control group defecated in the open field arena during testing ($\chi^2(1)=5.44$, $p=0.02$; Figure 3.4A). Comparing groups for all other measures, there were no differences between the two treatment groups in the total distance travelled ($t_{37}=1.07$, $p=0.293$; Figure 3.4B), the time spent moving ($t_{37}=0.92$, $p=0.364$; Figure 3.4C), the duration spent in the centre ($t_{37}=0.19$, $p=0.853$; Figure 3.4D), frequency in the centre ($t_{37}=0.80$, $p=0.431$; Figure 3.4E) or the velocity at which they travelled ($t_{37}=1.08$, $p=0.287$; Figure 3.4F).

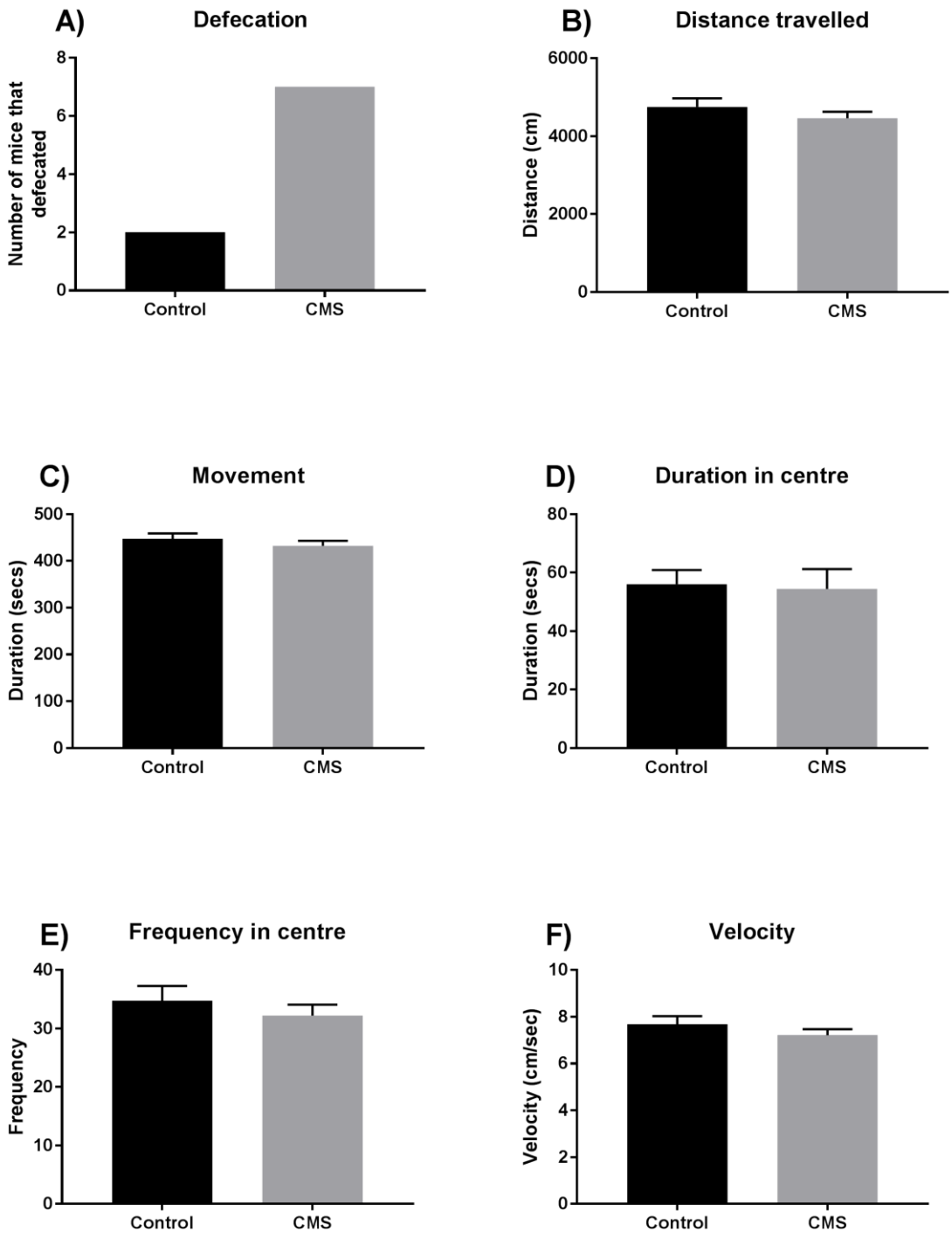


Figure 3.4: The behaviour of CMS and Control mice in the open field test. A. Total number of mice that defecated during testing. B. Mean (+SEM) distance travelled (cm) C. Mean (+SEM) length of time spent moving (s). D. Mean (+SEM) length of time spent in the centre (s) E. Mean (+SEM) number of crosses to the centre. F. Mean (+SEM) velocity when moving (cm/sec).

3.4.3 Body weight

There were no significant differences in bodyweight between animals in the CMS and Control groups at any time point ($F_{1,37}$=1.11, p=0.298; Figure 3.5A), with both groups steadily gaining weight across time, as would be expected for mice with access to food *ad libitum* ($F_{2.137,79.054}$=187.33, p<0.001). There was no significant difference in percentage weight gain between baseline (week 0) and week 11 between the two groups (t_{37}=1.79, p=0.082; Figure 3.5B), although there was a tendency, contrary to expectation, for CMS mice to have gained more weight than the Control mice.

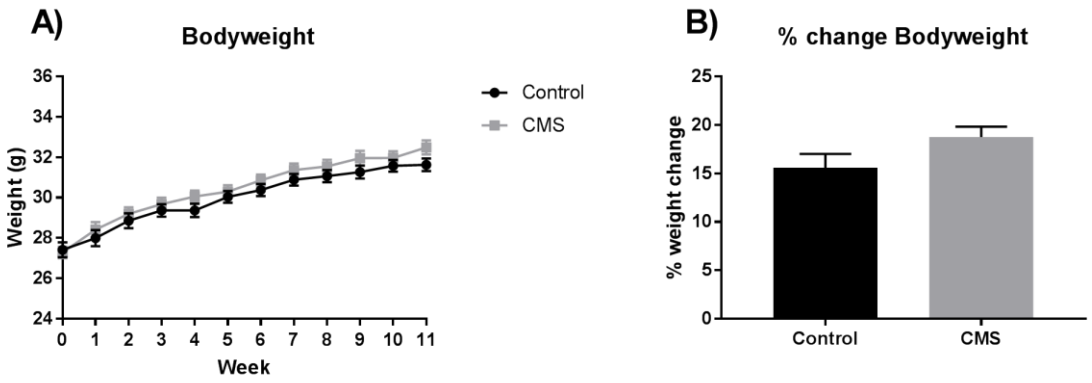


Figure 3.5: Bodyweights for CMS and Control mice. A. Mean (±SEM) body weight (g) for each week B. Mean relative percentage change (+SEM) from baseline.

3.5 Discussion

My study aimed to manipulate laboratory mice's affective state using the Chronic Mild Stress (CMS) regime to determine its effect on licking microstructure. However, my manipulation did not produce any significant changes indicative of a negative affective state in C57BL/6 mice. There was no evidence that the application of the CMS regime produced more anxiety-like behaviour in an open field test, reduced body mass, or lowered the amount of sucrose consumed. Therefore, given these findings, it is not surprising that I also found no effect of CMS on animal's hedonic responses that would be indicative of stress-induced anhedonia in their licking microstructure. Although I cannot make any firm conclusions as to why the

application of CMS was insufficient to cause measurable changes indicative of a negative affective state in these mice, I offer some potential explanations.

Although the CMS manipulation has been widely used and its effectiveness demonstrated across a wide range of studies (Willner 1997; Ducottet et al. 2003; Grippo et al. 2005; Goshen et al. 2008; Jindal et al. 2013; Malki et al. 2015; Yu et al. 2015), there is increasing concern regarding its reliability in producing measurable effects, both within and between laboratories (Willner 2017a; Willner 2017b). Since negative results and failures to replicate previous findings are less likely to be published, researchers need to be surveyed to fully ascertain the scale of the problem with using CMS to manipulate affective state. A recently conducted user survey found that most (75%) respondents reported that CMS worked reliably within their laboratory; however, this study still did not capture information from laboratories with unpublished findings, meaning it had a relatively small sample size (n=53) (Willner 2017a). When comparing responses from those 75% laboratories that were able to demonstrate the effectiveness of CMS with laboratories that could not demonstrate consistent findings, the methodologies used were very similar, and no factor(s) could be identified that might account for differences in effectiveness across laboratories (Willner 2017a). However, despite this, differences in reliability and reproducibility both within and between laboratories have been attributed to the overall severity of the stressors applied and individual differences in susceptibility to stress (Willner 2017b). Therefore, it is possible that other factors could have attributed to the null result found in my study, which I will discuss in turn.

There is variability across laboratories in the types of stressor and the methods by which they are employed, which could impact on the effectiveness of CMS to change affective state (Willner 2017a). In my study, I predominantly manipulated the mice's environment, including using components of the CMS regime that are reportedly widely used with mice: cage tilt, wet bedding, light/dark reversal, food and water deprivation (Willner 2017a). However, I did not include any stressors that disrupted normal social behaviour, like subjecting animals to social isolation or swapping their cage mates. This was because I did not want to potentially increase the aggression in male mice that could lead to adverse effects, for example, fight wounds and potentially death. Therefore, although the application of stressors in my study remained unpredictable in nature, it might be that by not including social stressors, I didn't include stressors that were the most effective at manipulating an

animal's state. Certainly, a number of studies demonstrating behavioural or physiological effects of CMS in laboratory mice have used social stressors within their CMS regime (Ducottet et al. 2004; Ducottet & Belzung 2004a; Yalcin et al. 2005; Schweizer et al. 2009). However, in the recent survey of researchers using the regime, the use of a social stressor, such as crowding, was only used in approximately 30% of regimes (Willner 2017a). Therefore, social stressors are not necessary to induce negative affective states in mice, and I think it unlikely that this fully accounts for the null result reported here.

The effectiveness of CMS, on both mice and rats, can also be dependent upon the strain used (Willner 2017b). A number of studies have reported a lower susceptibility of C57BL/6 mice to stress, determined by a number of different behavioural and physiological measures, compared to other inbred laboratory mouse strains (Anisman et al. 1998; Mineur et al. 2006; Parfitt et al. 2007). This is also seen in experiments using the CMS regime, where C57BL/6 mice appear more resilient to the effects of CMS than a more susceptible strain such as BALB/c (Griffiths et al. 1992; Ducottet & Belzung 2004b; Ducottet & Belzung 2005; Farley et al. 2012). These strain differences have been attributed to differences in anxiety and depressive-like behaviours, with a greater susceptibility to CMS being associated with higher levels of anxiety (Ducottet & Belzung 2004b; Ducottet & Belzung 2005). However, although C57BL/6 mice are considered to be less susceptible than other strains to the effects of stress, they are not completely resilient. A number of studies have been successful in inducing behavioural changes indicative of negative affect in this strain, following the application of the CMS regime (Schweizer et al. 2009; Zhu et al. 2014). Therefore, it is unlikely that my choice of strain alone could have accounted for the null effect, although I cannot rule out that the combination of a relatively mild stress regime with a mouse strain with low susceptibility to stress could explain these findings.

There were also methodological differences compared to previous studies in the animal's husbandry conditions that could have contributed to the CMS not being effective in changing the state of the mice in my experiment. The mice were maintained on a reversed light/dark cycle in order for the experiments to be carried out in their active period (i.e. in the dark phase), ensuring maximal consumption during the sucrose consumption tests. It is possible that the reversal of the light/dark schedule could have contributed to the lack of effect, at least initially. This is because

the majority of the stressors were initially implemented during the animal's active phase (i.e. in the dark) rather their inactive period (i.e. in the light), which may have been less stressful because they did not cause sleep disturbances. In a similar CMS study using rats, subjects only displayed behaviours indicative of depression and anxiety when the CMS protocol was applied during the light phase but not when it was applied during the dark phase (Aslani et al. 2014). For this reason, I started to implement stressors during the animals' inactive phase (i.e. in the light) after the first 4 weeks. However, given that I continued to see a lack of effect even when stressors were applied in the light phase of the cycle, it seems unlikely that this could explain why the CMS was ineffective in my study.

One final possibility for why I did not see an effect of the CMS was that the tests employed in my study were not sensitive to detect changes in affect. However, given the data obtained, this seems unlikely. This is because both the open field and sucrose consumption tests are standard, widely used and established behavioural tests for measuring negative affect in laboratory mice (Muscat & Willner 1992; Gould et al. 2009). I also didn't find any differences in absolute body weight, or body weight changes between my treatment groups across the 11 week state manipulation phase. All these measures are typically employed in the literature to measure changes in affective state and therefore can be considered good measures at detecting the effectiveness of CMS (Willner et al. 1992; Willner 1997; Schweizer et al. 2009; Gould et al. 2009). Therefore, it is unlikely for all these measures to be insensitive to changes in affect, and more parsimonious that the manipulation did not work. One measure did suggest that I had manipulated state, in that I did find that mice undergoing CMS defecated more in the open field test compared to controls, which could be taken to reflect a negative state (Hall 1934). However, this finding alone isn't sufficient to infer the CMS regime effective in inducing a negative affective state in my mice. This is because although increased defecation and/or urination can be considered to be indicative of increased stress and fear (Hall 1934), it cannot be taken to reflect longer-term changes in affective state on its own, as usually this is taken in combination with behavioural responses in the OF test for example (Gould et al. 2009).

To conclude, I found no evidence to suggest that the CMS regime induced a negative affective state in laboratory mice, and as such I was unable to evaluate the effectiveness of licking microstructure as a measure of affective state. Further work

is required to assess the effects that stress might have on licking microstructure and whether this can be used to infer the underlying affective state and welfare of laboratory mice. Future studies would benefit from using other well-established paradigms of inducing negative affect in laboratory mice, in order to assess the effects that stress has on consummatory behaviour, and whether licking microstructure offers a novel measure of their affective state.

Chapter 4: The effect of chronic administration of exogenous corticosterone on licking microstructure

4.1 Abstract

Given that the Chronic Mild Stress regime failed to elicit changes in affective state in Chapter 3, I sought another established method of inducing negative affect in laboratory mice to determine the effect of stress on licking microstructure. The chronic administration of corticosterone is a widely used method of inducing negative affective states in laboratory rodents. Therefore, the aim of this study was to induce negative affect in laboratory mice through the chronic administration of exogenous corticosterone in the animal's drinking water and determine if licking microstructure was able to detect these changes.

I found no evidence to suggest the chronic administration of exogenous corticosterone was sufficient to induce a negative affective state in the mice. I found no behavioural differences between mice administered corticosterone and their respective controls in the elevated plus maze, open field test, or any differences in sucrose consumption or lick cluster size. However, the administration of corticosterone did elicit some physiological changes, with mice administered corticosterone having a worsened coat state. This suggests that some physiological effects are detectable even in the absence of changes in affective state, and that physiological and behavioural measures of animal affect and welfare are dissociable from one another.

4.2 Introduction

The exposure of an animal to stressful events is known to negatively impact their wellbeing, and can have detrimental effects on their underlying affective state and welfare. Therefore, it is important to develop methods which can tell us more about how an animal might 'feel', in order to minimise the exposure to stressful experiences and ensure good welfare and positive affect (Dawkins 1990; Mendl, Burman, et al. 2010).

It is, of course, not easy to ascertain how an animal might 'feel' due their inability for self-report, and therefore the development of suitable proxies are required (Dawkins 1990; Mendl, Burman, et al. 2010). How an animal responds to reward (i.e. their hedonic responses) may offer a useful way of accessing their underlying affective state (Paul et al. 2005; Boissy et al. 2007; Mendl, Burman, et al. 2010; Nettle & Bateson 2012). This is because humans with Major Depressive Disorder (MDD) often report alterations in their perception of reward, where their sensitivity to reward often decreases, a symptom referred to as anhedonia (American Psychiatric Association 2014).

Traditionally, anhedonia has been assessed in rodents by assessing the amount of sucrose they consume. Specifically, rodents that are anhedonic and have a depressed-like phenotype will drink less of a rewarding sucrose solution than animals in a more positive affective state (Willner et al. 1987; Papp et al. 1991; Willner et al. 1992; Muscat et al. 1992; Monleon-Paolo et al. 1995). However, it is also possible to assess how rewarding an animal finds something by assessing their specific licking patterns, or their lick cluster size. When rodents drink, they produce fast, rhythmic sets of licks that can be grouped into clusters, where the average number of licks in a cluster can be taken to measure the animals hedonic response (Davis 1973; Davis 1989; Hsiao & Fan 1993; Dwyer 2012; Austen et al. 2016). Given that the Chronic Mild Stress (CMS) manipulation was ineffective in Chapter 3, and there was a large amount of experimental variability, I wanted to use another established rodent model of stress-induced depression to assess the effects of stress on sucrose consumption and licking microstructure.

The administration of exogenous corticosterone in rodents is a well-established methodology of inducing a depressed-like state (Magariños et al. 1998; Fairchild 2003; Ardayfio & Kim 2006; Murray et al. 2008; Gourley & Taylor 2009; Rainer et al. 2012; Mekiri et al. 2017). Rodents which receive chronically administered (up to 28 days) exogenous corticosterone (the equivalent of cortisol in humans; Gong et al. 2015), display a more anxious and depressive-like phenotype (Magariños et al. 1998; Fairchild 2003; Ardayfio & Kim 2006; Murray et al. 2008; Gourley & Taylor 2009; Rainer et al. 2012; Mekiri et al. 2017). This has been demonstrated behaviourally, in tests such as emergence from the light/dark box, the open field and sucrose preference test (Ardayfio & Kim 2006; Ali et al. 2015; Sturm et al. 2015; Weng et al. 2016). For example, rodents which have been chronically

administered corticosterone take longer to emerge into the light component of the light/dark box indicative of an anxiogenic like effect (Ardayfio & Kim 2006) and have a lower preference for sucrose in the sucrose preference test (Gourley & Taylor 2009; Ali et al. 2015; Weng et al. 2016) suggestive of a state analogous to an anhedonic depressed-like state. Physiological and pharmacological measures support the outcomes of behavioural tests (Fairchild 2003; Zhao et al. 2008; Murray et al. 2008; Ali et al. 2015; Weng et al. 2016), demonstrating the robustness of this manipulation.

In the present study, I chose to administer corticosterone in the drinking water to minimise the stress associated with repeated subcutaneous injections or surgical implantation of a slow release corticosterone pellet. Although this method offers less control as to the dosage of corticosterone compared to repeated subcutaneous injection, it is still well-established and widely used within the literature (Magariños et al. 1998; Fairchild 2003; Nacher et al. 2004; Ardayfio & Kim 2006; Gourley & Taylor 2009; Rainer et al. 2012; Mekiri et al. 2017). This model is preferred to other existing manipulations, such as the CMS paradigm, because it enables more control over circulating corticosterone levels (Johnson et al. 2006; Zhao et al. 2008).

This model is also advantageous because it also has high face validity to the disease aetiology in humans. It is well-known that the Hypothalamic-Pituitary-Adrenal (HPA) axis is dysregulated in human patients with mood disorders (Rubin et al. 1987; Holsboer & Barden 1996; Peeters et al. 2004). For example, patients presenting with Major Depressive Disorder (MDD) often have higher levels of circulating cortisol (Peeters et al. 2004), which together with the ongoing exposure to stressful situations, is thought to contribute to the aetiology of depression. The role of the HPA axis and its impact on affect, is further supported by the fact that patients who have problems with their HPA axis, such as patients with Cushing's disease (who have higher levels of circulating cortisol), also have a high prevalence of depression and anxiety (Starkman et al. 1981; Sonino et al. 1998). Taken together, the administration of exogenous corticosterone is considered a robust and reliable method of inducing a depressed-like state in laboratory rodents.

The aim of this study was to assess whether chronic administration of exogenous corticosterone in the drinking water was sufficient to induce a negative depressive-like state in laboratory mice, and if so whether licking microstructure was able to detect these changes. I predicted that mice treated with corticosterone for 28 days would show more anhedonic and depressive-like behaviour through lower

consumption of, and smaller lick cluster sizes to, rewarding sucrose solutions and more anxiety-like responses in the elevated plus maze and open field test. I also predicted that if these changes were indicative of a depressed-like phenotype then these changes would be reversed by the administration of the anti-depressant fluoxetine hydrochloride.

4.3 Methods and Materials

4.3.1 Ethical Statement

Experiments were conducted at Newcastle University under the Home Office Project license (PC6981D63). All work was conducted in line accordance with the EU Directive (2010/63/EU), ASPA (1986) and the NIH Guidelines for the care and use of animals for experimental procedures (National Institutes of Health 2011). All animals were checked daily and weighed at least three times per week; no adverse effects of the study were reported. At the end of the study, animals were killed humanely by exposure to a rising concentration of carbon dioxide gas, in line with Schedule 1 guidance.

4.3.2 Animals, housing and husbandry

Thirty-two male C57BL/6 mice (*Mus musculus*) were purchased from Charles River Laboratories, UK and were approximately 8 weeks of age upon arrival. They were tail marked using permanent marker for identification. Mice were free from all recognised pathogens and the health status of the colony was monitored following the FELASA health monitoring recommendations (Guillen 2012). Mice were pair housed in M2 cages (33cm (L) x 15cm (W) x 13cm (H), North Kent Plastics), with sawdust bedding, nesting material (4HK Aspen chips, NestPak and Sizzlepet nesting, Datesand Ltd, Manchester) and a clear Perspex handling tunnel (50mm diameter, 150mm length). Because of the welfare benefits from handling mice with a tunnel rather than by their tail (Hurst & West 2010; Gouveia et al. 2013; Clarkson et al. 2018), all mice were handled via the tunnel handling refinement for the duration of the study. All cages were subject to one full cage clean per week, and mice had access to food and water *ad libitum*. Animals were maintained on a reverse 12:12

hour light/dark cycle (10:00 until 22:00) and therefore all behavioural experiments described below were conducted during the animal's active (dark) phase under red light illumination. Animals were kept under standard laboratory conditions, at a temperature of approximately 21±4°C and with a relative humidity of 55±10°C (Home Office 2014).

4.3.3 Sucrose drinking tests

Following one week acclimation period, mice were trained and tested in eight custom made drinking chambers (refer to Chapter 2 for full details). These consisted of standard mice IVC home cages (34 (L) x 19 (W) x 14(D) cm) with clear perspex sides, a metal perforated floor and wire cage lid with two modified attachments to connect the sipper tubes to either the right or left hand side of the cage. Solutions were made available through drinking spouts attached to 50ml falcon tubes, presented on the left hand side of the cage. The drinking chambers were connected to contact sensitive Med Associates dual contact lickometers (Med Associates Inc., St. Albans, Vermont), which transmit the time of each lick to the nearest 0.01 second to a computer using MED-PC software. Custom-built software calculated the lick cluster sizes according to a range of interbout intervals, which is the length of time used to determine when licks can be considered to be in a single bout (Davis & Smith 1992; Davis & Perez 1993; Dwyer 2012; Dwyer et al. 2013). The data presented here use an interbout intervals of 250ms, 500ms or 1000ms, meaning that any duration of 250ms, 500ms or 1000ms or longer between two licks defined the end of one bout and the start of the next. However, the data were qualitatively and quantitatively the same for a range of different interbout intervals.

They were sub-divided into four groups of 8 (two mice from each treatment per group), which were tested in the same order at the same time each day. Testing was conducted after the lights went off to motivate the mice to consume the sucrose solutions during these tests. Mice were exposed to 4% (w/w) sucrose for 15 minutes per day for six days in order to obtain baseline values before compounds were administered. For the first three sessions, the spout was left to protrude into the cage to ensure engagement with the task. For the remaining sessions, the spout was positioned in line with the cage in order to reduce accidental contact. After the

compounds were administered, mice were tested three times per week, where they had access to 4% sucrose for 15 minutes per session, for the remainder of the study.

4.3.4 Drug administration and treatment groups

Following habituation to the laboratory and baseline training sessions in the drinking apparatus, the animals were randomly assigned to one of four treatment groups, with different drugs administered through the drinking water (n=8 for all groups: Corticosterone, Fluoxetine, Corticosterone and Fluoxetine, and Vehicle-only). These compounds were administered via the drinking water for 28 days in accordance with well-established protocols in published work (Fairchild 2003; Nacher et al. 2004; Ardayfio & Kim 2006; Rainer et al. 2012; Mekiri et al. 2017). This methodology was the preferred method of drug administration to avoid the stress associated with chronic injections or surgical implantation of subcutaneous corticosterone pellets. Due to the fact that mice were pair housed, mice in the same cage were allocated to the same drug treatment regime.

To make the solutions, corticosterone and fluoxetine hydrochloride were purchased from Sigma-Aldrich, UK (Corticosterone product number: 27840, Fluoxetine Hydrochloride European Pharmacopoeia (EP) Reference Standard: F0253000). Solutions were made so that compounds were administered at the final concentrations of 35µg/mL for corticosterone and 112.5µg/mL for fluoxetine hydrochloride. The concentrations were calculated from the animals' average water intake across three consecutive days, the mean bodyweight of the mice, and the desired dosage to be administered in line with published work (5mg/kg/day for corticosterone, and 18mg/kg/day for fluoxetine; David et al. 2009; Rainer et al. 2012). Corticosterone had to be dissolved in ethanol due to its insolubility, before being diluted to 0.3%, meaning all compounds were administered in a 0.3% ethanol vehicle which meant that my vehicle only treatment consisted of 0.3% ethanol. All solutions were available *ad libitum* in opaque drinking bottles (to protect the compounds from light), and were replaced every 3 days.

4.3.5 Coat State Scoring

On day 29, all mice were scored with regards to their coat state by three treatment-blind observers. The observers were two experts (one veterinary surgeon and one animal technician) and one non-expert (a researcher not working with mice). Mice were scored on a scale of 0-2 (half marks were possible) according to the following criteria: 0 - shiny, clean, smooth and well-groomed coat; 1 - coat dull, ungroomed and might be soiled; 2 - ruffled and untidy coat, which can be greasy and stick together. The mean coat state score of each mouse was calculated from the independent scores of the three observers.

4.3.6 Elevated Plus Maze

At the end of the experiment (day 29), mice underwent behavioural testing in an elevated plus maze. The maze arms were 30cm (L) x 5cm (W) with side walls of 15cm on the two closed arms, and was elevated 50cm from the ground. Mice were delivered to the centre of the maze facing an open arm and their behaviour was filmed from above for 5 minutes (Cube HD 1080, Y-cam). After testing, each mouse was returned to either a holding cage or their home cage depending on whether it was the first or last mouse to undergo testing from its cage. The maze was cleaned with 70% ethanol and dried with a paper towel between subjects. The order in which mice were tested was counterbalanced with respect to the treatment group across the testing day. They were filmed from above and later analysed using Observer XT (v11, Noldus, Virginia, USA). The time spent in either the open or closed arms (defined when all four paws were in the arm) was scored by an observer blind to treatment group. One animal jumped off the maze before the end of the test and was excluded from the statistical analysis. This meant that the sample sizes for this behavioural test were: Corticosterone n=8; Fluoxetine n=8, Corticosterone & Fluoxetine n=8, Vehicle n=7).

4.3.7 Open Field Test

I also conducted an open field test on day 30. Each mouse (n=8 for each treatment group) was individually placed using the tunnel in the centre of a

rectangular arena (54.5cm (L) x 35.5cm (W) x 17cm (H)) made of white plastic with a transparent perspex lid and allowed to freely explore for 10 minutes. The order in which mice were tested was counterbalanced with respect to treatment group across the testing day. Behaviour was filmed from above and later analysed using Ethovision XT (v 5.1, Noldus, Virginia, USA). This automatically calculated the total duration spent in the centre, crosses to the centre, total distance travelled, the total time spent moving and the velocity of movement for each mouse. Presence of defecation during the open field test was also recorded.

4.3.8 Statistical Analyses

All statistical analyses were conducted using IBM Corp. SPSS (v23, SPSS Inc, Chicago, USA). Datasets were tested for normality using the Shapiro-Wilk test and homogeneity of variance using and the Levene statistic test and where assumptions were not met, non-parametric statistical methods were used. Where significant main effects were found, Bonferroni post hoc tests were performed to investigate pairwise comparisons (Table 4.1). In all statistical tests differences were considered significant using a p value <0.05.

Table 4.1: Statistical tests for each data set with respective factors, experimental unit and sample size.

Data	Dependent Variable	Statistical Test	Factor(s)	Unit	Sample Size
Sucrose drinking	Consumption (g); Lick cluster size	Repeated measures ANOVA	Between subject Factors: Treatment group (4 levels) Within-subject Factors: Week (5 levels)	Mouse	n=8 Cort n=8 Fluox n=8 Cort & Fluox n=8 Vehicle
Elevated plus maze	Number of open arm entries; Duration on open arms; Percentage of time spent on open arms; Number of protected stretch attend postures	Kruskal-Wallis Test	Treatment Group (4 levels)	Mouse	n=8 Cort n=8 Fluox n=8 Cort & Fluox n=8 Vehicle
Open field test	Duration of movement; Duration in centre; Crosses to centre; Distance travelled; Mean velocity	Kruskal-Wallis Test	Treatment group (4 levels)	Mouse	n=8 Cort n=8 Fluox n=8 Cort & Fluox n=8 Vehicle
Coat State Score	Mean Coat State Score	Kruskal-Wallis Test	Treatment Group (4 levels)	Mouse	n=8 Cort n=8 Fluox n=8 Cort & Fluox n=8 Vehicle

4.4 Results

Mice in all treatment groups increased their consumption across the 5 weeks ($F_{2.004,56.114}$=13.03, p<0.001; Figure 4.1), but there was no interaction between treatment and week ($F_{6.012,56.114}$=1.61, p=0.16). There was a main effect of treatment group ($F_{3,28}$=3.11, p=0.042). However when assessing Bonferroni adjusted pairwise comparisons between the treatment groups, I found that although there was a tendency for mice administered corticosterone to drink less overall, this didn't reach significance when compared to each respective treatment group (all p values >0.05).

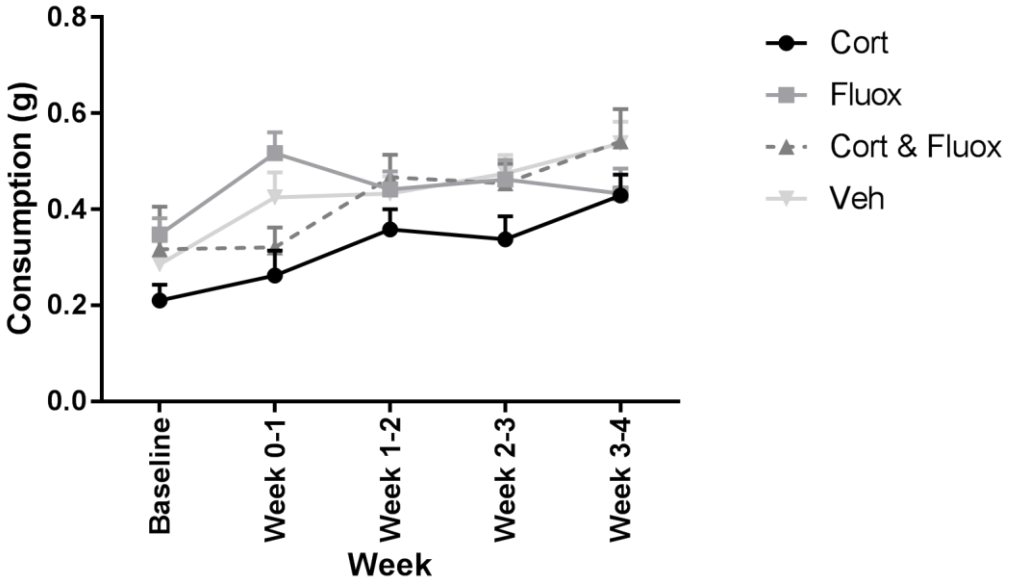


Figure 4.1: Mean (+SEM) consumption (g) of 4% sucrose for baseline and each week of the 28 day administration period.

I also assessed how the lick cluster size changed across time for the four groups. In line with the results for total consumption, lick cluster size increased across weeks (250ms: $F_{2.338,65.457}$=32.07, p<0.001; Figure 4.2A; 500ms: $F_{2.747,76.904}$=20.22, p<0.001; Figure 4.2B; 1000ms: $F_{2.673,74.835}$=13.54, p<0.001; Figure 4.2C), but there was no significant effect of treatment group on that increase (i.e. no significant interaction: 250ms: $F_{7.013,65.457}$=1.52, p=0.18, Figure 4.2; 500ms: $F_{8.240,76.904}$=1.32, p=0.25, Figure 4.2B; 1000ms: $F_{8.018,74.835}$=1.08, p=0.39, Figure 4.2C). However, in contrast to consumption, there was no significant effect of treatment group on lick cluster size (250ms: $F_{3,28}$=0.87, p=0.47; Figure 4.2A; 500ms: $F_{3,28}$=0.64, p=0.60; Figure 4.2B; 1000ms: $F_{3,28}$=0.43, p=0.74; Figure 4.2C).

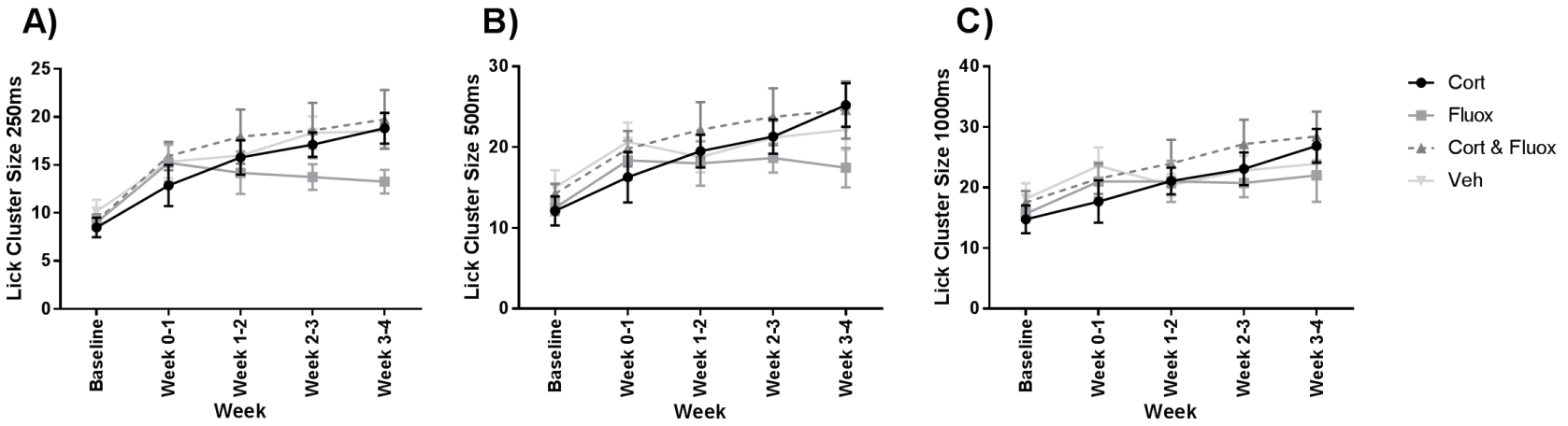


Figure 4.2: Mean (±SEM) lick cluster size of mice drinking 4% sucrose during baseline sessions and in each week of the administration period with an inter bout interval of: A. 250ms, B. 500ms, and C. 1000ms between two subsequent licks

4.4.1 Elevated Plus Maze

Whilst I expected that the administration of chronic corticosterone would increase anxiety-like behaviour, this was not evident. There was no difference between the four treatment groups in open arm entries (H(3)=1.58, p=0.66; Figure 4.3A), or time spent on the open arms (total duration: H(3)=2.66, p=0.45; Figure 4.3B; percentage duration: H(3)=2.58, p=0.46; Figure 4.3C). There was also no difference in the number of protected stretch attend (PSA) postures (H(3)=3.40, p=0.33; Figure 4.3D), considered to reflect a risk assessment behaviour (Rodgers & Dalvi 1997; Hurst & West 2010).

