The Welfare Significance of Inactivity in Captive Animals,
Using Mink as a Model

by
Rebecca Kathleen Meagher

A Thesis
presented to
The University of Guelph

In partial fulfilment of the requirements
for the degree of
Doctor of Philosophy
in
Animal and Poultry Science

Guelph, Ontario, Canada
© Rebecca K. Meagher, December, 2011
ABSTRACT

THE WELFARE SIGNIFICANCE OF INACTIVITY IN CAPTIVE ANIMALS, USING MINK AS A MODEL

Rebecca Kathleen Meagher
University of Guelph, 2011

Advisor: Professor Georgia Mason

Captive animals are sometimes very inactive, which can elicit concern for their welfare. However, inactivity is difficult to interpret in terms of welfare, since while some forms reflect chronic fear (hiding), apathy, or depression-like states, others reflect positive states (e.g. relaxation). This thesis aimed to determine whether high levels or particular sub-types of inactivity indicate poor welfare in fur-farmed mink (*Neovison vison*), and to identify the specific psychological states involved. These questions were addressed by studying individual differences within populations on three commercial farms, and comparing mink in standard, non-enriched cages to those in enriched cages. Two hypotheses were tested on farms: that the most inactive mink experience chronic stress, and that this would impair reproduction. Inactive females did have smaller litters, a difference that was not attributable simply to their greater body fat. However, there was no evidence of endocrine stress nor increased fear in “glove tests”, and their kits also grew more quickly. This suggests that inactive females do not experience more chronic stress than active females do. Tests of responsiveness to stimuli (measured in terms of contact and orientation) showed that, compared to mink in enriched cages, non-enriched mink were more responsive to all types of stimuli, especially neutral ones. This finding is inconsistent with the hypothesis that inactive individuals in these conditions are apathetic or depressed; instead, it supports the alternative hypothesis that non-enriched cages induce boredom. However, this boredom-like hyper-responsiveness did not co-vary with inactivity levels. Finally, non-enriched cages did not consistently elevate total inactivity. However, they did induce specific types: inactivity in the nest box, lying alert (vs. sleeping), and lying belly down rather than curled up were all more common than in enriched cages. Inactivity in the nest box may reflect hiding; it seemed linked to fearfulness in glove tests and to endocrine stress responses. In sum,
while non-enriched conditions induce poor welfare, they do not increase overall inactivity; furthermore, within populations, the welfare of highly inactive individuals is no more compromised than that of their more active counterparts. However, subtypes of inactivity provide more information about welfare than total inactivity.
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Georgia Mason, whose work led me to this field. Thank you for your faith in me, for supporting me and constantly challenging me, and for introducing me to mink. Thank you also to my advisory committee members, past and present: Ian Duncan, Derek Haley, Francesco Leri, and Suzanne Millman, for their advice and willingness to help. Special thanks go to Ian for all of his help and encouragement.

I am also very grateful to the farmers who agreed to work with us, Kirk Rankin and Ted and Lynn Parkinson, and to the staff at their farms, especially John, Jake and Brody: only one of the three of you actually helped directly with my research, but all of you definitely made my lunch breaks more enjoyable. Many thanks also go to Angelo Napolitano, Steve Bursian, Andrew Cohen-Barnhouse and the MSU Mink and Poultry Farm staff, who were unfailingly helpful, even if they thought the whole project was crazy. Bruce Hunter, too, kindly volunteered his time and expertise. I also received assistance in collecting and entering my data from a long list of people: Allison Bechard, Heather Smale, Kaitlin Bahlmann, Ryan Clarke, Deidre Jones and Bethany Vlaming.

I have benefitted a great deal from discussions with my lab-mates and all of the members of Behaviour Group. A few people, however, deserve special mention: María Díez León and Jamie Dallaire for their help and moral support while we were at MSU, Megan Jones for the same thing via a multitude of e-mail and Skype chats, and Ken Yau for prying the mink off my hand that one time.

Thank you finally to my friends and family for helping me stay (relatively) sane during this process. I am lucky enough to have parents who support me in everything I do, and even put up with their car smelling faintly of mink for a month or two. I would also like to thank my St. Joe’s girls specifically: you are awesome for many reasons, but the most relevant here is that you helped me to embrace nerdiness, which is a crucial step towards a PhD.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>vi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td><strong>Chapter One:</strong> Part 1. A review of the relationship between inactivity and welfare in captive animals</td>
<td>1</td>
</tr>
<tr>
<td><strong>Chapter One:</strong> Part 2. Project Outline</td>
<td>23</td>
</tr>
<tr>
<td><strong>Chapter Two:</strong> Who’s afraid of the big, bad glove? Assessing fear and its correlates in mink</td>
<td>24</td>
</tr>
<tr>
<td><strong>Chapter Three:</strong> Inactive mink have fewer but potentially higher quality offspring</td>
<td>56</td>
</tr>
<tr>
<td><strong>Chapter Four:</strong> Decreased litter size in inactive female mink (<em>Neovison vison</em>): mechanisms and implications for overall productivity</td>
<td>77</td>
</tr>
<tr>
<td><strong>Chapter Five:</strong> Psychological states induced by non-enriched housing and their relationship to inactivity</td>
<td>103</td>
</tr>
<tr>
<td><strong>Chapter Six:</strong> Effects of enrichment and loss of enrichment on inactivity</td>
<td>120</td>
</tr>
<tr>
<td><strong>Chapter Seven:</strong> General discussion</td>
<td>141</td>
</tr>
<tr>
<td><strong>References</strong></td>
<td>148</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1.1  Forms of inactivity and their characteristics  22
Table 2.1  Pairwise reliability comparisons for glove tests repeated once a day for five days  51
Table 2.2  Behaviour in validation tests: tests for unequal variances between temperament categories, as determined by minks’ immediate responses to the stimulus  51
Table 2.3  Comparison of results in the glove test vs. stick test applied to the same sample of mink  51
Table 3.1  Nest shape scale  73
Table 3.2  Dam inactivity as a predictor of kit weight and growth  73
Table 4.1  Ethogram  98
Table 4.2  Litter size path analysis results  98
Table 4.3  Kit weight path analysis results  98
Table 5.1  Responsiveness to different types of stimuli depending on psychological state  116
Table 5.2  Stimuli used, in order of presentation  116
Table 6.1  Ethogram  135
Table 6.2  Housing effects on inactivity and its subtypes in Experiment 1  136
Table 6.3  Summary of effects of enrichment  137
| Figure 2.1 | Population-level changes in temperament distribution over repeated tests (Experiment 1) | 52 |
| Figure 2.2 | Average total duration of contact with glove over 3 min test, grouped by immediate response | 53 |
| Figure 2.3 | Average total duration of ambivalence over 3 min test, grouped by immediate response | 53 |
| Figure 2.4 | Comparison of temperament distributions for stick and glove tests | 54 |
| Figure 2.5 | Population-level changes in temperament distribution over three tests (Experiment 3) | 54 |
| Figure 2.6 | Proportion of fearful vs. non-fearful mink that failed to reproduce | 55 |
| Figure 3.1 | Kit growth vs. maternal inactivity | 74 |
| Figure 3.2 | Latency to touch kits vs. inactivity | 75 |
| Figure 3.3 | Kit growth vs. retrieval time | 76 |
| Figure 4.1 | Partial residual plots of litter size vs. inactivity | 99 |
| Figure 4.2 | Kit weight vs. inactivity | 100 |
| Figure 4.3 | Litter size on PND 21 vs. body condition score | 101 |
| Figure 4.4 | Hypothesized path diagram for litter size, with non-significant paths removed for best fit | 102 |
| Figure 4.5 | Hypothesized path diagram for kit weight | 102 |
| Figure 5.1 | Latency to make contact with the stimulus, split by housing treatment and stimulus type | 117 |
| Figure 5.2 | Duration oriented to the stimulus, split by housing treatment and stimulus type | 118 |
| Figure 5.3 | Duration in contact with the stimulus, split by housing treatment and stimulus type | 118 |
| Figure 5.4 | Proportion of treats consumed by housing treatment | 119 |
| Figure 6.1 | Timeline of events in Experiment 2 |
| Figure 6.2 | Change in time spent lying alert after a change in housing treatment. |
| Figure 6.3 | Change in time spent inactive in the nest box after a change in housing treatment |
| Figure 6.4 | Change in lying time spent in each posture after a change in housing treatment. |
CHAPTER ONE

Part 1. A review of the relationship between inactivity and welfare in captive animals

Abstract

Inactivity is sometimes mentioned as a problem in captive animals, but often without an explicit rationale or any investigation of the subjective state associated with the behaviour. In some cases, humans might aim to decrease inactivity for reasons unrelated to the animals’ current welfare. Although inactivity is often adaptive, allowing energy conservation and reducing the risk of detection by predators, captive animals may have the opportunity to be so inactive that it is not only deemed “unnatural” but can also be detrimental to their physical health and reproductive success. From a practical perspective, these reproductive impairments are one of several reasons that high levels of inactivity tend to be undesirable to the animals’ human keepers. Neither of these perspectives offers direct information about whether very inactive animals are suffering, however. Interpreting inactivity in terms of animal welfare (how the animal feels) is much more challenging, for three reasons. First, inactivity is actually a heterogeneous category, ranging from typically ‘relaxed’ forms such as sleep to vigilant forms such as freezing in the presence of a potential threat, which are obviously associated with different affective states. When behaviour falls somewhere in between those two forms, the underlying affective state is often unclear, and it may not always be possible to infer without a priori knowledge of how the environment or situation being studied influences welfare. The second reason interpretation can be challenging is that some forms of inactivity have non-linear relationships with welfare. Finally, individuals and species sometimes differ qualitatively in the way they express a given subjective state, and in some cases this means that inactivity changes in opposite directions. For animal welfare researchers, it is therefore necessary to investigate sub-categories of inactivity that can be differentiated behaviourally for a given species and used as welfare indicators. For some forms of inactivity, it may also be necessary to categorize individuals according to ‘coping style’ or other behavioural traits, and look at alterations in inactivity within each group.
Introduction

Captive animals are often described as inactive, with the implicit (e.g. Broom, 1988) or explicit implication that this amount of inactivity is a problem (e.g. McPhee and Carlstead, 2010; Zanella et al., 1996). Too much inactivity can certainly be associated with negative affective states in humans: high levels are sometimes symptomatic of states such as physical illness and some forms of depression (Gold and Chrousos, 2002; Kelley et al., 2003; Roshanaei-Moghaddam et al., 2009), while imposed inactivity can lead to boredom (e.g. Berlyne, 1960; Heaman and Gupton, 1998). In Cowper’s words, “Absence of occupation is not rest / A mind quite vacant is a mind distress’d” (1997, p. 118). In other animals, however, the subjective states associated with high levels of inactivity are not well understood, in part because inactivity has rarely been the focus of behavioural studies. It is often considered simply a default state rather than a true “behaviour” (see e.g. Levitis et al., 2009; Lima et al., 2005). However, as will be discussed below, at least some forms of inactivity (e.g. lying or hiding) are motivated behaviours. Since motivation is likely to be closely linked with affective states (Fraser and Duncan, 1998), this suggests that inactivity can be connected to welfare.

The relationship between inactivity and animal welfare, however, is complex. Decreased welfare can lead to either an increase or a decrease in inactivity levels. This is due in part to the fact that inactivity is not a homogenous category of behaviour: there are many different forms, from a phenotypic perspective (see Table 1.1), and they are likely associated with very different affective states. Even when focusing on only a single form, such as sleep, it can be a challenge to determine what level is “normal” and what is too much or too little. The solution to this problem will depend on one’s perspective. The main focus of this paper will be animal welfare, by which I mean solely how the animal feels (Duncan, 1996; Spinka, 2006; cf. Kryiazakis and Tolkamp, 2011); however, levels of inactivity could also be considered too high for practical reasons, or due to concerns that they are “unnatural”. The significance of inactivity, both for the animals’ welfare and for their keepers, thus varies widely depending on context and even on characteristics of the animal. I will therefore briefly discuss how inactivity is assessed, why it evolved, its practical consequences in captivity, and how all of these factors relate to welfare.
concerns. I will not discuss forms of inactivity that are not likely to be influenced by affective states of interest to welfare researchers; such forms include basking and brooding.

**How is inactivity defined and measured?**

Although inactivity may seem fairly straightforward to define, there is some variation in exactly what the term encompasses. In some cases, the hypotheses under test relate to specific forms of inactivity, and so only those forms are assessed. However, in other cases, its operational definition seems to be a by-product of the methods used to assess activity. This usually occurs when inactivity is not the main focus of the study but is measured simply as a part of the time budget or to control for ‘activity’ levels because they influence the variable of interest. When it is studied in free-living wild animals, activity is usually assessed using radio-tracking or similar technology, and so any time not in locomotion is called inactivity. Some studies of captive animals similarly equate “activity” with locomotion, which is often the behaviour pattern of interest. Laboratory animal research, for example, often infers activity levels from proxy measures such as the number of entries into closed arms (and therefore locomotion) in the elevated plus maze, a test of anxiety (e.g. Louvart et al., 2005), while some agricultural studies use pedometers (e.g. O’Callaghan et al., 2003). Therefore, inactivity in those studies would include time spent stationary but performing purposeful movements such as grooming. However, in applied ethology research using video or live observation, any type of movement is typically also considered activity, even if the animal remains in one place; for example, kicking (e.g. Rushen et al., 2001) and eating (e.g. Burrell and Altman, 2006; Rochlitz et al., 1998) are not categorized as inactive. Thus, the most common definition of inactivity is being relatively motionless, and although it is rarely stated explicitly, this means no movement with an apparent function (e.g. grazing or chewing a bite of food) but would include other slight movements (e.g. turning the head or shifting positions). In some cases, the only form of inactivity assessed is resting behaviour; this typically means lying down, often assessed using automated devices that detect posture (e.g. Tucker et al., 2009). Finally, research on anticipatory behaviour sometimes discusses hyperactivity or decreases in activity during anticipation, but actually measures the number of behavioural transitions rather than total amount of time spent active or time spent
stationary (e.g. Van den Bos et al., 2003). Studies using any of these definitions of inactivity will be discussed here.

The natural and functional perspectives

All animals are, of course, inactive some of the time. One way of determining whether inactivity levels are too high in captive animals (i.e. abnormal and/or a welfare concern) is to compare their time budgets to those seen in the wild. For example, while lions in zoos spend most of their time inactive, so do their wild counterparts, sleeping approximately 20 hours per day (Altman et al., 2005); in contrast, inactivity in zoo-housed leopards (Panthera pardus) seems to exceed that observed in nature (68-90%: Mallapur and Chellam, 2002 vs. 48%: Clubb and Mason, 2007). Sloths are another interesting example: they are popularly believed to be extremely inactive, to the point that they were given a name synonymous with laziness, and therefore their high levels of inactivity in zoos do not attract much attention. However, early studies of sleep in sloths conducted under captive conditions may have overestimated how much of their time budget it occupied, since a recent study of wild sloths found much lower levels (9.6 h/day vs. 16 h/day reported in captivity; Rattenborg et al., 2008); thus, their “slothfulness” may be induced at least in part by captivity. Such comparisons are useful if promoting natural behaviour per se is important, as it can be for zoos (Newberry, 1995) or from certain ethical perspectives (see e.g. Fraser et al., 1997).