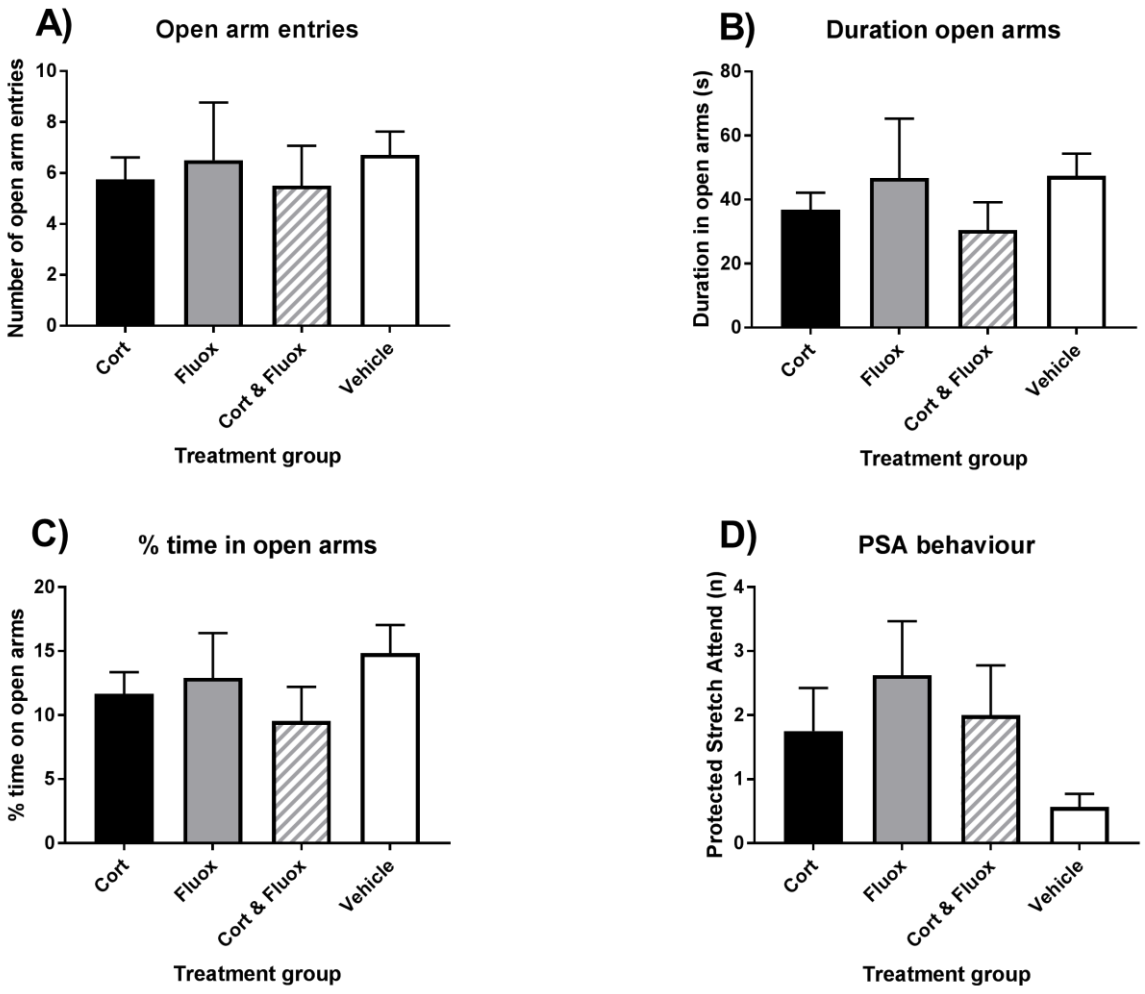


Figure 4.3: Behaviour in the elevated plus maze for each treatment group: A. Mean (+SEM) number of open arm entries; B. Mean (+SEM) duration on the open arms; C. Mean (+SEM) percentage of total time spent on the open arms; and D. Mean (+SEM) number of Protected Stretch Attend (PSA) postures.

4.4.2 Coat State Scoring

There was a significant effect of the drug treatment on the animals' coat scores (Kruskal wallis: $\chi^2(3)=10.12$, p=0.018; Figure 4.4), with a mean rank coat state score of 23.1 for Corticosterone, 12.5 for Fluoxetine, 16.1 for Corticosterone and Fluoxetine and 14.3 for vehicle treated animals. Bonferroni adjusted pairwise comparisons revealed that mice treated with corticosterone had a significantly greater coat state score compared to mice treated with fluoxetine (p=0.018) but not when compared to mice treated with corticosterone and fluoxetine (p=0.306) or vehicle (p=0.083).

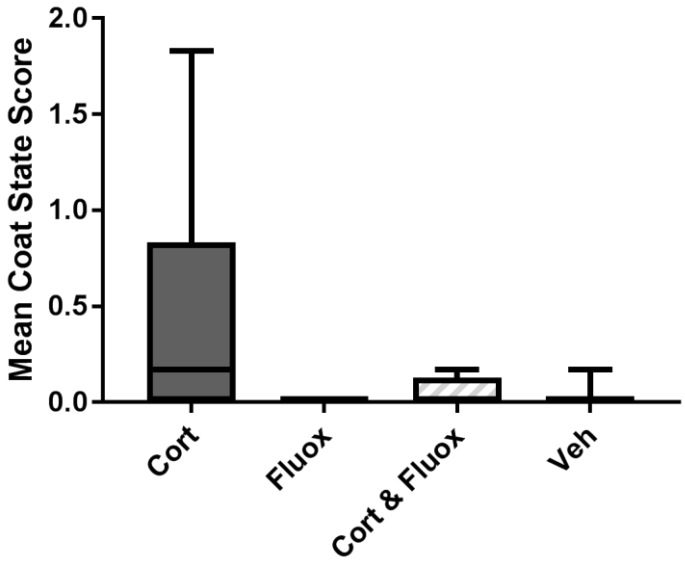


Figure 4.4: Mean (+SEM) coat state score in each treatment group.

4.4.3 Open Field Test

Mice from the Corticosterone group did not show higher levels of anxiety-like responses in the open field test. In line with the results from the elevated plus maze, there was no difference across our treatment groups in the time spent in the centre (Total time: H(3)=1.40,p=0.71; Figure 4.5A; Percentage of time: H(3)=1.76, p=0.63; Figure 4.5B) or in the periphery (H(3)=3.44, p=0.33; Figure 4.5C). The groups also

didn’t differ in the distance mice moved ($H(3)=1.43$, $p=0.70$; Figure 4.5D), the time they spent moving ($H(3)=1.20$, $p=0.75$; Figure 4.5E), or their velocity when moving ($H(3)=1.47$, $p=0.69$; Figure 4.5F). Therefore, there was no evidence that the administration of corticosterone had any effect on the affective state of the mice.

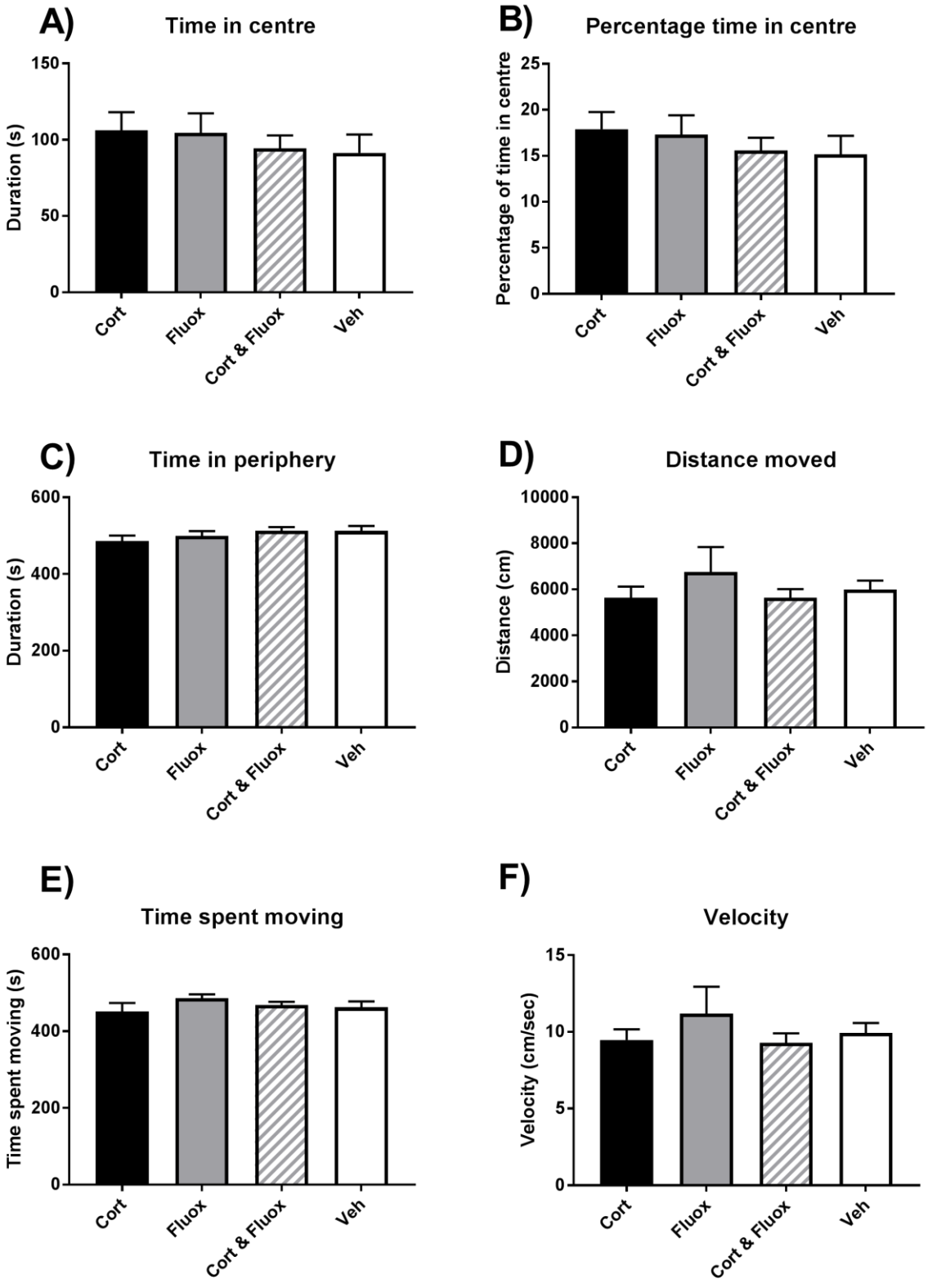


Figure 4.5: Behaviour in the open field test for each treatment group: A. Mean (+SEM) duration (s) spent in the centre of the open field; B. Mean (+SEM) percentage of total time spent in the centre of the open field; C. Mean (+SEM) duration (s) spent in the periphery of the open field; D. Mean (+SEM) distance (m) moved E. Mean(+SEM) time (s) spent moving F. Mean (+SEM) velocity of movement (m/s).

4.4.4 Retrospective calculations of drug dosages

Although I had aimed to deliver the desired doses based on the mass of the mice and the amounts that they drank, the introduction of the compounds to the animal's drinking water affected their intake. In order to keep administration consistent across time, rather than adjust the concentration each week, the same concentration was administered throughout the entire study and the dosages were calculated retrospectively. Although the desired dose of corticosterone of 5 mg/kg/day was achieved for the Corticosterone and Fluoxetine group, it was 10 mg/kg/day for the Corticosterone group. This higher dose for the Corticosterone group was due to this group increasing their intake following the administration of Corticosterone. The dose of fluoxetine was just slightly lower than the desired dose, being 16 mg/kg/day for the Fluoxetine group, and 17 mg/kg/day for the Corticosterone and Fluoxetine group.

4.5 Discussion

The chronic administration of corticosterone did not appear to induce a depressive or anxiety-like state in the mice in my study. There were no behavioural differences between any of the treatment groups in sucrose consumption or lick cluster size, nor on the elevated plus maze or in the open field test. However, the corticosterone treatment did appear to elicit some physiological changes, since mice chronically administered corticosterone had a worsened coat state. Therefore, whilst there appeared to be no change in affective state, the corticosterone manipulation was altering some physiological processes.

It was surprising that the corticosterone treatment did not alter affective state, since the administration of exogenous corticosterone has been well established, and is considered a robust and reliable method to manipulate state, compared to other manipulations, such as the chronic mild stress paradigm (Sterner 2010). A number of studies have reported clear behavioural differences indicative of anxiety and/or depression following the administration of exogenous corticosterone, including: increased immobility time in the forced swim test (Murray et al. 2008), reduced sucrose preference (Ali et al. 2015; Weng et al. 2016), reduced weight gain (Donner et al. 2012), reduced time spent on the open arms of the elevated plus maze

(Skórzewska et al. 2014) and reduced time spent in the centre of the open field (Rainer et al. 2012; Skórzewska et al. 2014). Since I used some of these tests in my study, it seems unlikely that my methods were simply unable to detect a negative affective state, but that the administration of chronic exogenous corticosterone was simply not sufficient to induce a negative affective state analogous to depression and/or anxiety in my mice.

One possible reason for this could be the method by which I administered the corticosterone, since many other studies have used subcutaneous injection. The reason I administered corticosterone via the animals drinking water was to eliminate any stress that could arise from repeated injections. Although this route of administration results in less control over the precise dosage, it does offer a less invasive route of administration and results in more flattened circulating corticosterone levels (Gasparini et al. 2016). Whilst it is a less common method, this route of administration is well-validated and considered robust in producing reliable results, and therefore it seems unlikely that this could explain the lack of effect (David et al. 2009; Gourley & Taylor 2009; Rainer et al. 2012; Mekiri et al. 2017). Furthermore, it is unlikely that the null effect was due to insufficient dosing. This is because corticosterone was administered at a concentration and dose in line with previous published work (David et al. 2009; Gourley & Taylor 2009; Rainer et al. 2012; Mekiri et al. 2017) where clear behavioural differences have been demonstrated (David et al. 2009; Rainer et al. 2012). For example at the same concentration and dose, Rainer et al., (2012) demonstrated that male mice administered corticosterone in their drinking water showed greater anxiogenic behaviour in the open field which could be reversed by the application of the antidepressant fluoxetine. Therefore, taken together, it is unlikely that the method of administration or associated dosage is sufficient in explaining the lack of effect.

However, there are some other differences between my study and those in the literature that may have meant that the corticosterone treatment did not alter affective state in the way I expected. The first is the lighting schedule on which the mice were kept, as they were maintained on a reversed light/dark cycle. Mice are nocturnal animals, and tend to do most of their eating and drinking within the first few hours of the dark phase (Millard et al. 1983). Since I was interested in their drinking behaviour, I conducted the sucrose drinking experiments towards the start of the dark phase to ensure maximal consumption of the solutions without the requirement for

food or water restriction. However, previous research using the chronic corticosterone model have maintained the animals under standard laboratory conditions under non-reversed lighting. Although I can only speculate, it is possible that the reversal of the light/dark cycle could have contributed to the lack of effect on the mice's affective states, since the phase of the light/dark cycle is known to impact on the efficacy of stress manipulations and can significantly affect findings (Aslani et al. 2014). In addition, this affects the timing of the behavioural tests relative to when the mice do most of their drinking. In this study, behavioural testing started shortly after (10-30 minutes) their lights went off. This means that it is likely that they underwent behavioural testing before they had time to drink the full dose of the compounds prior to testing. This differs to that of previous studies where behavioural tests were conducted during the animals' light phase, and therefore animals will have been tested <12hours after dosing, meaning they had more time to be affected by the full dose of the compounds. Therefore, although I cannot make any firm conclusions, it is possible that conducting behavioural tests in the dark phase rather than in the light phase (as in previous studies) could explain why I did not detect a difference among my groups.

The second methodological difference was the choice of handling method. In this study, all mice were handled using a tunnel rather than using the standard practice of tail handling (Deacon 2006; Leach & Main 2008), which is known to reduce the expression of behaviours associated with anxiety and depression (Hurst & West 2010; Gouveia et al. 2013; Clarkson et al. 2018). Since tunnel handling is a relatively recent husbandry refinement and not yet widely implemented across research institutions, it seems likely that most, if not all, previous studies have used tail handling. Unfortunately, existing publishing guidelines do not require the handling method to be specified, and I can only assume that previous studies reporting the depressogenic effects of chronic corticosterone employed tail handling to handle their mice (David et al. 2009; Gourley & Taylor 2009; Rainer et al. 2012; Mekiri et al. 2017). Since tunnel handling reduces stress and anxiety in mice (Hurst & West 2010; Gouveia et al. 2013; Clarkson et al. 2018), this might mitigate against the impact of other stressors and/or elevated corticosterone levels. My findings raise the question of whether, if tail handled, there would have been an effect of corticosterone and differences in behaviour between our treatment groups. More research is needed to

investigate this question, and better understand the effects that current handling practice has on rodent models of depression, stress and anxiety.

However, despite not being able to detect any clear behavioural differences indicative of depressive and/or anxiety-like states in my mice, there was some indication that it caused some physiological changes. In line with previous work, I found that mice administered corticosterone had a worsened coat state (Rainer et al. 2012). Therefore, it seems that some physiological effects are detectable even in the absence of any change in affective state. This could be because the effects of corticosterone on physiology and behaviour are dissociable. Similar results have been found in other studies; for example, Murray, Smith, & Hutson (2008) found that chronic administration of corticosterone was sufficient to induce cellular changes in the hippocampus, but not to detect a depressed-like phenotype in the forced swim test. Therefore, this highlights that in some instances behavioural and physiological measures may not be complementary and emphasises that an animal's affective state cannot be inferred from a single behavioural or physiological measure.

To conclude, I was unable to induce a negative affective state through the administration of chronic exogenous corticosterone, although some change in physiology was detected. Consequently, the work opens up some interesting research questions about whether routine laboratory practices, such as handling method, alters animals' resilience to additional stressors, and how existing physiological and behavioural measures of affective state are related to each other. Future research could explore the full impact of handling method on laboratory mice in scientific experiments.

Chapter 5: Handling method alters the hedonic value of reward in laboratory mice

** The work presented in this chapter has been published (Clarkson, Dwyer, Flecknell, Leach & Rowe (2018): Handling method alters the hedonic value of reward in laboratory mice. Scientific Reports 8: 2448; See Appendix A).*

5.1 Abstract

Given the inability of standard manipulations to induce negative affect in the previous two chapters, I explored the effect of routine handling practices on the affective state of laboratory mice. This is because recent studies have identified that the standard practice of handling laboratory mice by their tails increases behaviours indicative of anxiety, which can be overcome by handling mice using a tunnel. It is important to refine laboratory procedures and practices to ensure high standards of animal welfare and scientific data quality, particularly for mice given their widespread use. However, despite clear negative effects of handling mice by their tails, the refinement of tunnel handling has yet to be widely implemented across research institutions.

In this study, I provide the first evidence that tail handling also reduces mice's responses to reward. I found that tail handled mice showed more anhedonic responses in both measures compared to tunnel handled mice, indicative of a decreased responsiveness to reward and therefore a more depressive-like state. Tail handled mice drank less sucrose and the size of their licking bouts, or their lick cluster sizes, were also smaller. The findings have significant implications for the welfare of laboratory mice, as well as the design and interpretation of scientific studies, particularly those investigating or involving reward.

5.2 Introduction

Mice are the most widely used model species in biomedical research. Consequently, understanding the experiences of mice used in research is of significant importance in order to provide evidence-based improvements to housing

and husbandry that will bring welfare benefits to a large number of animals, and ensure that empirical findings are robust (Wurbel 2001; Benefiel et al. 2005).

Early work aimed at measuring and improving laboratory mouse welfare investigated the housing in which mice are kept: small cage sizes, lack of environmental enrichment, room temperatures and isolation can all negatively impact on mouse welfare, producing measurable changes in behaviour, physiology and/or affective state (Ortiz et al. 1985; Chapillon et al. 1999; Chourbaji et al. 2005; Gaskill et al. 2009). However, more recently, it has been proposed that the handling technique used by researchers and laboratory staff influences not only the welfare of mice, but also data obtained from behavioural studies (Hurst & West 2010; Gouveia et al. 2013; Ghosal et al. 2015; Gouveia et al. 2017). Handling represents the most common procedure experienced by laboratory mice, meaning that refinements to current practice that minimise stress, could have a profound effect.

The standard and most widely used practice of handling laboratory mice is to use their tails, referred to as tail handling and therefore likely to affect millions of mice used in research annually (Deacon 2006; Leach & Main 2008). However, tail handling appears to increase anxiety compared to being handled using a tunnel or by cupping mice on the open hand (Hurst & West 2010; Gouveia et al. 2013; Gouveia et al. 2017). For example, compared to tail handled mice, tunnel handled mice spend more time voluntarily interacting with a handler, and show less anxiety-related behaviour in standardised behavioural tests of anxiety such as the elevated plus maze (Hurst & West 2010; Gouveia et al. 2013; Gouveia et al. 2017). Tail handling can also reduce the performance of mice in cognitive tests (Gouveia et al. 2017). This was demonstrated by the impairment of tail handled mice to engage with a novel mouse odour in a habituation-dishabituation task (Gouveia et al. 2017) and therefore has a number of implications such as the requirement for larger sample sizes and longer training periods. Therefore in relation to the principle of the 3Rs, tunnel handling not only offers to make refinements to existing practices, but also the possibility of reducing the number of mice required for experiments. However, despite the evidence that tail handling can impair welfare and scientific data collection, it remains the main method used to handle mice, and refinements such as tunnel handling have yet to be widely implemented across research institutions.

Here, I tested if being handled by the tail or with a tunnel can affect the hedonic responses of mice towards a rewarding stimulus. Whilst previous studies

investigating the effects of handling method on mouse welfare have measured the animals' behaviour towards aversive experiences or punishments (such as being picked up by a handler or being placed in a novel test environment; Hurst & West 2010; Gouveia et al. 2013; Gouveia et al. 2017), measuring responses to positive experiences and rewards (hedonic responses) are also important for understanding the full impact of handling methods on the affective state of an animal (Boissy et al. 2007; Yeates & Main 2008; Mendl, Burman, et al. 2010). How an animal responds to both punishment and reward offers a way of accessing their enduring negative affective states, and therefore offer a way of determining their cumulative experiences and how these influence their welfare. This is because whilst anxiety and depression can both be characterized by greater expectation of punishment, depression is also associated with a reduced expectation of reward (Nettle & Bateson 2012). Therefore in order to assess the full extent of handling method on affective state and establish whether the stress associated with tail handling also has a depressogenic effect, information about how mice respond to reward is also required.

Stress has been widely documented to play a role in the development of a depressive-like state in both humans and animals (Papp et al. 1991; Hammen et al. 2009). Specifically, in rodent models, exposure to either a single severe (acute) stressor, or several mild (chronic) stressful experiences are sufficient in inducing a depressed-like state (Simson et al. 1986; Willner et al. 1992). This depressed-like state has been validated by assessing the amount of sucrose they consume, whereby a lower consumption is taken to reflect an anhedonic-like state (Willner et al. 1987; Muscat & Willner 1992; Willner 1997; Willner 2017b). Anhedonia is defined as the reduction or inability to experience pleasure from rewarding stimuli (Ribot 1897; Gorwood 2008) and is a core symptom of Major Depressive Disorder (MDD) in humans (American Psychiatric Association, 2014). Consequently, assessing anhedonia in mice has been important for developing and validating laboratory models of depression (Deussing 2006). Historically, hedonic state has been measured in rodents using voluntary consumption of sucrose solutions, under the assumption that anhedonia results in sucrose being perceived as less pleasant, which results in lower intake (Willner et al. 1987; Papp et al. 1991; Forbes et al. 1996; Willner & Healy 1996). For example, mice that have undergone established manipulations, such as the chronic mild stress paradigm, and show behavioural

symptoms of depression also drink less sucrose solution than control animals (Papp et al. 1991; Muscat & Willner 1992; Willner et al. 1992; Willner 1997). This effect can be reversed through the application of anti-depressants which has led to sucrose consumption being widely used as an indicator of affective state in rodents (Willner et al. 1987; Muscat & Willner 1992; Monleon et al. 1995; Forbes et al. 1996; Willner 1997; Willner 2017a).

Despite its widespread application and use, the sucrose consumption test is only an indirect indicator of hedonic state. This is because sucrose consumption is influenced by a number of factors: whilst the amount of sucrose solution a mouse drinks may be driven in part by how much it likes the taste, it may also be affected by motivational factors (Brennan et al. 2001) and the post-ingestive effects of the sucrose (see Chapter 1 for further discussion) (Booth et al. 1972; Warwick & Weingarten 1996). Alternative and more direct measures of hedonic responses towards tastants are based on a more detailed examination of how an animal drinks. The orofacial movements produced upon tasting a solution and the pattern of licks during consumption are both considered to be more direct measures of palatability and hedonic responses to sucrose consumption (Davis 1973; Davis 1989; Davis & Smith 1992; Berridge 2000; Dwyer 2012). In this study, I measured the effect of handling method not only on sucrose consumption, but also on a measure of their licking behaviour considered indicative of their hedonic response to reward. When rodents drink, the pattern of licks is not random (Davis 1973; Davis 1989; Davis & Smith 1992; Dwyer 2012). Instead they produce rhythmic sets of licks that can be grouped into clusters (Davis 1989; Davis & Smith 1992; Dwyer 2012). The number of licks in these clusters, known as 'lick cluster size', is positively related to the palatability of the tastant. For example, larger lick cluster sizes are elicited by more palatable solutions (Davis 1973; Davis 1989; Davis & Smith 1992; Dwyer 2012).

Therefore, the aim of this study was to test if the handling method affected the capacity of mice to experience pleasure (i.e. their hedonic responses) from reward. I predicted that tail handling, a known stressor, would produce measurable changes in mouse behaviour indicative of a depressed-like state. Specifically, I predicted that handling method would affect the hedonic experience of mice drinking sucrose whereby tail handled mice would have lower consumption of, and smaller lick cluster sizes towards, sucrose solutions compared to tunnel handled mice. Assessing both consumption and lick cluster size will enable a comparison between these two

measures and establish whether licking microstructure provides more information regarding the animal's hedonic state.

5.3 Methods and Materials

5.3.1 Ethical Statement

Experiments were conducted at Newcastle University following approval from the University's Animal Welfare and Ethical Review Body (AWERB Project ID: 540), and in accordance with the EU Directive (2010/63/EU), ASPA (1986) and the NIH Guidelines for care and use of animals for experimental procedures (National Institutes of Health 2011). All animals were checked daily, and no adverse effects were reported. At the end of the experiment, animals were humanely killed via exposure to a rising concentration of carbon dioxide gas in accordance to Schedule 1 guidance.

5.3.2 Animals, housing and husbandry

Thirty-two male C57BL/6J mice (*Mus musculus*) were purchased from Charles River Laboratories, UK and were approximately 7 weeks of age (Mean±SEM mass: 24.6±1.6g) on arrival. Mice were free from all recognised pathogens, and the health status of the colony was monitored following the FELASA health monitoring recommendations (Guillen 2012). Mice were pair-housed in M2 cages (33cm (L) x 15cm (W) x 13cm (H), North Kent Plastics), with sawdust bedding, nesting material (4HK Aspen chips, NestPak and Sizzlepet nesting, Datesand Ltd, Manchester) and a clear perspex home cage tunnel (50mm diameter, 150mm length). Cages were cleaned once per week. Animals had access to food (Special Diet Services, RM3E diet) and water *ad libitum*, except prior to drinking experiments (described below). Mice were maintained on a reverse 12:12 hour light/dark cycle (lights off 10:00 until 22:00) and experiments were conducted under red light illumination. Mice were kept under standard laboratory conditions, at a temperature of approximately 21±4°C and with a relative humidity of 55±10°C (Home Office 2014). In line with previous studies, mice were marked for identification using hair dye (Jerome Russel B Blonde, UK) which does not interfere with the response to handling (Hurst & West 2010; Gouveia et al. 2013).

5.3.3 Study Design and timeline

The mice were allowed to acclimatize to the laboratory for one week prior to the start of the experiment (they were not handled during this time). The study had two main phases: 'the handling phase' and 'the sucrose drinking phase' (see Figure 5.1). The handling phase aimed to manipulate the animals experiences and establish clear behavioural differences between tail and tunnel handled mice as previously reported (Hurst & West 2010; Gouveia et al. 2013), before investigating anhedonic behaviour in the sucrose drinking phase.

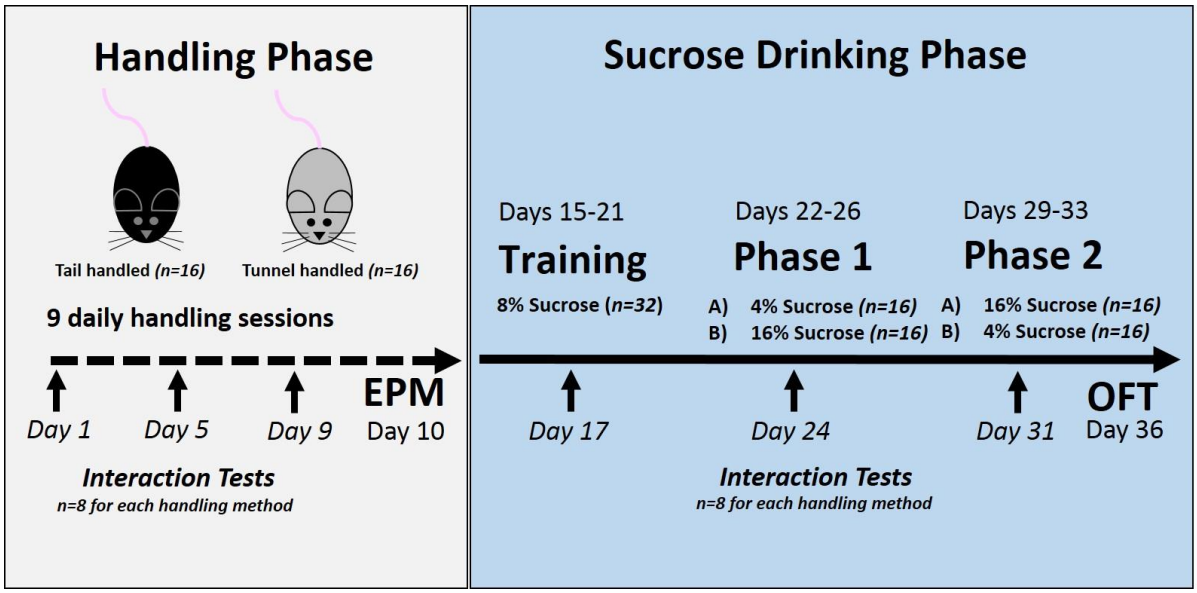


Figure 5.1: Study timeline to show the design and order of behavioural tests and sample sizes.

5.3.4 Handling Methods

Each cage of two mice were randomly assigned (via random number generator) to one of two handling treatment groups, tail or tunnel handled (n=16 mice per handling treatment group). Mice were then only handled by their designated method (tail or tunnel handled) by the same handler wearing nitrile gloves. The handler wore gloves that were rubbed in soiling bedding before each handling session (from mice of the same sex and strain) and a laboratory coat that was contaminated with mouse scent (Hurst & West 2010; Gouveia et al. 2013).

Tail handling involved grasping the mouse at the base of its tail using the thumb and forefinger, and then lifting onto the sleeve of the laboratory coat for 30 seconds before being returned to its home cage. For tunnel handling, the mouse was guided into the transparent Perspex tunnel, and lifted above the cage and held for 30 seconds. For the first two days, the handler's hands were loosely cupped over the ends of the tunnel to prevent escape.

During the initial handling phase mice were handled twice daily for 30 seconds, 60 seconds apart, for nine consecutive days. Prior to handling, the nesting material (care was taken not to disrupt the structure) and home cage tunnel were removed. This procedure was also conducted once weekly to coincide with the sucrose drinking phase (days 17, 24 and 31). For routine husbandry practices, such as cage cleaning, mice were captured and transferred using their designated handling method either on the sleeve for tail handled mice, or in the tunnel for tunnel handled mice. The same protocol was used when transferring mice to behavioural tests, i.e. the elevated plus maze, open field test and sucrose drinking chambers.

5.3.5 Voluntary interaction tests

On designated days during the handling phase (days 1, 5 and 9) and the sucrose drinking phase (days 17, 24 and 31), each cage of animals underwent 'voluntary interaction tests' to assess their responses to the handler (see Figure 5.1). These tests allowed a comparison of behaviour in anticipation of being handled compared to after the animals were handled on specified test days (Hurst & West 2010; Gouveia et al. 2013). Each test consisted of removing the cage lid, nesting material and home cage tunnel and the handler standing motionless in front of the cage for 60 seconds. A gloved hand (tail handled) or a gloved hand holding the home cage tunnel (tunnel handled) was held resting on the substrate in the front of the cage for 60 seconds to assess voluntary interaction. Each mouse in the cage was then handled twice for 30 seconds by their designated handling method described above, before voluntary interaction was assessed again. Behaviour was filmed (Cube HD 1080, Y-cam) and later analysed using Observer XT (v11, Noldus, Virginia, USA). Time spent interacting with the handler was measured for each mouse within a cage, from which an overall mean cage score was calculated as a percentage of the total test time. These were summed together for analyses for the two treatment groups (tail and tunnel handled). Therefore, for these tests, the experimental unit was

'cage' (n=8 for both groups). Interaction was defined as any of the following: sniffing (nose within 0.5cm), touching (including paw contact), climbing on or in the handling tunnel and/or the handler's hand. Due to the differences in how mice in the two treatments were handled during the interaction tests, the observer could not be blind to the treatment, but was blind to whether the interaction test was carried out before ('pre') or after ('post') handling.

5.3.6 Elevated Plus Maze

On day 10, mice underwent behavioural testing in an elevated plus maze (refer to general methods and materials section for full details). Mice were delivered to the centre of the maze (via their designated handling method) facing an open arm and filmed (Cube HD 1080, Y-cam) for 5 mins and returned to a holding cage or the home cage, depending on whether it was the first or last mouse to undergo testing from its cage. Between subjects, the maze was cleaned with 70% ethanol and the running order was counterbalanced with respect to handling method across the testing day. Time spent in the open or closed arms was scored by a treatment blind observer using Observer XT (v11, Noldus, Virginia, USA), where the time spent in an arm was defined as being when all four paws were in the arm. Three animals jumped off before the end of the test and were excluded from statistical analysis meaning the sample sizes were reduced (tail handled n=14; tunnel handled n=15).

5.3.7 Sucrose Drinking Phase

Mice were trained and tested in eight custom made drinking chambers (see Chapter 2 for full details). Solutions were delivered through drinking spouts attached to 50ml falcon tubes. Drinking chambers were connected to contact sensitive Med Associates dual contact lickometers (Med Associates Inc., St. Albans, Vermont), which transmitted the time of each lick to the nearest 0.01 second to a computer using MED-PC software. Custom-built software calculated the lick cluster sizes according to a range of interbout intervals, which is the length of time used to determine when licks can be considered to be in a single bout (Davis & Smith 1992; Davis & Perez 1993; Dwyer 2012; Dwyer et al. 2013). The data presented here use interbout intervals of 250ms, 500ms or 1000ms, meaning that any duration of 250ms,

500ms or 1000ms or longer between two licks defined the end of one bout and the start of the next.

Mice were separated into four groups of eight (four mice from each treatment per group) referred to as the 'testing group'. A random number generator assigned mice into testing groups according to their cage number, meaning both mice within a cage were assigned to the same testing group and were tested in the same order and time each day. Water bottles on the home cage were removed 2 hours prior to sucrose drinking trials, before the lights went off to encourage consumption. Mice were trained across seven consecutive days for 15 minutes each day (Days 15-21) to drink sucrose (8% (w/w) sucrose solution) from the spouts. During the first three training sessions the spout was left to protrude into the cage to ensure engagement with the task, after this the spout was flush with the cage lid in order to reduce accidental contact. Once all animals were engaged with the task and consistently drinking (>100 licks), the experimental phases of the sucrose drinking phase began (see Figure 5.1).