Comparisons to time budgets in free-living animals are also informative from a scientific perspective because their behaviour is shaped by natural selection and is therefore typically adaptive (see e.g. Alcock, 2001), at least in the environment in which it evolved. In other words, it increases lifetime reproductive success on average, often by increasing health and survival. Inactivity can be adaptive for two main reasons: remaining motionless conserves energy and reduces the risk of being detected by predators (Engel and Schmale, 1972; Hart, 1988; Rose 1981). Inactivity that occurs in long bouts and in relatively safe environments is likely to be primarily for energy conservation and is thus called “rest”, although it may also serve the second function of decreasing predation risk if safe sleep sites are available (see Lima et al., 2005). This includes deep and prolonged forms of rest such as hibernation and torpor. Sleep, another
subcategory of rest, is thought to have additional functions, although exactly what those are is still much debated and may differ between animals of different ages and taxa (see Siegel, 2008, 2009). Currently, non-REM sleep is hypothesized to be important for avoiding oxidative stress and facilitating neurogenesis, while REM sleep seems to play a role in brain development; REM sleep may also aid in maintenance of neurotransmitter systems (e.g. noradrenergic receptors) and, more speculatively, memory consolidation (Siegel, 2005).

Meanwhile, other forms of inactivity occur only when the animal is exposed to some type of threat. The evolution of these forms was presumably driven by reduced detection and therefore predation risk (reviewed by Nesse, 1999). This includes brief “freezing” responses (e.g. rodents: Blanchard and Blanchard, 1969; ungulates: Caro et al., 2004), and also tonic immobility or “death-feigning” (Gilman et al., 1950), which may be adaptive even after the animal has been captured because it can induce the predator to loosen its hold (Engel and Schmale, 1972; Sargeant and Eberhardt, 1975). Engel and Schmale (1972) described another form of inactivity that occurs in response to stress, which they called conservation-withdrawal. This is a broad category of stress-induced forms of inactivity that include decreased responsiveness to the environment (see Table 1.1), in contrast with freezing, which is associated with hypervigilance (Bracha, 2004). It is not specific to immediate risk of predation, but occurs any time environmental conditions are unfavourable (e.g. no resources available or too much stimulation), and can persist over a long period of time. As the name suggests, conservation-withdrawal is believed to be adaptive not only because it reduces predation risk but also because it allows the conservation of significant amounts of energy (Engel and Schmale, 1972).

Inactivity that is not threat-related and does not exceed species norms for resting behaviour under natural conditions can likely be assumed to be functional and perhaps even energetically necessary. Even higher levels might be adaptive when there is no superseding motivation such as the need to search for food. However, in captivity, these higher levels may become maladaptive (a term used here to mean that it either decreases evolutionary fitness or more loosely, has proximate costs to the animal, whether or not the ultimate fitness effects are known; cf. Alcock, 2001; Mills, 2003) because so little activity is demanded when food is freely available that animals can become obese (e.g. Johnson et al., 2004). This, in turn, can decrease
fitness because it often leads to reproductive impairments (e.g. Malmkvist and Palme, 2008; Mitchell et al., 2011) and poor health, which reduces longevity (reviewed by Byles, 2009; D’Eath, 2009; German et al., 2010). Stress-induced forms of inactivity, too, could reflect a failure to adapt to the captive environment if individuals continue to respond to stimuli that do not pose a real threat, such as the human caretakers or visitors. Not only would this be an indicator of chronic stress, which has deleterious consequences for fitness (see e.g. Bartolomucci, 2007; Hemsworth, 2003; Wingfield and Sapolsky, 2003), but it could theoretically even have fitness costs itself, since a strong motivation to hide might prevent the animal from performing other functional activities such as mating.

**The welfare perspective**

*Inactivity that reflects positive welfare states*

Animals are often inactive when they are in a familiar, safe environment where all of their immediate needs are met. Thus, higher levels of inactivity are interpreted as a sign of improved welfare when compared to animals exposed to a possible threat. For example, sheep exhibit less locomotion when socially housed or placed in a familiar setting than when in a novel environment or isolated (Cockram, 2004), while gorillas in zoos rest more when there are fewer visitors, and show signs of stress including stereotypic behaviour and aggression on days with high numbers of visitors (Wells, 2005). Meanwhile, a variety of species show increased activity, presumably reflecting a motivation to forage, when they are feed-restricted (e.g. calves: de Paula Vieira et al., 2008; mink: Bildsøe et al., 1991; Hansen and Møller, 2008). Satiation, which is the absence of the negative affective state called hunger (see e.g. D’Eath et al., 2009), is thus associated with increased inactivity. Having adequate space and freedom from competition can also be important in promoting increased lying (e.g. Fregonesi et al., 2007). In all of the above settings, inactivity indicates a lack of stress and/or negative affect.

Inactivity is also frequently used as an indicator of positive welfare when it is associated with greater comfort. In cows, lying is a highly motivated behaviour (Munksgaard et al., 2005), important for rumination, and sometimes inhibited in conventional housing systems by stall
design or hard flooring. Consequently, research into improved housing systems often tests for an increase in time spent lying (e.g. Phillips and Schofield, 1994; Tucker et al., 2009). Flooring types that increase lying time when cows are restricted to a given type also tend to be the types that are preferred when the cows are given a choice of stalls, confirming that the increased lying times likely correlate with improved welfare (Tucker et al., 2003). Pigs, too, show an increase in time spent lying down when given bedding than when in stalls with bare concrete floors (e.g. Fraser, 1975).

**Inactivity that reflects negative welfare states**

Inactivity is also elevated, however, in a variety of situations where welfare is believed to be poor. While all of the examples of “positive” inactivity above could likely be described as reflecting comfort and/or relaxation (i.e. the absence of pain or stress), it seems there are a variety of negative subjective states that can all give rise to inactivity. In some cases, the underlying state is ambiguous, while in others it is relatively easy to infer based on context or physiological correlates.

Illness and injury are among the most obvious welfare problems reliably associated with increased inactivity. Lethargy (i.e. increased inactivity associated with “sleepiness” and decreased responsiveness) is a well-established component of sickness behaviour (Hart, 1988), as well as being a common sign of chronic pain (e.g. Wiseman et al., 2001). Lame animals typically lie down more, since standing and walking are painful (e.g. dairy cattle: Calderon and Cook, 201; broilers: Weeks et al., 2000). Acute pain can also increase the proportion of time awake spent lying down (e.g. Leach et al., 2009; Leslie et al., 2010), although it also has the contradictory effect of causing restlessness in some cases (see Wright and Woodson, 1990).

Fear, too, is a state of compromised welfare that is easily identifiable, and often, although not always, leads to inactivity. In this case, the subjective state is primarily inferred based on the context in which the behaviour is observed. Immobility or ‘freezing’ is a common expression of fear in the face of an immediate (perceived) threat (reviewed by Boissy, 1995), and is used to quantify fearfulness in many behavioural tests (see e.g. Bouton and Bolles, 1980; Forkman et al.,
2007); this is an acute form of inactivity (Bracha, 2004). Temporary decreases in activity have also been linked to acute stress or fear in calves: inactivity is induced by both novel object tests, for individuals that exhibit physiological stress responses indicating fear of the object (Van Reenen et al., 2005), and by direct activation of the HPA (hypothalamic-pituitary-adrenal) axis by administration of adrenocorticotropic hormone (Negrao et al. 2010), suggesting this pathway might mediate fear-induced inactivity. Fear can also be manifest as hiding, over both short and long timescales (e.g. cats: Rochlitz et al., 1998). Thus, the location of the animal while inactive may be an important clue to the psychological state, with use of any kind of shelter or visual barrier potentially indicating fear. Location is already used as an indicator of fear in standardized tests such as the elevated plus maze, where spending more time in closed than open (exposed) arms is interpreted as fear. Clearly, this depends on there being a relatively “safe” location available: if animals are afraid and therefore motivated to hide or to escape but unable to do so, their welfare is likely to be even worse (e.g. Carlstead et al., 1993a; Nimon and Broom, 1999).

In other contexts where high levels of inactivity are observed, states such as ‘depression’ and ‘apathy’ have been postulated as explanations. In humans, depression can alter locomotor activity and rest in opposite directions depending on the sub-type of the disorder (see e.g. Gold and Chrousos, 2002), meaning that both increases and decreases in inactivity levels can be indicators of this type of poor welfare. However, in the non-human models of depression used by researchers interested in the human disorder, where activity levels are discussed as a symptom, inactivity seems to be elevated. It is involved in multiple validated animal models, notably the forced swim test and the learned helplessness paradigm. The forced swim test is the most widely used test used for screening potential antidepressants (Cryan et al. 2002), in which spending more time immobile and floating rather than actively struggling to escape is taken as an indicator of depression (Porsolt et al., 1977). Of course, this is not a measure of general inactivity; immobility in the forced swim test does not always correlate with activity levels under other conditions (e.g. Hilakivi and Lister, 1990) and actually often negatively correlates with activity in a novel environment, which is related to anxiety (e.g. Brenes et al., 2008). Learned helplessness, on the other hand, is a phenomenon typically accompanied by an overall decrease in activity (see e.g. Mineka and Henderson, 1985). The term “learned helplessness” referred initially to a deficit in avoidance learning induced by repeated exposure to uncontrollable shock conditions.
(reviewed by Maier, 1984), linked to depression both because the aetiology corresponds to theories of human depression and because it is associated with symptoms of depression including decreased libido, anhedonia and altered sleep patterns (reviewed by Henn and Vollmayr, 2005). It is thus believed to be associated with negative affect. The meaning of the label “learned helplessness” has now been expanded, as it is sometimes applied to any “passive” behaviour (i.e. quiescence or the absence of active responses to stress, such as escape attempts; cf. Oxford English Dictionary, 2005) that appears to result from exposure to uncontrollable stressors (Maier, 1984). Some researchers believe it might be applicable to many captive animals that seem very passive or inactive (Carlstead, 1996; Wemelsfelder, 1990). This hypothesis is supported by findings that animals reared in socially isolated and/or barren cages are more vulnerable to developing learned helplessness than those reared in more socially and physically complex, and presumably controllable, environments (Chourbaji et al., 2005; Seligman, 1972). While learned helplessness implies that a lack of control was the cause of the behaviour, there are other terms that may describe the same state but are free of assumptions about the aetiology. Most commonly, the word “apathy” has been applied to very similar behaviour patterns in suboptimal, inescapable environments (e.g. Van Putten, 1980; Visser et al., 2008; Wells et al., 2002), but is also used in other situations, such as for animals reacting to the loss of a mate (“grief”; e.g. Meyer-Holzapfel, 1968). Similarly, Fordham and colleagues (1991) describe a period of extreme inactivity and unresponsiveness they called “withdrawal” in sheep one week after they had been moved from pasture to indoor crates. The delayed onset of this withdrawal could perhaps reflect a transient initial response of exploration and/or escape attempts.

“Boredom” and similar states represent other potential correlates of inactivity. Thus, another contrasting explanation for the high levels of inactivity sometimes exhibited in barren or otherwise sub-optimal environments, compared to enriched ones, is that there are simply no activities that animals are motivated to perform. As mentioned above, the captive environment should provide everything needed for immediate survival, which superficially appears to be good for welfare; however, if this eliminates the need for foraging, predator evasion and other natural behaviour patterns without providing opportunities for other motivated activities such as reproduction, animals may be left in a state of “limbo” (McFarland, 1989). McFarland suggested that such an amotivational state might well cause suffering, because most species will not have
evolved methods of coping with such a situation. Others have invoked the concept of boredom to describe the psychological state experienced in barren environments, which would correspond with what he called “limbo”. For example, when giving pigs enrichment was found to increase their behavioural diversity and reduce time spent inactive, Wood-Gush and Beilharz (1983) believed this was an indication that boredom had been reduced. The same rationale could be applied to reductions in inactivity induced by many forms of enrichment, such as parrots that went from spending only 6% of their time off their perch to 26% when given a feeding enrichment, bringing them closer to the 40% or more of the time budget most wild parrot species are reported to spend foraging (Rozek et al., 2010). It is worth noting, however, that this hypothesis to explain inactivity in captive animals is complicated by the fact that boredom is not always associated with inactivity; although lethargy can be a long-term symptom in humans (Berlyne, 1960; Inglis, 1983), boredom is tightly linked to a motivation to explore, and can induce restlessness when exploration is not possible (reviewed by e.g. Berlyne, 1960). No single phenotype for boredom in animals has yet been established. It is thus rather difficult to determine which category in Table 1.1 would best describe the inactivity associated with boredom or any other “limbo” state. According to the model of Engel and Schmale (1972), stimulus deprivation (and therefore limbo) might induce a conservation-withdrawal state, and therefore decreased alertness; conversely, however, bored animals could also be expected to be alert because they are seeking stimulation. This topic therefore still needs research (see Chapter 5).

Finally, there are some situations that induce both elevated inactivity and more unambiguous or widely accepted signs of stress, such as high corticosteroid levels or prevalent abnormal behaviours, but that do not obviously induce any of the aforementioned states of fear, depression or boredom. For example, unpredictability of feeding makes capuchins more inactive and reduces social behaviour (Ulyan et al., 2006). Although the unpredictability was believed to cause poor welfare, a conclusion supported by increased cortisol concentrations, no specific subjective state was inferred. It could perhaps be attributed to frustration, when food was no longer presented at the expected times (i.e. absence of an expected reward; reviewed by Mineka and Henderson, 1985), but again, more research is needed to test this hypothesis.
General lessons for group-level welfare inferences

It is clear from the above that increases in inactivity can reflect either improved or diminished welfare, or to look at it another way, poor welfare can affect inactivity in either direction. Although very little systematic work has been done to investigate the reasons for these opposing findings, it is possible to discern some patterns. At the group level, the direction in which inactivity can be expected to change clearly depends on the type of stimulus or treatment applied (and thence the type of inactivity). However, the method of assessing inactivity, the timing of the observations, and the species could also be important.

The examples in the previous two sections highlight the importance of the specific type of welfare problem being investigated and correspondingly, the situations being compared. Environments that lack a comfortable substrate for lying will likely reduce time spent inactive, as in the cattle studies cited above. So too will feed restriction, and perhaps more generally, frustration of any motivation to perform a specific activity (deprivation: cf. Dawkins, 1988), which tends to induce stereotypic behaviour, repeated escape attempts or other “restless” behaviour (e.g. migratory birds when caged: Mewaldt and Rose, 1960; laying hens unable to nest: Duncan, 1970; mink blocked from access to swimming water: Vinke et al., 2008). Chronic, inescapable exposure to aversive stimuli or experiences, on the other hand, seems likely to produce depression-like states, which are often associated with increased inactivity. This is true whether the stress is physical, as in the standard learned helplessness paradigm, or psychological, as with chronic social stress (e.g. Fuchs and Flugge, 2002) or repeated or long-term isolation for many social species (e.g. rats: Hurst et al., 1997; goats: Siebert et al., 2011; horses: Visser et al., 2008; but see Cockram, 2004, who report opposite results in sheep). A lack of relevant behavioural opportunities also tends to lead to increased inactivity; thus, successful environmental enrichment often decreases inactivity in a wide range of species (e.g. Koistinen et al., 2009; Olsson and Sherwin, 2006; Rozek et al., 2010), the exceptions to this rule being types of physical enrichment that would primarily be expected to increase comfort or perceived safety, such as shelters (e.g. Würbel et al., 1998). Increased inactivity in non-enriched environments
may be particularly prevalent in animals with previous exposure to enrichment (e.g. Bolhuis et al., 2006).