The sucrose drinking phase had two experimental phases (see Figure 5.1), where mice were tested on all 5 days for each phase. Phase 1 (Days 22-26) consisted of half the animals (n=16; 8 from each handling method) receiving 4% (w/w) sucrose and half (n=16; 8 from each handling method) receiving 16% (w/w) sucrose for 15 minutes. This was balanced with regards to treatment group and across testing groups. Phase 2 (Days 29-33) reversed the sucrose concentration. I used two concentrations of sucrose to assess the responses to stimuli with differing hedonic properties. I measured the mass of sucrose solution consumed and the timing of each lick in every test trial; from this, I calculated the mean consumption of sucrose (g) and the mean lick cluster sizes for each animal across the five days at both concentrations for use in my analyses.

5.3.8 Open Field Test

On day 36, each mouse was individually placed via their designated handling method in the centre of a rectangular arena for 10 minutes (see Chapter 2 for full details). The order was counterbalanced with respect to handling method. Behaviour was filmed (Cube HD 1080, Y-cam) and analysed using Ethovision XT (v 5.1, Noldus, Virginia, USA), which automatically tracked the total duration spent in the

centre, crosses to the centre, total distance travelled, the total time spent moving and the velocity of movement for each mouse. Presence of defecation during the open field test was noted. Due to a technical error with the video cameras, videos were only scored for 14 out of the 16 tail handled mice.

5.3.9 Statistical Analyses

All statistical analyses were conducted using IBM Corp. SPSS (v23, SPSS Inc, Chicago, USA). Datasets were tested for normality using the Shapiro-Wilk test and homogeneity of variance using and the Levene statistic test. Where assumptions were not met, data were transformed or non-parametric statistical methods were used. Where significant main effects were found, Bonferroni post hoc tests were performed to investigate pairwise comparisons (Table 5.1). In all statistical tests differences were considered significant using a p value <0.05.

Table 5.1: Statistical tests for each data set with respective factors, experimental unit and sample size.

Data	Dependent Variable	Statistical Test	Factor(s)	Unit	Sample Size
Voluntary interaction tests	Percentage time spent interacting	Repeated measures ANOVA	Between subject factors: Handling method (2 levels) Within-subject factors: day (6 levels), time (2 levels: pre or post handling)	Cage	n=8 tail handled n=8 tunnel handled
Elevated plus maze	Number of open arm entries; Duration on open arms	Mann-Whitney U test	Handling method (2 levels)	Mouse	n=14 tail handled n=15 tunnel handled
	Number of protected stretch attend postures	Independent t-test	Handling method (2 levels)	Mouse	n=14 tail handled n=15 tunnel handled
	Number of mice that defecated	Binary logistic regression	Handling method (2 levels)	Mouse	n=14 tail handled n=15 tunnel handled
Sucrose drinking	Consumption (g); Consumption (ml/g bodyweight); Lick cluster size (log transformed)	Repeated measures ANOVA	Between subject Factor: Handling method (2 levels) Within-subject Factor: Sucrose concentration (2 levels)	Mouse	n=16 tail handled n=16 tunnel handled
Open field test	Duration of movement; Duration in centre; Crosses to centre; Distance travelled; Mean velocity	Independent t tests	Handling method (2 levels)	Mouse	n=14 tail handled n=16 tunnel handled
	Number of mice that defecated	Binary logistic regression	Handling method (2 levels)	Mouse	n=16 tail handled n=16 tunnel handled
Bodyweight (g)	Mean bodyweight	Independent t test	Handling method (2 levels)	Mouse	n=16 tail handled n=16 tunnel handled

5.4 Results

First I established whether the nine day handling regime was sufficient to cause the behavioural differences in voluntary interaction previously described (Hurst & West 2010; Gouveia et al. 2013). Consistent with previous studies (Hurst & West 2010; Gouveia et al. 2013), the repeated handling was sufficient to create significant differences between the two groups of mice in behavioural tests considered to be indicative of anxiety. There was a main effect of handling method, with tunnel handled mice spending significantly more time interacting with the handler compared to tail handled mice overall (ANOVA: $F_{1,14} = 1062.7$, $p<0.001$; Figure 5.2; Table 5.2). There was also a main effect of day (ANOVA: $F_{1,28} = 29.03$, $p<0.001$; Figure 5.2; Table 5.2) and a significant handling method and day interaction (ANOVA: $F_{1,28} = 23.67$, $p<0.001$; Figure 5.2; Table 5.2). Tunnel handled mice increased the time spent interacting with the handler after Day 1 (Bonferroni adjusted pairwise comparisons Day 1 versus Day 5 $p<0.001$; Day 1 versus Day 9 $p<0.001$), but no increase was evident for those mice handled by their tails (Day 1 versus Day 5 $p>0.99$; Day 1 versus Day 9 $p=0.49$).

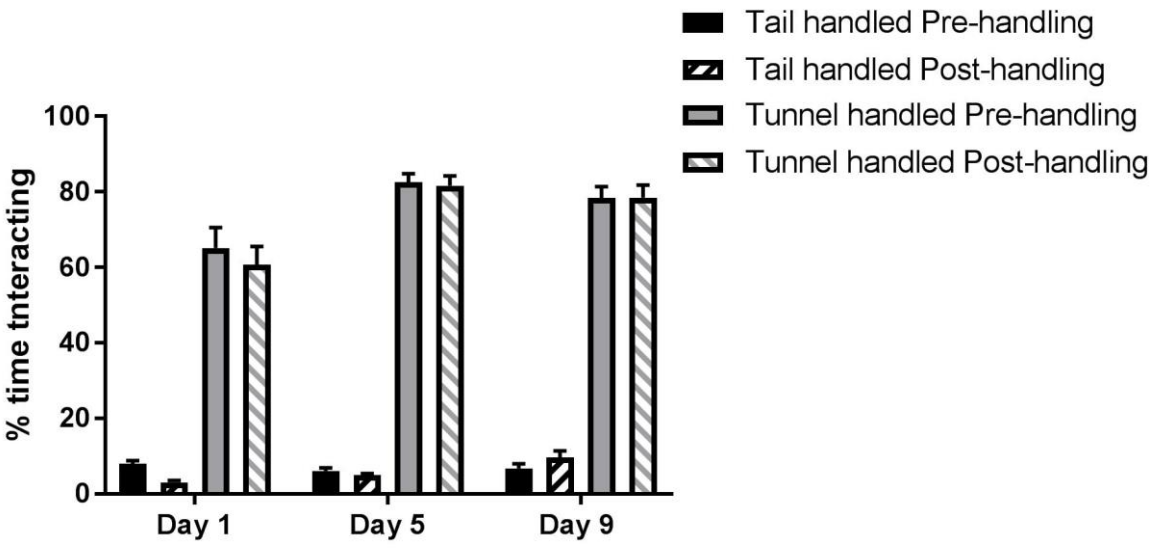


Figure 5.2: Mean (+SEM) percentage of time spent interacting with the handler in the voluntary interaction tests conducted on three different days in the handling phase. Interaction tests were conducted both before (pre-handling) and after (post-handling) the animals were handled via the tail or tunnel handling method.

Table 5.2: Statistical results for the repeated measures ANOVA conducted for the percentage of time spent voluntarily interacting with the handler in line with the handling phase on days 1, 5 and 9.

Factor	**F_{df}**	**p value**
Handling method	$F_{1,14} = 1062.7$	<0.001 ***
Day	$F_{1,28} = 29.03$	<0.001 ***
Time (pre or post handling)	$F_{1,14} = 0.60$	=0.45 ns
Handling method x Day	$F_{1,28} = 23.67$	<0.001 ***
Handling method x Time	$F_{1,14} = 0.04$	=0.85 ns
Day x Time	$F_{1,28} = 1.30$	=0.29 ns
Handling method x Day x Time	$F_{1,28} = 0.11$	=0.89 ns

Mice were also tested in an elevated plus maze on Day 10 to assess their anxiety levels. Consistent with the expectation that tail handling produces higher levels of anxiety than tunnel handling, tail handled mice showed fewer entries onto the open arms (Mann Whitney U= 174.5, p= 0.002; Figure 5.3A), and spent less time on them (Mann Whitney U= 175, p= 0.002; Figure 5.3B). However, although in the predicted direction, the number of protected stretch attend postures ($t_{27} = 1.718$, p= 0.097; Figure 5.3C) and defecation events ($\chi^2(1) = 0.36$, p= 0.552; Figure 5.3D) did not significantly differ between our two handling methods.

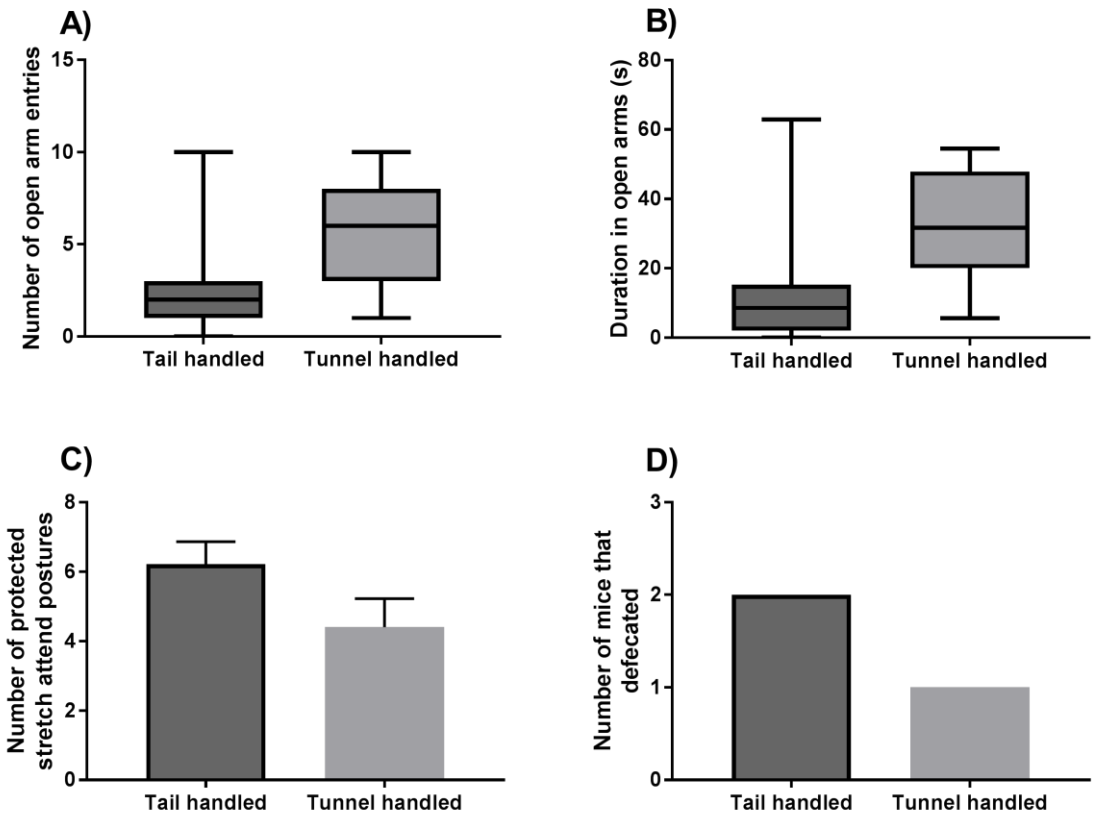


Figure 5.3: Results from tail handled and tunnel handled mice placed in the elevated plus maze on day 10. A. Mean (±SEM) number of entries into the open arms. B. Mean (±SEM) length of time spent (s) in the open arms. C. Mean (+SEM) number of stretch attend postures. D. Total number of mice that defecated during testing.

Once I had established I could detect the same behavioural differences between tail and tunnel handled mice previously described (Hurst & West 2010; Gouveia et al. 2013), I moved onto the sucrose drinking phase. I found that whilst there was a main effect of sucrose concentration (ANOVA: $F_{1,30} = 30.82$, $p<0.001$; Figure 5.4) i.e. both groups of mice drank more sucrose at the higher concentration, there was a also a main effect of handling method (ANOVA: $F_{1,30} = 7.14$, $p=0.012$; Figure 5.4), tunnel handled mice drank significantly more of both sucrose solutions than mice that were tail handled. However, there was no interaction between handling method and sucrose concentration on total consumption (ANOVA: $F_{1,30} = 0.1$, $p=0.754$; Figure 5.4).

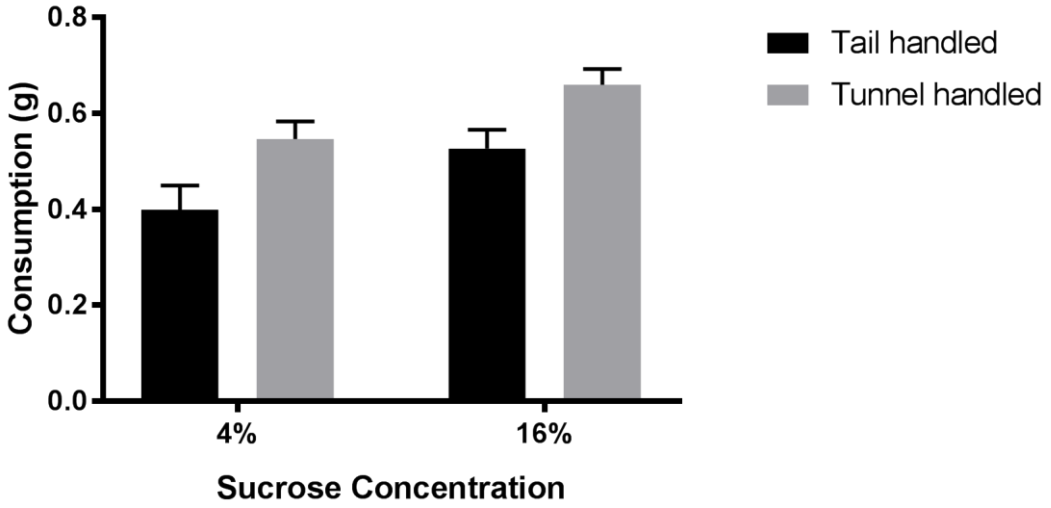


Figure 5.4: The mean (+SEM) amount consumed (g) of 4% and 16% sucrose for tail and tunnel handled mice during the sucrose drinking phase.

Although handling method had no significant effect on the mean body weight of the animals, there was a tendency for tail handled mice to be heavier than tunnel handled mice (t_{30} = 1.905, p=0.066; Figure 5.5A). Therefore, to be sure that bodyweight was not influencing the sucrose consumption data (e.g. by larger mice requiring more calories and being more motivated to drink), I re-analysed the consumption data whilst controlling for body weight (i.e analysing ml/g bodyweight; Figure 5.5B).The results were qualitatively the same: both tail and tunnel handled mice drank significantly more of the higher sucrose concentration ($F_{1,30}$ = 32.56, p<0.001; Figure 5.5B), with tunnel handled mice drinking significantly more sucrose irrespective of concentration compared to tail handled mice ($F_{1,30}$ = 13.43, p=0.001; Figure 5.5B). There was no interaction between handling method and sucrose concentration ($F_{1,30}$ = 0.02, p=0.887; Figure 5.5B). Therefore, any differences in the sizes of the mice were not influencing the findings.

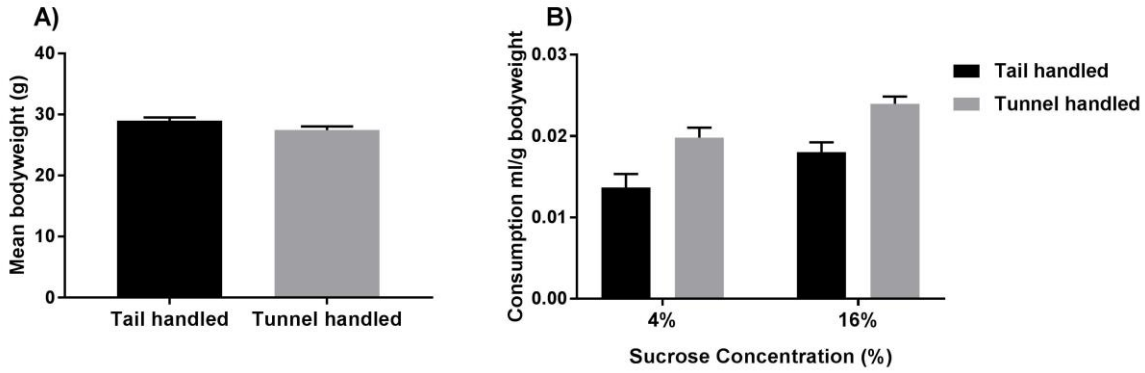


Figure 5.5: A. Mean (+SEM) body weight (g) of tail and tunnel handled mice during the sucrose consumption phase. The mean body weight was derived from the weekly body weights taken during the two phases of the sucrose consumption phase. B. Mean (+SEM) consumption of sucrose solutions during the sucrose drinking phase normalised for body weight (ml/g).

In addition to measuring the total amount consumed, I also assessed how handling method affected mice's licking microstructure. I found a main effect of sucrose concentration, lick cluster sizes were larger when mice drank the higher sucrose concentration (ANOVA: 250ms: $F_{1,30}$= 38.50, p<0.001; Figure 5.6A; 500ms: $F_{1,30}$= 70.12, p<0.001; Figure 5.6B; 1000ms: $F_{1,30}$= 60.25, p<0.001; Figure 5.6C) and a main effect of handling method, tunnel handled mice had larger lick cluster sizes overall (ANOVA: 250ms: $F_{1,30}$= 4.62, p=0.04; Figure 5.6A; 500ms: $F_{1,30}$= 4.16, p=0.05; Figure 5.6B; 1000ms: $F_{1,30}$= 5.78, p=0.02; Figure 5.6C). However, there was also a significant interaction between handling method and sucrose concentration (ANOVA: 250ms: $F_{1,30}$= 10.20, p=0.003; Figure 5.6A; 500ms $F_{1,30}$= 11.44, p=0.002; Figure 5.6B; 1000ms: $F_{1,30}$= 10.20, p=0.003; Figure 5.6C). This was because tunnel handled mice only had significantly larger lick cluster sizes than the tail handled mice when drinking the 4% sucrose solution (Bonferroni pairwise comparisons; all p<0.01) but not the 16% sucrose solution (all p>0.20). However, both tail and tunnel handled mice had larger lick cluster sizes for 16% sucrose compared to those at 4% (Bonferroni pairwise comparisons; tail handled mice; all p<0.001, tunnel handled mice; all p<0.05).

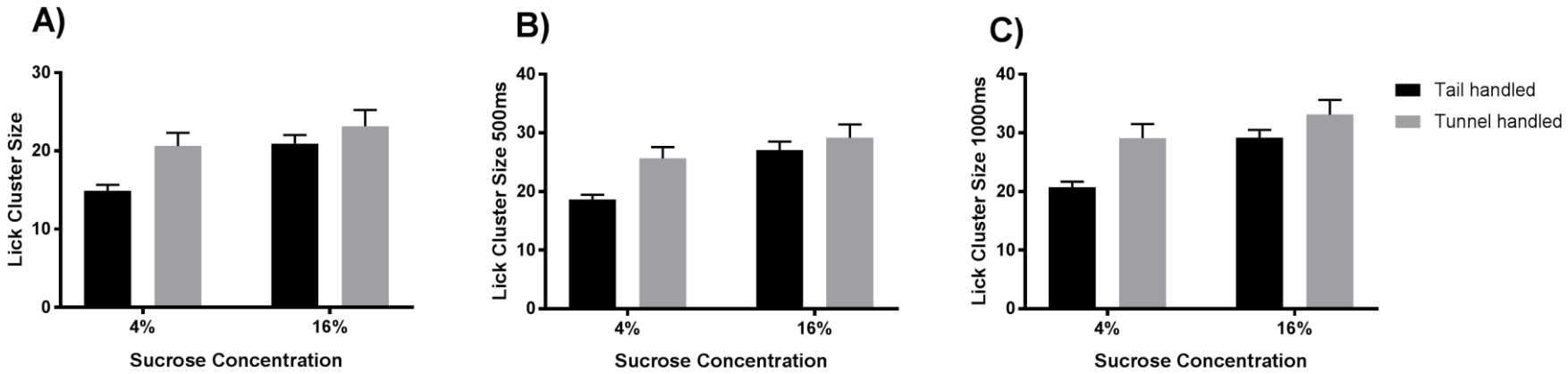


Figure 5.6: The mean (+SEM) lick cluster size for 4% and 16% sucrose for tail and tunnel handled mice during the sucrose drinking phase using A. an interbout interval of 250ms B. An interbout interval of 500ms C. An interbout interval of 1000ms.

In order to ensure that the established effects of tail and tunnel handling on measures of anxiety were still evident and did not diminish during this sucrose drinking phase, I gave mice three further voluntary interaction tests. I found a main effect of handling method, tail handled mice continued to interact significantly less with the handler compared to tunnel handled mice overall (ANOVA: $F_{1,14} = 462.34$, $p<0.001$; Figure 5.7; Table 5.3), although this did interact with day (ANOVA: $F_{1,28} = 19.73$, $p<0.001$; Figure 5.7; Table 5.3). Tail handled mice spent more time interacting with the handler over these later three tests (Bonferroni adjusted pairwise comparisons for days 24 and 31 relative to day 17, p values $p<0.001$; Figure 5.7; Table 5.3). There was also a signficant main effect of timing of the interaction test (i.e. pre or post handling: ANOVA: $F_{1,14} = 5.4$, $p=0.036$; Table 5.3) and a significant handling method and timing interaction (ANOVA: $F_{1,14} = 8.39$, $p=0.012$; Table 5.3). Bonferroni adjusted pairwise comparisons revealed that tail handled mice interacted significantly more with the handler after they had been handled (post handling) compared to before (pre handling) ($p=0.002$) which was not evident in the tunnel handled mice ($p=0.691$).

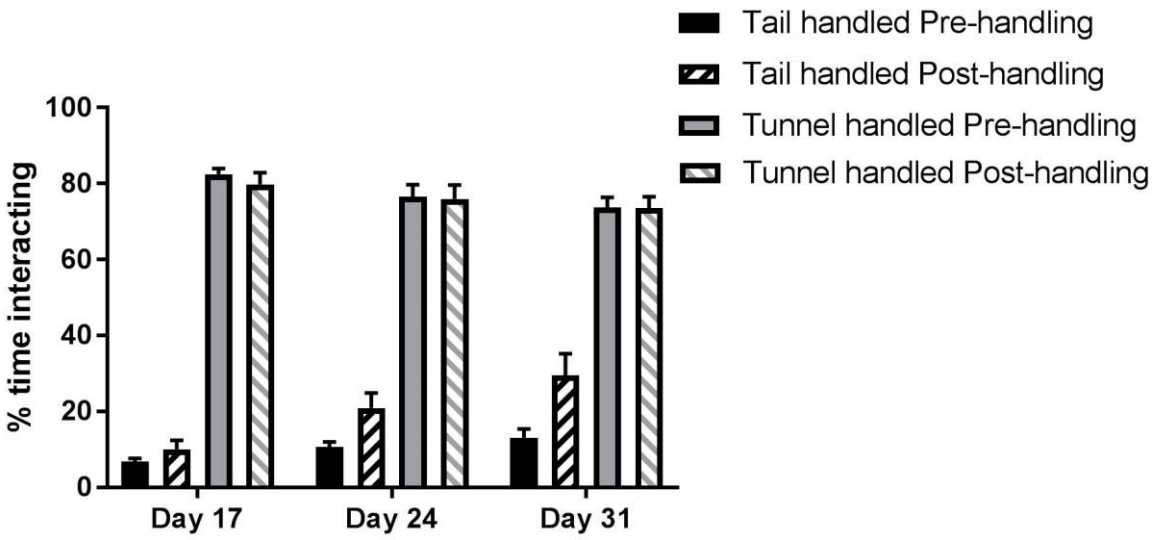


Figure 5.7: Mean (+SEM) time spent interacting with the handler in the voluntary interaction tests conducted on three different days during the sucrose drinking phase. Interaction tests were conducted both before (pre) and after (post) the animals were handled via either the tail or tunnel handling method.

Table 5.3: Statistical results for the repeated measures ANOVA conducted for the percentage of time spent voluntarily interacting with the handler in line with the sucrose drinking phase on days 17, 24 and 31.

Factor	**F_{df}**	**p value**
Handling method	$F_{1,14} = 462.34$	<0.001 ***
Day	$F_{1,28} = 1.36$	=0.27 ns
Time (pre or post handling)	$F_{1,14} = 5.40$	=0.036 *
Handling method x Day	$F_{1,28} = 19.73$	<0.001 ***
Handling method x Time	$F_{1,14} = 8.39$	=0.012 *
Day x Time	$F_{1,28} = 3.06$	=0.063 ns
Handling method x Day x Time	$F_{1,28} = 1.40$	=0.26 ns

Finally, at the very end of the sucrose drinking phase and before euthanasia, I also conducted an open field test as an independent measure of anxiety in both groups of mice. I found that tail handled mice showed significantly higher levels of anxiety. Although the time that mice spent moving did not significantly differ between the handing methods (t_{28}=0.86, p=397; Figure 5.8A), their patterns of movement were very different. Tunnel handled mice spent significantly longer in the centre of the arena (t_{28}=3.29, p=0.003; Figure 5.8B) and performed significantly more crosses into the centre (t_{28}=5.10, p<0.001; Figure 5.8C). They also travelled significantly further (t_{28}=4.36, p<0.001; Figure 5.8D) and had a significantly higher mean velocity when travelling (t_{28}=4.53, p<0.001; Figure 5.8E) than tail handled mice. Tail handled mice also showed a tendency to be more likely to defecate more ($\chi^2(1) = 2.22$, p= 0.136; Figure 5.8F), which is also an indication of stress (Henderson et al. 2004)

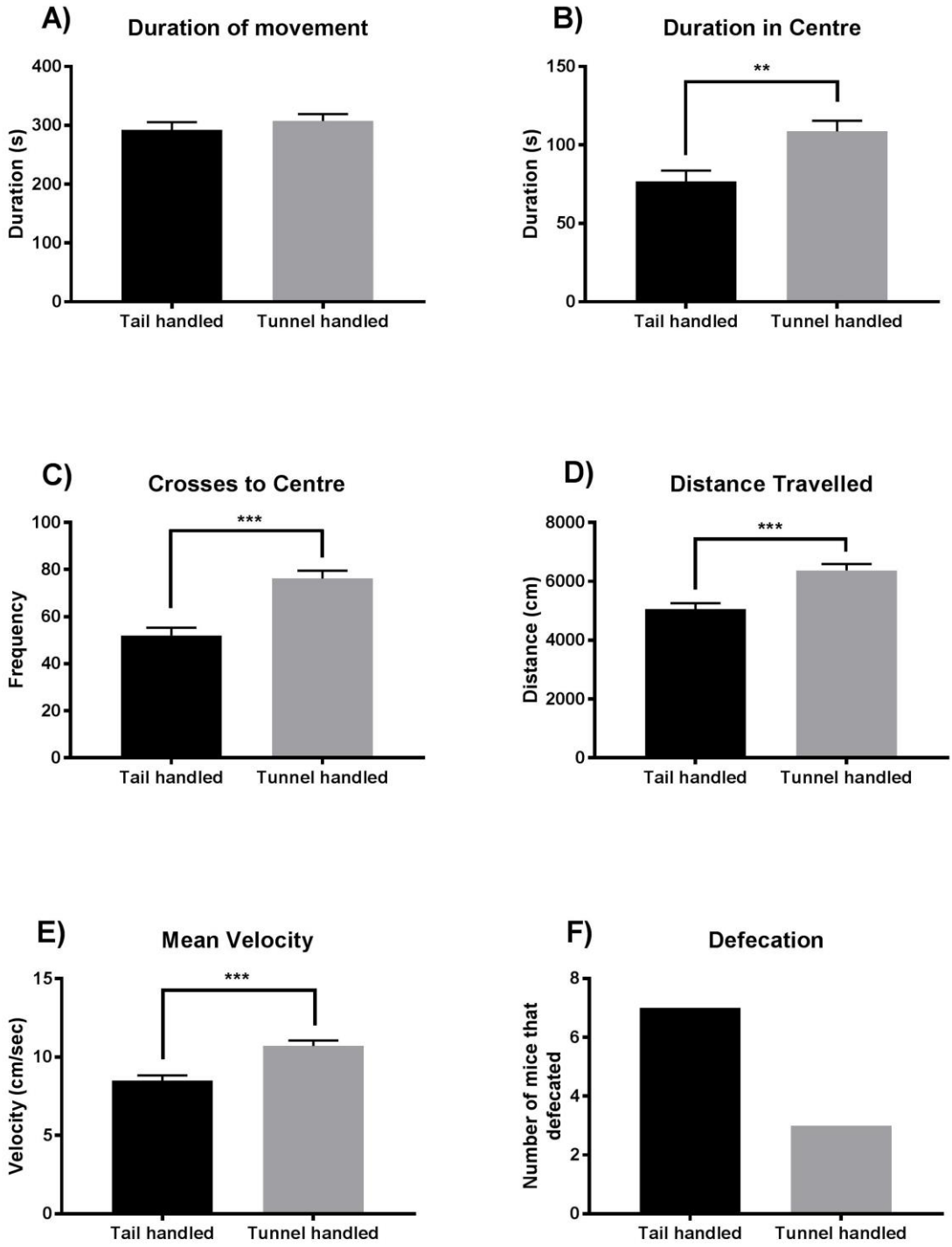


Figure 5.8: The behaviour of tail and tunnel handled mice in the open field test. A. Mean (+SEM) length of time spent moving (s). B. Mean (+SEM) length of time spent in the centre (s). C. Mean (+SEM) number of crosses to the centre. D. Mean (+SEM) distance travelled (cm). E. Mean (+SEM) velocity when moving (cm/sec). F. Total number of mice that defecated during testing. **p<0.01 and *** p<0.001.

5.5 Discussion

The aim of this study was to determine whether the handling method used to handle laboratory mice, was sufficient to cause changes in the animals ability to experience pleasure when presented with rewarding sucrose solutions. This study provides the first evidence that handling method affects how laboratory mice perceive and respond to positive rewarding stimuli. Tail handling not only makes mice more anxious compared to tunnel handled mice (Hurst & West 2010; Gouveia et al. 2013; Gouveia et al. 2017), but it also reduces their hedonic responses towards a sucrose reward. My data show that tail handled mice drank less sucrose at both concentrations and had lower lick cluster sizes overall, although smaller cluster sizes were only evident at the lower concentration (i.e. 4%). This result could be due to a ceiling effect, due to testing under mild water restriction, and further work would be needed in order to determine if the anhedonic effects of tail handling are moderated by the nature of the solution being consumed. However, taken together, my combined data indicate that tail handling makes mice more anhedonic and less responsive to reward compared to being handled using a tunnel. Since tail handling is the most widely used method to handle laboratory mice (Deacon 2006; Leach & Main 2008), this finding has significant implications for animal welfare and the refinement of current laboratory practices, as well as scientific data collection, particularly where protocols include or investigate reward.

The presence of anhedonia in our tail handled mice, combined with increased anxiety-like behaviours relative to tunnel handled mice, is indicative of a more depressive-like state and a more negative affective state compared to tunnel handled mice. Perhaps surprisingly, this difference seen between our handling treatments is similar to that for more severe manipulations which have been explicitly designed to induce depressive-like states in rodents, such as the chronic mild stress paradigm (Willner et al. 1992; Muscat & Willner 1992; Willner 2017b) or chronic restraint (Sun et al. 2015). However, since tail handling may mimic a predatory attack (Layne 1972; Hurst & West 2010; Shargal et al. 2017), it could be that this handling method is inherently more stressful than currently thought (Deacon 2006). At 4% sucrose, the reduction in tail handled mice's consumption compared to that of tunnel handled mice was 27%. From published studies using chronic manipulations to produce models of depression in C57BL/6 mice, I have estimated the reduction in sucrose consumption

relative to controls to be in the range of 33-57% (Pothion et al. 2004; Strekalova et al. 2004). Although these are only estimates, this does suggest that the mice could be subject to a similar depressive-like state following tail handling. This is something that could be explored further pharmacologically, by the application of anti-depressant compounds.

In addition to detecting effects of tail handling towards positive reward, I also found that tail handled mice interacted less with the handler and showed greater levels of anxiety in behavioural tests (elevated plus maze and open field test) compared to tunnel handled mice further supporting the findings of previous studies (Hurst & West 2010; Gouveia et al. 2013; Gouveia et al. 2017). This is the first study to replicate these findings at a different research institution. However, given that our study was longer than previous published work (days 17, 24 and 31 compared to just days 1, 5 and 9), I was also able to explore the effect of tail handling on mouse behaviour over a longer period of time. The results from our open field test carried out at the end of the experiment still showed a significant effect of handling method on behavioural measures of anxiety. Tail handled mice continued to spend significantly less time interacting with the handler than tunnel handled mice across the entire experiment, although I did detect some changes in their behaviour during the sucrose drinking phase (Figure 5.7). Tail handled mice increased the time they spent interacting with the handler over the last three voluntary interaction tests, and spent more time interacting with the handler once they had been handled (post handling). This may be due to simple habituation, the mice forming an association with the sucrose reward, or alternatively perhaps the mice were able to learn about the sequence of events in the repeated voluntary interaction tests. For example, it could be that the tail handled mice were more apprehensive of the handler when they were about to be picked up by their tail, compared to when they had already been picked up by their tail. This could be because they had learnt the sequence of events for example, once they had been handled they were unlikely to be handled again. However, currently it is impossible to draw any firm conclusions.

Taken together, my data clearly demonstrate that mice are more anxious and more anhedonic when they are handled by their tail rather than when using a tunnel. This finding adds to the increasing number of studies that show that tail handling is an aversive procedure (Hurst & West 2010; Gouveia et al. 2013; Ghosal et al. 2015; Gouveia et al. 2017), and that tail handling has a negative impact on the welfare of

laboratory mice. It also shows that sucrose consumption and licking behaviour can be used to assess the presence of positive experiences by measuring hedonic responses towards reward and applied in a welfare context. Recent papers have highlighted that to fully understand animal welfare, we need to measure positive experiences, such as pleasure (Fraser & Duncan 1998; Boissy et al. 2007; Yeates & Main 2008). This is because both positive and negative experiences can cumulatively influence an animal's affective state and welfare (Mendl, Burman, et al. 2010; Bateson 2016). The ability to assess how an animal responds to positive rewarding stimuli, such as sucrose solutions, offers the potential to evaluate the effects of routine laboratory conditions on their affective state and welfare, which formed one of the main aims of this thesis and will be explored further in Chapter 6.

Tail handling has been shown to affect data collection through reducing the likelihood that a mouse will engage with a cognitive task (Gouveia et al. 2017). Our data suggest that the effects of tail handling may be more complex than simply not engaging with a task but could affect how animals respond to rewards in behavioural and cognitive tasks. The vast majority of *in-vivo* work involving behavioural paradigms rely on the use of reward to train the animal to perform in the given task, for example, condensed milk is often used to train mice in spatial memory tasks (Lyon et al. 2011; Wilson et al. 2015) and sucrose pellets are used in operant conditioning tasks (Malkki et al. 2010; Sharma et al. 2012). If tail handling reduces mice's sensitivity for reward, this may result in longer training periods or reduced effect sizes leading to larger sample sizes and therefore tunnel handling offers the potential to address both refinement and reduction of the 3Rs. Tail handling may also be negatively affecting the neural circuitry underlying reward, which may mean that studies of reward pathways may not be using reliable or accurate models. I recommend that researchers consider the potential effect of tail handling on their results and interpretation of their findings.

Although the findings from this study provide the first evidence of an anhedonic-like state in tail handled mice, further work is needed in order to establish the full effects of tail handling on their experiences and ultimately how these shape their underlying affective state and welfare. In the present study, I was interested in assessing mice's sensitivity towards reward with differing hedonic properties. However, it is well known that an animal's perception of reward is not solely dependent on the reward itself, but can also be influenced by the animals prior

experiences (Flaherty & Largen 1975). For example, animals that have prior experience of a higher value reward will respond more negatively when exposed to a reward of lower value, compared to an animal that has only ever encountered a lower value reward (Flaherty 1996; Flaherty et al. 1998; Burman et al. 2008). A next step would be to ask how tail handling laboratory mice might affect their perception of reward relative to a reward of higher or lower value. Such a study could shed more insight as to the underlying mechanisms affected by tail handling. This would provide more information about how handling method might influence resilience to reward loss and/or experience pleasure from reward gain. It would therefore offer the potential to assess the animal's responses following a negative situation (i.e. shifting from a large to a small reward) and a positive situation (i.e. shifting from a small to a large reward) within the same behavioural paradigm.