The variation in methods of assessment discussed earlier, both in terms of how behaviour is categorized as active versus inactive, and the time and place in which that behaviour is assessed, also plays a role in the diversity of the effects observed. If inactivity only refers to sleep or rest (i.e. time not alert), it would be expected to decrease when animals are fearful or anxious (e.g. Abou-Ismail et al., 2008; Rattenborg et al., 1999), whereas if all stationary behaviour is considered inactive, the category will include vigilance; thus, inactivity might be expected to increase (e.g. Carlstead et al., 1993b). In terms of specific testing protocols, although proxy measures of activity obtained from a behavioural test conducted outside of the home cage, such as the elevated plus maze, typically correlate with general activity levels (e.g. Lister, 1987; Tang and Sanford, 2005), they do not always correspond with home cage behaviour (e.g. Morgan and Pfaff, 2001). Locomotion in the open field test is also sometimes used as an indicator of overall activity (e.g. Touma et al., 2008), but it is a poor one because it is also strongly influenced by anxiety; locomotion in the first five minutes in the open field does not correlate with that in the home cage (Tang and Sanford, 2005). Of course, the open field could still be an appropriate setting for assessing acute fear-induced inactivity (freezing). Such differences between measures of inactivity depending on the setting and duration of observations might explain why there are exceptions to the trend of chronic stress decreasing activity discussed above. For example, Pijlman and colleagues (2003) report that a repeated emotional stressor, witnessing a conspecific being shocked, increased rats’ locomotor activity in a later open field test. This is likely specifically an acute, anxiety-related response to the novel environment, and whether their activity in the home cage was altered, and in what direction, is unknown. In addition to the relative novelty or familiarity of the environment, the available options in a given environment will influence the behavioural response. To illustrate, Cooper and colleagues (1996) showed that voles responded to an unfamiliar sound by freezing if in an enriched environment where cover was available, but otherwise responded actively, by running or digging. Meanwhile, in an effect that was dependent solely on duration of the treatment and not on setting, pigs initially increased their activity levels when put in isolation, in contrast to the social isolation studies of other species cited above; however, they then appeared to have increased inactivity on
subsequent days as isolation continued (Ruis et al., 2001). The time of day when the observations are conducted can also influence the results of a study. For example, restricted feeding in mink decreased their inactivity before feeding but increased it earlier in the day (Hansen and Møller, 2008; see also strain differences in mice: Tang and Sanford, 2005).

Lastly, species and strains can differ qualitatively in the way they exhibit a given subjective state, including whether inactivity is a typical response. For example, Glickman and Hartz (1964) found that both guinea pigs and chinchillas exhibited behaviour suggestive of fear in an open field: guinea pigs tended to freeze, while chinchillas defecated more often than the other species tested and refrained from grooming. However, chinchillas exhibited high levels of locomotor activity, while guinea pigs were very inactive. Similar species differences in open field activity have been reported among muroid rodents, and could be explained by their ecology: field-dwelling species were the most likely to freeze, while arboreal species were most likely to climb (Wilson et al., 1976). Brashares and colleagues (2000) similarly found differences in hiding between antelopes depending on an aspect of their ecology: whether they were solitary or gregarious. The way activity levels change when animals are anticipating either a positive or a negative stimulus likewise seems to vary between species, in a way that may sometimes be attributable to ecological differences. While mink spend more time in the nest box (and thus are more inactive) when anticipating an aversive stimulus compared to their baseline behaviour (Hansen and Jeppesen, 2006), and they, along with foxes and rats, become more active when anticipating a positive stimulus (a food treat) than an aversive one (Hansen and Jeppesen, 2006; Moe et al., 2006; Van den Bos et al., 2003), cats and chickens instead become inactive when anticipating a positive stimulus (Van den Bos et al., 2003; Zimmerman et al., 2011). Van den Bos and colleagues (2003) attribute the difference between cats and rats to species-specific foraging strategies, since cats are ‘sit-and-wait’ predators while rats search for food (but see Zimmerman et al., 2011 for an alternative explanation). Whether species’ ecology or phylogeny can also explain other contradictory findings remains to be seen. For example, felids seem prone to hiding or becoming very inactive as a response to chronic stress (e.g. Carlstead et al., 1993a; Rochlitz et al., 1998; Wielebnowski et al., 2002), and also seem to have suppressed activity in zoos on days with high visitor density (Mallapur and Chellham, 2002), while the opposite effect of visitors has been reported for at least one primate species (Wells, 2005). Such differences in
stress responses might be predicted by variables that influence optimal anti-predator strategies, such as crypsis or flight speed.

*Welfare inferences at the individual level*

Up to this point, I have discussed only changes that would be observed at the group level with a given treatment; however, responses will also vary within the group depending on a variety of individual characteristics. Both age and sex are known to influence susceptibility to stress (reviewed by Anisman and Matheson, 2005), and therefore whether a response of any kind is likely to be observed. They may also influence the form of the response; for example, in humans, females are more vulnerable to depression (reviewed by Anisman and Matheson, 2005), and among people with depression, females are more likely to report fatigue and hypersomnia while males are likely to report agitation (reviewed by Antonijevic, 2006). From what little work has been done on sex differences in depression, it appears that this may also be true in rodent models (see e.g. Becker et al., 2007); however, in contrast, female rats may be less prone to becoming inactive in response to repeated shocks in a learned helplessness paradigm (reviewed by Shors, 1998). One example of age differences in response strategies is seen in red deer: juveniles of this species employ a hiding strategy, and thus freeze in response to threat, but as they age, they begin fleeing from some threats instead (Espmark and Langvatn, 1985). The influence of age and sex will likely vary between species. For example, more age differences presumably exist in species with altricial young, which differ more from adults in their abilities and thus behaviour, while sex-specific responses might be more likely in species which differ in home range size, since those with larger home ranges might pay a larger penalty for reduced locomotion. Sex differences in response to acute and perhaps even chronic stress might similarly be more likely where there is sexual dimorphism influencing crypsis or flight speed, and therefore whether inactivity or activity is the best anti-predator strategy. Hormonal states that vary with the reproductive cycle can also determine whether inactivity will increase under stressful conditions (Perrot-Sinal et al., 2000); thus, seasonal effects would be expected in many species, and might differ between sexes.
In addition to these simple physiological differences, several stable individual characteristics might influence the presence/absence and/or the direction of effects on inactivity. Learning ability will influence the specificity of fear-mediated changes in inactivity (e.g. Rushen et al., 1999) and their duration, as there will be changes in habituation rates. An individual’s experience will also influence stress responses; for example, captive-born individuals may be more likely to respond actively to a sub-optimal captive environment, developing stereotypic behaviour, while wild-caught individuals may be more likely to respond by hiding (e.g. Jones et al., 2011). Finally, personality or “coping style” will determine the form of an individual’s response, including whether they become inactive. Research on coping styles, a term applied to suites of traits describing both behavioural and physiological responses to stress, has categorized animals as either “passive copers” or “active copers”; passive individuals tend to become immobile when frightened and are thought to be less active or exploratory in novel environments, while active copers tend to either fight or flee (Koolhaas et al., 1999; Sih et al., 2004). Van Reenen and colleagues (2005) thus suggest that their failure to find a correlation between open field locomotion and other measures of response to novelty might be explained by the presence of different coping styles, such that some calves that were frightened of the open field responded with escape attempts, but others with immobility. They also suggest that these differences in strategy might reflect underlying differences in general activity levels. Coping styles are also associated with differences in personality traits such as shyness or boldness (Sih et al., 2004).

To what extent these individual differences reflect genuine differences in welfare is difficult to determine. While individuals that use passive rather than active coping strategies typically show higher HPA responses to stress, and thus might often be judged to have poorer welfare since corticosteroid levels are one of the most common welfare indicators, they also have decreased sympathetic responses (Koolhaas et al., 1999; cf. Duncan, 1980 for comparing welfare when manifestations of stress vary). However, Mason and Latham (2004) found in a meta-analysis that more often than not, within populations where stereotypic behaviour was prevalent, individuals that did not stereotype, or had relatively low levels of stereotypy, had poorer welfare than those that with high levels, according to a variety of welfare measures. Since non-stereotypic individuals are likely to be the most inactive individuals within a population (e.g.
Bildsøe et al., 1990; Würbel et al., 1998), this may indicate that passive or inactive responses to stressful conditions are more often associated with poor welfare. One interesting phenomenon to note is that even when inactivity is not the result of poor welfare, it may result in the individual having poorer welfare if group-housed; for example, inactive hens are more likely to be victims of feather pecking (Riber and Forkman, 2007). Thus, large inter-individual variation in inactivity might have practical consequences for farmers as well as for animals.

The practical perspective

Aside from concerns that high levels of inactivity may reflect poor welfare, there are several reasons why they might be problematic in captivity. First, as mentioned above, inactivity sometimes leads to obesity and reproductive impairments, which reduces profits on farms and interferes with zoo breeding programmes. In animal shelters, inactivity can reduce adoption rates for dogs and perhaps cats, because potential owners consider animals less attractive as pets if they are inactive and unresponsive (Gouveia et al., 2011; Wells et al., 2002). Inactive animals in zoos are similarly bad for business because they attract less visitor interest (Margulis et al., 2003), which might in turn reduce visitor learning (Altman, 1998). In Mills’ (2003) classification of problem behaviour, this would be categorized as: “actions which have adaptive value within the species but which are inconvenient for the keeper” (p. 270), as opposed to the maladaptive behaviour patterns discussed above, or what he called malfunctional behaviour, which reflects nervous system pathology. Finally, too much inactivity could decrease the success of reintroduction programmes for endangered species, either because the animals will lack specific locomotor skills required in their natural environment (e.g. Stoinski et al., 2003), or because they are generally in poor physical condition, although so far what little evidence exists does not support the latter hypothesis (Biggins, 1999). Of course, as with all wild animals, retaining adaptive forms of inactivity is still essential: individuals lacking normal hiding or freezing responses are unlikely to survive when reintroduced (McPhee, 2004).
Merging the perspectives

There is no simple link between behaviour that is ‘natural’ and/or adaptive and welfare states. Although inactivity is adaptive both when it is used to conserve energy and when its main function is to avoid detection, the first type might feel pleasant, whereas being frightened enough to freeze completely most likely does not. This difference is explained by the fact that resting is an adaptive response to positive contexts, and freezing to negative contexts. Conservation-withdrawal states, too, are adaptive and yet may involve suffering in many, if not all, cases. The conservation-withdrawal states induced by chronic social stress such as defeats or subordination are particularly interesting because they appear depression-like but may have special adaptive value, since some depressive behaviour can help to recruit social support (reviewed by Mills, 2003). The original paper which described conservation-withdrawal suggested that whenever it is elicited by psychological stressors, it would be mediated by ‘giving up’ and thus, the state would be associated with a depressed mood (Engel and Schmale, 1972). This dovetails with learned helplessness theory, and with evolutionary explanations of human depression (Nesse, 1999). Thus, even if they are the best possible way of coping with the situation, the long-term elevations in inactivity described in many non-enriched or otherwise stressful environments may well be associated with poor welfare. The converse is also true: being maladapative does not guarantee that a behaviour pattern reflects poor current welfare. For example, an animal may be inactive enough to become obese, thus increasing its risk of poor health or reduced fertility, but until illness develops, its welfare will be unaffected. Nonetheless, maladaptive behaviour will likely often be associated with poor welfare in the long term. Although comparisons with natural behaviour do not provide direct information about animal welfare, they may aid in identifying potential problems in individuals or groups, which then require direct study to determine whether the animals are actually suffering.

There is perhaps more overlap between the situations in which very high levels of inactivity cause practical problems and those in which the animals are likely to have poor welfare. Individuals that have not been active enough to learn appropriate skills before being reintroduced into the wild are likely to have poor welfare once released, since they may be unable to attain sufficient food or find shelter, and may be at higher risk of injury. If inactivity
prevents companion animals from being adopted, then they too are likely to have poorer welfare than those that find homes. When inactivity is problematic only because it tends to make the animals more boring zoo exhibits, it does not influence their welfare unless they are moved to poorer enclosures as a result; however, where inactivity levels are both a practical problem and reflect a dramatic departure from wild behaviour, their welfare is also likely to be compromised.

**Applications for animal welfare research**

Given that relationships between inactivity and welfare are so complex and variable, interpretations of inactivity in a given individual or population are impossible without detailed knowledge of the species’ ecology and behaviour patterns, or of the context being studied. As can be seen in Table 1.1, inactivity can be phenotypically quite similar in terms of commonly noted characteristics such as duration and level of alertness, and yet be associated with quite different affective states. For example, there are several forms that are defined by long-term changes to the time budget, and involve partial rather than complete immobility as well as decreased alertness. While sleep can be easily differentiated from the others because it involves reduced consciousness, there are no established reasons to categorize “depression” separately from any other conscious-but-low-alertness immobility, except our awareness that they are seen in very different situations. This can lead to a danger of circular reasoning, when inactivity is interpreted as evidence the animals are depressed in that situation, without independent validation. Similarly, freezing and a sit-and-wait hunting behaviour might look similar, and need to be differentiated based on the species’ niche and/or the antecedents: in the former case, a predator cue, versus a prey cue in the latter.

Before a form of inactivity can be used as a welfare indicator, it must be clearly defined phenotypically and the associated affective state or, at least, the valence of that state must be identified. The initial phenotypic category can be based on broad, universally applicable variables such as those mentioned in Table 1.1. One or more affective states that are likely candidates to elicit this behaviour can then be identified, based on these characteristics, the contexts in which it is seen, and, potentially, analogies to human behaviour. Validation of its use
as a welfare indicator can then be conducted accordingly. This will include looking at how it changes when welfare is manipulated, whether it co-varies with other indicators of the hypothesized underlying state, and potentially using pharmacological validation if possible (e.g. testing that it is reduced by anxiolytics if it is believed to reflect fear). At this stage, the phenotypic category can be further refined for a given species, by investigating whether the correlates differ depending on variables such as posture, time of day (e.g. during a species-typical active period or not) or even associated physiological response, and compared to that seen in other settings. This needs to be done on a species-by-species basis, since subtle differences such as posture will vary. Once a subtype has been identified that can be differentiated behaviourally and is specific to treatments associated with a welfare problem or a positive welfare state, then it can be used as a welfare indicator for that species in studies of less well-understood situations.