To conclude, my study supports advice that, wherever possible, mice should be handled using a tunnel and not by their tails. Tunnel handling is a simple yet effective refinement that has the potential to not only significantly improve animal welfare but also scientific data quality. Based on these findings, and those of others (Hurst & West 2010; Gouveia et al. 2013; Ghosal et al. 2015), research institutions should seek to evaluate and introduce tunnel handling as a refinement to their husbandry procedures, and that published protocols for handling mice are revised (Deacon 2006).

Chapter 6: Handling method alters how laboratory mice respond to reward loss and reward gain

6.1 Abstract

As outlined in Chapter 5, the standard method of handling laboratory mice by the tail induces negative affect, which can be reduced by handling them using a tunnel. The aim of this study was to further understand the impacts of tail handling on the affective state of mice and elucidate how the depressive-like state is mediated using changes in cognitive processing. I investigated if the method used to handle laboratory mice (by the tail or using a tunnel) changed their sensitivity to both reward loss and reward gain. This is because humans with depression change how they process and evaluate information, which can be measured by their responses towards changes in reward. Typically, individuals in negative affective states are more sensitive to reward loss (i.e. they show greater disappointment), and less responsive to reward gain (i.e. they show less elation).

Licking microstructure was better able to measure responses following shifts in reward value compared to measuring changes in consumption, and was able to detect 'elation' and 'disappointment' in laboratory mice. Although handling method did not impact how mice responded to a gain in reward, tunnel handled mice were more resilient to a loss in reward, showing less disappointment, compared to mice handled via standard practice of using their tails. My findings highlight the potential of this methodology for measuring responses to positive and negative events in the same paradigm and provides insight into how both the positive and negative valence of affect is affected by handling method.

6.2 Introduction

Making improvements to animal welfare requires a better understanding of how animals feel (Duncan 1981; Dawkins 1990; Fraser & Duncan 1998; Duncan 2006; Mendl, Burman, et al. 2010). As outlined in the previous chapter, the current method used to handle laboratory mice raises welfare concerns. The standard practice of handling laboratory mice by their tails has been previously associated with increased stress and anxiety, which can be overcome by handling mice using tunnels

(Hurst & West 2010; Gouveia et al. 2013). My work (Chapter 5) supported these findings, and additionally showed that tail handled mice were more anhedonic, and in a more depressive-like state compared to mice handled using a tunnel (Clarkson et al. 2018). However, despite the increasing evidence that argues against the use of tail handling (Hurst & West 2010; Gouveia et al. 2013; Ghosal et al. 2015; Gouveia et al. 2017; Clarkson et al. 2018), tunnel handling has yet to be widely implemented across research institutions.

An important part of being able to determine how an animal might feel involves understanding their underlying affective state to determine the impact of laboratory conditions on its welfare (Dawkins 1990; Mendl & Paul 2004; Duncan 2006; Boissy et al. 2007). This is because the absence of negative affective states is often taken to reflect good welfare (Boissy et al. 2007). In humans, an individual's affective state can influence a number of cognitive functions (Hinde 1985; Forgas 2000; Paul et al. 2005). People diagnosed with anxiety and/or depression often show changes in how they process and retrieve information, as well as in how they make decisions (Beck et al. 1979; Mathews & Macleod 2002). For example, they focus more on negative outcomes, have a reduced memory for positive events, and make more negative judgements about ambiguous stimuli and/or events (Beck et al. 1979; Mineka et al. 1998; Mathews & Macleod 2002; Paul et al. 2005; Burman et al. 2008). These well-documented cognitive effects have led to the development of 'judgement bias' tests across a wide range of species including rats (*Rattus norvegicus*), starlings (*Sturnus vulgaris*), dogs (*Canis lupus familiaris*), sheep (*Ovis aries*) and even honeybees (*Apis mellifera*) (Harding et al. 2004; Bateson & Matheson 2007; Mendl, Brooks, et al. 2010; Brilot et al. 2010; Doyle et al. 2010; Bateson et al. 2011). Generally, these tests involve an initial training period where the animals learn to associate one stimulus with a positive outcome (i.e. a reward), and another stimulus with a negative outcome (i.e. no reward or a punishment). Following this, the animals are then offered one or more ambiguous stimuli (stimuli that lie between the positive and negative stimuli) and the animals' responses are assessed. These types of test have been useful in determining an animal's affective state to make inferences about how their environment influences their welfare (Harding et al. 2004; Bateson & Matheson 2007; Mendl, Brooks, et al. 2010; Brilot et al. 2010; Doyle et al. 2010; Bateson et al. 2011). For example, rats maintained under unpredictable housing, which is considered sufficient to induce negative affect and poorer welfare, showed more

‘pessimistic’ responses when presented with an ambiguous stimulus. They made more negative judgements about ambiguous stimuli, and perceived them as more similar to the negative stimulus predicting no reward, compared to rats kept under standard housing conditions (Harding et al. 2004). Conversely, in another study, improvements made to rats’ housing conditions through environmental enrichments associated with increasing an animal’s welfare, led to animals showing more ‘optimistic responses’. Rats made more positive judgements about an ambiguous stimulus, perceiving it as being more similar to the positive stimulus predicting reward, and suggesting they had a more positive affective state than rats kept under standard laboratory conditions (Brydges et al. 2011). Although these judgment bias tasks can provide novel insights into the affective state in animals (Harding et al. 2004; Bateson & Matheson 2007; Mendl, Brooks, et al. 2010; Brilot et al. 2010; Doyle et al. 2010; Bateson et al. 2011), they aren’t without their limitations, since they often require long training and testing periods and tests need to be species-specific (Brilot et al. 2010; Bateson et al. 2011; Brydges et al. 2011; Monk et al. 2018). Currently, although there is one developed for rats (Harding et al. 2004) which has recently been shortened (Brydges & Hall 2017), there is no validated cognitive bias test for laboratory mice.

One existing methodology that could offer an effective test, is measuring how they respond to changes in reward value, and particularly how they respond following reward loss and/or gain (Crespi 1942; Flaherty 1982; Flaherty 1996; Flaherty et al. 1998; Burman et al. 2008). This idea comes from findings in humans with negative affective states, who show stronger responses to reward loss (i.e. they show greater disappointment) and lower responses to reward gain (i.e. they show less elation) compared to those in more positive affective states (Beck 1969; Wenzlaff & Grozier 1988; Tucker & Luu 2007). Therefore, assessing an animal’s sensitivity to changes in reward value could provide more insight into how positive and negative experiences are differentially affected by their underlying affective state (Flaherty et al. 1998; Burman et al. 2008; Mitchell et al. 2012; Riemer et al. 2016; Neville et al. 2017).

How an animal responds to a given reward is highly dependent upon a number of factors. One important factor is the size of the reward: responses will be stronger for large than small rewards. A good example of this is the speed with which rats run down a runway to access a sucrose reward: they will run faster to gain access to a 32% sucrose reward than a 4% sucrose reward (Crespi 1942; Flaherty

1982). Prior experience with that reward is also an important factor (Flaherty et al. 1998). This is because when exposed to a reward of a given value, an association forms, leading to the expectation of that reward value in that specific context (e.g. a rat learns that it will receive 32% sucrose at the end of the runway). Therefore, when the value of the reward changes, the difference between the actual and expected reward value leads to changes in the way that animals will respond (Crespi 1942; Flaherty 1982; Flaherty & Rowan 1986; Burman et al. 2008). For example, the running speed of rats will change if the reward value unexpectedly changes (Crespi 1942). Rats shifted from receiving a high to receiving a low reward (reward loss), run even slower than rats that had only ever received the low reward, a response referred to as a Successive Negative Contrast (SNC). Conversely, animals shifted from receiving a low to receiving a high reward (reward gain), can run faster than animals that have only ever experienced the high reward, a response referred to as a Successive Positive Contrast (SPC). These contrast effects are widely documented within the experimental psychology literature, and have played a large role in understanding the underlying processes of learning and memory (Crespi 1942; Flaherty & Largen 1975; Flaherty & Rowan 1986; Mustaca et al. 2000; Dwyer et al. 2011). However, it has also been suggested that assessing these types of responses may offer a novel way to explore animal affect (Flaherty et al. 1998; Burman et al. 2008; Mitchell et al. 2012; Riemer et al. 2016; Neville et al. 2017).

This is because contrast effects appear to be mediated by an emotional response to the shift in reward value (Flaherty et al. 1998), and evidence shows that an animal's long term affective state influences the magnitude and/or duration of the contrast effect (Flaherty et al. 1998; Burman et al. 2008; Mitchell et al. 2012; Riemer et al. 2016; Neville et al. 2017). For example, larger and/or longer SNC effects, or smaller and/or shorter SPC effects are observed in individuals with negative affective states (Flaherty et al. 1998). Strong evidence that these effects are driven by emotional responses comes from experiments where rats have been given anti-anxiolytics, such as benzodiazepines, which reduce negative affect, reduce the size of the SNC effect (Becker & Flaherty 1983; Flaherty et al. 1986), and rats bred for high-anxiety which have increased negative affect, show larger and longer SNC responses (Rosas et al. 2007). Furthermore, reducing positive affect, for example, through withdrawal from amphetamine, also reduces the size of an SPC effect (Vacca & Phillips 2005). Therefore, since manipulations of an animal's affective state

can influence both the magnitude and duration of a contrast effect, these contrast paradigms may offer the potential to provide objective measures of affective state in rodents (Becker & Flaherty 1983; Flaherty et al. 1998; Rosas et al. 2007; Burman et al. 2008; Mitchell et al. 2012).

Contrast effects have recently been used to investigate affective states in a welfare context. The first study used a successive negative contrast paradigm, and found that rats kept in barren and unenriched cages (typically associated with poorer welfare and a negative affective state), had a prolonged negative contrast compared to rats housed in enriched caging (Burman et al. 2008). A later study, using European Starlings, assessed the effects that early life feeding schedules (simulating early life adversity) had on the development of a depressive-like phenotype in adulthood. Although this study did not use standard SNC and SPC paradigms, and instead used a within-subjects design, it did show that the amount of food and the amount of begging effort a bird had to do early in life differentially affected their sensitivity to reward loss and gain (Neville et al. 2017). These studies highlight the potential for using the magnitude and/or duration of contrast effects to provide information about an animal's underlying affective state and applying it to make inferences about their welfare.

To build on my findings in Chapter 5, I asked if the method used to handle mice could produce changes in their sensitivity to reward loss and/or gain. Assessing whether tail handling not only results in a lack of sensitivity to reward generally, but also results in a greater susceptibility to reward loss and/or blunted responses to reward gain, offers the potential to elucidate how the depressive-like state is mediated in tail handled mice. I predicted that if tail handling is sufficient to induce an anxious, depressive-like state, then tail handled mice will show a larger or more prolonged negative contrast effect indicative of greater disappointment, and a smaller or less sustained positive contrast effect indicative of reduced elation. I will determine the presence of these contrast effects by measuring changes in total consumption of sucrose solutions in line with previous work (Flaherty & Largen 1975; Mustaca et al. 2000). Assessing the consumption of a high (32%) or a low (4%) concentration of sucrose solution following reward loss or gain has been widely used in the experimental psychology literature to determine the presence of a SNC or SPC in laboratory rodents (Flaherty & Largen 1975; Mustaca et al. 2000). However, I will also assess whether changes in licking microstructure offers a more sensitive method

to determine the changes in affect following reward loss and reward gain. This is because, as previously discussed, licking microstructure is considered to better reflect the hedonic influences driving consumption than solely assessing total intake (Dwyer 2012). Therefore, this study not only aims to further explore how mice's affective state is influenced by handling method, but also assess whether measuring changes in licking microstructure offers a useful method for determining the affective state of laboratory mice in order to make inferences about their welfare. Furthermore, given that there is yet to be compelling evidence about the effects of handling method on physiological markers indicative of stress, I also aimed to explore whether tail handling is sufficient to induce physiological changes indicative of chronic stress through measuring differences in adrenal and thymus gland weight. I predicted that the stress-inducing effects of tail handling would cause alterations in the activity of the Hypothalamic-Pituitary-Adrenal (HPA) axis and the immune system resulting in differences in adrenal and thymus gland mass (Nemeroff et al. 1992; Rubin et al. 1995; Živković et al. 2005).

6.3 Methods and Materials

6.3.1 Ethical Statement

Experiments were conducted at Newcastle University following approval from the Animal Welfare and Ethical Review Body (AWERB Project ID: 540), and in accordance with the EU Directive (2010/63/EU), ASPA (1986) and the NIH Guidelines for the care and use of animals for experimental procedures (National Institutes of Health 2011). All animals were checked daily, and no adverse effects were reported.

6.3.2 Animals, housing and husbandry

Sixty-four male C57BL/6 mice (*Mus musculus*) were purchased from Charles River Laboratories, UK and were approximately 7 weeks of age upon arrival. They arrived and were tested in two separate batches (32 mice in each batch; Batch 1 arrival date 09/05/2017; Batch 2 arrival date 04/07/2017). Mice were free from all recognised pathogens, and the health status of the colony was monitored following

the FELASA health monitoring recommendations (Guillen 2012). Mice were pair housed in M2 cages (33cm (L) x 15cm (W) x 13cm (H), North Kent Plastics), with sawdust bedding, nesting material (4HK Aspen chips, NestPak and Sizzlepet nesting, Datesand Ltd, Manchester) and a clear perspex home cage tunnel (50mm diameter, 150mm length). All cages were subject to one full clean per week. All animals had access to food (Special Diet Services, RM3E diet) and water *ad libitum*, except prior to training and testing for drinking experiments (described below). Animals were maintained on a reverse 12:12 hour light/dark cycle (lights off: 10:00 until 22:00). Therefore, all experimental procedures described below were conducted under red light illumination. Mice were housed under standard laboratory conditions at an optimal temperature of 21±4°C and a relative humidity 55±10% (Home Office 2014). Three days prior to the start of the study, mice were marked for identification on either the shoulder or rump using hair dye (Jerome Russel B Blonde, UK). This was conducted in line with previous studies, and does not interfere with their responses to handling (Hurst & West 2010; Gouveia et al. 2013; Clarkson et al. 2018).

6.3.3 Handling method manipulation

After one week of acclimatisation to the laboratory, each cage of two mice was randomly assigned to one of two treatment groups: tail or tunnel handled. From that point, mice were only handled by their designated method, following the methods of previous studies (Hurst & West 2010; Gouveia et al. 2013). For the tail handling manipulation, the base of their tail was grasped between thumb and forefinger, and the mouse was lifted onto the sleeve of the laboratory coat and held there for 30 seconds before the mouse was returned to its home cage. For the tunnel handling manipulation, the mouse was guided into a perspex tunnel, which was lifted above the cage and held for 30 seconds. On the first two days of tunnel handling, the handler's hands were loosely cupped over the ends of the tunnel to prevent escape before the mice became accustomed to this method of handling.

Mice were handled twice daily for 30 seconds, 60 seconds apart, for the first nine days, and also prior to the interaction tests (described below) on days 19 and 27 (Figure 6.1). Prior to handling, the nest material and home cage tunnel were removed for 60 seconds, and care was taken not to disrupt the nest structure. For routine

husbandry practices, such as cage cleaning, mice were captured and transferred using their designated handling method (either on the sleeve for tail handled mice, or in the tunnel for tunnel handled mice). The same handling method was also used when transferring mice for behavioural testing, i.e. when placed in the elevated plus maze, open field arena and sucrose drinking chambers.

6.3.4 Voluntary interaction tests

On designated days (Figure 6.1) during the handling manipulation (days 1, 5 and 9) and the sucrose drinking tests (days 19 and 27), each cage of animals underwent 'voluntary interaction tests' to assess their responses to the presence of a handler (Hurst & West 2010; Gouveia et al. 2013). Each test consisted of first removing the cage lid, nesting material and home cage tunnel, and then the handler stood motionless in front of the cage for 60 seconds. Next, a gloved hand (for tail handled mice) or a gloved hand holding the home cage tunnel (for tunnel handled mice) was then held resting on the substrate at the front right hand side of the cage for 60 seconds to assess voluntary interaction. The mice were then handled by the designated handling method described above, before the mice were returned to the cage and the hand replaced for 60 seconds to assess voluntary interaction after handling.

These tests were filmed from above (Cube HD 1080, Y-cam) and later analysed using Observer XT (v11, Noldus, Virginia, USA). Time spent interacting with the handler was measured for each mouse within a cage, and the two values were used to obtain an overall mean cage score, and this was then calculated as a percentage of the total test time. These were summed together for analyses for the two treatment groups (tail and tunnel handled). Therefore, for these tests, the experimental unit was 'cage' (n=16 for both groups). Interaction was defined as any of the following behaviours; sniffing (nose within 0.5cm), touching (including paw contact), climbing on or in the handling tunnel and/or the handler's hand. Due to differences in how mice in the two treatments were handled during the interaction tests, the observer could not be blind to the treatment, but was blind to whether the interaction test was carried out before (pre) or after (post) handling (Hurst & West 2010; Gouveia et al. 2013; Clarkson et al. 2018).

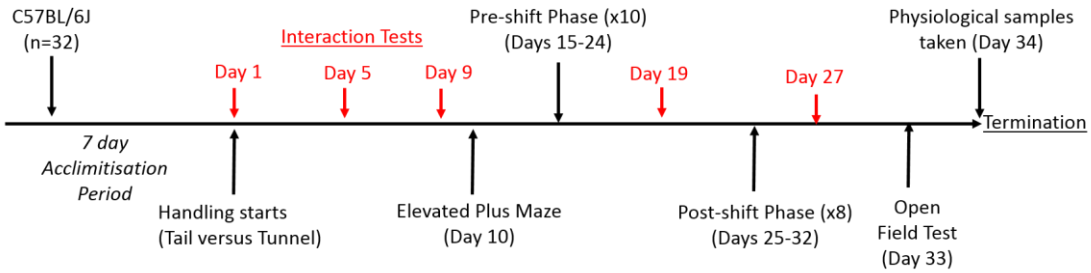


Figure 6.1: Schematic illustrating the study timeline, and showing when each of the behavioural tests were conducted.

6.3.5 Elevated Plus Maze

On day 10, mice underwent behavioural testing in an elevated plus maze (Figure 6.1), with arms measuring 30cm (L) x 5cm (W) and side walls of 15cm on the two closed arms, and elevated 50cm from the ground. Mice were delivered to the centre of the maze facing an open arm, and allowed to explore for a total duration of 5 minutes. Each mouse was then returned to either a holding cage or its home cage, depending on whether it was the first or last mouse to undergo testing from its cage. The maze was cleaned between subjects with 70% ethanol and dried with a paper towel. The order in which tail and tunnel handled mice were tested was counterbalanced across the testing day. This was filmed from above (Cube HD 1080, Y-cam) and later analysed using Observer XT (v11, Noldus, Virginia, USA). Six animals jumped off the maze before the end of the test, and one animal's data was lost due to a technical fault. Therefore, sample sizes were reduced for statistical analyses (tail handled, n=28; tunnel handled, n=29). The number of open arm entries (defined when all four paws were in the arm), time spent on the open arms, and the number of protected stretch attend postures were recorded.

6.3.6 Sucrose Drinking Tests

Mice were trained and tested in eight custom-built drinking chambers. These consisted of standard mice IVC home cages (34 (L) x 19 (W) x 14(D) cm) with clear perspex sides, a metal perforated floor and wire cage lid with two modified attachments to connect the sipper tubes. Solutions were made available through

drinking spouts attached to 50ml falcon tubes, presented on the left hand side of the cage. The drinking chambers were connected to contact sensitive Med Associates dual contact lickometers (Med Associates Inc., St. Albans, Vermont), which transmit the time of each lick to the nearest 0.01 second to a computer using MED-PC software. Custom-written software (courtesy of Prof Dominic Dwyer) calculated the lick cluster sizes according to a range of interbout intervals (IBI), which is the length of time used to determine when licks can be considered to be in a single bout (Davis & Smith 1992; Davis & Perez 1993; Dwyer 2012; Dwyer et al. 2013). The data presented here use interbout intervals of 250ms, 500ms or 1000ms in order to ensure data were robust across criteria. This means that, for example, when an interbout interval of 250ms was applied, any duration of 250ms or longer between two licks defined the end of one bout and the start of the next.

Mice were randomly sub-divided into groups for testing in the drinking apparatus referred to as 'testing cohorts', where eight mice were tested simultaneously. Each testing cohort included two mice from each treatment group. All cohorts were run in the same order and at the same time each day. Water bottles on the home cage were removed 2h prior to sucrose drinking trials, and before the lights went off to motivate the mice to consume the sucrose solutions during these tests (Millard et al. 1983).

Sucrose drinking tests were split into two distinct phases; the pre-shift phase and the post-shift phase (Figure 6.1). The pre-shift phase consisted of ten consecutive days where mice were transferred to the drinking chamber and had access to sucrose for 15 minutes. The number of pre-shift sessions followed previous literature looking at contrast effects in mice, to ensure that the mice had sufficient experience with the sucrose solution prior to the shift in the post-shift phase (Mustaca et al. 2000). For the first three pre-shift sessions the spout was left to protrude into the cage to ensure engagement with the task, but for the remaining sessions, the spout was positioned in line with the cage in order to reduce accidental contact. Depending on the treatment group, mice had access to either 4% or 32% (w/w) sucrose (Table 6.1). The post-shift phase lasted for eight consecutive days due to extinction of contrast effects being relatively short in the literature (Mustaca et al. 2000; Burman et al. 2008). Again the mice had access to sucrose solutions for 15 minutes and the concentration depended on the treatment group (Table 6.1) where mice were given access to either a low (4% w/w sucrose) or high (32% w/w sucrose)

reward. For the post shift phase, half the groups (allocated using a random number generator) were shifted from one concentration to the other. These groups formed either the loss or the Successive Negative Contrast (SNC) condition, where animals were shifted from high-to-low reward and compared to a unshifted control group that remained on the low reward, or the gain or Successive Positive Contrast (SPC) condition, where animals were shifted from low-to-high reward, and compared to a matched unshifted control group that remained on the high concentration throughout (Table 6.1). The responses of mice undergoing a SNC or SPC needed to be compared with their respective controls, i.e. mice that were unshifted, to identify any effect of contrast condition.

Table 6.1: The treatment groups and respective sample sizes for Successive Negative Contrast (SNC) and Successive Positive Contrast (SPC) conditions.

Treatment group	**Pre-shift Phase sucrose concentration**	**Post-shift Phase sucrose concentration**	**Handling Method**	**Sample size**
Successive Negative Contrast (SNC)	32%	4%	Tail Tunnel	8 8
Unshifted (SNC) Control	4%	4%	Tail Tunnel	8 8
Successive Positive Contrast (SPC)	4%	32%	Tail Tunnel	8 8
Unshifted (SPC) Control	32%	32%	Tail Tunnel	8 8

6.3.7 Open Field Test

Following completion of the sucrose drinking tests, I conducted an open field test on Day 33. Each mouse (n=32 for each handling method) was individually placed by their designated handling method in the centre of a rectangular arena (54.5cm (L) x 35.5cm (W) x 17cm (H)) made of white plastic with a transparent perspex lid and allowed to freely explore for 10 minutes (see Chapter 2 for full details). The order in which tail and tunnel handled mice were tested was counterbalanced across the testing day. Behaviour was filmed from above (Cube HD 1080, Y-cam) and later analysed using Ethovision XT (v 5.1, Noldus, Virginia, USA). This calculated the total duration spent in the centre, crosses to the centre, total distance travelled, the total time spent moving and the velocity of movement for each mouse.

6.3.8 Physiological measures

At the end of the study, animals were humanely killed via cervical dislocation in line with Schedule 1 guidance, and the adrenal and the thymus glands were excised and weighed from the unshifted control animals across both batches of mice (Tail handled n=16; Tunnel handled n=16). This was done in order to compare the masses of these glands and explore any stress-related physiological differences between tail and tunnel handled mice.

6.3.9 Statistical analyses

All statistical analyses were conducted using SPSS (v23, SPSS Inc, Chicago, USA). All datasets were tested for normality using the Shapiro-Wilk test and homogeneity of variance using and the Levene statistic test; where assumptions were violated, non-parametric analyses were used. Where significant main effects were found, Bonferroni post hoc tests were performed to investigate pairwise comparisons (Table 6.2). In all statistical tests, differences were considered significant using a p value <0.05.

Table 6.2: Statistical tests conducted for each data set with respective factors, experimental unit and sample size.

Data	Dependent Variable	Statistical Test	Factor(s)	Unit	Sample Size
Voluntary interaction tests	Percentage time spent interacting	Repeated measures ANOVA	Between subject factor: Handling method (2 levels) Within-subject factors: day (5 levels), time (2 levels: pre or post handling)	Cage	n=16 tail handled n=16 tunnel handled
Elevated plus maze	Number of open arm entries; Duration on open arms; Percentage of time on open arms, Protected Stretch Attend postures	Independent t-test	Handling method (2 levels)	Mouse	n=28 tail handled n=29 tunnel handled
Sucrose drinking	Consumption and Lick cluster size (at the pre-shift phase)	Repeated measures ANOVA	Between subject factor: Handling method (2 levels) Within-subject factor: Sucrose concentration (2 levels)	Mouse	n=32 4% sucrose n=32 32% sucrose
	Lick Cluster Size (at the post-shift phase for 250ms; 500ms; 1000ms)	Repeated measures ANOVA	Between subject factor: Post-shift Phase (2 levels) Within-subject factor: Contrast Condition (2 levels)	Mouse	Conducted separately for tail and tunnel handled mice: n=16 (SPC) n=16 (SPC Controls) n=16 (SNC) n=16 (SNC Controls)
Open field test	Duration of movement; Duration in centre; Crosses to centre; Distance travelled; Mean velocity	Independent t tests	Handling method (2 levels)	Mouse	n=32 tail handled n=32 tunnel handled
Physiology	Adrenal weight Thymus weight	Independent t-tests	Handling method (2 levels)	Mouse	n=16 tail handled n=16 tunnel handled

6.4 Results

6.4.1 Voluntary interaction tests

In order to establish that the differential handling produced the clear behavioural differences between tail and tunnel handled mice previously reported (Hurst & West 2010; Gouveia et al. 2013; Clarkson et al. 2018), I conducted voluntary interaction tests in line with the handling manipulation (days 1, 5 and 9) and sucrose drinking tests (days 19 and 27).

During the handling manipulation, I found that tunnel handled mice spent more time interacting with the handler than mice handled via the tail ($F_{1,30}$=625.1, p<0.001; Figure 6.2; Table 6.3). However, the timings of these tests were also important: there was a significant main effect of day ($F_{2,60}$=7.84, p=0.001; Table 6.3), a handling method by day interaction ($F_{2,60}$=4.21, p=0.020; Table 6.3), and a three-way interaction of handling method, time and day ($F_{1,60}$=3.45, p=0.038; Table 6.3). Pairwise comparisons between days (Bonferroni adjusted for multiple comparisons) showed that tunnel handled mice interacted significantly more with the handler on days 5 and 9 compared to day 1, both during the pre-handling and post-handling interaction tests (all p values <0.05). This increase in interaction after day 1 was not seen for tail handled mice (all p values >0.05). However, tail handled mice did increase the time they spent interacting with the handler but this was only evident on day 9 after they had been handled (post-handling) compared to before they had been handled on day 9 (pre-handling, p=0.03) which was not seen on either day 1 or day 5 (all p values >0.05).

Table 6.3: Statistical results for the repeated measures ANOVA conducted for the percentage of time spent voluntarily interacting with the handler in line with the handling manipulation on days 1, 5 and 9. *p<0.05 and ***p<0.001.

Factor	**F_{df}**	**p value**
Handling method	$F_{1,30}$ = 625.1	<0.001 ***
Day	$F_{2,60}$ = 7.84	0.001***
Time (pre or post handling)	$F_{1,30}$ = 2.07	0.161 ns
Handling method x Day	$F_{2,60}$ = 4.21	0.020*
Handling method x Time	$F_{1,30}$ = 0.02	0.885 ns
Day x Time	$F_{2,60}$ = 0.06	0.944 ns
Handling method x Day x Time	$F_{2,60}$ = 3.45	0.038 *

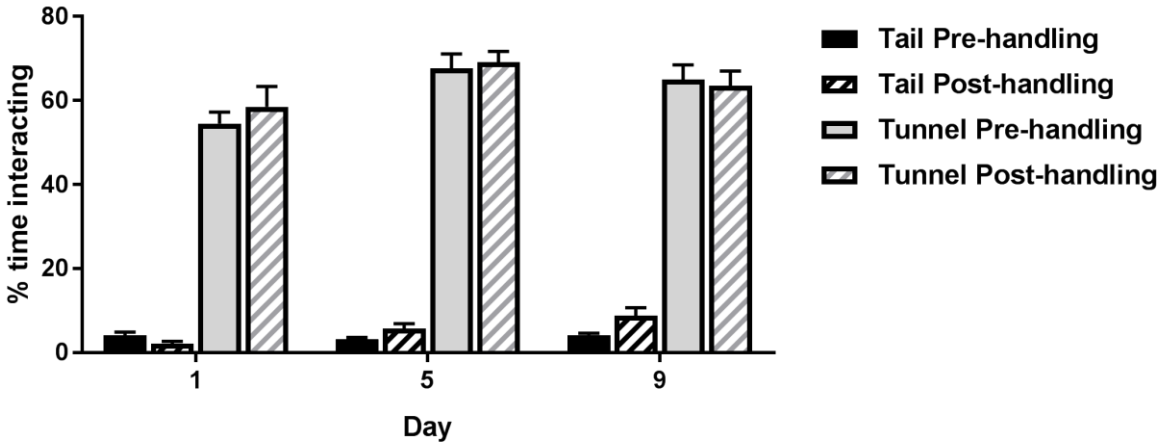


Figure 6.2: Mean (+SEM) percentage of time spent interacting with the handler in the voluntary interaction tests conducted on three different days during the handling manipulation. Voluntary interaction tests were conducted both before (pre-handling) and after (post-handling) the animals were handled via the tail or tunnel handling method on each day.

When comparing the time spent interacting with the handler in the voluntary interaction tests conducted in line with the sucrose drinking tests (days 19 and 27), clear differences between tail and tunnel handled mice were still evident. Tunnel handled mice still interacted for longer with the handler than mice handled via their tails ($F_{1,30}$=244.65, p<0.001; Figure 6.3; Table 6.4), although there was also a significant interaction between handling method and time ($F_{1,30}$=18.59, p<0.001). Pairwise comparisons (Bonferroni adjusted for multiple comparisons) revealed that tail handled mice spent more time interacting with the handler after being handled compared to before they were handled (p=0.013), whereas the tunnel handled mice spent less time interacting with the handler after being handled compared to before handling (p=0.002). There were no other significant main effects or interactions (Table 6.4).

Table 6.4: Statistical results for the repeated measures ANOVA conducted for the percentage of time spent voluntarily interacting with the handler on days 19 and 27 in line with the sucrose drinking tests. Where ***p<0.001.

Factor	F_{df}	p value
Handling method	$F_{1,30}$ = 244.65	0.001 ***
Day	$F_{1,30}$ = 0.03	0.597 ns
Time (pre or post handling)	$F_{1,30}$ = 0.36	0.555 ns
Handling method x Day	$F_{1,30}$ = 0.15	0.705 ns
Handling method x Time	$F_{1,30}$ = 18.59	0.001 ***
Day x Time	$F_{1,30}$ = 0.86	0.362 ns
Handling method x Day x Time	$F_{1,30}$ = 0.03	0.861 ns

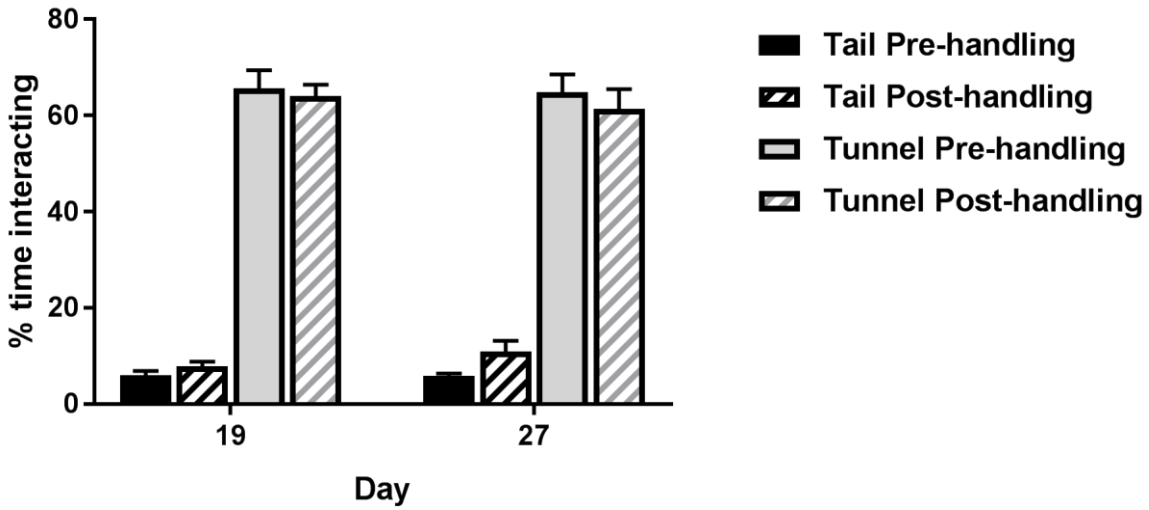


Figure 6.3: Mean (+SEM) percentage of time spent interacting with the handler in the voluntary interaction tests conducted on two different days in line with the sucrose drinking tests. Interaction tests were conducted both before (pre-handling) and after (post-handling) the animals were handled via the tail or tunnel handling method.

6.4.2 Elevated Plus Maze

On day 10, mice underwent testing on the elevated plus maze to assess anxiety-like behaviour. In line with previous findings (Hurst & West 2010; Gouveia et al. 2013; Clarkson et al. 2018), tail handled mice entered the open arms of the elevated plus maze less often (t_{61}=3.36, p=0.001; Figure 6.4A), and spent less time on the open arms (t_{61}=3.95, p<0.001; Figure 6.4B). Furthermore, tail handled mice performed more protected stretch attend postures onto these open arms (t_{61}=3.30, p=0.002; Figure 6.4C), which is thought to reflect an increase in risk-assessment

behaviour (Pellow et al. 1985; Rodgers & Dalvi 1997). Taken together, the results confirm that tail handled mice were more anxious than tunnel handled mice in the elevated plus maze.

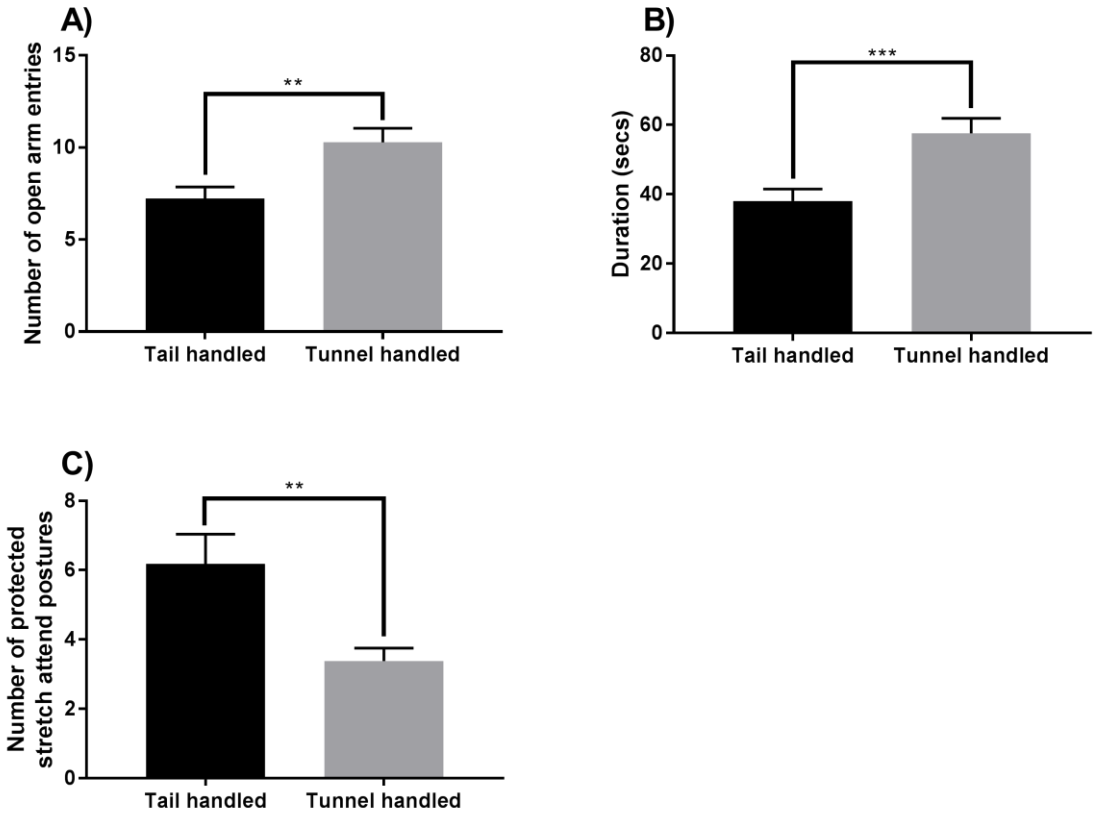


Figure 6.4: Results from tail handled and tunnel handled mice placed in the elevated plus maze on day 10. A. Mean (+SEM) number of open arm entries. B. Mean (+SEM) duration spent (s) in the open arms. C. Mean (+SEM) number of protected stretch attend (PSA) postures. **p<0.01 and ***p<0.001.

6.4.3 Sucrose drinking tests

6.4.4 Pre-shift: Consumption Data

In order to determine whether the expected difference in hedonic rating towards the low and high concentrations of sucrose reward were present, I included data from all mice for the pre-shift phase. In contrast to expectations that mice should drink more of the more concentrated sucrose solution, they in fact drank significantly less of the high concentration solution compared to the low concentration solution ($F_{1,60}$=4.24, p=0.044; Figure 6.5). There was no difference in sucrose consumption between mice that were tunnel handled or tail handed ($F_{1,60}$=1.62, p=0.208), and no

significant interaction between handling and sucrose concentration ($F_{1,60}$=0.16, p=0.696). The fact that the mice did not drink more of the higher reward means that clear interpretation of a negative or positive contrast effect was impossible. This is because in order to test a contrast effect, the higher reward should be valued more, i.e. the mice should drink more of the high than the low reward solution. Since this did not occur, I was unable to look for contrast effects using the consumption data.

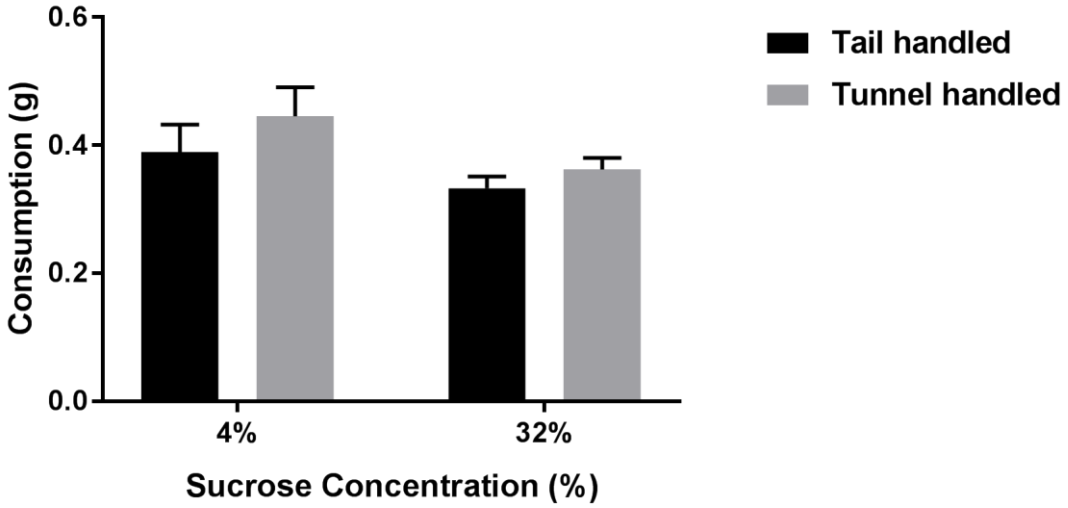


Figure 6.5: Mean (+SEM) consumption (g) for both low (4%) and high (32%) sucrose for tail and tunnel handled mice during the pre-shift period.

6.4.5 Pre-shift: Lick Cluster Size

The results differed slightly according to the interbout interval criterion used, and so I report the findings for all three criteria. Overall, mice tended to have larger lick cluster sizes to the high concentration sucrose solution compared to the low concentration sucrose solution (Table 6.5; Figure 6.6). Furthermore, tunnel handled mice tended to have larger lick cluster sizes compared to tail handled mice (Table 6.5; Figure 6.6). However, these factors did not interact (Table 6.5; Figure 6.6).

Table 6.5: Statistical results for the full results of the repeated measures ANOVA conducted for lick cluster sizes according to the three interbout intervals (IBIs) for the pre-shift phase. *p<0.05 and **p<0.01.

Interbout Interval	**Factor**	**F_{df}**	**p value**
250ms	Handling method	$F_{1,60} = 3.13$	p=0.082 ns
	Concentration	$F_{1,60} = 4.27$	p=0.043*
	Handling method x Concentration	$F_{1,60} = 0.83$	p=0.366 ns
500ms	Handling method	$F_{1,60} = 6.82$	p=0.011*
	Concentration	$F_{1,60} = 4.10$	p=0.047*
	Handling method x Concentration	$F_{1,60} = 4.27$	p=0.243 ns
1000ms	Handling method	$F_{1,60} = 9.84$	p=0.003**
	Concentration	$F_{1,60} = 2.06$	p=0.156 ns
	Handling method x Concentration	$F_{1,60} = 2.19$	p=0.144 ns

Taken together, these analyses reveal that unlike sucrose consumption, lick cluster size accurately reflects the difference in hedonic value between the low and high sucrose solutions, and can be used to investigate a contrast effect following a reward loss or gain. Whilst a difference in lick cluster size was not unexpected (see Chapter 5), this means that the subsequent analyses in the post-shift phase were conducted separately for tail and tunnel handled mice.

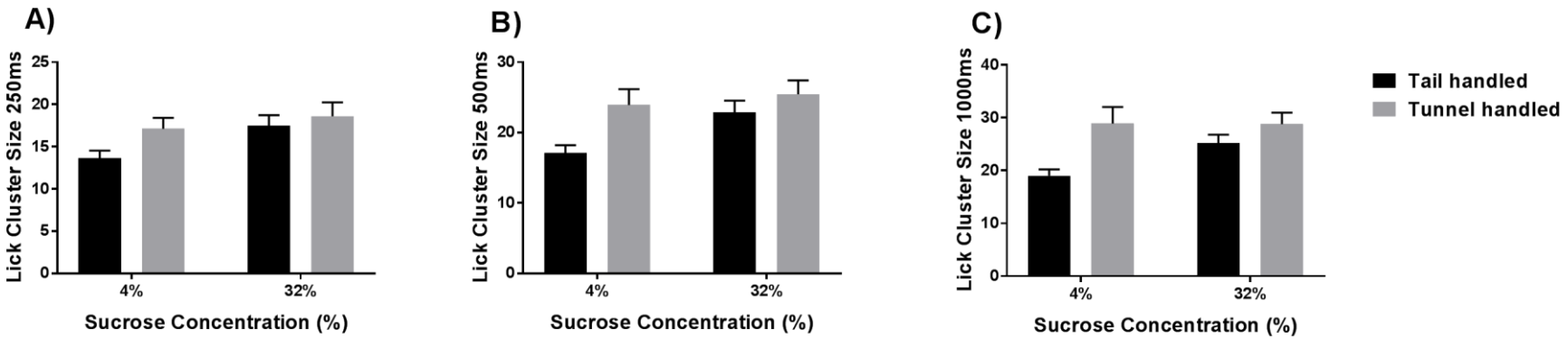
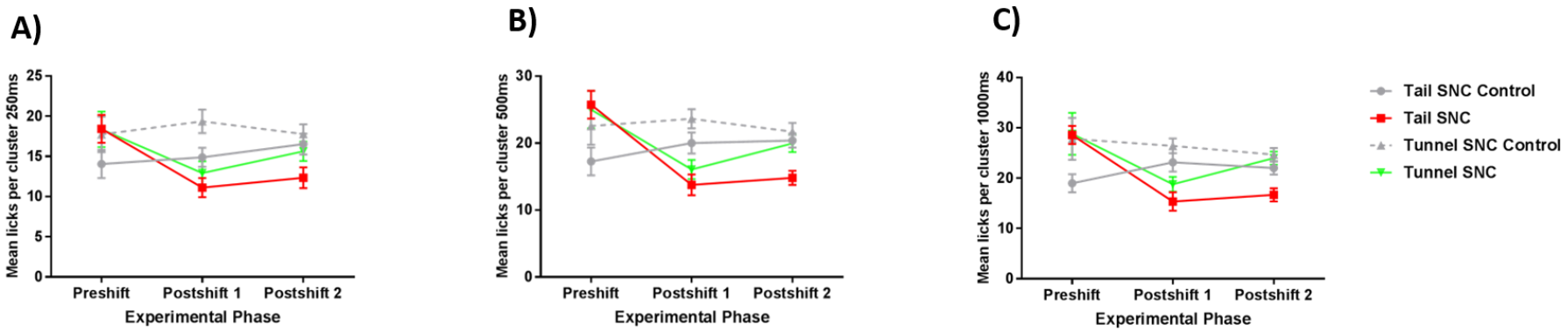


Figure 6.6: The mean (+SEM) lick cluster sizes for the pre-shift phase at the low (4%) and high (32%) sucrose, for tail and tunnel handled mice using three different interbout (IBI) criteria A. Mean (+SEM) lick cluster size using IBI of 250ms B. Mean (+SEM) lick cluster size using IBI of 500ms C. Mean (+SEM) lick cluster size using IBI of 1000ms.

6.4.6 Post-shift: Successive Negative Contrast

Because the tail and tunnel handled mice valued the sucrose differently in the pre-shift phase, I decided to look at the contrast effects separately by handling method. Since the results had slight qualitative differences according to which interbout interval (IBI) criteria (250ms, 500ms and 1000ms) was used to classify the start of one bout and the end of another, I present the data separately for all three criteria.

In the SNC condition, I found that handling method affected the way that mice responded to reward loss. Both tail and tunnel handled mice that were shifted from a high to a low reward demonstrated a negative contrast effect. Mice shifted from high to low sucrose had smaller lick cluster sizes compared to mice that had remained on the lower 4% sucrose solution throughout; there was a main effect of contrast condition at all three IBIs for tail and tunnel handled mice (Figure 6.7A-F). There was no main effect of the post-shift phase (i.e. post-shift 1 or 2) for either tail or tunnel handled mice for any IBI (Figure 6.7D-F). However, there was a significant interaction between contrast and post-shift phase in the tunnel handled mice (Figure 6.7A-F), which was absent in the tail handled mice (Figure 6.7A-F). Pairwise comparisons (adjusted for multiple comparisons using Bonferroni correction) revealed that tunnel handled mice only had lower lick cluster sizes than their controls at the first post-shift phase (all $p<0.05$) but not at the second post-shift phase (all $p>0.05$). Taken together, this shows that the tail handled mice undergoing a successive negative contrast showed a longer lasting negative contrast effect than mice handled via the tunnel refinement.



D)

Handling method	Factor	F_{df}	P
Tail	Contrast	$F_{1,14}$=6.01	0.028*
	Post-shift phase	$F_{1,14}$=4.32	0.057
	Contrast x Post-shift phase	$F_{1,14}$=0.09	0.766
Tunnel	Contrast	$F_{1,14}$=7.78	0.014*
	Post-shift phase	$F_{1,14}$=0.25	0.622
	Contrast x Post-shift phase	$F_{1,14}$=3.35	0.080

E)

Handling method	Factor	F_{df}	P
Tail	Contrast	$F_{1,14}$=11.67	0.004**
	Post-shift phase	$F_{1,14}$=0.84	0.375
	Contrast x Post-shift phase	$F_{1,14}$=0.20	0.664
Tunnel	Contrast	$F_{1,14}$=10.35	0.006**
	Post-shift phase	$F_{1,14}$=0.61	0.447
	Contrast x Post-shift phase	$F_{1,14}$=5.28	0.037*

F)

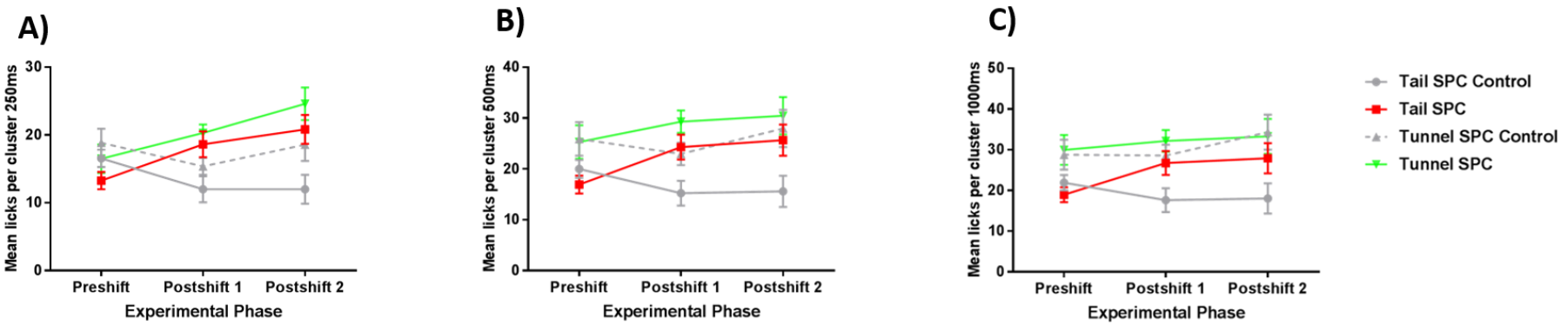
Handling method	Factor	F_{df}	P
Tail	Contrast	$F_{1,14}$=10.95	0.005**
	Post-shift phase	$F_{1,14}$=0.01	0.932
	Contrast x Post-shift phase	$F_{1,14}$=1.31	0.271
Tunnel	Contrast	$F_{1,14}$=9.30	0.009*
	Post-shift phase	$F_{1,14}$=1.49	0.243
	Contrast x Post-shift phase	$F_{1,14}$=5.99	0.028*

Figure 6.7: The mean (±SEM) lick cluster sizes for the Successive Negative Contrast (SNC) during the post-shift phase, for tail and tunnel handled mice using three different interbout (IBI) criteria A. Mean (±SEM) lick cluster size using IBI of 250ms B. Mean (±SEM) lick cluster size using IBI of 500ms C. Mean (±SEM) lick cluster size using IBI of 1000ms. Statistical results for the full results of the repeated measures ANOVA for lick cluster sizes according to the three IBIs D. IBI of 250ms E. IBI of 500ms F. IBI of 1000ms. *p<0.05 and **p<0.01.

6.4.7 Post-shift: Successive Positive Contrast

Once again the data were analysed separately for tail and tunnel handled mice that underwent the SPC condition. I found that tail handled mice had significantly larger lick cluster sizes than their respective controls following a gain in reward value, with significant or near-significant effects of contrast condition at all IBIs (Figure 6.8A-F). This appeared to be driven not only by an increase in lick cluster size in the reward gain condition, but also a decrease lick cluster size in the control mice (Figure 6.8A-F). There was no effect of the post-shift phase, and no interactions at any IBI for tail handled mice (Figure 6.8A-F), showing that this difference was unchanged across the whole post-shift phase.

In contrast, there was only a significant effect of contrast condition and post-shift phase at the 250ms IBI criterion, and not at the other two IBI criteria for tunnel handled mice (Figure 6.8D-F). There were also no significant interactions at any IBI for tunnel handled mice. Therefore, whilst there was an indication that tunnel handled mice in the SPC condition had larger lick cluster sizes than their respective controls, this was not consistent across all IBI criteria. This could be due to ceiling effects, since the tunnel handled mice had higher lick cluster sizes than the tail handled mice, which makes any difference between the low 4% and high 32% sucrose more difficult to detect.



D)

Handling method	Factor	F_{df}	P
Tail	Contrast	$F_{1,14}$=7.66	0.015*
	Post-shift phase	$F_{1,14}$=2.57	0.131
	Contrast x Post-shift phase	$F_{1,14}$=2.56	0.132
Tunnel	Contrast	$F_{1,14}$=6.40	0.024*
	Post-shift phase	$F_{1,14}$=5.31	0.037*
	Contrast x Post-shift phase	$F_{1,14}$=0.113	0.741

E)

Handling method	Factor	F_{df}	P
Tail	Contrast	$F_{1,14}$=6.71	0.021*
	Post-shift phase	$F_{1,14}$=0.42	0.527
	Contrast x Post-shift phase	$F_{1,14}$=0.14	0.713
Tunnel	Contrast	$F_{1,14}$=2.62	0.128
	Post-shift phase	$F_{1,14}$=1.33	0.268
	Contrast x Post-shift phase	$F_{1,14}$=0.03	0.874

F)

Handling method	Factor	F_{df}	P
Tail	Contrast	$F_{1,14}$=4.52	0.052
	Post-shift phase	$F_{1,14}$=0.25	0.625
	Contrast x Post-shift phase	$F_{1,14}$=0.06	0.815
Tunnel	Contrast	$F_{1,14}$=0.08	0.787
	Post-shift phase	$F_{1,14}$=2.70	0.123
	Contrast x Post-shift phase	$F_{1,14}$=1.23	0.286

Figure 6.8: The mean (±SEM) lick cluster sizes for the Successive Positive Contrast (SPC) during the post-shift phase, for tail and tunnel handled mice using three different interbout (IBI) criteria A. Mean (±SEM) lick cluster size using IBI of 250ms B. Mean (±SEM) lick cluster size using IBI of 500ms C. Mean (±SEM) lick cluster size using IBI of 1000ms. Statistical results for the full results of the repeated measures ANOVA for lick cluster sizes according to the three IBIs. D. IBI of 250ms E. IBI of 500ms F. IBI of 1000ms. *p<0.05.

6.4.8 Open Field Test

On day 33, individual mice were placed in an open field arena for 10 minutes and their behaviour was filmed from above. In line with my previous findings (Clarkson et al. 2018), tail handled mice spent significantly less time in centre of the open field compared to tunnel handled mice (Duration in centre: t_{62}=2.94, p=0.005; Figure 6.9A; Frequency in centre: t_{62}=2.36, p=0.021; Figure 6.9B). Whilst mice from both groups spent similar amounts of time moving (Movement: t_{62}=0.11, p=0.910; Figure 6.9C), tunnel handled mice tended to move further and faster, although this was not significant (Distance travelled: t_{62}=1.86, p=0.068, Figure 6.9D; Velocity: t_{62}=1.76, p=0.084, Figure 6.9E).

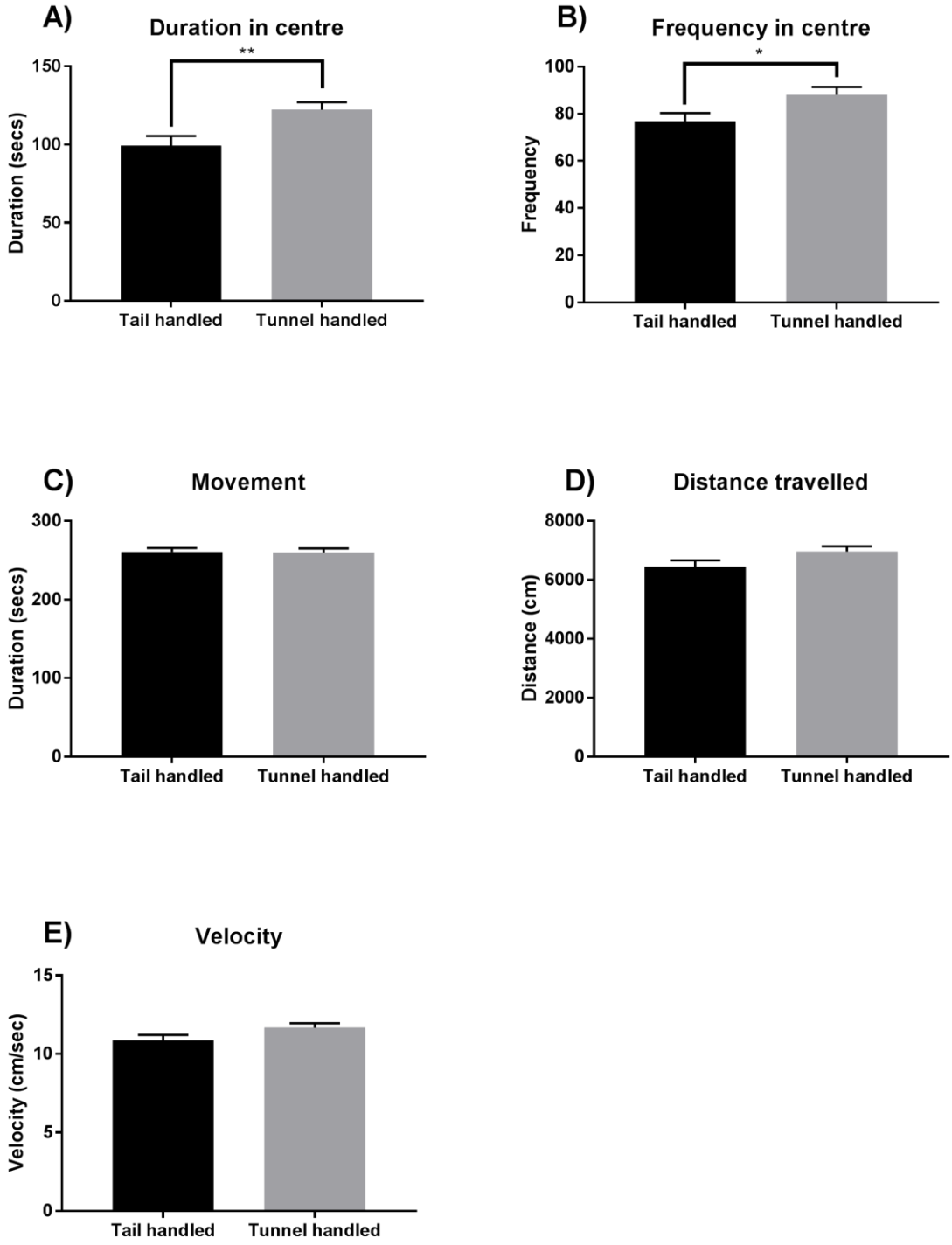


Figure 6.9: The behaviour of tail and tunnel handled mice in the open field test. A. Mean (+SEM) length of time spent in the centre (s). B. Mean (+SEM) number of crosses to the centre. C. Mean (+SEM) length of time spent moving (s). D. Mean (+SEM) distance travelled (cm). E. Mean (+SEM) velocity when moving (cm/sec). *p<0.05 and **p<0.01.

6.4.9 Physiological measures

At the end of the study, the control (unshifted) animals were euthanised and their adrenal and thymus glands excised and weighed. Although there were no significant differences in body weight (tail handled: Mean ± SEM=25.81±0.42; tunnel handled: Mean ± SEM=26.48±0.38; t_{30}=1.19, p=0.243), I nonetheless controlled for individual variation in body weight when comparing the adrenal and thymus gland weights. The handling method had a significant effect on the weight of adrenal glands: tail handled mice had significantly larger adrenal glands compared to tunnel handled mice (t_{30}=2.65, p=0.013; Figure 6.10A). Thymus mass was similar between the two groups (t_{30}=1.53, p=0.136; Figure 6.10B).

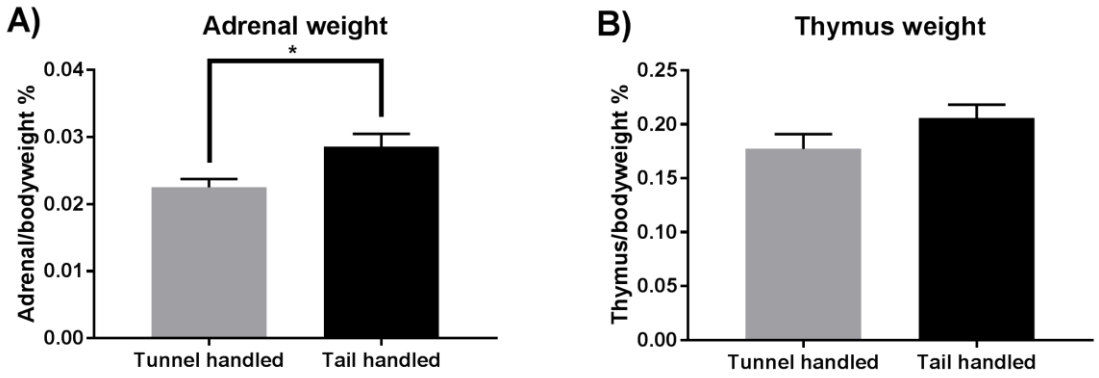


Figure 6.10: The physiological measures obtained for tail and tunnel handled mice from the control (unshifted) conditions, calculated as a percentage of their final body weight (%) A. Mean (+SEM) Adrenal weight relative to body weight (%) B. Mean (+SEM) thymus weight relative to body weight (%). *p<0.05.

6.5 Discussion

This study shows that the method used to handle laboratory mice can alter how they respond to changes in reward value. Following a decrease in reward value, from 32% to 4% sucrose, tail handled mice showed a more sustained Successive Negative Contrast (SNC) effect compared to tunnel handled mice. Whilst all mice demonstrated a SNC, irrespective of how they were handled, only the tail handled

mice showed a sustained SNC effect in measures of their licking microstructure. Shifted mice handled via their tails continued to have lower hedonic responses towards the 4% sucrose relative to their controls even in the second post-shift phase, whilst the shifted tunnel handled mice did not. This finding builds on previous work (Hurst & West 2010; Gouveia et al. 2013; Clarkson et al. 2018) to provide the first evidence that tail handling laboratory mice is sufficient to induce cognitive changes that are indicative of negative affect. This is because the SNC is considered to result from an emotional response, such as frustration or disappointment, when the actual reward doesn't match the expected reward (Crespi 1942; Flaherty & Largen 1975; Flaherty et al. 1998; Burman et al. 2008), which is worse when in a negative affective state (Flaherty et al. 1998). This has been highlighted in previous work whereby a greater or more prolonged SNC is often seen in animals considered to be in a negative affective state and indicative of poorer welfare (Burman et al. 2008; Mitchell et al. 2012; Riemer et al. 2016; Neville et al. 2017).

Whilst I found an effect of handling method on the duration of the SNC, it did not affect the size of the SNC as might have been expected (Crespi 1942; Flaherty 1996; Flaherty et al. 1998; Neville et al. 2017) but see (Burman et al. 2008) for a similar result. It has been suggested that the initial response to a drop in reward value is a non-emotional process, and less likely to be sensitive to changes in underlying affect (Flaherty 1996; Burman et al. 2008). Therefore, it is possible that I did not detect a difference in the size of the SNC between groups because it was unlikely to differ between tail and tunnel handled mice. In contrast, the duration of the SNC effect might be more meaningful for inferring animal affect because it might reflect an individual's recovery to the change in reward value (Burman et al. 2008). In support of this idea, humans who are in negative affective states often show long lasting negative cognitions and affect following exposure to negative events (Gunthert et al. 2007; Burman et al. 2008). Therefore, the prolonged SNC effect observed in my study could be attributed to the fact that tail handled mice take longer to recover from the drop in reward and could also have a lower expectation that the reward will return to its original size, suggestive of more pessimistic cognitive biases than tunnel handled mice (Paul et al. 2005). Although I am unable to draw firm conclusions as to the exact mechanisms underlying the prolonged SNC, the fact that

tail handled mice show a greater sensitivity to reward loss is suggestive of negative affect and poorer welfare when handled by the tail.

Although there was evidence that handling method affected the duration of the SNC, this wasn't evident in the SPC data. I found no difference between tail and tunnel handled mice in the size or duration of a SPC effect. However, I did find evidence that lick cluster size was capable of detecting a SPC in both tail and tunnel handled mice, although this was not consistently found across all IBIs. This novel finding suggests that mice may experience a state akin to 'elation' or 'joy' (Crespi 1942; Flaherty 1996), and to my knowledge, is the first evidence of positive affect in laboratory mice.

My ability to detect a SPC using licking microstructure highlights this method as a potential valuable way of assessing positive affect and positive contrast effects in laboratory rodents. This is important because far less research has focused on how an animal responds to reward gain, largely because SNC effects are easier to detect than SPC effects (Flaherty & Largen 1975; Flaherty 1996; Papini et al. 2001). This is mainly due to two reasons. The first is that a drop in reward value is considered more salient than a gain in reward value (Flaherty & Largen 1975; Flaherty 1996; Papini et al. 2001), and so has remained the predominant contrast paradigm. The second is that measures of SPC (e.g. runway speed or the total amount of sucrose drank) are more susceptible to ceiling effects. For example, a rat has an upper limit of how fast it can run, or how much it can consume in a given trial. The ability to provide a more effective method for measuring positive contrast effects in laboratory rodents has important implications for animal welfare science. This is because it is now widely recognised that, for an animal to be considered in a good welfare state, it not only has to have the absence of negative affective states but should also be capable of experiencing positive states which, to date, have remained relatively understudied (Boissy et al. 2007). Therefore, my work highlights the potential use of licking microstructure to measure positive affect in laboratory mice and shed more light on their emotional lives and their welfare.

Contrary to my expectations, I failed to see a consistent SPC effect across all IBIs in tunnel handled mice, yet saw a more consistent SPC effect across IBIs in tail handled mice. However, this could also be explained due to a number of factors. The

first is the lack of stability in the lick cluster sizes obtained from the tail handled control mice. Despite remaining on the same sucrose concentration, their lick cluster sizes decreased. Therefore, the difference between the contrast group and its control did not solely reflect an increase in lick cluster size in tail handled mice undergoing the SPC but also a decrease in their respective controls. The second factor might be differences in the animals' valuations of reward. As previously shown (outlined in Chapter 5), tunnel handled mice have a higher hedonic rating of low concentration sucrose solutions compared to mice handled by their tails. Therefore, the expected difference in valuation of the low (4%) and high (32%) reward were different between our two treatments, with the magnitude of difference between the two concentrations being lower for tunnel handled mice. Therefore, it is possible that tunnel handled mice perceive the gain in reward value as less rewarding compared to tail handled mice, because of their increased hedonic value of the low sucrose concentration. Finally, an alternative explanation for an inconsistent result for SPC effects could be explained by the single shift paradigm used. Previous research has suggested that a SPC is more likely seen following multiple shifts in reward value, where animals must have first encountered an adverse state to fully appreciate a gain in reward (Maxwell et al. 1976). The fact that the mice in my study were only shifted from 4% to 32% (i.e. a single shift), and did not experience an adverse state before experiencing the gain condition, might help explain the inconsistent SPC effect. Therefore, although I cannot be sure, perhaps more shifts in reward value, or a longer post-shift phase allowing mice to return to the level of their controls, might have helped to detect differences between tail and tunnel handled mice in their SPC effects. Future experiments would benefit from longer post-shift phases in order to determine fully how handling method might influence mice's sensitivity to reward gain.

My study captures the animals' responses to both a SNC and a SPC in a single study, which is important for assessing both positive and negative affect simultaneously and inferring an animal's position in core affect space. Previous studies exploring how these contrast effects are mediated by an animal's affective state have predominantly focused on the animals responses following a drop in reward value (Burman et al. 2008; Mitchell et al. 2012; Riemer et al. 2016), and the one study that attempted to combine both did not use traditional contrast paradigms (Neville et al. 2017). Measuring responses to both reward loss and gain in a single

study allowed me to establish how handling method affected positive and negative changes in affective state simultaneously. This has important implications for elucidating the mechanisms contributing to the negative affective state in tail handled mice, and in particular, whether these changes are mediated by negative events being perceived more negatively and/or positive events being perceived as less positive. This is of particular importance when trying to understand an animal's position in core affect space. The integrative framework developed by Mendl and colleagues (Mendl, Burman, et al. 2010) is made up of four quadrants, based upon valence and arousal (Figure 1.1, p7). Where an animal might sit in core affect space can be ascertained by how they respond to both positive and negative events and stimuli. For example, Mendl suggested that an animal in a depressed-like state would show decreased expectation of positive events or stimuli (such as reward), whereas an animal in an anxious state would show enhanced expectation of negative events or stimuli (such as punishment) (Mendl, Burman, et al. 2010). Therefore, the ability to assess how an animal responds to reward is fundamental if we are to determine its position in core affect space. In this study, I only found evidence that tail handling caused changes in the mice's perception towards reward loss, but not to reward gain. Therefore, this suggests that tail handled mice are in a sad or depressed-like state. However, given that my post-shift phase may not have been sufficiently long enough to detect differences between the groups in their response towards reward gain, I am unable to make firm conclusions as to how handling method might affect their capacity to experience positive affective states. However, this study clearly demonstrates the potential of this methodology for testing the integrative framework suggested by Mendl and colleagues (Mendl, Burman, et al. 2010). Testing an animal's responses to both reward loss and gain in a single study can assess positive and negative affect simultaneously, and help determine its position in core affect space.

I not only found that handling method impacted on the animal's evaluation of reward and show that tail handling is sufficient to induce anxiety and negative affect, in line with previous findings (Hurst & West 2010; Gouveia et al. 2013; Clarkson et al. 2018), but also conclusively showed for the first time that tail handling causes physiological effects associated with chronic stress. Although I found no evidence of handling method influencing immune system function through the size of the thymus

glands (Živković et al. 2005), I did find an effect on the hypothalamic-pituitary-adrenal (HPA) axis. Tail handled mice had larger adrenal glands compared to tunnel handled mice, indicative of hyperactivity of the hypothalamic-pituitary-adrenal (HPA) axis due to long-term exposure to stress (Nemeroff et al. 1992; Rubin et al. 1995). Although a previous study found differences in physiological measures of blood glucose and circulating corticosterone levels between tail and cup handled mice, the findings could be explained by habituation to being handled rather than the handling method itself: tail handled mice received less handling compared to cup handled animals (Ghosal et al. 2015). This provides the first evidence that the stress-inducing effects of tail handling are likely chronic in nature, and argues against any suggestion that increased anxiety in the OF and EPM could be an acute effect of tail handling immediately before the behavioural test. This finding has huge implications for biomedical research. Mice are used in modelling a vast array of different human diseases, but also for the development of novel pharmacological compounds: it is possible that the stress associated with tail handling could affect scientific outcomes. For example, it could have unwanted effects on well-established *in-vivo* disease models or cause differences in metabolism that change the outcomes of pharmacokinetic studies leading to less translatable findings to human clinic (Shanks et al. 2009; Perrin 2014). Future research would benefit from exploring the mechanisms underpinning this change in order for researchers to make evidence-based decisions about their choice of handling method.

To conclude, this study highlights the potential of using licking microstructure as a novel methodology to test existing theoretical frameworks for understanding animal affect and applying these to improve animal welfare. I found that lick cluster size better reflected the hedonic responses of laboratory mice towards rewarding sucrose solutions, and as such provided an effective method for measuring their responses following changes in reward value and highlighted its use for contrast paradigms. I provide more data to support informed decisions about the choice of handling method, and strengthen the case for the implementation of tunnel handling on both welfare and scientific grounds. Tunnel handling is a simple yet effective refinement, however, more research is needed to shed more insight into how underlying physiological mechanisms are affected in order for researchers to make

informed decisions and predictions about how their choice of handling method might influence existing *in-vivo* models.

Chapter 7: General discussion and conclusions

At the start of this thesis, I set out to fulfil three aims. In this chapter, I will evaluate how my empirical work has addressed these key aims, and highlight possible future directions to help advance our understanding of animal affect and improve animal welfare.