One form of inactivity for which this process has been carried out in multiple species is hiding. It is easily defined operationally as remaining stationary and out of sight or camouflaged. As described above for cats, it is increased by various situations that induce stress and co-varies with physiological indicators of stress, such as corticosteroid levels (e.g. Rochlitz et al., 1998); for rodents, it can similarly be induced by stressors such as predator odours (e.g. Dielenberg et al., 1999). It has also been validated pharmacologically as an indicator of fear in rodents, being reduced by anxiolytics (e.g. Dielenberg et al., 1999). Thus, this behaviour pattern can now be used as a welfare indicator across a variety of settings. While the other fear-related forms of inactivity, freezing and tonic immobility, have likewise been well studied, most other forms of inactivity still need to be better characterized before they can be used as welfare indicators in this way. According to the criteria (e.g. alertness) described here, for example, “apathetic” inactivity cannot reliably be differentiated phenotypically from positive forms of rest. Future research should investigate what other variables should be measured in conjunction with inactivity to make it a valid indicator of apathy (e.g. whether there are certain inactive postures that are associated with unresponsiveness, and reduced by antidepressants), and if possible, identify specific situations in which apathy is a likely response.
If a given form of inactivity does not co-vary unidirectionally with welfare, additional information necessary. This is the case with sleep, since either too much or too little sleep can be associated with stress and/or depression, as mentioned above. A starting point would be to refine its use by assessing the two main categories of sleep (REM and non-REM) separately when possible, since they are believed to have different functions and may be differentially affected by stress (cf. Antonjievic, 2008; Siegel, 2005). However, even for just one of these forms, the direction of its change may vary between individuals when exposed to stress. As a result, even if a treatment does induce stress, researchers may find no overall effect of treatment on sleep because the opposite responses cancel each other out if they are approximately equally prevalent in the population. Where this type of bidirectional response is suspected to exist, it may still be possible to use the behaviour pattern as a welfare indicator if there is prior knowledge of the individuals being tested that would allow them to be classified a priori as likely to increase sleep or decrease sleep in response to stress. Bolhuis and colleagues (2006) employed this approach, categorizing pigs as “high resisting” or “low resisting” in a test of “coping style” and then examining the responses of each group to a housing treatment separately. This is effective if such tests have been developed for the species, so that “coping” or response styles can be identified based on a characteristic other than activity level; otherwise, there is some danger of circular reasoning.

Conclusions

It is clear that inactivity is a heterogeneous category of behaviour, in terms of its functions or eliciting factors and its associated subjective states. With so many variables at play, it is not surprising that no one has yet attempted to delineate general patterns in its relationship to welfare, so that forms of inactivity could be used more successfully as welfare indicators. However, when divided into phenotypic subtypes, and combined with knowledge about the species and individuals, it can provide important information about welfare states. Even at the group level, knowledge of the individuals is useful, since the direction in which inactivity will change when conditions are altered will likely depend on the distribution of ages, sexes and response styles within the population. The form of inactivity that should be assessed will depend on the type of potential stressor or positive treatment applied. Understanding how all of these
variables influence inactivity would allow us not only to better use it as a welfare indicator, but also to manipulate it and thus alleviate the many practical problems it can cause on farms, in shelters and in zoos.
Table 1.1. Forms of inactivity that might be influenced by welfare and their characteristics.

<table>
<thead>
<tr>
<th>Timescale</th>
<th>Immobility level</th>
<th>Alertness</th>
<th>Label used</th>
<th>Context</th>
<th>Other distinguishing features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic</td>
<td>Partial</td>
<td>Very low</td>
<td>Sleep</td>
<td>Any</td>
<td>Altered EEG, metabolic rate, etc. Species-specific postures</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic</td>
<td>Partial</td>
<td>Low</td>
<td>Rest or relaxation</td>
<td>Typically, absence of threat and physical comfort Physical illness</td>
<td>Species-specific postures?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Elevated cytokines</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Co-varies with symptoms such as anhedonia</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>May involve specific postures e.g. “dog-sitting” in sows</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Broad category, possibly including all above forms but usually only applied to sleep or rest if stress-induced</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic or acute</td>
<td>Partial</td>
<td>High</td>
<td>Hiding</td>
<td>Presence of potential threat</td>
<td>Location: sheltered or camouflaged</td>
</tr>
<tr>
<td>Acute</td>
<td>Partial</td>
<td>Low</td>
<td>Fainting “Shock”</td>
<td>Pain, illness or threat</td>
<td>Parasympathetic nervous system activation (e.g. bradychardia)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sympathetic nervous system activation (tachycardia)</td>
</tr>
<tr>
<td>Acute</td>
<td>Complete</td>
<td>High</td>
<td>Freezing</td>
<td>Presence of potential threat</td>
<td>More common in prey species Likely “tense” posture (signs of alertness such as erect ears; species-specific)</td>
</tr>
</tbody>
</table>

1 Acute = immediate response to a stimulus or situation, usually with a maximum observation period of a few hours; chronic = behaviour patterns or changes to time budget persisting over days or weeks.

2 Complete = totally motionless; partial = no locomotion or goal-directed behaviour.

3 Engel and Schmale (1972) also include tonic immobility in this category, i.e. they considered it applicable for acute as well as chronic behavioural responses to challenge.
Part 2. Project Outline

The aim of this project was to improve our understanding of the relationships between inactivity and welfare in American mink (*Neovison vison*, formerly *Mustela vison*). This species is kept in very large numbers on fur farms, allowing for large sample sizes, and can be used as a model for other captive carnivores. Mink are known to exhibit a great deal of individual variation in behaviour, with very high levels of inactivity being the apparent alternative to stereotypic behaviour as a response to captive conditions (e.g. Svendsen et al., 2007). The welfare of fur-farmed animals is also a subject of great public concern.

I began with the following hypotheses: (1a) extreme inactivity can reflect captivity-induced depression, chronic fear, or a combination thereof; (1b) if this is the case, their fitness may be impaired; (2) these states are more common in sub-optimal environments; and (3) forms of inactivity that reflect good vs. poor welfare differ subtly (e.g. in posture, location, or alertness).

In order to test these hypotheses, it was necessary to have a valid and sensitive method for assessing fear. Thus, in Chapter Two, I developed and validated a temperament test in which mink could be categorized as fearful or curious, aggressive or unresponsive. This was based on an existing test used in Scandinavia that did not detect sufficient levels of fear in these populations. Chapters Three and Four describe the results of a project conducted on commercial farms in Ontario. These experiments were designed to test Hypotheses 1a and 1b. Individual levels of inactivity were correlated with reproductive success, to determine whether inactive females had impaired fitness. Chronic stress or fear and body fat were examined as possible mechanisms for any reproductive impairment, to determine whether the effects observed were relevant to the welfare of the animals. In Chapter Five, I developed a way of operationalizing depression, apathy and boredom, in order to determine which of these subjective states was most likely to be induced by farm conditions and thus potentially underlie the elevated inactivity levels. Lastly, Chapter Six tests Hypothesis 3. It describes the effects of environmental enrichment, which was assumed to improve welfare, on total inactivity and on specific subtypes that might differ in their welfare significance.
CHAPTER TWO
Who’s afraid of the big bad glove? Testing for fear and its correlates in mink

Rebecca K. Meagher, Ian Duncan, Allison Bechard, Georgia J. Mason


Abstract

Fear in farm animals is a welfare and economic concern. For Scandinavian mink, the “stick test” is common for assessing fearfulness: a spatula is inserted into the cage and minks’ immediate responses are noted. However, on Ontario farms, fearfulness in the stick test was very rare and aggressive responses were prevalent, rendering this test poor for testing hypotheses related to fear and welfare. We therefore developed a modified version, the “glove test”, where the finger of a handling glove is inserted into the cage. This proved more sensitive than the stick test for detecting fearfulness in Ontario mink (20% vs. 2.5%, P<0.0001); and successfully reduced aggressive responding (22% vs. 41%, P<0.0001). When test-retest reliability was assessed conventionally, it was only moderate (e.g. 37.5% mink behaved the same in three tests). However, it is biologically realistic to expect habituation over repeated trials (e.g. reduced fearfulness), and treating such changes as acceptable, results were reliable for 76% of mink over three tests. Reliability could be further improved by taking location into account, since some mink were unresponsive only if in the nest box, half-asleep (Expt. 3: kappa of 0.38 if never in nest box vs. kappa of 0.002 if were). Glove tests had construct validity: mink immediately classified as “fearful” spent more time exhibiting other fear-related behaviours (‘ambivalence’:...
mean 30 s vs. 4 s, \( P=0.009 \)), while mink immediately classified as “curious” then spent more time investigating the glove (mean 103 s vs. 57 s, \( P<0.0001 \)). Glove tests also revealed expected temperament differences between Black and Pastel colour-types, with Pastels being less fearful (\( P=0.001 \)). Finally, we tested whether fearfulness in the glove test is associated with decreased reproductive success. Pastel mink that were fearful during the presumed gestation period were less likely to reproduce (\( P=0.006 \)). Like stick tests, glove tests are thus practical, valid and reliable for assessing fearfulness in farmed mink, but better for detecting fearfulness in populations where fear levels are low.

**Introduction**

Fear in farm animals has serious economic, practical and welfare implications (Hemsworth, 2003; Jones, 1997). Attention has therefore been paid to decreasing fear through altered management (e.g. Hemsworth et al., 1989) or selective breeding (e.g. Hansen, 1996). Methods of fear assessment vary, however, and many have been insufficiently validated (Forkman et al., 2007). Farmed mink (\textit{Neovison vison}) exhibit some fear of humans, since they are less domesticated than other livestock (Nimon and Broom, 1999), and are handled only sporadically (e.g. for relocating or vaccinating) by stockpersons wearing thick gloves to protect against bites (European Commission, 2001). This handling is acutely stressful (Hansen and Damgaard, 1991; Korhonen et al., 2000). The fear mink experience under typical farm practices is thus a potential welfare issue.

The most common test for fearfulness is the “stick test” (e.g. Hansen, 1996; Kirkden et al., 2010), which categorizes mink as fearful, curious (sometimes called “confident”), or
aggressive based on their immediate response to a wooden spatula held through the mesh of the cage. This test, developed and validated in Scandinavia, has proved very successful (see e.g. Kirkden et al., 2010): it is rapid, averaging 9 s per subject; can have good inter-observer reliability (Hansen and Møller, 2001); and has been used to select lines of fearful or confident mink which then differ consistently in physiological stress responses (e.g. Hansen, 1997) and their behaviour in diverse other fear tests (Malmkvist and Hansen, 2002). We therefore wished to use the stick test on Ontario farms, to test welfare-relevant hypotheses about relationships between fearfulness and stereotypic behaviour, reproductive success, and environmental enrichment. However, informal pilots soon revealed that very few Ontario mink showed fearful responses, giving the stick test minimal usefulness for distinguishing between individuals. A more sensitive test (i.e. able to detect fear in more animals) was therefore needed. Aggressive responses were also very common: a problem for our research because it is unclear how to interpret these in term of welfare.

On Scandinavian farms, the “Trapezov hand test” is the recommended alternative test with higher sensitivity to fear, being more threatening: the cage lid is opened, a gloved hand is reached in, and if possible, the mink is touched; the mink’s response is scored (Malmkvist and Hansen, 2002; Trapezov, 2000). However, this test was impossible on our farms because cage design differed from that typical of European farms: the nest box is placed at the back of the cage rather than the front; it sits within a hole in the cage roof; and the only way to open the cage is to lift this nest box out. Reaching in to conduct a hand test would thus be very slow; would involve flushing out any mink in nest boxes; and most importantly, would leave the animal with no clear place for retreat (the nest box is gone, plus the hand is entering from the back of the cage, but the human tester is standing at the front), making fearful responses challenging to score. We thus
needed to develop and validate a variant on the stick test that would be more appropriate for North American farms.

We therefore modified the stick test by using a well-used handling glove – a leather gauntlet used on-farm for catching animals - in place of the stick. We hypothesised that this would be more aversive than a stick, thence better at eliciting fear, because the mink may associate the glove with past handling events; it also smelled of the scent that mink release when frightened (Dunstone, 1993). This was expected to increase the probability of an active response, including fearful ones, by increasing arousal and vigilance (cf. Zalaquett and Thiessen, 1991). The second aim of using a glove was to reduce aggressive responses that might reflect predation or play. The spatula used in the stick test is similar in length to minks’ average prey (15-20 cm: Dunstone and Sinclair, 1978; although only approximately half of the stick is inside the cage); biting it often involves attempts to pull it into the cage (Kirkden et al. 2010); and mink are known to chew and hoard small objects (Axelsson et al. 2009; personal observation). This suggests that at least some ‘aggressive’ stick-biting may reflect predatory or even playful motivations to obtain it for carrying, chewing and hoarding. We hypothesised that the finger of a glove might be less likely to elicit such responses because more aversive than the stick, and part of an object larger than typical prey.

Here we present three experiments exploring this new test’s utility. In Experiment 1, we assessed its test-retest reliability (repeatability), and its construct validity (its ability to measure what it is meant to be measuring: Cronbach and Meehl, 1955; Meagher, 2009). In Experiment 2, we replicated our previously-unrecorded pilot study, directly comparing glove and stick tests to determine whether the former was indeed more sensitive to fear and less prone to eliciting
aggression. In Experiment 3, we used the glove test to assess whether fearfulness predicts reduced reproductive success, as it does in some livestock (e.g. Hemsworth et al., 1989, Hemsworth 2003).

**Experiment 1: Validation of the glove test**

*Methods*

*Subjects and timing*

All mink were adult females of two colour-types, Black and Pastel, on a commercial farm, individually housed in standard cages, in several different sheds. These cages are arranged in long rows with opaque partitions between them, and furnished with a wooden nest box. Tests were conducted in September, and immediately after feeding to ensure that as many animals as possible were awake. A total of 187 mink (113 Blacks, 80 Pastels) were tested multiple times to assess test-retest reliability, while an additional 46 (18 Blacks, 28 Pastels) were tested once in a longer test for assessing construct validity.

*General test protocol*

The experimenter held the glove against the cage-front with one empty finger extended through the wire. So that all subjects had the same pre-test exposure to the experimenter, the test was always first performed on the mink in the neighbouring cage, even if that animal was not included in the study, and a well-used glove was used so that it would not shift from odour-free to odorous during the course of testing. The mink was exposed to the stimulus for up to 10 s, or 30 s if it apparently failed to notice the glove (see below).
As is standard in stick tests, mink were categorized according to their immediate reaction (e.g. Hansen 1996; Kirkden et al. 2010; Korhonen et al. 2002). The categories used, modelled on those of Hansen (1996) were: (1) fearful, if the mink retreated (note that withdrawing into the nest box from a lying position with the head out of the box was considered a retreat); but in contrast to Hansen (1996), we also included in this category mink that remained standing still at the far side of the cage and oriented to the stimulus for at least 10 s, because no retreat was possible for these subjects and their ‘freezing’ suggested fear (cf. Blanchard and Blanchard 1969; Malmkvist 2001); (2) curious, if it approached and made contact with the glove without a hard bite, i.e. without closing its teeth; (3) aggressive, if it gave a hard bite to the glove, clamping its teeth together fully; or (4) other, if the mink was alert and initially oriented to the front of the cage but did not respond in any of the above ways, i.e. it either exhibited no obvious behavioural response (thus called “unresponsive”), or performed some other activity such as stereotypic behaviour. Mink that remained stationary and oriented to the stimulus fell into the “other” category if they were not standing at the far back of the cage, as described for “fearful”, because it was less clear whether this was a ‘freezing’ response.

Any mink that could not be tested and thus assigned to a category was excluded from analyses. This occurred if an individual was out of sight in the nest box (i.e. lying with the head down, in which case they were not visible through the opening) or resting/sleeping either in the nest box or the open cage such that it did not appear to notice the glove within 30 s on two separate attempts. In some cases, the decision that mink were untestable was also made at the group level: if more than half of the remaining animals in a section of a shed were resting, it was
assumed that they had entered their post-feeding rest period (Hansen and Møller, 2008) and testing ceased in that section.