7.1 Aim 1: How does stress affect the hedonic responses of laboratory mice towards reward?

The exposure to stressors is known to impact the amount of sucrose that laboratory mice consume, where a lower consumption is taken to reflect an anhedonic and depressed-like state (Willner et al. 1987; Willner et al. 1992; Monleon et al. 1995; Willner 1997; Willner 2017b). However, it has been suggested that measuring changes in sucrose consumption may not solely reflect an animal's hedonic response, because it can be highly influenced by other factors, particularly motivation (Dwyer 2012). Instead, measuring changes in the way that rodents drink may offer a more sensitive measure of their hedonic responses (Dwyer 2012; Austen et al. 2016). Consequently, my initial experiments were all aimed at manipulating mice's affective state using known stressors in order to establish their effect on their licking microstructure, and explore whether it can measure changes in their hedonic responses towards reward.

Although the Chronic Mild Stress paradigm (Chapter 3) and the chronic administration of corticosterone (Chapter 4) are both established protocols for inducing negative affective states in mice (Willner et al. 1987; Willner et al. 1992; Ardayfio & Kim 2006; Rainer et al. 2012), they were ineffective in doing so in my experiments. Consequently, I was unable to test if stress had any measurable effect on the licking microstructure and hedonic responses of laboratory mice towards rewarding sucrose solutions. However, I did successfully manage to induce negative affective states in mice through another known stressor: the method by which mice were handled (Chapters 5 and 6). I found that mice handled by the tail were less sensitive to rewarding sucrose solutions, and had more anhedonic-like responses

compared to mice handled using a tunnel. Not only did tail handled mice drink less sucrose than tunnel handled mice, but the size of their licking bouts (their lick cluster sizes) were also smaller. Therefore, I fulfilled the first aim of my thesis by successfully demonstrating that stress negatively affected laboratory mice's hedonic responses towards reward through changes in their licking microstructure. Licking microstructure can be used to measure the effects of stressors on the affective state of laboratory mice.

It was perhaps surprising to find such a clear effect of tail handling on licking microstructure given the null results reported in Chapters 3 and 4, which used well-established paradigms considered effective in inducing negative affect in laboratory mice. However, problems with the reliability, reproducibility and efficacy of the Chronic Mild Stress (CMS) paradigm have been relatively well documented within the literature, which prompted a user survey just last year (Willner 2017a). Although a similar body of evidence questioning the efficacy of chronic corticosterone as a stress manipulation doesn't exist, it does not necessarily mean that it is a reliable and robust method of inducing negative affect. It is possible that negative results are commonly found across laboratories, but are not reported, leading to a publication bias. Whilst existing stress-inducing manipulations used in this thesis may vary in their reliability and reproducibility, tail handling appears to produce reliable and reproducible effects. This highlights the possibility that tail handling and other routine laboratory practices may be more stressful than previously thought.

7.2 Aim 2: Is licking microstructure a better and more robust measure of animal affect?

The second aim of my thesis was to establish whether assessing changes in licking microstructure offered a better and more robust method of accessing the affective state of laboratory mice compared to using the standard measure of sucrose consumption. I predicted that assessing changes in laboratory mice's licking microstructure would better reflect how much they like a tastant, and provide a better measure of the hedonic responses of laboratory mice compared to assessing total consumption. In Chapter 5, I found a slight difference in the results between total

consumption and lick cluster size. Specifically, I found that lick cluster size was affected by ceiling effects at the higher sucrose concentration, whereas consumption was not. Therefore, although results were qualitatively the same using both measures, I found no evidence to suggest that lick cluster size was advantageous over assessing changes in consumption.

However, I did find evidence to suggest that licking microstructure was advantageous in Chapter 6. Here, I found evidence to suggest that licking microstructure offered a novel methodology that gave a greater ability to assess mice's affective states following changes in reward value. This was because the total amount of sucrose consumed did not reflect the predicted hedonic difference between the high (32%) and low (4%) sucrose solutions, and could not be used to investigate a Successive Negative Contrast (SNC) or Successive Positive Contrast (SPC) effect. This is likely to be, at least in part, explained to the bell-shaped relationship that exists between consumption and sucrose concentration (Figure 1.2, p14), where low and high sucrose concentrations can lead to the same amount being consumed. Therefore, the work from this chapter clearly demonstrated the benefits of assessing changes in licking microstructure to measure mice's hedonic responses, and make inferences regarding their underlying affective state.

The work in Chapter 6 also highlighted the use of this methodology for measuring the resilience to reward loss and sensitivity to reward gain and therefore measure both negative and positive affect in a single behavioural paradigm. This is not possible with other behavioural tests such as the Open Field (OF) and Elevated Plus Maze (EPM), which work on the basis that the exposure to the test themselves inflict stress and fear, and therefore only predominantly measure the presence of negative affective states. This has important implications for research on animal emotion and welfare because, given that there are no established tests to measure cognitive biases in laboratory mice, it offers a novel methodology of accessing the cognitive changes associated with changes in affect that are sensitive to the valence of emotion (Mendl & Paul 2004; Paul et al. 2005; Mendl, Burman, et al. 2010). I propose that assessing contrast effects using licking microstructure may provide a novel way to further explore the integrative framework on animal emotions proposed by Mendl and colleagues (Mendl, Burman, et al. 2010) (Figure 1.1, p7). This is

because it enables both positive and negative affect to be explored in a single experimental paradigm. Therefore, in addition to being a more direct measure over consumption with reduced dependence on motivation, contrast paradigms using licking microstructure have the potential to provide information to help identify an animal's position in core affect space.

My findings also demonstrate that licking microstructure can be used to detect changes indicative of negative but also positive affect. In Chapter 6, perhaps for the first time, I found evidence of positive affect in laboratory mice. This is of significance as it highlights the potential of licking microstructure to measure not only the absence of negative affective states but also the presence of positive ones, both of which are considered important for longer term positive affect and good welfare (Boissy et al. 2007).

Despite the clear benefits, it is also important not to overlook possible limitations of this methodology compared to other tests. Lickometers are expensive compared to performing more traditional sucrose consumption tests, which perhaps limits their application. In addition, animals need to be trained and habituated to the novel taste of sucrose and being moved to the drinking equipment itself, which takes additional time and resources compared to simply placing bottles of sucrose on the animal's home cage. Because of these potential limitations, future work could focus on measuring changes in licking microstructure using 'intellicage' systems which would offer a less time consuming methodology to measure the affective state of laboratory mice across longer time periods.

7.3 Aim 3: Can licking microstructure be used to make improvements to laboratory mouse welfare?

My experiments highlight how licking microstructure can be successfully used to measure the impact of standard husbandry practices and suggest evidence-based improvements to current legislation and improve animal welfare. In Chapters 5 and 6, I showed that the standard practice of tail handling laboratory mice induced a negative affective state which was not restricted to anxiety and fear as previously documented (Hurst & West 2010; Gouveia et al. 2013). Therefore, my work built on

existing work conducted by Hurst and colleagues (Hurst & West 2010; Gouveia et al. 2013) and was the first to show that tail handling causes alterations in mice's hedonic responses leading to an anhedonic and therefore depressed-like state in laboratory mice (Chapter 5). I also demonstrated that these changes in hedonic responses, and reward sensitivity, were also dependent upon prior experience with reward (Chapter 6), where tail handled mice were more sensitive to reward loss. This finding suggests that tail handled mice perceive negative events worse than mice handled using a tunnel, which has significant implications for understanding how their experiences shape their affective state, and vice versa. Taken together, my work not only provides more support for the implementation of tunnel handling as a refinement for handling laboratory mice, but also raises some important ethical considerations.

The findings presented in Chapters 5 and 6 provide evidence to suggest that tail handling is inherently more stressful than previously thought, and potentially more stressful than well-validated stressors such as CMS and chronic corticosterone administration. In fact, when comparing the consumption of tail handled mice to tunnel handled mice, they drank 27% less, which is comparable to mice undergoing CMS who show a 33-57% reduction in total sucrose consumption compared to control mice (Pothion et al. 2004; Strekalova et al. 2004). Therefore, the fact that tail handling produces a comparable reduction in sucrose intake similar to studies that have proven CMS effects, means it can be considered as stressful as an effective CMS regime. This is of particular importance because manipulating an animal's affective state using the CMS regime or via corticosterone administration are both deemed to be invasive and are regulated procedures that require a Home Office license in the UK. According to the Animals (Scientific Procedures) Act (ASPA) 1986, a regulated procedure is a "procedure applied to a protected animal for a qualifying purpose which may have the effect of causing the animal a level of pain, suffering, distress or lasting harm equivalent to, or higher than, that caused by the introduction of a needle in accordance with good veterinary practice" (Home Office 1986). Therefore, given the evidence presented here and by others (Hurst & West 2010; Gouveia et al. 2013; Ghosal et al. 2015), it poses the question of whether tail handling should also be considered a regulated procedure. According to ASPA, pain, suffering or lasting harm can be taken to include anything that affects an animal's "physical, mental and social wellbeing" and can include "physiological or

psychological discomfort" (Home Office 1986). My data not only demonstrate an effect of tail handling on physiological wellbeing (larger adrenal glands are taken to be indicative of a greater stress response), but also adds to the increasing evidence of the psychological impact of tail handling. Tail handled mice demonstrate behavioural changes indicative of stress, anxiety and for the first time, a depressive-like state (Clarkson et al. 2018). Therefore, when considering all the evidence, I believe that there is now sufficient data to explore if tail handling should be considered a regulated procedure. However, perhaps to be fully compelling and change legislation, future work would benefit from the direct comparison between tail handling with the definition of a regulated procedure, i.e. the introduction of a needle.

My work in Chapter 6 also highlights the wider implications of tunnel handling for mitigating the effects of stress associated with negative experiences. Crucially, the data demonstrate the resilience of tunnel handled mice to reward loss, and therefore it is possible that handling method might also influence the animal's perception of negative experiences encountered in a laboratory setting. This is important because mice held in laboratories not only encounter stress associated with being kept in artificial environments, but are often used in experimental procedures which commonly induce some degree of pain, suffering or lasting harm. Therefore, the possibility that tunnel handling might mitigate some of the stress associated with these experimental procedures has significant implications for improving mouse welfare for those millions of mice used for scientific purpose worldwide.

At present, despite the building evidence of the welfare benefits of tunnel handling, resistance to its uptake remains, both across the scientific community but also across technical staff. From a scientific standpoint, there is the concern that, like with any new refinement, it could affect scientific data quality. It is often deemed necessary to maintain consistency across and within scientific studies, and so changing handling practices during successive or longitudinal studies may raise concern with regards to its effect on the reliability of scientific data collected from mice handled differently. From a technical standpoint, the main concern seems to be with regards to timing. Research institutions often house thousands of mice, which all

need basic husbandry and regular cage changes. Therefore there is concern that tunnel handling will take longer to perform these basic husbandry practices.

However, although the effects on scientific data quality remain unknown at this stage, I suggest that implementing tunnel handling not only has implications for making refinements, but also offers the potential to reduce the animals required for experiments. There is the possibility that tunnel handling could produce more valid results, larger effect sizes and ultimately reduce the sample size required (Gouveia et al. 2017; Clarkson et al. 2018). This is because stress has wide-ranging effects on an animals' behaviour and physiology (Moberg 2000; Antonia et al. 2012; Beery & Kaufer 2015; Starcevic et al. 2016) and has the potential to impact on most, if not all, scientific measures. This is of particular importance when considering the use of 'control' groups for comparison, the potential to limit the amount of stress that control animals are exposed to offers the potential for more meaningful comparisons, larger effect sizes and consequently smaller sample sizes.

7.4 Future work

I have clearly demonstrated that lick cluster size can be used to measure the affective state of laboratory mice. Further work would benefit from measuring lick cluster size in a home cage setting. This would not only limit the requirement for training and the novelty of the drinking cage, but would also provide longitudinal data in order to assess whether measuring changes in licking microstructure could be a useful, non-invasive tool for determining the affective state of mice undergoing experimental procedures, in order to make evidence-based refinements to these procedures. For example, the ability to assess the licking microstructure of mouse models of human disease or animals that have undergone a particular severe experimental procedure, such as surgery, would enable us to assess their affective state 24hrs a day. The ability to measure how a laboratory mice's affective state might change across time, could help to refine experimental procedures by making revisions to existing interventions such as analgesic treatments, or suitable end-points.

With regards to the handling method of laboratory mice**,** although there is now increasing evidence stacking up against the standard practice of tail handling, non-aversive handling methods, such as tunnel handling, are yet to be widely implemented across research institutions. This is because there is resistance to change from both the scientific community, but also technical staff. Therefore, more work is needed to further explore the scientific and practical barriers to the implementation of tunnel handling. In particular, future work would benefit from exploring whether tunnel handling offers the potential to mitigate the stressful effects associated with a number of scientific procedures (e.g. the administration of compounds via injection). Such studies would shed more light on the cumulative experiences of laboratory mice, as currently little is known about how the repeated exposure to negative events shape their underlying affective state. Furthermore, from a more scientific perspective, further work is required in order to investigate the underlying physiological mechanisms underpinning the stress response associated with tail handling. For example, it would provide useful to understand how the stress associated with tail handling affects fundamental biological processes such as neural processing in the brain underpinning learning and memory, or how it affects the development, progression and/or treatment of certain diseases. I believe this is fundamental for determining the effect of tail handling on existing *in-vivo* models, and for researchers to make evidence-based decisions about their choice of handling method.

7.5 Concluding remarks

Taken together, the work presented in this thesis clearly demonstrates that licking microstructure is a valuable method of accessing mice's underlying affective state, and demonstrates its potential to be applied to make evidence-based improvements to the welfare of laboratory mice. My thesis provides fundamental evidence to support the refinement of existing handling methods for laboratory mice, which if implemented, has the potential to improve the welfare of millions of mice worldwide.

References

Ali, S.H. et al., 2015. Resveratrol ameliorates depressive-like behavior in repeated corticosterone-induced depression in mice. *Steroids*, 101, pp.37–42.

American Psychiatric Association, 2014. *Diagnostic and Statistical Manual of mental disorders fifth edition*,

Anisman, H. et al., 1998. Stressor-induced corticotropin-releasing hormone, bombesin, ACTH and corticosterone variations in strains of mice differentially responsive to stressors. *Stress (Amsterdam, Netherlands)*, 2(3), pp.209–20.

Antkiewicz-Michaluk, L. et al., 2014. Antidepressant-like Effect of Tetrahydroisoquinoline Amines in the Animal Model of Depressive Disorder Induced by Repeated Administration of a Low Dose of Reserpine: Behavioral and Neurochemical Studies in the Rat. *Neurotoxicity Research*, 26(1), pp.85–98.

Antonia, K., Anastasia, A. & Tesseromatis, C., 2012. Stress can affect drug pharmacokinetics via serum/tissues protein binding and blood flow rate alterations. *European Journal of Drug Metabolism and Pharmacokinetics*, 37(1), pp.1–7.

Ardayfio, P. & Kim, K.-S., 2006. Anxiogenic-Like Effect of Chronic Corticosterone in the Light–Dark Emergence Task in Mice. *Behav Neurosci*, 120(6), pp. 1267.

Armario, A., Montero, J.L. & Balasch, J., 1986. Sensitivity of corticosterone and some metabolic variables to graded levels of low intensity stresses in adult male rats. *Physiology & behavior*, 37(4), pp.559–61.

Aslani, S. et al., 2014. Day and night: diurnal phase influences the response to chronic mild stress. *Frontiers in behavioral neuroscience*, 8, p.82.

Austen, J.M. & Sanderson, D.J., 2016. Contexts control negative contrast and restrict the expression of flavor preference conditioning. *Journal of experimental psychology. Animal learning and cognition*, 42(1), pp.95–105.

Austen, J.M., Strickland, J.A. & Sanderson, D.J., 2016. Memory-dependent effects on palatability in mice. *Physiology and Behavior*, 167, pp.92–99.

Axelrod, J. & Reisine, T.D., 1984. Stress hormones: their interaction and regulation. *Science (New York, N.Y.)*, 224(4648), pp.452–9.

Bachmanov, A.A., Tordoff, M.G. & Beauchamp, G.K., 2001. Sweetener Preference of C57BL/6ByJ and 129P3/J Mice. *Chemical Senses*, 26(7), pp.905–913.

Baird, J.-P., St John, S.J. & Nguyen, E.A.-N., 2005. Temporal and qualitative dynamics of conditioned taste aversion processing: combined generalization testing and licking microstructure analysis. *Behavioral neuroscience*, 119(4), pp.983–1003.

Bardo, M.T. & Bevins, R.A., 2000. Conditioned place preference: what does it add to our preclinical understanding of drug reward? *Psychopharmacology*, 153(1), pp.31–43.

Barr, A.M. & Markou, A., 2005. Psychostimulant withdrawal as an inducing condition in animal models of depression. *Neuroscience & Biobehavioral Reviews*, 29(4–5), pp.675–706.

Barr, A.M., Markou, A. & Phillips, A.G., 2002. A 'crash' course on psychostimulant withdrawal as a model of depression. *Trends in Pharmacological Sciences*, 23(10), pp.475–482.

Barrett, L.F. et al., 2007. The Experience of Emotion. *Annual Review of Psychology*, 58(1), pp.373–403.

Bateson, M. et al., 2011. Agitated honeybees exhibit pessimistic cognitive biases. *Current biology : CB*, 21(12), pp.1070–3.

Bateson, M., 2016. Cumulative stress in research animals: Telomere attrition as a biomarker in a welfare context? *BioEssays*, 38(2), pp.201–212.

Bateson, M. & Matheson, S.M., 2007. Performance on a categorisation task suggests that removal of environmental enrichment induces "pessimism" in captive European starlings (Sturnus vulgaris). *Animal Welfare*, 16(S), pp.33–36.

Bateson, P. & Martin, P., 2013. *Play, playfulness, creativity and innovation*, Cambridge University Press.

Beck, A.T. et al., 1979. *Cognitive Therapy of DepressionNew York: Guilford Press,*

1979, Guildford Press.

Beck, A.T., 1969. *Depression: Clinical, Experimental and Theoretical Aspects*, Taylor & Francis.

Becker, H.C. & Flaherty, C.F., 1983. Chlordiazepoxide and ethanol additively reduce gustatory negative contrast. *Psychopharmacology*, 80(1), pp.35–37.

Beery, A.K. & Kaufer, D., 2015. Stress, social behavior, and resilience: insights from rodents. *Neurobiology of stress*, 1, pp.116–127.

Benefiel, A.C. et al., 2005. Mandatory “Enriched” Housing of Laboratory Animals: The Need for Evidence-based Evaluation. *ILAR Journal*, 46(2), pp.95–105.

Berridge, C.W. & Dunn, A.J., 1989. Restraint-stress-induced changes in exploratory behavior appear to be mediated by norepinephrine-stimulated release of CRF. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 9(10), pp.3513–21.

Berridge, K.C., 1996. Food reward: brain substrates of wanting and liking. *Neuroscience and biobehavioral reviews*, 20(1), pp.1–25.

Berridge, K.C., 2000. Measuring hedonic impact in animals and infants: Microstructure of affective taste reactivity patterns. *Neuroscience and Biobehavioral Reviews*, 24(2), pp.173–198.

Boissy, A. et al., 2007. Assessment of positive emotions in animals to improve their welfare. *Physiol Behav*, 92(3), pp.375–397.

Booth, D.A., Lovett, D. & McSherry, G.M., 1972. Postingestive modulation of the sweetness preference gradient in the rat. *Journal of Comparative and Physiological Psychology*, 78(3), pp.485–512.

Bower, G.H., 1981. Mood and memory. *American Psychologist*, 36(2), pp.129–148.

Brennan, K. et al., 2001. Individual differences in sucrose consumption in the rat: motivational and neurochemical correlates of hedonia. *Psychopharmacology*, 157(3), pp.269–276.

Brilot, B.O., Asher, L. & Bateson, M., 2010. Stereotyping starlings are more ‘pessimistic.’ *Animal Cognition*, 13(5), pp.721–731.

Brydges, N.M. et al., 2011. Environmental enrichment induces optimistic cognitive bias in rats.

Brydges, N.M. & Hall, L., 2017. A shortened protocol for assessing cognitive bias in rats. *Journal of Neuroscience Methods*, 286, pp.1–5.

Bućan, M. & Abel, T., 2002. The mouse: genetics meets behaviour. *Nature Reviews Genetics*, 3(2), pp.114–123.

Burgdorf, J. & Panksepp, J., 2001. Tickling induces reward in adolescent rats. *Physiology & Behavior*, 72(1–2), pp.167–173.

Burman, O. et al., 2008. Sensitivity to reward loss as an indicator of animal emotion and welfare. *Biology letters*, 4(4), pp.330–3.

Carstens, E. & Moberg, G.P., 2000. Recognizing Pain and Distress in Laboratory Animals. *ILAR Journal*, 41(2), pp.62–71.

Chabot, D., Gagnon, P. & Dixon, E.A., 1996. Effect of predator odors on heart rate and metabolic rate of wapiti (Cervus elaphus canadensis). *Journal of Chemical Ecology*, 22(4), pp.839–868.

Chapillon, P. et al., 1999. Rearing Environmental Enrichment in Two Inbred Strains of Mice: 1. Effects on Emotional Reactivity. *Behavior Genetics*, 29(1), pp.41–46.

Chiba, S. et al., 2012. Chronic restraint stress causes anxiety- and depression-like behaviors, downregulates glucocorticoid receptor expression, and attenuates glutamate release induced by brain-derived neurotrophic factor in the prefrontal cortex. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 39(1), pp.112–119.

Chourbaji, S. et al., 2006. IL-6 knockout mice exhibit resistance to stress-induced development of depression-like behaviors. *Neurobiology of Disease*, 23(3), pp.587–594.

Chourbaji, S. et al., 2005. Social and structural housing conditions influence the development of a depressive-like phenotype in the learned helplessness paradigm in male mice. *Behavioural Brain Research*, 164(1), pp.100–106.

Chu, X. et al., 2016. 24-hour-restraint stress induces long-term depressive-like

phenotypes in mice. *Scientific Reports*, 6(1), p.32935.

Clarkson, J.M. et al., 2018. Handling method alters the hedonic value of reward in laboratory mice. *Scientific Reports*, 8(1), p.2448.

Crespi, L., 1942. Quantitative variation of incentive and performance in the white rat. *The American Journal of Psychology*, 55(4), pp.467–517.

David, D.J. et al., 2009. Neurogenesis-Dependent and -Independent Effects of Fluoxetine in an Animal Model of Anxiety/Depression. *Neuron*, 62(4), pp.479–493.

Davis, J.D., 1973. The effectiveness of some sugars in stimulating licking behavior in the rat. *Physiol Behav*, 11(1), pp.39–45.

Davis, J.D., 1989. The Microstructure of Ingestive Behavior. *Annals of the New York Academy of Sciences*, 575(1), pp.106–121.

Davis, J.D. & Perez, M.C., 1993. The acquired control of ingestive behavior in the rat by flavor-associated postingestional stimulation. *Physiology & behavior*, 54(6), pp.1221–1226.

Davis, J.D. & Smith, G.P., 1992. Analysis of the Microstructure of the Rhythmic Tongue Movements of Rats Ingesting Maltose and Sucrose Solutions. *Behav Neurosci*, 106(1), pp.217–228.

Dawkins, M.S., 1990. From an animal's point of view: Motivation, fitness, and animal welfare. *Behavioral and Brain Sciences*, 13(01), pp.1–9.

Dawkins, M.S., 2008. The Science of Animal Suffering. *Ethology*, 114(10), pp.937–945.

Deacon, R.M.J., 2006. Housing, husbandry and handling of rodents for behavioral experiments. *Nature Protocols*, 1(2), pp.936–946.

Deussing, J.M., 2006. Animal models of depression. *Drug Discovery Today: Disease Models*, 3(4), pp.375–383.

Donner, N.C. et al., 2012. Chronic non-invasive corticosterone administration abolishes the diurnal pattern of tph2 expression. *Psychoneuroendocrinology*, 37(5), pp.645–61.

Doyle, R.E. et al., 2010. Release from restraint generates a positive judgement bias in sheep. *Applied Animal Behaviour Science*, 122(1), pp.28–34.

Ducottet, C., Aubert, A. & Belzung, C., 2004. Susceptibility to subchronic unpredictable stress is related to individual reactivity to threat stimuli in mice. *Behavioural brain research*, 155(2), pp.291–9.

Ducottet, C. & Belzung, C., 2004a. Behaviour in the elevated plus-maze predicts coping after subchronic mild stress in mice. *Physiology & behavior*, 81(3), pp.417–26.

Ducottet, C. & Belzung, C., 2004b. Behaviour in the elevated plus-maze predicts coping after subchronic mild stress in mice. *Physiology and Behavior*, 81(3), pp.417–426.

Ducottet, C. & Belzung, C., 2005. Correlations between behaviours in the elevated plus-maze and sensitivity to unpredictable subchronic mild stress: evidence from inbred strains of mice. *Behavioural Brain Research*, 156(1), pp.153–162.

Ducottet, C., Griebel, G. & Belzung, C., 2003. Effects of the selective nonpeptide corticotropin-releasing factor receptor 1 antagonist antalarmin in the chronic mild stress model of depression in mice. *Progress in neuro-psychopharmacology & biological psychiatry*, 27(4), pp.625–31.

Dudink, S. et al., 2006. Announcing the arrival of enrichment increases play behaviour and reduces weaning-stress-induced behaviours of piglets directly after weaning. *Applied Animal Behaviour Science*, 101(1–2), pp.86–101.

Duncan, I.J., 2006. The changing concept of animal sentience. *Applied Animal Behaviour Science*, 100(1–2), pp.11–19.

Duncan, I.J.H., 1981. Animal Rights - Animal Welfare: A Scientist's Assessment. *Poultry Science*, 60(3), pp.489–499.

Dwyer, D.M. et al., 2009. A learned flavor preference persists despite the extinction of conditioned hedonic reactions to the cue flavors. *Learn Behav*, 37(4), pp.305–310.

Dwyer, D.M., 2012. EPS Prize Lecture. Licking and liking: the assessment of hedonic

responses in rodents. *Q J Exp Psychol (Hove)*, 65(3), pp.371–394.

Dwyer, D.M., 2009. Microstructural analysis of ingestive behaviour reveals no contribution of palatability to the incomplete extinction of a conditioned taste aversion. *Quarterly Journal of Experimental Psychology*.

Dwyer, D.M., Gasalla, P. & Lopez, M., 2013. Nonreinforced flavor exposure attenuates the effects of conditioned taste aversion on both flavor consumption and cue palatability. *Learn Behav*, 41(4), pp.390–401.

Dwyer, D.M., Lydall, E.S. & Hayward, A.J., 2011. Simultaneous contrast: evidence from licking microstructure and cross-solution comparisons. *Journal of experimental psychology. Animal behavior processes*, 37(2), pp.200–10.

European Commission, 2013. *Seventh Report on the Statistics on the Number of Animals used for Experimental and other Scientific Purposes in the Member States of the European Union*,

Fairchild, G., 2003. Acute and chronic effects of corticosterone on 5-HT1A receptor-mediated autoinhibition in the rat dorsal raphe nucleus. *Neuropharmacology*, 45(7), pp.925–934.

Farley, S. et al., 2012. Increased expression of the Vesicular Glutamate Transporter-1 (VGLUT1) in the prefrontal cortex correlates with differential vulnerability to chronic stress in various mouse strains: Effects of fluoxetine and MK-801. *Neuropharmacology*, 62(1), pp.503–517.

Fava, M. & Kendler, K.S., 2000. Major Depressive Disorder. *Neuron*, 28(2), pp.335–341.

Finlayson, K. et al., 2016. Facial Indicators of Positive Emotions in Rats E. Hillmann, ed. *PLOS ONE*, 11(11), p.e0166446.

Flaherty, C.F., 1982. Incentive contrast: A review of behavioral changes following shifts in reward. *Animal Learning & Behavior*, 10(4), pp.409–440.

Flaherty, C.F., 1996. *Incentive Relativity*, Cambridge University Press.

Flaherty, C.F., Grigson, P. & Rowan, G., 1986. Chlordiazepoxide and the determinants of negative contrast. *Animal Learning & Behavior*, 14(3), pp.315–

321.

Flaherty, C.F. & Largen, J., 1975. Within-subjects positive and negative contrast effects in rats. *Journal of comparative and physiological psychology*, 88(2), pp.653–64.

Flaherty, C.F., Martin, J. & Leszczuk, M., 1998. Relationship of negative contrast to animal models of fear and anxiety. *Animal Learning & Behavior*, 99826(4), pp.397–407.

Flaherty, C.F. & Rowan, G. a, 1986. Successive, simultaneous, and anticipatory contrast in the consumption of saccharin solutions. *Journal of experimental psychology. Animal behavior processes*, 12(4), pp.381–93.

Forbes, N.F. et al., 1996. Chronic Mild Stress and Sucrose Consumption: Validity as a Model of Depression. *Physiology & Behavior*, 60(6), pp.1481–1484.

Forgas, J.P., 2000. Introduction: the role of affect on social cognition. In *Feeling and Thinking: the role of affect on social cognition*. Cambridge University Press, pp. 1–28.

Fraser, D. & Duncan, I., 1998. 'Pleasures', 'Pains' and Animal Welfare: Toward a Natural History of Affect. *WEL*.

Garb, J.L. & Stunkard, A.J., 1974. Taste aversions in man. *The American journal of psychiatry*, 131(11), pp.1204–7.

Gärtner, K. et al., 1980. Stress response of rats to handling and experimental procedures. *Laboratory Animals*, 14(3), pp.267–274.

Gaskill, B.N. et al., 2009. Some like it hot: Mouse temperature preferences in laboratory housing. *Applied Animal Behaviour Science*, 116(2–4), pp.279–285.

Gasparini, S.J. et al., 2016. Continuous corticosterone delivery via the drinking water or pellet implantation: A comparative study in mice. *Steroids*, 116, pp.76–82.

Ghosal, S. et al., 2015. Mouse handling limits the impact of stress on metabolic endpoints. *Physiology & behavior*, 150, pp.31–7.

Gong, S. et al., 2015. Dynamics and correlation of serum cortisol and corticosterone under different physiological or stressful conditions in mice. *PloS one*, 10(2),

p.e0117503.

Gorwood, P., 2008. Neurobiological mechanisms of anhedonia. *Dialogues in clinical neuroscience*, 10(3), pp.291–9.

Goshen, I. et al., 2008. Brain interleukin-1 mediates chronic stress-induced depression in mice via adrenocortical activation and hippocampal neurogenesis suppression. *Molecular psychiatry*, 13(7), pp.717–28.

Gould, T.D., Dao, D.T. & Kovacsics, C.E., 2009. The Open Field Test.

Gourley, S.L. & Taylor, J.R., 2009. Recapitulation and reversal of a persistent depression-like syndrome in rodents. *Current protocols in neuroscience*, Chapter 9, p.Unit 9.32.

Gouveia, K. et al., 2017. Optimising reliability of mouse performance in behavioural testing: the major role of non-aversive handling. *Scientific Reports*, 7, p.44999.

Gouveia, K. et al., 2013. Reducing Mouse Anxiety during Handling: Effect of Experience with Handling Tunnels E. M. Mintz, ed. *PLoS ONE*, 8(6), p.e66401.

Gray, R.W. & Cooper, S.J., 1995. Benzodiazepines and palatability: Taste reactivity in normal ingestion. *Physiology & Behavior*, 58(5), pp.853–859.

Griffiths, J., Shanks, N. & Anisman, H., 1992. Strain-specific alterations in consumption of a palatable diet following repeated stressor exposure. *Pharmacology, biochemistry, and behavior*, 42, pp.219–227.

Grill, H. & Norgren, R., 1978a. The taste reactivity test. I. Mimetic responses to gustatory stimuli in neurologically normal rats. *Brain research*, 143(2), pp.263–79.

Grill, H. & Norgren, R., 1978b. The taste reactivity test. I. Mimetic responses to gustatory stimuli in neurologically normal rats. *Brain Research*, 143(2), pp.263–279.

Grippo, A.J. et al., 2005. Neuroendocrine and cytokine profile of chronic mild stress-induced anhedonia. *Physiology & behavior*, 84(5), pp.697–706.

Guillen, J., 2012. FELASA guidelines and recommendations. *Journal of the American Association for Laboratory Animal Science : JAALAS*, 51(3), pp.311–21.

Gunthert, K.C. et al., 2007. Depression and Next-day Spillover of Negative Mood and Depressive Cognitions Following Interpersonal Stress. *Cognitive Therapy and Research*, 31(4), pp.521–532.

Gustavsson, A. et al., 2011. Cost of disorders of the brain in Europe 2010. *European Neuropsychopharmacology*, 21(10), pp.718–779.

Haemisch, A., Guerra, G. & Furkert, J., 1999. Adaptation of corticosterone-but not *β* -endorphin-secretion to repeated blood sampling in rats. *Laboratory Animals*, 33(2), pp.185–191.

Halberg, F., Albrecht, P.G. & Bittner, J.J., 1959. Corticosterone rhythm of mouse adrenal in relation to serum corticosterone and sampling. *American Journal of Physiology-Legacy Content*, 197(5), pp.1083–1085.

Hall, C. & Ballachey, E., 1932. A study of the rat's behavior in a field. A contribution to method in comparative psychology. *University of California Publications in Psychology.*, 6, pp.1–12.

Hall, C.S., 1934. Emotional behavior in the rat. I. Defecation and urination as measures of individual differences in emotionality. *Journal of Comparative Psychology*, 18(3), pp.385–403.

Hammen, C. et al., 2009. Chronic and acute stress and the prediction of major depression in women. *Depression and Anxiety*, 26(8), pp.718–723.

Haney, M. et al., 1997. Alprazolam increases food intake in humans. *Psychopharmacology*, 132(3), pp.311–314.

Harding, E.J., Paul, E.S. & Mendl, M., 2004. Cognitive bias and affective state. *Nature*, 427(6972), pp.312–312.

Heinla, I. et al., 2018. Behavioural characterization of C57BL/6N and BALB/c female mice in social home cage – Effect of mixed housing in complex environment. *Physiology & Behavior*, 188, pp.32–41.

Henderson, N.D. et al., 2004. QTL Analysis of Multiple Behavioral Measures of Anxiety in Mice. *Behavior Genetics*, 34(3).

Hennessy, M.B. & Levine, S., 1978. Sensitive pituitary-adrenal responsiveness to

varying intensities of psychological stimulation. *Physiology & Behavior*, 21(3), pp.295–297.

Higgs, S. & Cooper, S.J., 1998. Effects of benzodiazepine receptor ligands on the ingestion of sucrose, intralipid, and maltodextrin: An investigation using a microstructural analysis of licking behavior in a brief contact test. *Behavioral Neuroscience*, 112(2), pp.447–457.

Hinde, R.A., 1985. Was 'The expression of the emotions' a misleading phrase? *Animal Behaviour*, 33(3), pp.985–992.

Hirschfeld, R.M.A., 2001. The Comorbidity of Major Depression and Anxiety Disorders: Recognition and Management in Primary Care. *Primary care companion to the Journal of clinical psychiatry*, 3(6), pp.244–254.

Holmstrup, M. et al., 2010. Interactions between effects of environmental chemicals and natural stressors: A review. *Science of The Total Environment*, 408(18), pp.3746–3762.

Holsboer, F. & Barden, N., 1996. Antidepressants and Hypothalamic-Pituitary-Adrenocortical Regulation. *Endocrine Reviews*, 17(2), pp.187–205.

Home Office, 1986. Animals (Scientific Procedures) Act 1986.

Home Office, 2014. Code of Practice for the Housing and Care of Animals Bred, Supplied or Used for Scientific Purposes.

Home Office, 2017. *Statistics of Scientific Procedures on living animals, Great Britain 2016*,

Hsiao, S. & Fan, R.J., 1993. Additivity of taste-specific effects of sucrose and quinine: Microstructural analysis of ingestive behavior in rats. *Behav Neurosci*, 107(2), pp.317–326.

Humphrey, N.K. & Keeble, G.R., 1974. The reaction of monkeys to 'fearsome' pictures. *Nature*, 251(5475), pp.500–502.