Reliability

The test was repeated on five consecutive days to investigate patterns of change. The order in which the sheds were tested was reversed on alternate days (thus no group was always tested first or last). Traditional test-retest reliability is based on the proportion of subjects showing identical responses from one trial to the next. However, some decrease in fear due to habituation was expected over repeated tests (e.g. Kirkden et al. 2010). This has been observed in the stick test (Malmkvist and Hansen, 2001), and more generally, decreased responsiveness is expected with repeated presentations of a stimulus that does not either provide a reward or cause harm (see e.g. Domjan, 2003). Thus, we predicted that with re-retesting, some initially fearful mink would become curious, while fearful, curious or aggressive mink might become unresponsive. We judged that only changes that did not fit these expected patterns would indicate a problem with the test that would justify calling it “unreliable”. Therefore, in addition to testing for test-retest reliability in the orthodox manner, we also assessed reliability in a more biologically-relevant way taking these predicted changes in response into account.

Such habituation effects might complicate the interpretation of test-retest reliability measures, but it is important to determine whether changes over time are due to habituation or are apparently random. In the latter case, the use of even a single test would have questionable biological relevance. In the former, a single test should instead reflect temperament; furthermore, repeated testing could be used to increase the data per individual and thus the statistical power
(by summing results; c.f. Kenttämies et al., 2002), and/or to reduce noise by eliminating those individuals whose responses seem to vary due to chance effects (see Experiment 3).

Construct validity

A separate, naïve group of 46 mink was screened in one glove test. This test was carried out as above except that the test duration was 3 min instead of 10 s or less. The aim was to check that the immediate response used to classify the mink predicted other aspects of behaviour consistent with that classification. Quantitative data were recorded, using a stopwatch, for durations of “ambivalent behaviour” and total interaction with the stimulus. Ambivalent behaviour was defined as alternating approach and withdrawal while oriented to the glove, an “oscillation” associated with conflicts between motivations to avoid and explore (Gray, 1987; Miller, 1944); it was thus expected to be most common in mink categorized as fearful based on their immediate response. Those mink instead categorized as curious based on their immediate response were expected to show more prolonged interaction with the glove over the 3 min, compared with fearful mink or those initially classified as “unresponsive”.

Statistical analysis

For the reliability tests, kappa statistics were calculated (Cohen’s kappa for comparisons between two tests; Fleiss’ kappa for more than two tests). Because many mink were not alert every day and could therefore not be tested five times, these statistics were calculated across the first three tests, as well as being calculated across all five tests in the subset of animals for which that was possible. Mink that could not be tested at least three times were excluded from the reliability
analysis. In addition, to account for expected changes due to habituation, we calculated the percent of mink that followed this predicted habituation pattern. Before this, McNemar’s tests were used to check for population-level changes confirming this type of habituation. Differences between colour-types were assessed using Pearson’s chi-square tests on data from the first test. The above analyses were conducted using Minitab 14 (Minitab Inc., PA, USA).

Behavioural differences between temperament categories, as assigned according to immediate response, were analysed first using General Linear Models to include colour-type. Colour-type was not significant in these validation tests and was therefore removed so that simple Analysis of Variance (ANOVA) could be employed. Normality was assessed using Anderson-Darling tests, and Levene’s tests for equal variances were conducted. Where variances were not equal (P<0.05), Welch’s ANOVAs were used to accommodate differences in variance (Welch, 1951); planned contrasts were assessed in the same way. These analyses were carried out in JMP 8 (SAS Institute Inc., NC, USA).

**Results**

*Behaviour in the five reliability tests*

Twenty-seven mink could not be tested at least three times, and so were excluded. Of the remaining 160, 75 yielded data from all five tests. In the full sample of 160, 20% were fearful on the first test; 62%, curious; 10%, aggressive; and 8% were categorized as “other”. Mink categorized as “other” were all behaviourally unresponsive (i.e. exhibited little to no locomotion during the test), and so are called “unresponsive” henceforth.
Temperament distributions on the first test varied between colour-types ($\chi^2=14.9$, d.f.=3, $P=0.002$). Blacks were more likely than Pastels to be fearful (29% vs. 7%; $\chi^2=11.8$, d.f.=1, $P=0.001$), while Pastels were more likely to be unresponsive (13% vs. 4%; $\chi^2=4.1$, d.f.=1, $P=0.04$).

Test-retest reliability

Test-retest reliability was fairly poor, assessed in the traditional way without taking predicted habituation patterns into account (such that all changes in temperament category, even expected ones, were counted as unreliable). Thus for the 160 mink for which reliability could be calculated over three tests, only 60 (37.5%) behaved consistently. This produced a Fleiss’ kappa of 0.33, indicating only “fair agreement” (Dohoo et al., 2003). For the subset tested the full five times, 20 individuals (26.7%) were consistent across all tests, and Fleiss’ kappa was 0.40. Reliability seemed to increase somewhat over repeated tests, as shown in Table 2.1, with pairwise kappas between the later tests being over 0.40, thus indicating “moderate reliability” (Dohoo et al., 2003).

However, population-level changes over repeated testing showed that mink were following the pattern predicted by habituation (Figure 2.1). Thus, fewer of these 75 mink tended to be fearful by the last test (McNemar’s test, $\chi^2=3.0$, $P=0.08$), while significantly more were unresponsive by the last test (McNemar’s test, $\chi^2=10.7$, $P=0.001$). More mink also became aggressive by the last test (McNemar’s test, $\chi^2=7.4$, $P=0.007$). When the two types of change expected due to habituation (becoming less fearful/more unresponsive) were taken into account, a more optimistic view of reliability emerged: 122 mink (76.2%) were consistent or changed in
predicted directions over three tests. Over five tests, however, this value was lower, with only 35 of 75 (46.7%) fitting these criteria.

*Construct validity*

Sample sizes, means and standard errors for both validation measures are presented in Table 2.2. For ambivalence duration, the assumption of equal variances was not met; this variable was also non-normal (ambivalence AD=7.6, P<0.005), as were the residuals obtained from ANOVAs, nor could this be corrected by transformation. However, the residuals appeared approximately normal by inspection even for untransformed data. Because the equivalent non-parametric test (Kruskal-Wallis) is not advised if variances are non-homogenous (Maxwell and Delaney, 2004), and ANOVAs are robust to deviations from normality (see e.g. Cochran and Cox, 1957) we therefore used Welch’s ANOVAs to test the differences in the means of the untransformed data for those two variables.

Mink classified into different temperament categories based on their immediate responses differed in both duration of contact and ambivalence (Table 2.2). Curious mink spent more time in contact with the glove than did fearful or unresponsive mink (contrast of curious vs. other categories: $F_{1,43}=23.34$, $P<0.0001$; Figure 2.2). Fearful mink, meanwhile, displayed more ambivalence than did mink in the other two categories (Figure 2.3; contrast $F_{1,9}=11.06$, $P=0.009$).
Discussion

Categorizing mink via their immediate response to the glove in a single test had construct validity: mink categorized as fearful on this basis then exhibited conflict between approach and avoidance motivations during more in-depth testing, while mink categorized as curious then exhibited the most sustained interest in the glove. The test also revealed expected differences between colour-types: Black mink, generally considered more nervous and restless than Pastels (European Commission, 2001), were more likely to be fearful. The test’s repeatability over consecutive days appeared low; by comparison, Hansen and Møller (2001) present data from two pairs of stick tests from which we can calculate kappas of 0.58 and 0.77 for test-retest reliability. However, much of the apparent ‘unreliability’ in the current experiment was successfully explained by habituation and therefore did not invalidate the test. This issue of habituation has not yet been raised with regard to the stick test, likely because when multiple tests have been conducted, they were spread over a long period of time, which reduces habituation effects (Domjan, 2003). However, in some experiments, temperament must be determined over a relatively short time period to test causal relationships. Repeated testing is useful in those situations (see Methods), and so habituation effects are important to consider. The one change observed that was not predicted a priori from habituation was the increase of aggression over time, addressed in the General Discussion.

Several times, mink were asleep immediately prior to the test: a phenomenon that we suspect adds noise to the test results. When sleeping mink awoke and saw the glove, they were tested, but were often awake only briefly, and seemed more likely than others to be classified as “unresponsive”. Mink that were in this ‘drowsy’ state on one test day but not on the following
one might therefore be expected to change apparent temperament categories, not because of progressive habituation but because of the chance events of being awake or asleep as the tester approached – chance events that might then reduce the test’s apparent reliability. We suspect that not only is the “unresponsive” category subject to chance events like this, but also that it is motivationally heterogeneous: the lack of an active response (whether contacting the glove or withdrawing from it) may thus sometimes reflect temporary ‘drowsiness’, as above, but in other cases it could be due either to a genuine lack of interest, or in other cases still, to fear causing freezing (mink labelled “unresponsive” did display some ambivalent behaviour during the longer tests of the validation phase). The idea that “unresponsive” may be a heterogeneous category is returned to in the General Discussion.

Experiment 2: Comparison of glove and stick tests.

Methods

Subjects and timing

This experiment aimed to formally replicate our informal pilot observations on the low sensitivity of the stick test for detecting fear in Ontario mink. Subjects were 121 adult female Black mink at another commercial farm. Tests were conducted in December, one to two weeks after pelting had been completed on the farm.
3.1.2. Test protocol

Based on Experiment 1’s findings, we modified the test protocol in two ways. First, we added a statistical control for whether mink were in the nest box, where most sleeping occurs, and thus likely to be ‘drowsy’. Second, we increased the time allowed for mink to respond from 10 s to 30 s. This was to increase the likelihood that when individuals were categorized as “unresponsive”, it reflected a genuine temperament trait rather than a temporary ‘drowsiness’.

All tests were conducted in the morning, beginning approximately three hours before feeding time. On the first day, half of the mink were exposed to the stick test, and half to the glove test. The following day, this was reversed to control for effects of test order. The protocol from Experiment 1 was followed for both stimuli, except that, as discussed above, mink were allowed 20 s longer to respond. Location (whether in or out of the nest box) was also now noted at the start of each test. As before, if a sleeping individual did not wake up when the stimulus was placed in the cage on two 30-second attempts, no category was assigned.

The proportions of mink categorized as fearful and aggressive in each test were compared using Fisher’s exact tests, or McNemar’s tests where repeated measures from the same individuals were involved. Kappa statistics for agreement between the two tests were calculated, both overall and split by day.
Results

Temperament distributions

The proportions of mink assigned to each category are shown in Figure 2.4, and comparisons on an individual level in Table 2.3. One individual was excluded for remaining asleep during attempted tests, leaving 120 mink in the study. Significantly more mink were fearful in the glove than the stick test (McNemar’s test, $\chi^2=19.2$, $P<0.0001$). Test order affected fearfulness in the glove test: mink were more likely to exhibit fear if they were glove-tested on the first than on the second day (Fisher’s exact test, $P=0.0004$). Results of the two tests were therefore also compared separately for each day. The proportion of mink that were fearful was significantly higher using the glove test than the stick test on Day 1 (Fisher’s exact test, $P<0.0001$); the effect was in the same direction on Day 2, albeit not significantly (Fisher’s exact, $P=0.119$). No significant effect of test day on fearfulness was detected for the stick test (Fisher’s exact, $P>0.10$), but statistical power was very low (see Table 2.3) and the only three individuals that exhibited a fearful response to the stick were tested on the first rather than the second day.

Meanwhile, significantly more mink were aggressive in the stick test than the glove test (McNemar’s test, $\chi^2=18.2$, $P<0.0001$). As above, test day affected the glove test: more mink tended to exhibit aggression in this test if this was the second test to which they were exposed ($\chi^2=3.1$, $P=0.076$). Again, test day did not significantly affect the stick test (Fisher’s exact, $P>0.10$).
Correlation between tests

Table 2.3 compares the two tests’ results on an individual basis. Of 120 mink, 68 (57%) were assigned the same category using both tests. This produced a kappa of 0.2729. For agreement in simply categorizing mink as fearful or not, the kappa value was very low at 0.0969, considered poor reliability. Relatively few mink were in the nest box at the beginning of the test, and so reliability was not improved by excluding them (kappa=0.2669, N=111).

Discussion

Our previous impressions from pilot trials were confirmed: the stick test categorized hardly any Ontario mink as fearful, but many as “aggressive”. In contrast, the lowest reported prevalence of fear in adult females in Scandinavian studies using the stick test was 22% (Møller and Hansen, 2001) compared to our 2.5%. The prevalence of aggression was also low, e.g. 12% (Møller and Hansen, 2001) versus our 41%. Experiment 2 also supported our hypothesis that the glove test would be more sensitive than the stick test to fear, and less prone to eliciting aggression. Contrary to our expectations, the mink’s initial location did not seem relevant in this experiment, but this was likely because only nine individuals were in the nest box at the start of either test, and most were awake so unlikely to be ‘drowsy’.

Because of the shifts in the number of mink exhibiting fear or aggression, agreements between stick and glove tests were poor, especially for categorizing fear. This is somewhat surprising given that past research suggests that fear in the stick test generalizes to other tests (Malmkvist and Hansen, 2002). However, the floor effects on the stick test data here (which
classed only three of 120 mink as fearful) made it statistically impossible to determine whether individuals that were fearful in that test were also fearful in the glove test. A second possible explanation for the lack of agreement between tests is that both the test person and general testing procedure were completely novel on the first test day, but not the second. The glove test seemed especially sensitive to such ‘order effects’, as discussed below.

The glove test’s sensitivity to fear appeared affected by the novelty of the situation (e.g. the unfamiliarity of the test person): more mink were fearful if given this test on the first rather than the second day. The reverse was true for aggression. These findings are rather consistent with the habituation data from Experiment 1, and suggest that being naïve rather than having been previously tested (in any test) makes mink more fearful and less aggressive in glove tests. In Experiment 2 (unlike Experiment 1), the subjects were also naïve to the experimenter on the first day: unfamiliarity of the tester may thus interact with the stimulus presented to affect the sensitivity of a fear test. A longer interval between tests might reduce this ‘day effect’, but if maximum sensitivity to fear is needed for the purposes of research or animal selection, then these data suggest results of the first glove test should be used.

**Experiment 3: Reproductive correlates of fear in the glove test**

This final experiment aimed to determine whether fearfulness predicted poor reproductive success, since reproductive suppression is a common consequence of chronic stress (Wingfield and Sapolsky, 2003), and fear of humans specifically can reduce reproductive output (see Introduction).
Methods

Subjects and locations

Subjects were 546 female mink in their second or third breeding seasons, spread across three farms. All had been mated in late February. Of these, on Farm 1 there were 147 Blacks and 148 Pastels, on Farm 2, 140 Blacks, and on Farm 3, 111 Pastels. Farms 1 and 2 were those used in Experiments 1 and 2, respectively. One researcher worked at Farm 1, another at Farms 2 and 3.

Test protocol and measures of reproductive success

Due to concerns about reliability, we continued testing each mink three times. This allowed us to eliminate individuals that changed responses seemingly randomly, since we were not confident that these responses reflected underlying temperament. The results of the sensitive first test were then used to test our hypotheses for all remaining mink. The protocol was thus the same as in Experiment 2, but repeated three times over a four-day period; and mink categorized according to their response on the first test, but only if they were then consistent across all three tests or if any change was in a direction predicted by habituation (see Experiment 1). Mink not meeting this criterion were excluded from analyses. To control for the possible confound of ‘drowsiness’, reliability was then compared between individuals that were sometimes in the nest box during tests and those that were not. All testing occurred in early April, during the period between mating and being moved to the cages where they would give birth.

Reproductive success was assessed via failure to have a litter (hereafter termed ‘barrenness’), litter size counted on the day after birth, and kit mortality in the first three weeks,
assessed as a proportion of original litter size. Body fat was scored by the farmers on a three-point scale just after mating, because this influences reproduction (Baekgaard et al., 2007; Malmkvist and Palme, 2008).

**Statistical analysis**

Glove test reliability was assessed again using kappa statistics, overall and split by farm. Population level changes in habituation were assessed using McNemar’s tests comparing Tests 1 and 3 for each category, split by colour; for simplicity of presentation, farms were pooled unless the relationship differed between them. Likelihood of barrenness was compared between fearful mink and all others within each colour-type using Pearson’s Chi-square tests, or Fisher’s exact tests if any cells had counts of less than five. Because litter size and infant mortality data were not normally distributed according to Shapiro-Wilk tests, generalized linear models were used to test whether fearfulness predicted either variable. For testing whether fearfulness predicted litter size, we used a model with a Poisson distribution and a log link function. The relationship of infant mortality with fearfulness was assessed in a binomial logistic regression. Both models controlled for farm, colour-type and, in the case of litter size, body fat score. All analyses were conducted in JMP 8 (SAS Institute Inc., NC, USA).

**Results**

**Reliability of the glove test and habituation effects**

Sixteen mink were excluded because they were not tested three times due to remaining asleep. For the 530 others, the overall kappa was 0.33. However this value was affected by whether mink
were in the nest box when tested: thus for the 405 individuals that were outside the nest box
during all three glove tests, the kappa statistic for agreement among all three was 0.38
(“moderate”), compared to just 0.002 for the 125 that were in the nest box at least once.
Reliability also differed between farms: overall kappa values were 0.40, 0.21 and 0.28 at Farms
1, 2 and 3, respectively.

As in Experiment 1, population-level changes in temperament from the first to the last test
were consistent with habituation (Figure 2.5). Fear decreased from 13% to 4% in Blacks
(McNemar’s $\chi^2=13.4$, P<0.0001), and no Pastels were fearful on Test 3 while 5% had been on
Test 1. Unresponsiveness increased from 5% to 31% in Blacks and from 10% to 31% in Pastels
(McNemar’s $\chi^2=66.4$, 43.8 respectively; both P<0.0001). Aggression increased only at Farm 1,
in both colour-types (Blacks: 24% to 38%, McNemar’s $\chi^2=13.4$, P=0.0002; Pastels: 36% vs.
49%, McNemar’s $\chi^2=8.5$, P=0.004).

Colour and farm effects on fear

Overall, 12% of categorizable mink were fearful in the first test; 45%, curious; 37%, aggressive,
and 7% unresponsive. As in Experiment 1, Blacks were more likely than Pastels to be fearful on
Farm 1, the only farm with both colour-types: $\chi^2=13.4$, d.f.=1, P=0.0002). Fearfulness also
differed between farms: a chi-square test between farm by colour groups was significant
($\chi^2=12.8$, d.f.=3, P=0.005), due to a difference between Pastels at Farms 1 and 3 ($\chi^2=8.9$, d.f.=1,
P=0.003).
Fear vs. reproductive success

Among Pastels, fearful mink were more likely to be barren (Fisher’s exact, P=0.006; Figure 2.6). When split by farm, this difference appeared only at Farm 3 (Fisher’s exact P=0.04; 44% of 9 fearful individuals vs. 14% of 86 non-fearful individuals); but this could reflect low statistical power since only four Pastels were barren at Farm 1 (and Farm 2 had no Pastels). There was no significant difference between fearful and non-fearful mink among Blacks. Amongst mink that successfully produced kits, no significant relationship between fearfulness on litter size or kit mortality were detected (P>0.05).

Discussion

Previous studies provided evidence of relationships between fearfulness and reproduction in mink, but its manifestations have varied. Malmkvist and colleagues (1997) found that mink from a line selected for fearfulness mated later in the season than did those from a confident line, suggesting reduced motivation to mate; however, once mated, their reproductive success was unaffected. Korhonen and colleagues (2002), by contrast, found no difference in mating dates, perhaps because mating attempts began later in this experiment and thus both groups were ready to mate by the time of testing. In this experiment, however, females who were fearful in stick tests gave birth to smaller litters than did confident females. Among primiparous females, rates of infertility were also higher in fearful individuals, as in Pastels in the current study, but this difference was not statistically significant. Fearfulness has likewise been linked to infertility or decreased rate of breeding in a variety of other captive species, from chickens (Shabilina, 1984) to cheetahs (Wielebnowski, 1999). The lack of significant relationships between fearfulness and
litter size or kit mortality may be due to the low power of the tests rather than a true lack of relationship, since the fearfulness data are qualitative, and relatively few individuals were fearful. The General Discussion discusses potential mechanisms mediating the link between barrenness and fearfulness in our mink.

**General discussion**

In stick tests, our Ontario mink did not react as expected from published Scandinavian data, being far less fearful and more aggressive. It is unclear why – differences in age, genetics, husbandry, and subtle aspects of the test such as season and the height and sex of testers could all play roles – but our need for a test with greater resolution and sensitivity, applicable to North American-style cages that preclude using the Trapezov hand test, led to us successfully modifying the stick test into a glove test. This combined the basic stimulus of the Trapezov hand test with the advantages of the stick test, being quick and simple because conducted from outside the cage, and therefore able to be used on large numbers of animals and by testers inexperienced at handling mink. Furthermore, the glove test did indeed prove more sensitive than the stick test for investigating fear, increasing the number of mink classified as fearful in our tests by about eightfold. This is important for populations where fearfulness is relatively rare, and so where tests using more neutral stimuli might not detect fear in enough animals to assess its correlates reliably.

The glove test thus adds to a range of existing temperament tests that differ in how challenging they are for mink, ranging from the Trapezov hand test at the most stressful end of the spectrum, best for identifying individual differences within very confident populations.
(Kirkden et al., 2010), through to the least threatening “food” or “tit-bit” tests (where the human offers food; see Hansen and Møller, 2001) and “walk by” tests (where the tester simply passes the cage; e.g. Harri et al., 1998) most suitable for very fearful populations where other tests would yield ceiling effects. As an additional nuance, our data showed the glove test to be more fear-inducing on the first test, and likely even more so when conducted by an unfamiliar tester; the sensitivity of the glove test can thus be modified according to how exactly it is applied. In Experiment 3, for instance, classifying mink according to sustained reactions instead of just the first test would have left us with only three fearful individuals (although these were undoubtedly very reliably fearful). Both which test and what method of application is ideal will therefore depend on the population characteristics, and also on the test’s aims. For example, the glove test might be useful in genetically selecting for very calm animals at the population level, and where fear levels are already fairly low; in this case, maximum sensitivity and speed are desirable, and farmers might only conduct the test once. For research, however, it is more crucial that the classification accurately represents underlying temperament, and so multiple tests may be preferable in order to reduce noise by eliminating individuals whose responses change unpredictably (see Methods).

Glove tests also proved to have good construct (convergent) validity, and practical usefulness. Construct validity was good when assessed formally (Experiment 1), and the observed differences between colour-types provided further, albeit less formal, support for the test’s validity. The glove test’s validity was also supported by its usefulness in predicting (or detecting in advance) a biologically relevant variable, barrenness. This observed relationship between barrenness and greater fearfulness could be mediated in several different ways, which can now be empirically investigated. Fearfulness could reduce reproductive success through the
activation of the HPA axis, which adversely affects reproductive hormones (Sapolsky et al., 2000; von Borell et al., 2007). Fearful individuals may also copulate less frequently, due to decreased receptivity to males or because the difficulty of handling these individuals results in farmers making fewer attempts to breed them. In the current study, mink were tested for fearfulness after they should have been pregnant; it is thus also possible that any that had not successfully conceived were in a different hormonal state from pregnant females, with this then affecting responses to the test. Until this possibility has been investigated, it is uncertain whether fearfulness is truly predictive of reproductive success. As a final testable hypothesis, those that had many unsuccessful mating attempts might have been handled more often, and so been more likely to develop fear of humans with handling gloves. The test’s decision validity (ability to predict performance, i.e. its usefulness for making management decisions) remains to be evaluated, for example by determining whether it can predict barrenness when assessed before the mating season, or be used in genetic selection for temperament.

Categorical tests such as the glove or stick test have both advantages and disadvantages for scientific use. They differentiate between qualitatively different responses in a way that quantitative measures may not; for example, latency to approach is commonly used as a measure of fear in novel object tests, but does not distinguish adequately between fear and unresponsiveness (a true lack of interest in the stimulus). However, as mentioned above, qualitative data do not provide high statistical power. When sample sizes are small, it may therefore be necessary to combine qualitative analysis of behaviour with quantitative measures such as minimum or mean distance from the stimulus (e.g. Malmkvist and Hansen, 2002).
Reliability of the glove test was not ideal: most mink did not show identical responses each day when tested daily. Nevertheless, this is not grounds for discarding the glove test in favour of other tests. First, if conventional reliability tests for the other widely-used tests for mink temperament have ever been conducted, they have not been published in accessible literature, so their reliabilities are essentially unknown. Second, most changes in reaction seemed to be due to normal processes of habituation. Third, reliability appeared to increase over repeated tests, as if responses after the first or second test are more stable. Fourth, there seems scope for improving reliability through training the tester, in that a possible explanation for lower reliabilities at Farms 2 and 3 in Experiment 3 was that the researcher here was less experienced than that working at Farm 1. Fifth, reliability was greatly improved by excluding mink that were awake but in the nest box at testing, and might be improved yet further still by additional refinements in categorization as suggested below.

Our data highlighted “unresponsiveness” (sometimes called “undecided”) as a category needing further research. We suspect that this category encompasses mink that are genuinely uninterested in the stimulus; mink that are scared and so ‘freezing’; and mink that are only half-awake. Refining future methods might therefore involve better controls for each mink’s ‘drowsiness’ at the time of the test, for example by excluding only mink with eyes closed or head tucked prior to the test if this is visible in the nest box. Our method of identifying ‘freezing’ in mink that are standing as far as possible from the tester by their immobility and sustained attention to the stimulus (see Methods) could also be extended to mink in any location. Future research could assess sympathetic responses (e.g. via telemetry devices) to determine whether mink suspected to be ‘freezing’ are indeed frightened by the stimulus. “Aggression” is another category whose motivational basis and welfare significance remain unclear. From our data, it is
evident that aggression levels depend on the test stimulus. The lower aggression in the glove test than the stick test does not seem to be due to greater fear of the glove inhibiting aggression, since most mink that were aggressive only towards the stick were curious, not fearful, in the glove test. This is consistent with our suggestion in the Introduction that “aggression” towards the stick may reflect something other than a response to the human tester, such as predation or exploration directed at the small object. This hypothesis could be tested by comparing test results with their responses to being handled and with their feeding motivation, as well as testing whether aggression increases if they are always allowed to carry the stick away at the end of the test. Similarly, the increased aggression over repeated tests in Experiments 1 and 3 cannot be explained by initial fear: mink that became aggressive had typically begun as curious. Some sensitization may have occurred with repeated disturbances by a human, a hypothesis that could now be tested by comparing results of a single test between animals that had been frequently disturbed by humans in other ways over the preceding days.

Thus in the glove test, as in other temperament tests, further research could refine how mink are classified, and also yield a better fundamental understanding of what some responses mean. However, even without these additional data, when used carefully by trained researchers, the glove test is a useful, valid tool for assessing mink fearfulness.

Conclusions

The glove test is a promising and valid test of temperament, including fearfulness, for farmed mink, especially useful for populations where very few mink exhibit fear in the stick test. Some further refinement is needed to improve its repeatability and thereby reduce data loss from
individuals that are inconsistent across tests. Using the current method, however, we were able to detect expected temperament differences between colour-types, and a negative relationship between fearfulness and reproductive success. This suggests that reducing fearfulness through selective breeding or changes in management may improve mink productivity as well as welfare.

Acknowledgements

Many thanks are due to the owners of these farms, Kirk Rankin, Ted Parkinson and Lyn Parkinson, for their cooperation. We would also like to thank Mike Walker and anonymous reviewers for their comments on earlier drafts of the manuscript, and NSERC and the Canada Mink Breeders Association for funding.
Table 2.1. Pairwise reliability comparisons for glove tests repeated once a day for five days.

<table>
<thead>
<tr>
<th>Tests compared</th>
<th>Kappa (N=160)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 vs 2</td>
<td>0.3184 ± 0.0567</td>
</tr>
<tr>
<td>2 vs 3</td>
<td>0.3623 ± 0.0535</td>
</tr>
<tr>
<td>3 vs 4</td>
<td>0.4904 ± 0.0553</td>
</tr>
<tr>
<td>4 vs 5</td>
<td>0.4574 ± 0.0805</td>
</tr>
</tbody>
</table>

Values are kappa ± SE. Kappa values greater than 0.40 indicate “moderate reliability” (Dohoo et al., 2003)

Table 2.2. Behaviour in validation tests: tests for unequal variances between temperament categories, as determined by minks’ immediate responses to the stimulus.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Curious (N=22)</th>
<th>Fearful (N=9)</th>
<th>Unresponsive (N=15)</th>
<th>Levene’s test P-value(a)</th>
<th>Difference in means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration of contact (s)</td>
<td>103.0 ± 6.0</td>
<td>58.0 ± 6.2</td>
<td>57.1 ± 10.5</td>
<td>0.08</td>
<td>(F_{2,43}=12.10, P&lt;0.0001)</td>
</tr>
<tr>
<td>Duration of ambivalence (s)</td>
<td>1.0 ± 3.0</td>
<td>30.6 ± 4.7</td>
<td>9.4 ± 3.6</td>
<td>0.0002</td>
<td>(F_{2,14.5}=14.16, P=0.004)</td>
</tr>
</tbody>
</table>

Data presented are means ± SE; test duration was 180 s.
\(a\) Test for equal variances among the three categories.

Table 2.3. Comparison of results in the glove test vs. stick test applied to the same sample of mink.

<table>
<thead>
<tr>
<th>Glove Test:</th>
<th>Stick Test:</th>
<th>Aggressive</th>
<th>Curious</th>
<th>Fearful</th>
<th>Unresponsive</th>
<th>Totals for glove test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggressive</td>
<td>23</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Curious</td>
<td>23</td>
<td>42</td>
<td>1</td>
<td>0</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Fearful</td>
<td>3</td>
<td>18</td>
<td>2</td>
<td>1</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Unresponsive</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Totals for stick test</td>
<td>49</td>
<td>66</td>
<td>3</td>
<td>2</td>
<td>120</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.1. Population-level changes in temperament distribution over repeated tests (Experiment 1).
Proportions of mink falling into each temperament category, including only the 75 mink for which five tests could be completed. Note that no directional prediction was made for the observed change in aggression.
Figure 2.2. Average total duration of contact with glove over 3 min test, grouped by immediate response.
Data are means ± standard error.

Figure 2.3. Average total duration of ambivalence over 3 min test, grouped by immediate response.
Data are means ± standard error of time spent alternating between approach and avoidance of the glove.
Figure 2.4. Comparison of temperament distributions for stick and glove tests. N=120, all exposed to both tests.