Hurst, J.L. & West, R.S., 2010. Taming anxiety in laboratory mice. *Nature*, 7(10), pp.825–826.

Jeong, J.Y., Lee, D.H. & Kang, S.S., 2013. Effects of chronic restraint stress on body

weight, food intake, and hypothalamic gene expressions in mice. *Endocrinology and metabolism (Seoul, Korea)*, 28(4), pp.288–96.

Jindal, A., Mahesh, R. & Bhatt, S., 2013. Etazolate rescues behavioral deficits in chronic unpredictable mild stress model: Modulation of hypothalamic-pituitary-adrenal axis activity and brain-derived neurotrophic factor level. *Neurochemistry International*, 63(5), pp.465–475.

Johnson, a. W. et al., 2010. An analysis of licking microstructure in three strains of mice. *Appetite*, 54(2), pp.320–330.

Johnson, S.A., Fournier, N.M. & Kalynchuk, L.E., 2006. Effect of different doses of corticosterone on depression-like behavior and HPA axis responses to a novel stressor. *Behavioural Brain Research*, 168(2), pp.280–288.

June, H.L. & Gilpin, N.W., 2010. Operant self-administration models for testing the neuropharmacological basis of ethanol consumption in rats. *Current protocols in neuroscience*, Chapter 9, p.Unit 9.12.1-26.

Katz, R.J., 1982. Animal model of depression: Pharmacological sensitivity of a hedonic deficit. *Pharmacology Biochemistry and Behavior*, 16(6), pp.965–968.

Kilkenny, C. et al., 2010. Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research. *PLoS Biology*, 8(6), p.e1000412.

Krams, I. & Krama, T., 2002. Interspecific reciprocity explains mobbing behaviour of the breeding chaffinches, Fringilla coelebs. *Proceedings of the Royal Society B: Biological Sciences*, 269(1507), pp.2345–2350.

Langford, D.J. et al., 2010. Coding of facial expressions of pain in the laboratory mouse. *Nature Methods*, 7(6), pp.447–449.

Layne, J.N., 1972. Tail Autotomy in the Florida Mouse, Peromyscus floridanus. *Journal of Mammalogy*, 53(1), p.62.

Leach, M. & Main, D., 2008. An assessment of laboratory mouse welfare in UK animal units. *Animal Welfare*, 17(2), pp.171–187.

Lepicard, E. et al., 2006. Spatio-temporal analysis of locomotion in BALB/cByJ and C57BL/6J mice in different environmental conditions. *Behavioural Brain*

Research, 167(2), pp.365–372.

Lepicard, E.M. et al., 2000. Balance control and posture differences in the anxious BALB/cByJ mice compared to the non anxious C57BL/6J mice. *Behavioural brain research*, 117(1–2), pp.185–95.

Lewis, S.R. et al., 2005. Inbred mouse strain survey of sucrose intake. *Physiology & Behavior*, 85(5), pp.546–556.

Lyon, L. et al., 2011. Fractionation of spatial memory in GRM2/3 (mGlu2/mGlu3) double knockout mice reveals a role for group II metabotropic glutamate receptors at the interface between arousal and cognition. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, 36(13), pp.2616–28.

MacLeod, A.K. & Byrne, A., 1996. Anxiety, depression, and the anticipation of future positive and negative experiences. *Journal of abnormal psychology*, 105(2), pp.286–9.

Magariños, A.M., Orchinik, M. & McEwen, B.S., 1998. Morphological changes in the hippocampal CA3 region induced by non-invasive glucocorticoid administration: a paradox. *Brain research*, 809(2), pp.314–8.

Malakoff, D., 2000. The rise of the mouse, biomedicine's model mammal. *Science (New York, N.Y.)*, 288(5464), pp.248–53.

Malki, K. et al., 2015. Pervasive and opposing effects of Unpredictable Chronic Mild Stress (UCMS) on hippocampal gene expression in BALB/cJ and C57BL/6J mouse strains. *BMC genomics*, 16(1), p.262.

Malkki, H.A.I. et al., 2010. Appetitive operant conditioning in mice: heritability and dissociability of training stages. *Frontiers in behavioral neuroscience*, 4, p.171.

Mason, G.J. & Latham, N.R., 2004. Can't stop, won't stop: is stereotypy a reliable animal welfare indicator? *Animal Welfare*, 13, pp.57–69.

Mathews, A. & Macleod, C., 2002. Induced processing biases have causal effects on anxiety. *Cognition and Emotion*, 16(3), pp.331–354.

Mathews, A. & MacLeod, C., 1994. Cognitive Approaches to Emotion and Emotional

Disorders. *Annual Review of Psychology*, 45(1), pp.25–50.

Maxwell, F.R. et al., 1976. Positive and negative successive contrast effects following multiple shifts in reward magnitude under high drive and immediate reinforcement. *Animal Learning & Behavior*, 4(4), pp.480–484.

McManus, S. et al., 2016. *Adult Psychiatric Morbidity Survey: Survey of Mental Health and Wellbeing, England, 2014*,

McNamara, G.I. et al., 2016. Behavioural abnormalities in a novel mouse model for Silver Russell Syndrome. *Human molecular genetics*, 25(24), pp.5407–5417.

Mekiri, M. et al., 2017. Chronic corticosterone administration effects on behavioral emotionality in female c57bl6 mice. *Experimental and Clinical Psychopharmacology*, 25(2), pp.94–104.

Mendl, M. et al., 2009. Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Applied Animal Behaviour Science*, 118(3–4), pp.161–181.

Mendl, M., Brooks, J., et al., 2010. Dogs showing separation-related behaviour exhibit a 'pessimistic' cognitive bias. *Current Biology*, 20(19), pp.R839–R840.

Mendl, M., Burman, O.H.P. & Paul, E.S., 2010. An integrative and functional framework for the study of animal emotion and mood. *Proceedings. Biological sciences / The Royal Society*, 277(1696), pp.2895–2904.

Mendl, M. & Paul, E., 2004. Consciousness, emotion and animal welfare: insights from cognitive science. *Animal Welfare*, 13(1), pp.17–25.

MentalHealth.gov, 2017. What Is Mental Health? | MentalHealth.gov. Available at: http://www.mentalhealth.gov/basics/what-is-mental-health [Accessed August 13, 2018].

Millard, W.J., And, I. & Dole, V.P., 1983. *BRIEF COMMUNICATION Intake of Water and Ethanol by C57BL Mice: Effect of an Altered Light-Dark Schedule*,

Mineka, S. & Sutton, S.K., 1992. Cognitive Biases and the Emotional Disorders. *Psychological Science*, 3(1), pp.65–69.

Mineka, S., Watson, D. & Clark, L.A., 1998. Cormorbidity of anxiety and unipolar

mood disorders. *Annual Review of Psychology*, 49(1), pp.377–412.

Mineur, Y.S., Belzung, C. & Crusio, W.E., 2006. Effects of unpredictable chronic mild stress on anxiety and depression-like behavior in mice. *Behavioural Brain Research*, 175(1), pp.43–50.

Mitchell, E.N. et al., 2012. Evaluation of an operant successive negative contrast task as a method to study affective state in rodents. *Behavioural Brain Research*, 234(2), pp.155–160.

Moberg, G.P., 2000. Biological response to stress: implications for animal welfare. In *The biology of animal stress: basic principles and implications for animal welfare.* Wallingford: CABI, pp. 1–21.

Monk, J.E. et al., 2018. Towards a more practical attention bias test to assess affective state in sheep. *PloS one*, 13(1), p.e0190404.

Monleon-Paolo, S. et al., 1995. Attenuation of sucrose consumption in mice by chronic mild stress and its restoration by imipramine. *Psychopharmacology*, 117, pp.453–457.

Monleon, S. et al., 1995. Attenuation of sucrose consumption in mice by chronic mild stress and its restoration by imipramine. *Psychopharmacology*, 117(4), pp.453–457.

Montgomery, K.C., 1955. The relation between fear induced by novel stimulation and exploratory behavior. *Journal of comparative and physiological psychology*, 48(4), pp.254–60.

Morgan, K.N. & Tromborg, C.T., 2006. Sources of stress in captivity. *Applied Animal Behaviour Science*, 102(3–4), pp.262–302.

Morley-Fletcher, S. et al., 2003. Environmental enrichment during adolescence reverses the effects of prenatal stress on play behaviour and HPA axis reactivity in rats. *The European journal of neuroscience*, 18(12), pp.3367–74.

Morton, D. et al., 1993. Removal of blood from laboratory mammals and birds: First Report of the BVA/FRAME/RSPCA/UFAW Joint Working Group on Refinement. *Laboratory Animals*, 27(1), pp.1–22.

Mouse Genome Sequencing Consortium et al., 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature*, 420(6915), pp.520–562.

Moy, S. et al., 2007. Mouse behavioral tasks relevant to autism: Phenotypes of 10 inbred strains. *Behavioural Brain Research*, 176(1), pp.4–20.

Murray, F., Smith, D.W. & Hutson, P.H., 2008. Chronic low dose corticosterone exposure decreased hippocampal cell proliferation, volume and induced anxiety and depression like behaviours in mice. *European Journal of Pharmacology*, 583(1), pp.115–127.

Muscat, R., Papp, M. & Willner, P., 1992. Reversal of stress-induced anhedonia by the atypical antidepressants, fluoxetine and maprotiline. *Psychopharmacology*, 109(4), pp.433–438.

Muscat, R. & Willner, P., 1992. Suppression of sucrose drinking by chronic mild unpredictable stress: A methodological analysis. *Neuroscience & Biobehavioral Reviews*, 16(4), pp.507–517.

Mustaca, A.E., Bentosela, M. & Papini, M.R., 2000. Consummatory Successive Negative Contrast in Mice. *Learning and Motivation*, 31, pp.272–282.

Nacher, J., Gomez-Climent, M.A. & McEwen, B., 2004. Chronic non-invasive glucocorticoid administration decreases polysialylated neural cell adhesion molecule expression in the adult rat dentate gyrus. *Neuroscience letters*, 370(1), pp.40–4.

National Institutes of Health, 2011. *Guide for the Care and Use of laboratory animals Eighth Edition Committee for the Update of the Guide for the Care and Use of Laboratory Animals Institute for Laboratory Animal Research Division on Earth and Life Studies*,

Nemeroff, C.B. et al., 1992. Adrenal gland enlargement in major depression. A computed tomographic study. *Archives of general psychiatry*, 49(5), pp.384–7.

Nestler, E.J. & Hyman, S.E., 2010. Animal models of neuropsychiatric disorders. *Nature neuroscience*, 13(10), pp.1161–9.

Nettle, D. & Bateson, M., 2012. The evolutionary origins of mood and its disorders.

Current Biology, 22(17), pp.R712–R721.

Neville, V. et al., 2017. Dissociating the effects of alternative early-life feeding schedules on the development of adult depression-like phenotypes. *Scientific Reports*, 7(1), p.14832.

Nygren, T.E. et al., 1996. The Influence of Positive Affect on the Decision Rule in Risk Situations: Focus on Outcome (and Especially Avoidance of Loss) Rather Than Probability. *Organisational behaviour and human decision processes*, 66(1), pp.59–72.

Ortiz, R. et al., 1985. Post-weaning crowding induces corticoadrenal hyperreactivity in male mice. *Physiology & behavior*, 34(6), pp.857–60.

Panksepp, J. & Burgdorf, J., 2000. 50-kHz chirping (laughter?) in response to conditioned and unconditioned tickle-induced reward in rats: effects of social housing and genetic variables. *Behavioural Brain Research*, 115(1), pp.25–38.

Panlilio, L. V & Goldberg, S.R., 2007. Self-administration of drugs in animals and humans as a model and an investigative tool. *Addiction (Abingdon, England)*, 102(12), pp.1863–70.

Papini, M.R. et al., 2001. Apparent Incentive Contrast Effects in Autoshaping with Rats. *Learning and Motivation*, 32, pp.434–456.

Papp, M., Willner, P. & Muscat, R., 1991. An animal model of anhedonia: attenuation of sucrose consumption and place preference conditioning by chronic unpredictable mild stress. *Psychopharmacology*, 104(2), pp.255–259.

Parfitt, D. et al., 2007. Early life stress effects on adult stress-induced corticosterone secretion and anxiety-like behavior in the C57BL/6 mouse are not as robust as initially thought. *Hormones and Behavior*, 52(4), pp.417–426.

Paul, E.S., Harding, E.J. & Mendl, M., 2005. Measuring emotional processes in animals: the utility of a cognitive approach. *Neurosci Biobehav Rev*, 29(3), pp.469–491.

Peciña, S., Smith, K.S. & Berridge, K.C., 2006. Hedonic Hot Spots in the Brain. *The Neuroscientist*, 12(6), pp.500–511.

Peeters, F., Nicolson, N.A. & Berkhof, J., 2004. Levels and variability of daily life cortisol secretion in major depression. *Psychiatry Research*, 126(1), pp.1–13.

Pelchat, M.L. et al., 1983. Quality of acquired responses to tastes by Rattus norvegicus depends on type of associated discomfort. *Journal of Comparative Psychology*, 97(2), pp.140–153.

Pellow, S. et al., 1985. Validation of open:closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *Journal of neuroscience methods*, 14(3), pp.149–67.

Pellow, S. & File, S.E., 1986. Anxiolytic and anxiogenic drug effects on exploratory activity in an elevated plus-maze: a novel test of anxiety in the rat. *Pharmacology, biochemistry, and behavior*, 24(3), pp.525–9.

Perlman, R.L., 2016. Mouse models of human disease: An evolutionary perspective. *Evolution, medicine, and public health*, 2016(1), pp.170–6.

Perrin, S., 2014. Preclinical research: Make mouse studies work. *Nature*, 507(7493), pp.423–425.

Porsolt, R.D., Bertin, A. & Jalfre, M., 1977. Behavioral despair in mice: a primary screening test for antidepressants. *Archives internationales de pharmacodynamie et de therapie*, 229(2), pp.327–36.

Pothion, S. et al., 2004. Strain differences in sucrose preference and in the consequences of unpredictable chronic mild stress. *Behavioural brain research*, 155(1), pp.135–46.

Prus, A.J., James, J.R. & Rosecrans, J.A., 2009. *Conditioned Place Preference*, CRC Press/Taylor & Francis.

Quirce, C.M. & Maickel, R.P., 1981. Alterations of biochemical parameters by acute and repetitive stress situations in mice. *Psychoneuroendocrinology*, 6(1), pp.91–7.

Rainer, Q. et al., 2012. Beneficial behavioural and neurogenic effects of agomelatine in a model of depression/anxiety. *The International Journal of Neuropsychopharmacology*, 15(03), pp.321–335.

Ramos, A. & Mormède, P., 1998. Stress and emotionality: a multidimensional and genetic approach. *Neuroscience and biobehavioral reviews*, 22(1), pp.33–57.

Reeder, D.M. & Kramer, K.M., 2005. Stress in free-ranging mammals: integrating physiology, ecology, and naturl history. *Journal of Mammalogy*, 86(2), pp.225–235.

Ribot, T., 1897. The psychology of emotions. *Walter Scott Publishing Co*, 31.

Richter, C.P. & Campbell, K.H., 1940. Taste Thresholds and Taste Preferences of Rats for Five Common Sugars. *The Journal of Nutrition*, 20(1), pp.31–46.

Riemer, S. et al., 2016. A reappraisal of successive negative contrast in two populations of domestic dogs. *Animal Cognition*, 19(3), pp.471–481.

Riley, V., 1981. Psychoneuroendocrine influences on immunocompetence and neoplasia. *Science (New York, N.Y.)*, 212(4499), pp.1100–9.

Rodgers, R.J. & Dalvi, A., 1997. Anxiety, defence and the elevated plus-maze. *Neuroscience and biobehavioral reviews*, 21(6), pp.801–10.

Rosas, J.M. et al., 2007. Successive negative contrast effect in instrumental runway behaviour: A study with Roman high- (RHA) and Roman low- (RLA) avoidance rats. *Behavioural Brain Research*, 185(1), pp.1–8.

Roughan, J. V. et al., 2014. The Conditioned Place Preference Test for Assessing Welfare Consequences and Potential Refinements in a Mouse Bladder Cancer Model B. Taylor, ed. *PLoS ONE*, 9(8), p.e103362.

Rubin, R.T. et al., 1995. Adrenal gland volume in major depression. Increase during the depressive episode and decrease with successful treatment. *Archives of general psychiatry*, 52(3), pp.213–8.

Rubin, R.T. et al., 1987. Neuroendocrine aspects of primary endogenous depression III. Cortisol secretion in relation to diagnosis and symptom patterns. *Psychological Medicine*, 17(03), p.609.

Rushen, J., 1991. *Commentary Problems associated with the interpretation of physiological data in the assessment of animal welfare*,

Rushen, J., 1986. Some problems with the physiological concept of stress. *Australian*

Veterinary Journal, 63(11), pp.359–361.

Russell, J.A., 2003. Core affect and the psychological construction of emotion. *Psychological Review*, 110(1), pp.145–172.

Russell, W. & Burch, R., 1959. The principles of humane experimental technique. *Universities Federation for Animal Welfare*.

Rygula, R. et al., 2005. Anhedonia and motivational deficits in rats: Impact of chronic social stress. *Behavioural Brain Research*, 162(1), pp.127–134.

Rygula, R., Pluta, H. & Popik, P., 2012. Laughing Rats Are Optimistic. *PLoS ONE*, 7(12).

Sankoorikal, G.M. V. et al., 2006. A Mouse Model System for Genetic Analysis of Sociability: C57BL/6J Versus BALB/cJ Inbred Mouse Strains. *Biological Psychiatry*, 59(5), pp.415–423.

Schmidt, M. V., Wang, X.-D. & Meijer, O.C., 2011. Early life stress paradigms in rodents: potential animal models of depression? *Psychopharmacology*, 214(1), pp.131–140.

Schweizer, M., Henniger, M. & Sillaber, I., 2009. Chronic Mild Stress (CMS) in Mice: Of Anhedonia, 'Anomalous Anxiolysis' and Activity B. Baune, ed. *PLoS ONE*, 4(1), p.e4326.

Shanks, N., Greek, R. & Greek, J., 2009. Are animal models predictive for humans? *Philosophy, ethics, and humanities in medicine : PEHM*, 4, p.2.

Shargal, E. et al., 2017. Ecological and histological aspects of tail loss in spiny mice (Rodentia: Muridae, Acomys) with a review of its occurrence in rodents.

Sharma, S., Hryhorczuk, C. & Fulton, S., 2012. Progressive-ratio responding for palatable high-fat and high-sugar food in mice. *Journal of visualized experiments : JoVE*, (63), p.e3754.

Shishkina, G.T. & Dygalo, N.N., 2017. The glucocorticoid hypothesis of depression: History and prospects. *Russian Journal of Genetics: Applied Research*, 7(1), pp.128–133.

da Silva Cordeiro, A.F. et al., 2013. Understanding Vocalization Might Help to Assess

Stressful Conditions in Piglets. *Animals : an open access journal from MDPI*, 3(3), pp.923–34.

Simon O'Brien, E. et al., 2011. Fluoxetine, desipramine, and the dual antidepressant milnacipran reduce alcohol self-administration and/or relapse in dependent rats. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, 36(7), pp.1518–30.

Simson, P.G. et al., 1986. Reversal of behavioral depression by an infusion of an alpha-2 adrenergic agonist into the locus coeruleus. *Neuropharmacology*, 25(4), pp.385–389.

Skórzewska, A. et al., 2014. The effect of chronic administration of corticosterone on anxiety- and depression-like behavior and the expression of GABA-A receptor alpha-2 subunits in brain structures of low- and high-anxiety rats. *Hormones and Behavior*, 65(1), pp.6–13.

Smith, G.P., 2001. John Davis and the meanings of licking. *Appetite*, 36(1), pp.84–92.

Sonino, N. et al., 1998. Clinical Correlates of Major Depression in Cushing's Disease. *Psychopathology*, 31(6), pp.302–306.

Sotocina, S.G. et al., 2011. The Rat Grimace Scale: A Partially Automated Method for Quantifying Pain in the Laboratory Rat via Facial Expressions. *Molecular Pain*, 7(1), pp.1744-8069-7–55.

Spector, A.C. et al., 1984. A Detailed Analysis of Sucrose Drinking in the Rat. *Physiology & Behavior*, 33, pp.127–136.

Spector, A.C. & St. John, S.J., 1998. Role of taste in the microstructure of quinine ingestion by rats. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 274(6), pp.R1687–R1703.

Spector, A.C., Klumpp, P.A. & Kaplan, J.M., 1998. Analytical issues in the evaluation of food deprivation and sucrose concentration effects on the microstructure of licking behavior in the rat. *Behav Neurosci*, 112(3), pp.678–694.

Spiteri, T. & Le Pape, G., 2000. *What is learned during place preference*

conditioning? A comparison of food-and morphine-induced reward,

Starcevic, A. et al., 2016. Glucocorticoid levels after exposure to predator odor and chronic psychosocial stress with dexamethasone application in rats. *The Kaohsiung Journal of Medical Sciences*, 32(5), pp.235–240.

Starkman, M.N., Schteingart, D.E. & Schork, M.A., 1981. *Depressed Mood and Other Psychiatric Manifestations of Cushing's Syndrome: Relationship to Hormone Levels*,

Sterner, E.Y., 2010. Behavioral and neurobiological consequences of prolonged glucocorticoid exposure in rats: Relevance to depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 34(5), pp.777–790.

Steru, L. et al., 1985. The tail suspension test: a new method for screening antidepressants in mice. *Psychopharmacology*, 85(3), pp.367–70.

Strekalova, T. et al., 2004. Stress-Induced Anhedonia in Mice is Associated with Deficits in Forced Swimming and Exploration. *Neuropsychopharmacology*, 29(11), pp.2007–2017.

Sturm, M. et al., 2015. Effect of chronic corticosterone application on depression-like behavior in C57BL/6N and C57BL/6J mice. *Genes, Brain and Behavior*, 14(3), pp.292–300.

Sun, N., Sha, L. & Xu, Q., 2015. Establishment of a chronic restraint stress model of depression with C57BL/6J mice. *Zhongguo yi xue ke xue yuan xue bao. Acta Academiae Medicinae Sinicae*, 37(1), pp.8–11.

Tanimoto, H., Heisenberg, M. & Gerber, B., 2004. Experimental psychology: Event timing turns punishment to reward. *Nature*, 430(7003), pp.983–983.

Tannenbaum, J. & Bennett, B.T., 2015. Russell and Burch's 3Rs then and now: the need for clarity in definition and purpose. *Journal of the American Association for Laboratory Animal Science : JAALAS*, 54(2), pp.120–32.

Tarantola, M. et al., 2016. Effects of abrupt housing changes on the welfare of Piedmontese cows. *Italian Journal of Animal Science*, 15(1), pp.103–109.

Taylor, K. et al., 2008. Estimates for worldwide laboratory animal use in 2005. *ATLA*

Alternatives to Laboratory Animals, 36(3), pp.327–342.

Touma, C., Palme, R. & Sachser, N., 2004. Analyzing corticosterone metabolites in fecal samples of mice: a noninvasive technique to monitor stress hormones. *Hormones and behavior*, 45(1), pp.10–22.

Tsigos, C. et al., 2000. *Stress, Endocrine Physiology and Pathophysiology*, MDText.com, Inc.

Tucker, D.M. & Luu, P., 2007. Neurophysiology of Motivated Learning: Adaptive Mechanisms Underlying Cognitive Bias in Depression. *Cognitive Therapy and Research*, 31(2), pp.189–209.

Urani, A., Chourbaji, S. & Gass, P., 2005. Mutant mouse models of depression: Candidate genes and current mouse lines. *Neuroscience and Biobehavioral Reviews*, 29(4–5), pp.805–828.

Vacca, G. & Phillips, A.G., 2005. Inhibition of Successive Positive Contrast in Rats Withdrawn from an Escalating Dose Schedule of D-amphetamine. *International Journal of Comparative Psychology*, 18(4), pp.298–306.

Vanderschuren, L., Niesink, R. & Van Ree, J., 1997. The Neurobiology of Social Play Behavior in Rats. *Pergamon Neuroscience and Biobehavioral Reviews*, 21(3), pp.309–326.

Vos, T. et al., 2015. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, 386(9995), pp.743–800.

Walf, A.A. & Frye, C.A., 2007. The use of the elevated plus maze as an assay of anxiety-related behavior in rodents. *Nature protocols*, 2(2), pp.322–8.

Walling, C.A. et al., 2004. Predator Inspection Behaviour in Three-Spined Sticklebacks (Gasterosteus aculeatus): Body Size, Local Predation Pressure and Cooperation. *Behavioral Ecology and Sociobiology*, 56, pp.164–170.

Warwick, Z.S. & Weingarten, H.P., 1996. Flavor-postingestive consequence associations incorporate the behaviorally opposing effects of positive

reinforcement and anticipated satiety: implications for interpreting two-bottle tests. *Physiology & behavior*, 60(3), pp.711–5.

Weary, D.M., 2014. What is Suffering in Animals? *Dilemmas in Animal Welfare*, 188.

Webster, A.B. & Fletcher, D.L., 2004. Assessment of the aversion of hens to different gas atmospheres using an approach-avoidance test. *Applied Animal Behaviour Science*, 88(3–4), pp.275–287.

Weng, L. et al., 2016. Apigenin reverses depression-like behavior induced by chronic corticosterone treatment in mice. *European Journal of Pharmacology*, 774, pp.50–54.

Wenzlaff, R.M. & Grozier, S.A., 1988. Depression and the magnification of failure. *Journal of Abnormal Psychology*, 97(1), pp.90–93.

Willner, P. et al., 1987. Reduction of sucrose preference by chronic unpredictable mild stress, and its restoration by a tricyclic antidepressant. *Psychopharmacology*, 93, pp.358–364.

Willner, P., 2017a. Reliability of the chronic mild stress model of depression: A user survey. *Neurobiology of stress*, 6, pp.68–77.

Willner, P., 2017b. The chronic mild stress (CMS) model of depression: History, evaluation and usage. *Neurobiology of Stress*, 6, pp.78–93.

Willner, P., 1997. Validity, reliability and utility of the chronic mild stress model of depression: a 10-year review and evaluation. *Psychopharmacology*, 134(4), pp.319–29.

Willner, P. & Healy, S., 1996. Decreased hedonic responsiveness during a brief depressive mood swing. *Journal of Affective Disorders*, 32(1), pp.13–20.

Willner, P., Muscat, R. & Papp, M., 1992. Chronic mild stress-induced anhedonia: A realistic animal model of depression. *Neuroscience & Biobehavioral Reviews*, 16(4), pp.525–534.

Wilson, J.J. et al., 2015. Spatial learning by mice in three dimensions. *Behavioural brain research*, 289, pp.125–32.

Wright, W.F. & Bower, G.H., 1992. *Mood Effects on Subjective Probability*

Assessment,

Wurbel, H., 2001. Ideal homes? Housing effects on rodent brain and behaviour. *TRENDS in Neurosciences*, 24(4).

Yalcin, I., Aksu, F. & Belzung, C., 2005. Effects of desipramine and tramadol in a chronic mild stress model in mice are altered by yohimbine but not by pindolol. *European journal of pharmacology*, 514(2–3), pp.165–74.

Yeates, J.W. & Main, D.C.J., 2008. Assessment of positive welfare: A review. *The Veterinary Journal*, 175(3), pp.293–300.

Yu, X.-B. et al., 2015. Knockdown of hippocampal cysteinyl leukotriene receptor 1 prevents depressive behavior and neuroinflammation induced by chronic mild stress in mice. *Psychopharmacology*.

Zhao, Y. et al., 2008. A mouse model of depression induced by repeated corticosterone injections. *European Journal of Pharmacology*, 581(1–2), pp.113–120.

Zhu, S. et al., 2014. Unpredictable chronic mild stress not chronic restraint stress induces depressive behaviours in mice. *NeuroReport*, 25(14), pp.1151–1155.

Živković, I. et al., 2005. The effects of chronic stress on thymus innervation in the adult rat. *Acta Histochemica*, 106(6), pp.449–458.

Zupana, M. et al., 2015. Assessing positive emotional states in dogs using heart rate and heart rate variability. *Physiology & Behavior*, pp.102–111.

Appendix A

SCIENTIFIC REPORTS

OPEN

Handling method alters the hedonic value of reward in laboratory mice

Jasmine M. Clarkson[1], Dominic M. Dwyer[2], Paul A. Flecknell[3], Matthew C. Leach[4] & Candy Rowe[1]

Received: 21 November 2017
Accepted: 23 January 2018
Published online: 05 February 2018

Mice are the most widely used model species for drug discovery and scientific research. Consequently, it is important to refine laboratory procedures and practices to ensure high standards of welfare and scientific data quality. Recent studies have identified that the standard practice of handling laboratory mice by their tails increases behaviours indicative of anxiety, which can be overcome by handling mice using a tunnel. However, despite clear negative effects on mice's behaviour, tunnel handling has yet to be widely implemented. In this study, we provide the first evidence that tail handling also reduces mice's responses to reward. Anhedonia is a core symptom of clinical depression, and is measured in rodents by assessing how they consume a sucrose solution: depressed mice consume less sucrose and the size of their licking bouts when drinking (their 'lick cluster sizes') also tend to be smaller. We found that tail handled mice showed more anhedonic responses in both measures compared to tunnel handled mice, indicative of a decreased responsiveness to reward and potentially a more depressive-like state. Our findings have significant implications for the welfare of laboratory mice as well as the design and interpretation of scientific studies, particularly those investigating or involving reward.

Laboratory environments can negatively impact on the behaviour, physiology, health and welfare of animals[1–4] and considerable effort is made to regulate and improve the welfare of animals used in research laboratories around the world[5,6]. Mice are the most widely used species in biomedical research globally; in 2016 they were used in 73% of all procedures in the UK alone[7]. Consequently, understanding the experiences of mice used in research is of significant importance in order to be able to provide evidence-based improvements to housing and husbandry that will bring welfare benefits to a large number of animals, and ensure that empirical findings are robust[3,8].

Much early work examined the housing in which mice are kept. This revealed that small cage sizes, lack of environmental enrichment, room temperatures and isolation can all negatively impact on mouse welfare, producing measurable changes in behaviour, physiology or affective state[9–12]. However, more recently, it has been proposed that the handling technique used by researchers and laboratory staff influences both the welfare of mice, and the data obtained from behavioural studies[13–16]. The standard practice of handling laboratory mice using their tails has been shown to increase anxiety compared to being handled with a tunnel or by cupping mice on the open hand[13–15]: tunnel handled mice spend more time voluntarily interacting with a handler, and show less anxiety-related behaviour in standardised behavioural tests of anxiety such as the elevated plus maze[13–15]. In addition, tail handling can reduce performance in cognitive tests, for example it has been shown to reduce the engagement of mice with a novel mouse odour in a habituation task, resulting in the impairment of the subsequent dishabituation test[15]. However, despite the evidence that tail handling can impair welfare and scientific data collection, it remains the main method used to handle mice and other more refined methods such as tunnel handling have yet to be widely implemented across research institutions.

Here, we extend previous work on handling methods to test if being handled by the tail or with a tunnel can affect the hedonic responses of mice towards a rewarding stimulus. Whilst previous studies investigating the effects of handling method on mouse welfare have measured the animals' behaviour towards aversive experiences or punishments (such as being picked up by a handler or being placed in a novel test environment[13–15]), measuring responses to positive experiences and rewards (hedonic responses) are also important for understanding the

[1]Centre for Behaviour and Evolution, Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, NE2 4HH, United Kingdom. [2]School of Psychology, Cardiff University, Cardiff, CF10 3AT, United Kingdom. [3]Comparative Biology Centre, Medical School, Newcastle University, Newcastle upon Tyne, NE2 4HH, United Kingdom. [4]School of Natural and Environmental Sciences, Agriculture Building, Newcastle University, Newcastle upon Tyne, NE1 7RU, United Kingdom. Correspondence and requests for materials should be addressed to J.M.C. (email: j.clarkson@ncl.ac.uk)

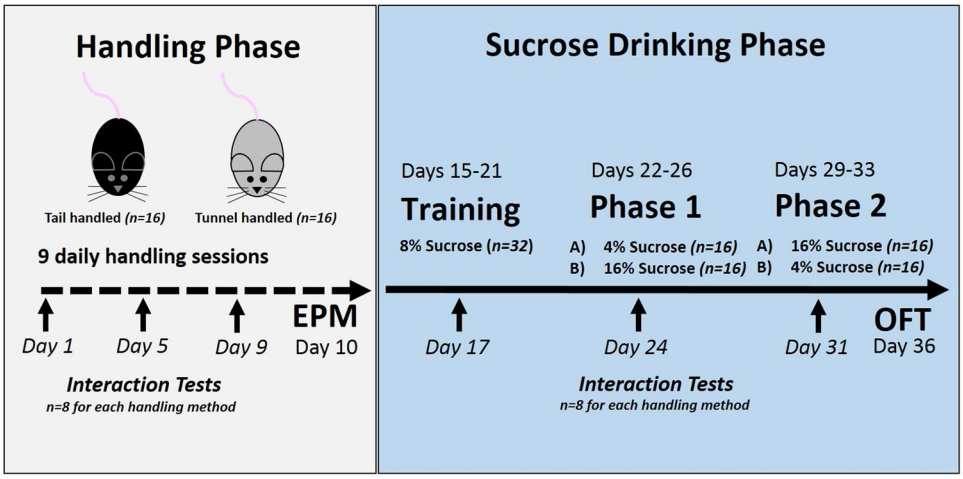


Figure 1. Study timeline to show the design and order of behavioural tests and sample sizes. EPM refers to Elevated Plus Maze and OFT refers to Open Field Test.

full impact of handling methods on the affective state of an animal[17–19]. How an animal responds to both punishment and reward offers a way of accessing their enduring negative affective states: whilst anxiety and depression can both be characterized by greater expectation of punishment, depression is also associated with a reduced expectation of reward[20]. Stressful life events have been implicated in the aetiology of depression in both humans and animals[21,22]. Specifically, rodent models have shown that exposure to either a single severe (acute) stressor, or several mild (chronic) stressful experiences are sufficient in inducing a depressogenic effect[23,24]. Therefore, we predicted that a threatening and stressful stimulus, such as tail handling, could also lead to measurable changes in mouse behavior indicative of a depressed-like state.

Whilst the main approach to studying animal welfare has been to measure the presence or absence of negative affective states, recent papers have highlighted the importance of measuring welfare from their positive experiences, such as pleasure[17,18,25]. Responses towards punishing and rewarding stimuli are dissociable, and the neural pathways that deal with punishment and reward are distinct[26,27]. Therefore, the aim of this study was to test if the handling method affects the capacity of mice to experience pleasure from reward.

The reduction or inability to experience pleasure from rewarding stimuli is known as anhedonia[28,29]. Anhedonia is a core symptom of Major Depressive Disorder (MDD) in humans[29,30], and consequently, assessing anhedonia in mice has been important for developing and validating laboratory models of depression[31]. Historically, hedonic state has been measured using voluntary consumption of sucrose solutions, under the assumption that anhedonia results in sucrose being perceived as less pleasant and consequently mice would be less motivated to drink it[22,32–34]. Animals that have undergone the chronic mild stress manipulation and show behavioural symptoms of depression, will drink less of a sweet solution than control animals[22,24,35,36]. This can be reversed through the application of anti-depressants, and has led to sucrose consumption being widely used as an indicator of affective state in rodents[32,36–40].

Despite its widespread application and use, the sucrose consumption test is only an indirect indicator of hedonic state. This is because sucrose consumption is influenced by a number of factors: whilst the amount of sucrose solution a mouse drinks may be driven in part by how much it likes the taste, it may also be affected by motivational factors[41] and the post-ingestive effects of the sucrose[42,43]. Therefore, researchers have sought alternative and more direct measures of hedonic responses towards tastants that are based on a more detailed examination of how an animal drinks. The orofacial movements produced upon tasting a solution and the pattern of licks during consumption are both considered to be more direct measures of palatability and hedonic responses to sucrose consumption[44–48].