Figure 2.5. Population-level changes in temperament distribution over three tests (Experiment 3). Proportion of all mink tested, with colour-types pooled (N=530).
Figure 2.6. Proportion of fearful vs. non-fearful mink that failed to reproduce. N= 237 Black (23 fearful), 236 Pastel (10 fearful).
CHAPTER THREE

Inactive mink have fewer but potentially higher quality offspring

Rebecca K. Meagher, Allison Bechard, Georgia J. Mason

A version of this paper was submitted to Ethology, July 28, 2011

Abstract

Activity levels have been linked to other personality traits (e.g. boldness), and can affect fitness, including offspring performance. We compared reproductive strategies in very high and low activity phenotypes in an annual breeder, American mink (*Neovison vison*). Fur-farmed populations yield large sample sizes and allow ready investigations into maternal behaviour, reproductive success, offspring performance and their inter-relationships. Very inactive individuals on farms generally have smaller litters, and this held in our study populations (*F*₁,₁₇₈=4.49, *P*=0.035). We tested two competing hypotheses to explain this effect: 1) inactive individuals experienced chronic stress and/or depression-like ‘apathy’, which predicts female-skewed litters, less maternal care and poorly performing kits (slower growth/higher mortality); and 2) inactive individuals employ an alternate reproductive strategy, providing more care to each kit and having improved individual kit performance. Pre-parturition, inactive females’ nests contained more fur (*F*₂,₃₈₉=3.47, *P*=0.033), although nests were also less structured (colour-type*nest shape *F*₇,₄₄₆=2.97, *P*=0.005). Their kits, especially sons, grew faster (both sexes *P*<0.001), even after statistically controlling for litter-size and dam body condition (inactive dams being fatter). In kit retrieval tests, inactive females were more responsive to their sons than to their daughters: both more likely to touch them (*P*=0.034), and amongst mothers that did so,
faster than active dams to reach their sons (F_{1,153}=7.05, P=0.009). Kit growth rates and dam latencies to touch them also co-varied, suggesting consistent differences in maternal style. By 21 days, inactive and active dams did not differ in total kit biomass, despite the former’s smaller litters. Hypothesis 2 was thus supported: inactive females favour offspring quality over quantity, investing more resources in fewer kits, particularly males. This potentially boosts their sons’ adult fitness. More broadly for lab-based studies, possible ‘captivity effects’ on the fitness correlates of activity are discussed.

**Introduction**

This study tests two competing hypotheses, respectively from applied animal welfare and human psychology, and from fundamental behavioural ecology on personality or behavioural syndromes, to investigate why very inactive captive mink often have small litters. Fur-farmed American mink (*Neovison vison*) illustrate the extreme variation in individual activity levels that is potentially present within populations. At the most active times of year (winter) and day (before the daily meal; Hansen and Møller, 2008; Mason, 1993), some mink are active for 80% or more of the time, largely performing stereotypic behaviours like repetitive pacing (Bildsøe et al., 1991; Svendsen et al., 2007a), while others spend 90% of this period recumbent in their nest boxes (Meagher and Mason, 2008). These individual differences predict 24h time-budgets: mink that are more or less active at these times are correspondingly more or less active overall (Mason, 1992; Svendsen et al., 2007b). They are also consistent across long time periods (Mason, 1993) and different housing conditions (M. Diez-León and G. Mason, pers. comm.), and are heritable (e.g. Hansen, 1993). The phenotypes also differ in other traits: stereotypic (active) mink show increased routine-proneness compared to inactive conspecifics (Dallaire et al., 2011. They can
also show reproductive differences in terms of annual litter size, with highly stereotypic, active females, producing relatively large litters but very inactive females, relatively small ones (Jeppesen et al., 2004; Mason, 1992; Meagher et al., 2010).

For captive animals on farms and in zoos, such differential activity is often studied and interpreted from an animal welfare perspective. Inspired by laboratory rodent and human clinical findings that chronic stress can induce ‘apathetic’ depression-like states (cf. Anisman and Matheson, 2005; Willner, 1997), very inactive subjects have been hypothesised to be apathetic (e.g. Cronin, 1985; Wielebnowski et al., 2002), ‘lethargic’ (e.g. Wemelsfelder et al., 2009; Zeller, 1991), or ‘depressed’ (Carlstead, 1996; Carlstead et al., 1999). Their inactivity can also reflect hiding, resulting from chronic fear (e.g. Carlstead et al., 1993; Rochlitz et al., 1998). In contrast, the stereotypic behaviours common in highly active subjects, despite indicating poor welfare at the population level, at the individual level seem often linked with lower stress levels than those of more inactive conspecifics, suggesting better abilities to cope with sub-optimal housing (Mason and Latham, 2004). There is some evidence this might be the case in farmed mink: when fearfulness co-varies with activity (demonstrated in some but not all studies; e.g. no relationship found in Korhonen et al., 2002), inactive mink are more likely to be timid (Hansen and Jeppesen, 2006). The small litters of inactive female mink are also potentially consistent with this view, because across multiple species, chronic stress decreases average litter sizes (Hemsworth, 2003; Wingfield and Sapolsky, 2003).

Behavioural ecology research on personality and behavioural syndromes offers another perspective, however. Differences in activity might reflect personality traits that are not directly related to stress, or to the overall quality of the individual. Personality refers to a set of
behavioural characteristics in an individual that are consistent over time and across contexts (Gosling, 2001; Réale et al., 2007). Inter-correlated personality traits are sometimes termed ‘behavioural syndromes’ (Réale et al., 2007); for example, activity, exploratory behaviour and boldness are frequently co-associated in a single behavioural syndrome (Sih et al., 2004).

Boldness has also been linked with increased aggression (Koolhaas et al., 1999) and tendencies to form routines (Verbeek et al., 1994); and low exploration with tendencies to show immobility and little aggression under stress (Koolhaas et al., 1999). Given the stability of minks’ individual differences in activity, this trait could therefore be considered an aspect of personality, and perhaps part of a broader behavioural syndrome. In the wild, the stable existence of alternative personalities or behavioural syndromes is thought maintained by a range of evolutionary processes. Sometimes they reflect variation in fitness, with relatively unfit (e.g. subordinate) individuals ‘making the best of a bad job’ via timid, non-aggressive phenotypes (Dall et al., 2004). In other cases, however, they represent alternative strategies with similar fitness, perhaps differentially favoured according to current circumstances, e.g. trait frequency within the population (Réale et al., 2007) or food abundance (Boon et al., 2007). In such cases, the different syndromes may diverge in trade-offs between the various life history traits that affect fitness. To illustrate, activity/boldness has been linked to trade-offs between mortality and fecundity or growth (Biro and Stamps, 2008), with active, bold animals showing greater growth and fecundity, but shorter lifespans. A related hypothesis is that individuals differ in “pace of life” (Careau et al., 2008; Réale et al., 2010). Here, bold or explorative individuals are considered to have a “live fast and die young” strategy (Careau et al., 2010), producing many offspring quickly but having shorter lifespans (Wolf et al., 2007), while timid individuals are expected to live longer and reproduce more slowly, allocating more maternal care to these young (Réale et al., 2010) – trade-offs resembling those seen across species (Jeschke and Kokko, 2007). Reproductive strategies co-
varying with personality occur in both human and non-human animals; in women of the same socioeconomic class, for example, the balance between investment in offspring quantity versus quality relates to maternal personality, women scoring lower in ‘neuroticism’ having fewer but larger (likely better nourished) offspring (Alvergne et al., 2010). This suggests an alternative hypothesis for mink: that active and inactive phenotypes do not differ in quality, but instead are adopting different reproductive strategies. There is preliminary evidence for this hypothesis: kits raised by inactive/non-stereotypic foster-mothers grow faster than do those fostered to active/stereotypic dams (Mason et al., 1995).

These two hypotheses – that profound inactivity reflects chronic stress, or that these behavioural phenotypes reflect alternate life history strategies – make competing, testable predictions. The ‘stress hypothesis’ predicts compromised reproductive function (Hemsworth, 2003; Wingfield and Sapolsky, 2003), including impaired offspring growth (as occurs in mink kits whose dams are experimentally stressed [Dobson et al., 2008], and sometimes in the children of depressed mothers [Stewart, 2007]), and increased offspring mortality rates (Janczak et al., 2003). Chronic stress can do this physiologically, disrupting reproductive endocrinology (Wingfield and Sapolsky, 2003) or behaviourally, reducing maternal responsiveness (e.g. Bahr et al., 1998); depressed women, for example, may show decreased attention to their children (Stewart, 2007). Furthermore, females in poor condition (e.g. stressed individuals) should also produce female-biased litters, assuming that other factors (e.g. sex-specific philopatry) are not at work (Trivers and Willard, 1973; Faust and Thompson, 2000), as is the case for mink. This prediction, originally made by Trivers and Willard (1973), is based on the assumption that maternal investment affects male offspring reproductive success in adulthood (see e.g. Wells, 2003), which is more variable than female reproductive success in polygynous or promiscuous
species. Supporting evidence has been found in several species where such assumptions hold: stressed dams produce more female than male offspring (e.g. silver foxes, *Vulpes vulpes*: Bakken, 1995, 1998). If inactive female mink have smaller litters because chronically stressed or depressed, one should therefore expect slower kit growth, higher mortality rates, reduced maternal care (e.g. less responsiveness to distressed kits), and female-skewed sex ratios.

In contrast, the ‘alternate strategies’ hypothesis does not predict any of the above effects. Instead, it predicts trade-offs between offspring number and quality, with inactive mothers showing more maternal care, and their kits showing enhanced growth and reduced mortality.

**Methods**

**Subjects**

A total of 549 females entering their second or third breeding season, were observed across three farms in south-western Ontario. At Farm 1, we observed mink of two colour-types (lines bred for fur colours, which also differ in temperament and average activity level): 149 Blacks and 144 Pastels. The other farms had only one colour-type each: we sampled 143 Blacks at Farm 2, and 113 Pastels at Farm 3. Mink were individually housed in wire-mesh cages with a wooden nest box. Nesting material was provided beginning shortly before any mink gave birth; at Farm 1, wheat straw was provided, while Farms 2 and 3 instead used aspen shavings. Exact cage-sizes varied but met Canadian Codes (being 61 cm x 19 cm x 46 cm or greater). Mink had access to water *ad libitum* and were fed once daily. The University of Guelph’s Animal Care Committee approved the research.
Pre-mating behaviour and maternal characteristics

Subjects were screened for inactivity in February. Live behavioural scanning was conducted daily from 09:00 until feeding began (c. 13:00-13:30), for nine or ten days. Individual variation is likely to be most evident at that time since it is when most mink peak in activity. Live scanning has been validated by correlation with video data (Bildsöe et al., 1990; Svendsen et al., 2007b). Analyses were based on percent of scans spent inactive (lying down/out of sight in nest box), and stereotypic behaviour (defined as a sequence repeated at least three times consecutively) was recorded separately from ‘normal’ activity.

Inactive mink tend to be fatter (Hansen and Møller, 2008; Jeppesen et al., 2004), which could have a confounding effect on offspring sex ratios (Cameron, 2004) and infant growth. To control for this, farmers scored animals for body condition on a 5-point scale, from very thin to very heavy, where three represented the ‘ideal’ (cf. Hynes et al., 2008). To attain acceptable reliability, scores were pooled into three simpler categories: thin, normal and overweight (Meagher et al., submitted provides details).

Litter characteristics and kit performance

Kits were counted, sexed, and weighed on post-natal days (PND) 1 and 21: a period when kits depend fully on mothers for food (Dunstone, 1993; Mason, 1994) and thermoregulation (Harjunpaa and Rouvinen-Watt, 2004). Weights were averaged for each litter, split by sex. PND 1 sexing can be unreliable; sex ratios were therefore calculated at PND 21. Growth rates were calculated as differences in average kit weight between PND 1 and 21, divided by 20 days, for
each sex. Kit mortality distinguished between deaths during the first 24 hours, including
stillbirths, and mortality after that time. Farmers cross-fostered some kits, or rescued kits that had
been chilled, according to their standard farm practices. However, litters in which this occurred
excluded from the litter size and sex ratio analyses.

Maternal behaviour

Three characteristics of nest-building were scored: the presence/absence of nest material, shape
that material was formed into, and amount of fur incorporated. In rabbits, where this is best
studied, using fur in nest-building is a maternal behaviour controlled by progesterone and
prolactin (Zarrow et al., 1961) and improves offspring growth and survival (Marai and Rashwan,
2003). Wild mustelids similarly line nests with fur (King and Powell, 2007), as do domestic
ferrets (Roberts, 1977), although mechanisms and effects on their young are unknown. Fur was
scored on a 3-point scale: 0) no fur; 1) a few clumps; 2) fur throughout nest. The presence of
nesting material (e.g. straw) influences mink kit growth and survival (Malmkvist and Palme,
2008). Furthermore, the shape of structure created affects offspring performance in other species
(Andersen et al., 2005; Deacon, 2006; Pedersen et al., 2006). Nest-shape was therefore rated from
1 to 7, where 1 signified that all nesting material was ejected, with none present around kits, and
7 signified that a fully enclosed nest (details in Table 3.1; cf. Malmkvist and Palme, 2008). The
maximum possible score was only 5 at Farms 2 and 3 because the nesting material provided
(shavings) could not be woven to form a cover as straw can. Nests were scored every second day
from late April until a score was obtained for each mink one-two days before giving birth and
one-two days after.
Kit retrieval tests (Malmkvist and Houbak, 2000) were used to assess maternal responsiveness in a subset of females selected via February inactivity levels; 25 of the most inactive and 25 of the least inactive females from each farm and colour-type were included (excluding females with fewer than four kits, to reduce potentially confounding effects of litter size): 200 in total. Tests occurred in the mornings of PND 5. One kit per litter was selected at random, its sex noted because this has been demonstrated to influence retrieval time (Clausen et al., 2008), and then placed in the cage centre, head facing the nest box. The time taken for the mother to touch it and return it to the nest box were recorded (to a maximum of 180 seconds to prevent kits becoming chilled).

**Statistical analysis**

Analyses were conducted in JMP 8, or SAS 9 (SAS Institute Inc., NC, USA). Females that did not produce litters were excluded from all analyses. Relationships between inactivity and temperament, body condition, litter size, sex ratio and growth rates were analysed using general linear models (GLMs), controlling for farm, colour-type, and, where appropriate, litter size. Relationships of interest between dependent variables were analysed similarly. Farm was a random effect, and restricted maximum likelihood (REML) methods of estimation were employed as recommended for mixed models (Searle et al., 1992). Exceptions were sex ratio models, which would only run if farm was treated as fixed; and models for kit growth including both sexes, where farm and colour-type were combined into a single fixed effect labelled ‘group’ to allow nesting of individual within group. Effects of farm, colour-type and litter size are not presented here unless interacting with variables of interest. Since nest scores were ordinal, inactivity was used as the dependent variable to allow use of GLMs. Tests were two-tailed.
Normality and homogeneity of variance were assessed by inspection of residuals (Grafen and Hails, 2002); Bartlett’s test for equal variances was used for categorical effects. Data were transformed where needed to meet assumptions of parametric tests (e.g. inactivity was logit-transformed in nest score analyses, while latencies to touch/retrieve kits were log-transformed). Where data could not be so transformed, non-parametric generalized linear mixed models were used (e.g. for relationships between kit mortality and dam inactivity; nest scores; and retrieval times). Two extreme outliers (probably measurement errors) were removed from male growth analyses.