Therefore, in this study, we measured the effect of handling method not only on the amount of sucrose drunk by mice, but also on a measure of their licking behaviour considered indicative of their hedonic response to reward. When rodents drink, the pattern of licks is not random[44,46–48]. Instead they produce rhythmic sets of licks that can be grouped into clusters[44,47,48]. The number of licks in these clusters, known as 'lick cluster size', is positively related to the palatability of the tastant: larger lick cluster sizes are elicited by more palatable solutions[44,46–48]. Moreover, lick cluster size is also affected by the experience and physiology of the animal. This has been demonstrated through manipulations such as conditioned taste aversion which directly devalues flavours (e.g. ref.[49]) and genetic or stress manipulations thought to reduce general hedonic tone (e.g. refs[44,50]) resulting in the reduction of lick cluster sizes elicited by otherwise palatable solutions. We predicted that if handling method affects the hedonic experience of mice from drinking sucrose, then tail handled mice would have lower consumption of, and smaller lick cluster sizes towards, sucrose solutions compared to tunnel handled mice.

Results

After habituation to the laboratory, mice underwent a 'handling phase' followed by a 'sucrose drinking phase' (for full details see Fig. 1). The handling phase aimed to manipulate their experiences and establish clear behavioural

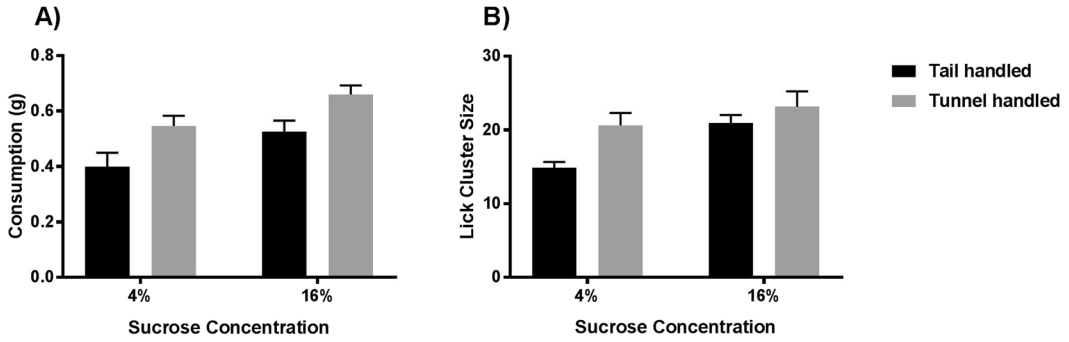


Figure 2. The drinking behaviour of tail and tunnel handled mice during the sucrose drinking phase. (**A**) Mean (±1 SEM) amount consumed (g). Both tail and tunnel handled mice drank more sucrose solution when given 16% compared to 4% solution ($F_{1,30} = 30.817$, $p < 0.001$), but at each concentration, tunnel handled mice drank significantly more sucrose than mice that were tail handled ($F_{1,30} = 7\,0.141$, $p = 0.012$). There was no interaction between handling method and sucrose concentration ($F_{1,30} = 0.100$, $p = 0.754$). (**B**) Mean (±1 SEM) Lick Cluster Size. Both tail and tunnel handled mice had larger lick cluster sizes for 16% sucrose compared to 4% sucrose ($F_{1,30} = 38.5$, $p < 0.001$) and the tunnel handled mice had significantly larger lick cluster sizes than tail handled mice overall (250 ms: $F_{1,30} = 4.6$, $p = 0.04$). However there was a significant interaction between handling method and sucrose concentration ($F_{1,30} = 10.2$, $p = 0.003$). Tunnel handled mice only had larger lick cluster sizes than tail handled mice at the lower (4%) sucrose concentration ($p = 0.003$) and not at the higher concentration ($p = 0.469$).

differences between tail and tunnel handled mice as previously reported[13,14], before investigating anhedonic behaviour in the sucrose drinking phase.

Consistent with previous studies[13,14], the repeated handling was sufficient to create significant differences between the two groups of mice in behavioural tests considered to be indicative of anxiety. We conducted 'voluntary interaction tests' on days 1, 5 and 9[13,14], both before and after mice were handled. Tunnel handled mice spent significantly more time interacting with the handler compared to tail handled mice (ANOVA: $F_{1,14} = 1062.7$, $p < 0.001$; Figure S1; Table S1), and also showed habituation to the handler (Bonferroni adjusted pairwise comparisons day 1 versus day 5 $p < 0.001$; day 1 versus day 9 $p < 0.001$), which was absent for those handled by their tails (day 1 versus day 5 $p > 0.99$; day 1 versus day 9 $p = 0.49$) (see Supplementary Information for full details; Figure S1; Table S1). Mice were also tested in an elevated plus maze on day 10 to assess their anxiety levels. Consistent with the expectation that tail handling produces higher levels of anxiety than tunnel handling, tail handled mice showed fewer entries onto the open arms (Mann Whitney $U = 174.5$, $p = 0.002$; Figure S2A), and spent less time on them (Mann Whitney $U = 175$, $p = 0.002$; Figure S2B) (see Supplementary Information for full details; Figure S2). Taken together, these tests demonstrated that our handling manipulation was successful in eliciting differences in anxiety-like behaviours towards potentially threatening events.

Mice then entered the sucrose drinking phase, where they received daily drinking trials in custom-built test chambers. Whilst both groups of mice drank more sucrose at the higher concentration (ANOVA: $F_{1,30} = 30.82$, $p < 0.001$; Fig. 2A), tunnel handled mice drank significantly more of both sucrose solutions than mice that were tail handled (ANOVA: $F_{1,30} = 7.14$, $p = 0.012$; Fig. 2A). There was no interaction between handling method and sucrose concentration on consumption (ANOVA: $F_{1,30} = 0.1$, $p = 0.754$; Fig. 2A). These findings were qualitatively the same when controlling for the animals' body weights (see Supplementary Information; Figure S3B), which did not significantly differ between tail and tunnel handled mice (Figure S3A).

Lick cluster sizes were also larger when mice drank the higher sucrose concentration (ANOVA: $F_{1,30} = 38.50$, $p < 0.001$; Fig. 2B) and were handled with a tunnel (ANOVA: $F_{1,30} = 4.62$, $p = 0.04$; Fig. 2B), however, there was also a significant interaction between handling method and sucrose concentration (ANOVA: $F_{1,30} = 10.20$, $p = 0.003$; Fig. 2B). This was because tunnel handled mice only had significantly larger lick cluster sizes than the tail handled mice when drinking the 4% sucrose solution (Bonferroni pairwise comparisons; $p = 0.003$) but not the 16% sucrose solution ($p = 0.469$). It is possible that we only detected the effects of handling on hedonic responses at the lower concentration because of ceiling effects at the higher concentration.

During the sucrose drinking phase, we gave mice three further voluntary interaction tests to ensure that established effects of tail and tunnel handling on measures of anxiety were still evident and had not diminished during this phase. We found that tail handled mice continued to interact significantly less with the handler compared to tunnel handled mice (ANOVA: $F_{1,14} = 462.34$, $p < 0.001$; Figure S4; Table S2), although their time spent interacting with the handler did increase over these three tests (Bonferroni adjusted pairwise comparisons for days 24 and 31 relative to day 17, p values $p < 0.01$) (see Supplementary Information, Figure S4; Table S2). At the end of the sucrose drinking phase, we also conducted an open field test as an independent measure of anxiety in both groups of mice. We found that tail handled mice showed significantly higher levels of anxiety, spending significantly less time in the centre of the open field compared to tunnel handled mice (Unpaired t-test: $t_{28} = 3.291$, $p = 0.003$; Figure S5B) (see Supplementary Information, Figure S5).

Data	Dependent Variable	Statistical Test	Factors	Experimental Unit	Sample Size
Voluntary interaction tests	Percentage time spent interacting	Repeated measures ANOVA	Between subject factors: Handling method (2 levels) Within-subject factors: day (6 levels), time (2 levels: pre or post handling)	Cage	n = 8 tail handled n = 8 tunnel handled
Elevated plus maze	Number of open arm entries; Duration on open arms	Mann-Whitney U test	Handling method (2 levels)	Mouse	n = 14 tail handled n = 15 tunnel handled
	Number of protected stretch attend postures	Unpaired t-test	Handling method (2 levels)	Mouse	n = 14 tail handled n = 15 tunnel handled
	Number of mice that defecated	Binary logistic regression	Handling method (2 levels)	Mouse	n = 14 tail handled n = 15 tunnel handled
Sucrose drinking	Consumption (g); Lick cluster size (log transformed)	Repeated measures ANOVA	Between subject Factors: Handling method (2 levels) Within-subject Factors: Sucrose concentration (2 levels)	Mouse	n = 16 tail handled n = 16 tunnel handled
Open field test	Duration of movement; Duration in centre; Crosses to centre; Distance travelled; Mean velocity	Unpaired t tests	Handling method (2 levels)	Mouse	n = 14 tail handled n = 16 tunnel handled
	Number of mice that defecated	Binary logistic regression	Handling method (2 levels)	Mouse	n = 16 tail handled n = 16 tunnel handled

Table 1. Statistical tests for each data set with respective factors, experimental unit and sample size.

Discussion

Our study provides the first evidence that handling method affects how laboratory mice perceive and respond to positive rewarding stimuli. Tail handling not only makes mice more anxious compared to tunnel handled mice[13–15], but it also reduces their hedonic responses towards a sucrose reward. Our data show that tail handled mice drank less sucrose at both concentrations and had lower lick cluster sizes overall, although smaller cluster sizes were only evident at the lower concentration (i.e. 4%). This result could be a ceiling effect due to testing under mild water restriction, and further work would be needed in order to determine if the anhedonic effects of tail handling are moderated by the nature of the solution being consumed. However, taken together, our combined data indicate that tail handling makes mice more anhedonic and less responsive to reward compared to being handled using a tunnel. Since tail handling is the most widely used method to handle laboratory mice[51,52], this finding has significant implications for animal welfare and the refinement of current laboratory practices, as well as scientific data collection, particularly where protocols include or investigate reward.

The presence of anhedonia in our tail handled mice, combined with increased anxiety-like behaviours relative to tunnel handled mice, is indicative of a more depressive-like state, thus suggestive of a more negative affective state in tail handled mice. This pattern of results is similar to treatments which have been explicitly designed to induce depressive-like states in rodents such as the chronic mild stress paradigm[24,35,53] or chronic restraint[54,55]. It may therefore be surprising that tail handling could also produce results similar to these more severe experimental manipulations. However, since tail handling may mimic a predatory attack[13,56,57], it could be that this handling method is inherently more stressful than currently thought[51]. Indeed, the degree of difference in consumption for sucrose found in mouse models of depression compared to controls, are very similar to those seen between our tunnel handled and our tail handled mice. At 4% sucrose, the reduction in tail handled mice's consumption compared to that of tunnel handled mice was 27%. From published studies using chronic manipulations to produce models of depression in C57BL/6 mice, we have estimated the reduction in sucrose consumption relative to controls to be in the range of 33–57%[58,59]. Although these are only estimates, this does suggest that the mice could be subject to a similar depressive-like state following tail handling, whether by acute or chronic effect. This could be explored further pharmacologically by the application of anti-depressant compounds.

In addition to detecting effects of tail handling towards positive reward, we also confirmed Hurst and colleagues' previous work and replicated their findings at a different research institution. We also found that tail handled mice interacted less with the handler and showed greater levels of anxiety in behavioural tests compared to tunnel handled mice[13–15]. However, given that our study was longer than previous published work (days 17, 24 and 31 compared to just days 1, 5 and 9), we were also able to explore the effect of tail handling on mouse behaviour over a longer period of time. The results from our open field test carried out at the end of the experiment still showed a significant effect of handling method on behavioural measures of anxiety, demonstrating that there were still clear behavioural differences. Tail handled mice continued to spend significantly less time interacting with the handler than tunnel handled mice across the entire experiment, although we did detect some changes in their behaviour during the sucrose drinking phase (see supplementary Figure S4). Tail handled mice increased the time they spent interacting with the handler over the last three voluntary interaction tests, and spent more time interacting with the handler once they had been handled. Although we are unable to draw any firm conclusions from this, there is the possibility that it may be due to forming an association with the sucrose reward, or alternatively perhaps the mice were able to learn about the sequence of events in the repeated voluntary interaction tests. Perhaps mice were more apprehensive of the handler when they were about to be picked up by their tail, compared to when they had already been picked up by their tail because they had learnt that once they had been handled, they were unlikely to be handled again.

Taken together, our data clearly demonstrate that mice are more anxious and more anhedonic when they are handled by their tail rather than by using a tunnel. This finding adds to the increasing number of studies that show that tail handling is an aversive procedure[13–16], and that tail handling has a negative impact on the welfare of laboratory mice. It also shows that sucrose consumption and licking behaviour can be used to assess the presence of positive experiences by measuring hedonic responses towards reward. Recent papers have highlighted that to fully understand animal welfare, we need to measure positive experiences, such as pleasure[17,18,25]. The ability to assess how an animal responds to rewarding sucrose solutions offers the potential to evaluate the effects of routine laboratory conditions on their affective state and welfare.

Tail handling has been shown to affect data collection through reducing the likelihood that a mouse will engage with a cognitive task[15]. Our data suggest that the effects of tail handling may be more complex than simply not engaging with a task, but could affect how animals respond to rewards in behavioural and cognitive tasks. The vast majority of *in vivo* work involving behavioural paradigms rely on the use of reward to train the animal to perform in the given task, for example, condensed milk is often used to train mice in spatial memory tasks[60,61] and sucrose pellets are used in operant conditioning tasks[62,63]. If tail handling reduces mice's sensitivity for reward, this may result in longer training periods or reduced effect sizes leading to larger sample sizes being required. Tail handling may also be negatively affecting the neural circuitry underlying reward, which may mean that studies of reward pathways may not be using reliable or accurate models. We recommend that researchers consider the potential effect of tail handling on their results and interpretation of their findings.

In conclusion, we would strongly advocate that, wherever possible, mice should be handled using a tunnel and not by their tails. Tunnel handling is a simple yet effective refinement that has the potential to not only significantly improve animal welfare but also scientific data quality. Based on our own findings and those of others[13–16], we recommend that research institutions should seek to introduce and widely implement tunnel handling as a refinement to their husbandry procedures, and that published protocols for handling mice are revised[51].

Methods and Materials

Ethical Statement. Experiments were conducted at Newcastle University following approval from the universities Animal Welfare and Ethical Review Body (AWERB Project ID: 540), and in accordance with the EU Directive (2010/63/EU), ASPA (1986) and the NIH Guidelines for care and use of animals for experimental procedures. All animals were checked daily, and no adverse effects were reported. At the end of the experiment, animals were humanely killed via exposure to a rising concentration of carbon dioxide gas in accordance to Schedule 1 guidance.

Animals, housing and husbandry. Thirty-two male C57BL/6 J mice were purchased from Charles River Laboratories, UK and were approximately 7 weeks of age (Mean ± SEM: 24.6 ± 1.6 g) on arrival. Mice were free from all recognised pathogens, and the health status of the colony was monitored following the FELASA health monitoring recommendations[64]. Mice were pair-housed in M2 cages (33 cm (L) × 15 cm (W) × 13 cm (H), North Kent Plastics), with sawdust bedding, nesting material (4HK Aspen chips, NestPak and Sizzlepet nesting, Datesand Ltd, Manchester) and a clear perspex home cage tunnel (50 mm diameter, 150 mm length) and were cleaned once per week. Animals had access to food (Special Diet Services, RM3E diet) and water *ad libitum*, except prior to drinking experiments (described below). Mice were maintained on a reverse 12:12 hour light/dark cycle (10:00 until 22:00) and experiments were conducted under red light illumination. They received relatively constant temperature (Mean ± SEM: 20.8 ± 0.7 °C) and relative humidity (Mean ± SEM: 29.3 ± 6.5%). In line with previous studies, mice were marked for identification using hair dye (Jerome Russel B. Blonde, UK) which does not interfere with the response to handling[13,14].

Handling Methods. After habituation to the laboratory (the animals were not handled during this time), each cage of two mice were randomly assigned (via random number generator) to one of two treatment groups, tail or tunnel handled (n = 16 per group). Mice were then only handled by their designated method (tail or tunnel handled) by the same handler wearing nitrile gloves, which were were rubbed in soiling bedding before each handling session (from mice of the same sex and strain) and a laboratory coat that was contaminated with mouse scent[13,14]. Tail handling involved grasping a mouse at the base of its tail using the thumb and forefinger, and then lifting onto the sleeve of the laboratory coat for 30 seconds before being returned to its home cage. For tunnel handling, the mouse was guided into the Perspex tunnel, and lifted above the cage and held for 30 seconds. For the first two days, the handler's hands were loosely cupped over the ends of the tunnel to prevent escape. Mice were handled twice daily for 30 seconds, 60 seconds apart, for the first nine days. Prior to handling, the nesting material (care was taken not to disrupt the structure) and home cage tunnel were removed. This procedure was also conducted once weekly to coincide with the drinking experiments (days 17, 24 and 31). For routine husbandry practices, such as cage cleaning, mice were captured and transferred using their designated handling method either on the sleeve for tail handled mice, or in the tunnel for tunnel handled mice. The same protocol was used when transferring mice to behavioural tests, i.e. the elevated plus maze, open field test and sucrose drinking chambers.

Voluntary interaction tests. During the handling phase (days 1, 5 and 9) and the sucrose drinking phase (days 17, 24 and 31), each cage of animals underwent 'voluntary interaction tests' to assess their responses to the handler. These tests allowed a comparison of behaviour in anticipation of being handled compared to after being handled on specified test days[13,14]. Each test consisted of removing the cage lid, nesting material and home cage tunnel and the handler standing motionless in front of the cage for 60 seconds. A gloved hand (tail handled) or a gloved hand holding the home cage tunnel (tunnel handled) was held resting on the substrate in the front of the cage for 60 seconds to assess voluntary interaction. Each mouse in the cage was then handled twice for 30 seconds by their designated handling method described above, before voluntary interaction was assessed again. Behaviour

was filmed and later analysed using Observer XT (v11). Time spent interacting with the handler was measured for each mouse within a cage, from which an overall mean cage score was calculated. These were summed together for analyses for both tail and tunnel handled mice. Therefore, for these tests, the experimental unit was 'cage' (n = 8 for both groups). Interaction was defined as: sniffing (nose within 0.5 cm), touching (including paw contact), climbing on or in the handling tunnel and/or the handler's hand. Due to the differences in how mice in the two treatments were handled during the interaction tests, the observer could not be blind to the treatment, but was blind to whether the interaction test was carried out before or after handling.

Elevated Plus Maze. On day 10, mice underwent behavioural testing in an elevated plus maze (dimensions: arms 30 cm (L) × 5 cm (W) with side walls of 15 cm on the two closed arms, elevated 50 cm from the ground). Mice were delivered to the centre of the maze (via their designated handling method) facing an open arm and filmed for 5 mins and returned to a holding cage or the home cage, depending on whether it was the first or last mouse to undergo testing from its cage. Between subjects, the maze was cleaned with 70% ethanol and the running order was counterbalanced with respect to handling method across the testing day. Time spent in the open or closed arms was scored by a treatment blind observer using Observer XT (v11), where the time spent in an arm was defined as being when all four paws were in the arm. Three animals jumped off before the end of the test and were excluded from statistical analysis meaning the sample sizes were reduced (tail handled n = 14; tunnel handled n = 15).

Sucrose Drinking Phase. Mice were trained and tested in eight custom made drinking chambers. These were standard mice IVC home cages (34 (L) × 19 (W) × 14 (D) cm) with clear Perspex sides, a metal perforated floor and wire cage lid with modified attachments to connect the sipper tube to the left hand side of the cage. Solutions were delivered through drinking spouts attached to 50 ml falcon tubes. Drinking chambers were connected to contact sensitive Med Associates dual contact lickometers (Med Associates Inc., St. Albans, Vermont), which transmitted the time of each lick to the nearest 0.01 second to a computer using MED-PC software. Custom-built software calculated the lick cluster sizes according to a range of interbout intervals, which is the length of time used to determine when licks can be considered to be in a single bout[44,48,65,66]. The data presented here use an interbout interval of 250 ms, meaning that any duration of 250 ms or longer between two licks defined the end of one bout and the start of the next. However, the finding were robust to the interbout interval used (see Supplementary Information; Figure S6).

Mice were separated into four groups of eight (four mice from each treatment per group) referred to as the 'testing group'. A random number generator assigned mice into testing groups according to their cage number, meaning both mice within a cage were assigned to the same testing group and were tested in the same order and time each day. Water bottles on the home cage were removed 2 hours prior to sucrose drinking trials, before the lights went off to encourage consumption. Mice were trained across seven consecutive days for 15 minutes each day (Days 15–21) to drink sucrose (8% (w/w) sucrose solution) from the spouts. During the first three training sessions the spout was left to protrude into the cage to ensure engagement with the task, after this the spout was flush with the cage lid in order to reduce accidental contact. Once all animals were engaged with the task and consistently drinking (>100 licks), the sucrose drinking phase began.

Sucrose drinking testing had two phases, where mice were tested on all 5 days for each phase. Phase 1 (Days 22–26) consisted of half the animals (n = 16; 8 from each handling method) receiving 4% (w/w) sucrose and half (n = 16; 8 from each handling method) receiving 16% (w/w) sucrose for 15 minutes. This was balanced with regards to treatment group and across testing groups. Phase 2 (Days 29–33) reversed the sucrose concentration. We used two concentrations of sucrose to assess the responses to stimuli with differing hedonic properties. We measured the mass of sucrose solution consumed and the timing of each lick in every test trial; from this, we calculated the mean consumption of sucrose (g) and the mean lick cluster sizes for each animal across the five days at both concentrations for use in our analyses.

Open Field Test. On day 36 each mouse was individually placed via their designated handling method in the centre of a rectangular arena, (54.5 cm (L) × 35.5 cm (W) × 17 cm (H)) made of white plastic with a transparent Perspex lid for 10 minutes. The order was counterbalanced with respect to handling method. Behaviour was filmed and analysed using Ethovision XT (v 5.1) which automatically tracked the time spent in the centre, relative to the periphery. Presence of defecation during the open field test was noted and later analysed. Due to a technical error, videos were only scored for 14 out of the 16 tail handled mice.

Statistical Analyses. All statistical analyses were conducted using IBM Corp. SPSS (v23, SPSS Inc, Chicago, USA). Datasets were tested for normality and homogeneity of variance, where assumptions were not met data were transformed or non-parametric statistical methods used. Where significant main effects were found, Bonferroni post hoc tests were performed to look at pairwise comparisons between variables. Refer to Table 1 for full statistical analyses.

Data availability. All data generated or analysed during this study are freely available on the Zenodo repository: https://doi.org/10.5281/zenodo.1157907.

References

1. Podberscek, A. L., Blackshaw, J. K. & Beattie, A. W. The behaviour of laboratory colony cats and their reactions to a familiar and unfamiliar person. *Appl. Anim. Behav. Sci. Elsevier Sci. Publ. B.V* **31**, 119–130 (1991).
2. Gunn, D. & Morton, D. B. Inventory of the behaviour of New Zealand White rabbits in laboratory cages. *Appl. Anim. Behav. Sci.* **45**, 277–292 (1995).
3. Wurbel, H. Ideal homes? Housing effects on rodent brain and behaviour. *TRENDS Neurosci.* **24**, (2001).

4. Clarke, A. S., Juno, C. J. & Maple, T. L. Behavioral effects of a change in the physical environment: A pilot study of captive chimpanzees. *Zoo Biol.* **1**, 371–380 (1982).
5. Toth, L. A., Kregel, K., Leon, L. & Musch, T. I. Environmental enrichment of laboratory rodents: the answer depends on the question. *Comp. Med.* **61**, 314–21 (2011).
6. Newberry, R. C. Environmental enrichment: Increasing the biological relevance of captive environments. *Appl. Anim. Behav. Sci.* **44**, 229–243 (1995).
7. Home Office. *Statistics of Scientific Procedures on living animals, Great Britain 2016*. At https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/627284/annual-statistics-scientific-procedures-living-animals-2016.pdf (2017).
8. Benefiel, A. C. *et al*. Mandatory 'Enriched' Housing of Laboratory Animals: The Need for Evidence-based Evaluation. *ILAR J.* **46**, 95–105 (2005).
9. Chapillon, P., Manneché, C., Belzung, C. & Caston, J. Rearing Environmental Enrichment in Two Inbred Strains of Mice: 1. Effects on Emotional Reactivity. *Behav. Genet.* **29**, 41–46 (1999).
10. Chourbaji, S., Zacher, C., Sanchis-Segura, C., Spanagel, R. & Gass, P. Social and structural housing conditions influence the development of a depressive-like phenotype in the learned helplessness paradigm in male mice. *Behav. Brain Res.* **164**, 100–106 (2005).
11. Gaskill, B. N., Rohr, S. A., Pajor, E. A., Lucas, J. R. & Garner, J. P. Some like it hot: Mouse temperature preferences in laboratory housing. *Appl. Anim. Behav. Sci.* **116**, 279–285 (2009).
12. Ortiz, R., Armario, A., Castellanos, J. M. & Balasch, J. Post-weaning crowding induces corticoadrenal hyperreactivity in male mice. *Physiol. Behav.* **34**, 857–60 (1985).
13. Hurst, J. L. & West, R. S. Taming anxiety in laboratory mice. *Nature*, https://doi.org/10.1038/NMETH.1500 (2010).
14. Gouveia, K. *et al*. Reducing Mouse Anxiety during Handling: Effect of Experience with Handling Tunnels. *PLoS One* **8**, e66401 (2013).
15. Gouveia, K., Hurst, J. L., Bolker, B., Walker, S. & Wurbel, H. Optimising reliability of mouse performance in behavioural testing: the major role of non-aversive handling. *Sci. Rep.* **7**, 44999 (2017).
16. Ghosal, S. *et al*. Mouse handling limits the impact of stress on metabolic endpoints. *Physiol. Behav.* **150**, 31–7 (2015).
17. Boissy, A. *et al*. Assessment of positive emotions in animals to improve their welfare. *Physiol Behav* **92**, 375–397 (2007).
18. Yeates, J. W. & Main, D. C. J. Assessment of positive welfare: A review. *Vet. J.* **175**, 293–300 (2008).
19. Mendl, M., Burman, O. H. P. & Paul, E. S. An integrative and functional framework for the study of animal emotion and mood. *Proc. Biol. Sci.* **277**, 2895–2904 (2010).
20. Nettle, D. & Bateson, M. The evolutionary origins of mood and its disorders. *Curr. Biol.* **22**, R712–R721 (2012).
21. Hammen, C., Kim, E. Y., Eberhart, N. K. & Brennan, P. A. Chronic and acute stress and the prediction of major depression in women. *Depress. Anxiety* **26**, 718–723 (2009).
22. Papp, M., Willner, P. & Muscat, R. An animal model of anhedonia: attenuation of sucrose consumption and place preference conditioning by chronic unpredictable mild stress. *Psychopharmacology (Berl).* **104**, 255–259 (1991).
23. Simson, P. G., W~rss, J. M., Hoffman, L. J. & Ambrose, M. J. Reversal of behavioral depression by an infusion of an alpha-2 adrenergic agonist into the locus coeruleus. *Neuropharmacology* **25**, 385–389 (1986).
24. Willner, P., Muscat, R. & Papp, M. Chronic mild stress-induced anhedonia: A realistic animal model of depression. *Neurosci. Biobehav. Rev.* **16**, 525–534 (1992).
25. Fraser, D. & Duncan, I. 'Pleasures', 'Pains' and Animal Welfare: Toward a Natural History of Affect. *WEL* At, http://animalstudiesrepository.org/acwp_awap/37 (1998).
26. Guitart-Masip, M. *et al*. Go and no-go learning in reward and punishment: Interactions between affect and effect. *Neuroimage* **62**, 154–166 (2012).
27. McCabe, C., Woffindale, C., Harmer, C. J. & Cowen, P. J. Neural Processing of Reward and Punishment in Young People at Increased Familial Risk of Depression. *Biol. Psychiatry* **72**, 588–594 (2012).
28. Ribot T. The psychology of emotions. *Walter Scott Publ. Co* **31**, (1897).
29. Gorwood, P. Neurobiological mechanisms of anhedonia. *Dialogues Clin. Neurosci.* **10**, 291–9 (2008).
30. American Psychiatric Association. *Diagnostic and Statistical Manual of mental disorders fifth edition*. (2014).
31. Deussing, J. M. Animal models of depression. *Drug Discov. Today Dis. Model.* **3**, 375–383 (2006).
32. Forbes, N. F., Stewart, C. A., Matthews, K. & Reid, I. C. Chronic Mild Stress and Sucrose Consumption: Validity as a Model of Depression. *Physiol. Behav.* **60**, 1481–1484 (1996).
33. Willner, P. & Healy, S. Decreased hedonic responsiveness during a brief depressive mood swing. *J. Affect. Disord.* **32**, 13–20 (1994).
34. Willner, P., Towell, A., Sampson, D., Sophokleous, S. & Muscat, R. Psychopharmacology Reduction of sucrose preference by chronic unpredictable mild stress, and its restoration by a tricyclic antidepressant. *Psychopharmacology (Berl).* **93**, 358–364 (1987).
35. Muscat, R. & Willner, P. Suppression of sucrose drinking by chronic mild unpredictable stress: A methodological analysis. *Neurosci. Biobehav. Rev.* **16**, 507–517 (1992).
36. Willner, P. Validity, reliability and utility of the chronic mild stress model of depression: a 10-year review and evaluation. *Psychopharmacology (Berl).* **134**, 319–29 (1997).
37. Willner, P., Towell, A., Sampson, D., Sophokleous, S. & Muscat, R. Reduction of sucrose preference by chronic unpredictable mild stress, and its restoration by a tricyclic antidepressant. *Psychopharmacology (Berl).* **93**, (1987).
38. Monleon, S. *et al*. Attenuation of sucrose consumption in mice by chronic mild stress and its restoration by imipramine. *Psychopharmacology (Berl).* **117**, 453–457 (1995).
39. Muscat, R., Papp, M. & Willner, P. Reversal of stress-induced anhedonia by the atypical antidepressants, fluoxetine and maprotiline. *Psychopharmacology (Berl).* **109**, 433–438 (1992).
40. Willner, P. Reliability of the chronic mild stress model of depression: A user survey. *Neurobiol. Stress* **6**, 68–77 (2017).
41. Brennan, K., Roberts, D., Anisman, H. & Merali, Z. Individual differences in sucrose consumption in the rat: motivational and neurochemical correlates of hedonia. *Psychopharmacology (Berl).* **157**, 269–276 (2001).
42. Warwick, Z. S. & Weingarten, H. P. Flavor-postingestive consequence associations incorporate the behaviorally opposing effects of positive reinforcement and anticipated satiety: implications for interpreting two-bottle tests. *Physiol. Behav.* **60**, 711–5 (1996).
43. Booth, D. A., Lovett, D. & McSherry, G. M. Postingestive modulation of the sweetness preference gradient in the rat. *J. Comp. Physiol. Psychol.* **78**, 485–512 (1972).
44. Dwyer, D. M. EPS Prize Lecture. Licking and liking: the assessment of hedonic responses in rodents. *Q J Exp Psychol* **65**, 371–394 (2012).
45. Berridge, K. C. Measuring hedonic impact in animals and infants: Microstructure of affective taste reactivity patterns. *Neurosci. Biobehav. Rev.* **24**, 173–198 (2000).
46. Davis, J. D. The effectiveness of some sugars in stimulating licking behavior in the rat. *Physiol Behav* **11**, 39–45 (1973).
47. Davis, J. D. The Microstructure of Ingestive Behavior. *Ann. N. Y. Acad. Sci.* **575**, 106–121 (1989).
48. Davis, J. D. & Smith, G. P. Analysis of the Microstructure of the Rhythmic Tongue Movements of Rats Ingesting Maltose and Sucrose Solutions. *Behav Neurosci* **106**, 217–228 (1992).
49. Dwyer, D. M., Boakes, R. A. & Hayward, A. J. Reduced palatability in lithium- and activity-based, but not in amphetamine-based, taste aversion learning. *Behav. Neurosci.* **122**, 1051–1060 (2008).

50. McNamara, G. I., Davis, B. A., Dwyer, D. M., John, R. M. & Isles, A. R. Behavioural abnormalities in a novel mouse model for Silver Russell Syndrome. *Hum. Mol. Genet.* **25**, 5407–5417 (2016).
51. Deacon, R. M. J. Housing, husbandry and handling of rodents for behavioral experiments. *Nat. Protoc.* **1**, 936–946 (2006).
52. Leach, M. & Main, D. An assessment of laboratory mouse welfare in UK animal units. At, http://www.ingentaconnect.com/contentone/ufaw/aw/2008/00000017/00000002/art00009 (2008).
53. Willner, P. The chronic mild stress (CMS) model of depression: History, evaluation and usage. *Neurobiol. Stress* **6**, 78–93 (2017).
54. Chiba, S. *et al*. Chronic restraint stress causes anxiety- and depression-like behaviors, downregulates glucocorticoid receptor expression, and attenuates glutamate release induced by brain-derived neurotrophic factor in the prefrontal cortex. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **39**, 112–119 (2012).
55. Galea, L. A. *et al*. Sex differences in dendritic atrophy of CA3 pyramidal neurons in response to chronic restraint stress. *Neuroscience* **81**, 689–697 (1997).
56. Shargal, E., Rath-Wolfson, L., Kronfeld, N. & Dayan, T. Ecological and histological aspects of tail loss in spiny mice (Rodentia: Muridae, Acomys) with a review of its occurrence in rodents. At, https://www.cambridge.org/core/services/aop-cambridge-core/content/view/S0952836999010080 (2017).
57. Layne, J. N. Tail Autotomy in the Florida Mouse, Peromyscus floridanus. *J. Mammal.* **53**, 62 (1972).
58. Strekalova, T., Spanagel, R., Bartsch, D., Henn, F. A. & Gass, P. Stress-Induced Anhedonia in Mice is Associated with Deficits in Forced Swimming and Exploration. *Neuropsychopharmacology* **29**, 2007–2017 (2004).
59. Pothion, S., Bizot, J.-C., Trovero, F. & Belzung, C. Strain differences in sucrose preference and in the consequences of unpredictable chronic mild stress. *Behav. Brain Res.* **155**, 135–46 (2004).
60. Wilson, J. J. *et al*. Spatial learning by mice in three dimensions. *Behav. Brain Res.* **289**, 125–32 (2015).
61. Lyon, L. *et al*. Fractionation of spatial memory in GRM2/3 (mGlu2/mGlu3) double knockout mice reveals a role for group II metabotropic glutamate receptors at the interface between arousal and cognition. *Neuropsychopharmacology* **36**, 2616–28 (2011).
62. Sharma, S., Hryhorczuk, C. & Fulton, S. Progressive-ratio responding for palatable high-fat and high-sugar food in mice. *J. Vis. Exp.* e3754, https://doi.org/10.3791/3754 (2012).
63. Malkki, H. A. I. *et al*. Appetitive operant conditioning in mice: heritability and dissociability of training stages. *Front. Behav. Neurosci.* **4**, 171 (2010).
64. Guillen, J. FELASA guidelines and recommendations. *J. Am. Assoc. Lab. Anim. Sci.* **51**, 311–21 (2012).
65. Davis, J. D. & Perez, M. C. The acquired control of ingestive behavior in the rat by flavor-associated postingestional stimulation. *Physiol. Behav.* **54**, 1221–1226 (1993).
66. Dwyer, D. M., Gasalla, P. & Lopez, M. Nonreinforced flavor exposure attenuates the effects of conditioned taste aversion on both flavor consumption and cue palatability. *Learn Behav* **41**, 390–401 (2013).

Acknowledgements

We would like to thank Paul Clarkson for building the elevated plus maze, the staff of the Comparative Biology Centre (Newcastle University) for their excellent technical support and Jack Leslie for additional technical assistance. We would also like to thank Melissa Bateson for her comments on the manuscript. This work was supported by a BBSRC DTP Grant (BB/J014516/1).

Author Contributions

C.R., M.C.L. and P.A.F. conceived the idea and acquired funding for the study, J.M.C. and C.R. designed the study and protocols, J.M.C. carried out experiments and analysed the data, D.D. provided training and analysis software for J.M.C., J.M.C. and C.R. drafted the manuscript, and all authors commented on the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-20716-3.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.