Kit retrieval tests showed ceiling effects because some mothers did not retrieve within the allotted time. Survival analyses were considered but their assumptions were violated: the active/inactive groups’ survival curves crossed (Kleinbaum and Klein, 2005). Probabilities of failure to respond were therefore analysed separately using chi-square tests, and non-retrieving mothers excluded from further analyses of latencies.

Results

Dam behaviour and litter characteristics

Subjects varied greatly in activity; the least active spent up to 98% of the observation period inactive, while the most active were never recorded inactive. As expected, body condition and inactivity co-varied ($F_{2,486}=28.0$, $P<0.0001$). Litter sizes averaged 5.9 kits (range: one to ten).
As expected, inactivity predicted small birth litter sizes (N= 343, F_{1,178}=4.49, P=0.035).

There was a non-significant trend for an interaction between colour-type and inactivity predicting sex ratio (proportion of litter that is male; F_{1,315}=2.90, P=0.090). Split by colour, a non-significant trend for a negative correlation with inactivity existed in Pastels only (F_{1,143}=3.19, P=0.076)\(^1\).

**Kit weight, growth and mortality rates**

Inactivity did not predict individual kit weight on PND 1 (Table 3.2). For kit growth rates, dam inactivity and kit sex interacted (F_{1,800}=6.68, P=0.010). Split by kit sex, inactivity significantly predicted enhanced growth in both males and females, but for males, this interacted with litter size (see Table 3.2), and the slope of the inactivity main effect was steeper for males than for females (Figure 3.1). These results control for litter size, and thus are not simply by-products of inactive dams having fewer kits. Controlling for body condition did not eliminate this effect (males: F_{1,381}=42.2, P<0.0001; females: F_{1,382}=16.0, P<0.0001). Inactivity tended to predict relatively small total litter biomasses on PND1 (F_{1,293}=3.28, P=0.071), but due to these dams’ faster kit growth, this effect disappeared by PND21 (P>0.10). Inactivity levels did not predict kit mortality at either stage (both P>0.10).

\(^1\) Overall, there was a non-significant trend for slightly female-biased litters, with a mean proportion of 0.48 male kits per litter (Wilcoxon signed-rank test vs. expected mean of 0.5: T=-1.99, N=328, P=0.066).
Maternal behaviour

In kit retrieval tests, response failures did not differ between active and inactive groups (Blacks: 11% of inactive vs. 14% of active dams, \(\chi^2_1=0.167, P=0.682\); Pastels: 14% of inactive vs. 10% of active dams, \(\chi^2_1=0.308, P=0.579\)). However, inactive mothers were more likely to touch their sons than their daughters (5.7% vs. 20.5% ignored respectively, Fisher’s exact \(P=0.033, N=97\)) – an effect that was non-significant in active mothers (6.4% for sons vs. 18.6% for daughters; Fisher’s exact \(P=0.109, N=90\)). Response failures did not predict kit mortality or growth (\(P>0.10\)).

For mothers that did respond, there was no effect of inactivity or kit sex on total retrieval time (both \(P>0.10\)). However, for latency to first touch, inactivity and kit sex interacted (\(F_{1,153}=5.29, P=0.023\)), with males being touched faster by inactive than active dams (Figure 3.2). Litter size was unrelated to latencies to touch or retrieve kits (both \(P>0.10\)). Again, neither latency predicted kit mortality (both \(P>0.10\)). However, latency to touch predicted kit growth rates, averaged across sexes (Figure 3; \(F_{1,159}=7.96, P=0.005\)). Latency to retrieve also predicted kit growth in Pastels (\(F_{1,67}=5.63, P=0.021\)), with shorter latencies correlating with increased growth.

The presence of nest fur before parturition positively co-varied with inactivity (\(F_{2,389}=3.47, P=0.033\)), although such effects vanished post parturition. No dams ejected all nesting material from nest boxes, so no analyses were needed for simple presence versus absence of nesting material. However, nest shape scores before the day of parturition correlated
negatively with inactivity, an effect strongest among Blacks (colour*shape $F_{7,446}=2.97$, $P=0.005$). Nest shape post partum was, however, unrelated to dam activity levels ($F_{6,426}=0.81$, $P=0.558$).

**Discussion**

Farmed mink yield unusually large sample-sizes while also allowing direct observation of maternal behaviour, in a manner not feasible in the wild or even laboratory conditions. We used this model to test alternate hypotheses to explain reduced litter size in inactive females, by investigating other aspects of maternal investment: litter sex ratios, offspring growth rates and maternal care. This allowed us to determine, for the first time, how all these variables inter-relate in a single system.

We tested two potential explanations for the smaller annual litters of highly inactive mink: the ‘chronic stress’ and ‘alternate strategies’ hypotheses. Only one finding was potentially consistent with the stress hypothesis: inactive females had lower nest shape scores before parturition. However, our use of nest quality to indicate maternal care was rather exploratory, and like other authors (Møller, 1990), we found no evidence that this score is biologically relevant for farmed mink: it did not predict kit mortality or growth. Furthermore, inactive dams did not have elevated cortisol nor elevated fear (see Chapter 4); thus, physiological stress could not explain their reduced litter sizes.

The ‘alternate strategies’ hypothesis was far more convincingly supported. Kit weight, growth rate and retrieval data revealed that inactive females invested more resources per kit. Both effects were significant even after controlling for litter size: thus increased growth in kits with
inactive mothers did not just reflect trade-offs between litter size and body size of individual offspring due to finite resources (e.g. milk; Sikes and Ylönen, 1998); nor were retrieval results mere side effects of small litters resulting in fewer kits in the nest box competing for the mother’s attention (Clausen et al., 2008). Furthermore, although inactive dams were plumper, suggesting larger energy reserves to support lactation, dam inactivity levels predicted faster kits’ growth rates even when this was statistically controlled for. Combined with our finding that shorter latencies to retrieve were statistically related to faster kit growth, this strongly suggests that inactive dams’ enhanced maternal responsiveness plays a major role in their kits’ faster growth. Past research failed to find differences in the amount of time stereotypic and inactive mothers spend with their kits (Mason et al., 1995): all dams are very inactive during this period, showing negligible stereotypy and spending much of their time in the nest box. This suggests the explanation must be sought elsewhere, perhaps in milk quality or dams’ readiness to let milk down (nest fur data suggesting that prolactin levels could differ between the two phenotypes [cf. Zarrow et al., 1961]: one avenue for future research). Whatever the mechanisms, inactive dams’ pattern of smaller litters but faster kit growth and greater responsiveness to sons indicates no less investment overall, but instead the favouring of “quality over quantity”: an increased investment per kit consistent with the ‘alternate strategies’ hypothesis.

That inactive females’ post-natal maternal investment was biased towards males was unexpected, or at least not predicted by either hypothesis. The opposite effect was previously found in this species in the kit retrieval test (Clausen et al., 2008); however, this finding was unexpected and no evolutionary explanation has yet been established. At the population level, maternal care is often biased towards males in many species, likely because of the greater variance in male reproductive success than female and the influence of maternal attention on this
success (see Introduction). In wild mink, larger adult body sizes benefit males during competition for mates (Dunstone, 1993); while in farmed mink, large males copulate for longer (Thom et al., 2004), and in mate choice tests, larger, relatively inactive males win more copulations than do light bodied, highly stereotypic males (Díez-León et al., in prep.). Because inactive dams’ sons are likely to inherit low activity levels and thence heavy bodyweights (Mason et al., 1995; Hansen et al., 2010), they are thus inherently more likely to succeed with females in the mating season. It is thus unsurprising that these potentially successful breeders are favoured by their mothers. Inactive dams thus may be boosting their sons’ future success yet further, both by producing small litters, which our data show particularly benefits male kit growth, and, independently, by devoting more maternal care to them. These hypothesised links between inactive dams’ behaviour, their sons’ growth rates during the suckling period, and these males’ copulatory success in adulthood, now need direct test.

The evolutionary framework of the “pace of life” idea, and other adaptive explanations for behavioural syndromes underlying our ‘alternative strategies’ hypothesis, along with the sexual selection ideas presented above, both indicate the great potential value of applying functional, evolutionary approaches to farmed animals. However, they also raise little-explored questions about both the types of selective pressures such populations are under, and about the effects that captivity can have on the correlates of personality traits. Captivity can influence behaviour in multiple ways, and our subjects certainly showed extremes of activity likely absent in free-living wild conspecifics. We estimate that our least active mink curled up in their nest boxes for 23 hours or more daily (cf. the 70-86% of every day spent in dens by wild conspecifics: Dunstone, 1993), while at the other extreme, our most stereotypic mink paced and head-twirled abnormally for hours, and extrapolating from their average speed, some paced over 12 km/day –
further than the typical distances covered by wild mink (Gerell, 1970). One effect of captivity is that animals show immediate behavioural differences from wild conspecifics caused by the constraints placed on some activities (e.g. flying, ranging widely) and the loss of need for others (e.g. food provisioning removes needs to forage). If captive-bred, animals can also show further behavioural and physiological changes caused by unnatural rearing conditions (e.g. fertility problems, premature mortality, and more stereotypic behaviour: Clubb et al., 2009; Jones et al., 2011; Latham and Mason, 2004). In addition, altered selection pressures over multiple generations cause further behavioural changes (e.g. McDougall et al., 2006; McPhee, 2004; Price, 1999). Captivity will thus often affect the range of behavioural phenotypes expressed, and potentially which traits cluster in syndromes. It may also affect the correlates of a given phenotype, for example by altering the proximate fitness-related outcomes of behaviour in terms of mortality or resource acquisition. Although many fundamental, evolution-oriented studies of animal personality are laboratory-based, some using subjects captive-bred for generations and/or small, non-enriched cages likely to induce stereotypic behaviour or extreme inactivity (e.g. Careau et al., 2008; Wilson et al., 2009), these important issues and their implications for data generality seem little acknowledged as yet in such research (see Smith and Blumstein, 2008 and Archard and Braithwaite, 2010 for two welcome exceptions).

For farmed mink, we consequently cannot tell whether the observed phenomena reflect evolutionary legacies from their wild ancestors (this species only has been farmed for c.110 years; Hansen 1996), selective pressures acting within farms, or both. Thus if inactive females truly have slower-paced life histories, for instance, their average lifespans should be longer (Réale et al., 2010) – but whether this is true on farms and/or in these animals’ wild forebears remains to be tested. Captive farmed mink do still face evolutionary pressures: for instance,
sexual selection can occur because females are given opportunities to mate with multiple males annually, and some males are rejected, while the successful males are kept for breeding the following year; while relatively high kit mortality rates (Dunstone, 1993) indicate the potential for natural selection to improve maternal behaviour too. Studies of wild mink are therefore now necessary to determine whether the relationships observed here between behavioural phenotypes and reproductive variables resemble those in wild populations. Mink could also prove a good model for future research comparing fitness correlates of personality traits in captive, non-enriched conditions, in the wild, and perhaps in more naturalistic forms of captive housing (e.g. large enriched enclosures).

**Acknowledgements**

We thank the farm owners and staff for their help, advice and cooperation, and NSERC and the Canada Mink Breeders Association for funding. Thank you also to William Sears and Margaret Quinton for statistical advice, and Allan Edelsparre for comments on the manuscript, and Megan Jones for helpful discussion.


Table 3.1. Nest shape scale.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No nesting material</td>
</tr>
<tr>
<td>2</td>
<td>Some nesting material; flat, unstructured</td>
</tr>
<tr>
<td>3</td>
<td>All nesting material remaining; flat, unstructured</td>
</tr>
<tr>
<td>4</td>
<td>Saucer-shaped indentation</td>
</tr>
<tr>
<td>5</td>
<td>Round hole with sides higher than the dam when lying down</td>
</tr>
<tr>
<td>6</td>
<td>Round hole with partial ‘ceiling’</td>
</tr>
<tr>
<td>7</td>
<td>At least 75% of nest covered by ‘ceiling’</td>
</tr>
</tbody>
</table>

Table 3.2. Dam inactivity as a predictor of kit weight and growth.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Male kits</th>
<th>Female kits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight day 1</td>
<td>Inactivity*colour: F(<em>{1,400})=3.70, P=0.055 (Positive in Blacks: split by colour, Inactivity: F(</em>{1,223})=3.69, P=0.056)</td>
<td>Inactivity<em>litter size</em>colour: F(<em>{1,419})=3.28, P=0.071 (Positive in Pastels with small litters only: split by colour, Inactivity*litter size: F(</em>{1,191})=3.12, P=0.079)</td>
</tr>
<tr>
<td>Growth</td>
<td>Inactivity: F(<em>{1,397})=42.39, P&lt;0.0001 (Positive; involved in interaction) Inactivity*litter size: F(</em>{1,396})=4.91, P=0.027 (Stronger in small litters)</td>
<td>Inactivity: F(_{1,405})=22.57, P&lt;0.0001 (Positive)</td>
</tr>
</tbody>
</table>

Direction of significant relationships indicated in italics.
Figure 3.1. Kit growth between day 1 and day 21 vs. maternal inactivity in winter.
Figure 3.2. Latency to touch kits vs. inactivity. Values are back-transformed least squares means. Error bars indicate confidence intervals. Statistically significant contrasts (P<0.05) are labelled. N=161.
Figure 3.3. Kit growth vs. retrieval time.
N=168. Growth values are averaged across male and female kits.
CHAPTER FOUR

Decreased litter size in inactive female mink (*Neovison vison*): mechanisms and implications for overall productivity

Rebecca Meagher, Allison Bechard, Rupert Palme, María Díez-León, Bruce Hunter, Georgia Mason

A version of this paper was submitted to Can. J. Anim. Sci., October, 2011

Abstract

Farmed mink vary dramatically and consistently in activity levels: very inactive individuals rarely leave the nest box, while others devote hours per day to stereotypic behaviour (SB) and show more general activity. SB is typically associated with increased reproductive output, the opposite being true of inactivity. Using data collected over two breeding seasons, we compared inactivity and SB as predictors of reproductive output, and examined three hypothesized mechanisms for inactive dams’ reduced litter sizes (LS): 1) excess fat; 2) chronic stress (potentially underlying inactivity because fear motivates hiding); 3) health problems. Hypothesis 2 was not supported: inactive females were no more fearful in ‘glove tests’ (P>0.10) and excreted lower, not higher, levels of faecal cortisol metabolites than active females (P=0.033). Hypothesis 3 was also rejected: veterinary exams found no significant clinical problems. As predicted by Hypothesis 1, inactive females had higher body condition (P<0.0001), which predicted decreased LS (P=0.040); however, path analysis determined this was unlikely to mediate the inactivity-LS relationship. Inactivity also proved a more consistent predictor of both LS and kit weight (with which its relationship was positive: P≤0.037) than SB or normal activity. Therefore, decreasing inactivity in farmed mink, rather than increasing their SB or decreasing their body condition should most improve productivity.
Introduction

Mink (*Neovison vison*) on fur farms show dramatic, consistent individual differences in activity levels (e.g. Mason, 1993), with the most inactive doing little but lie in the nest box, while the most active spend much time out in the cage, often performing stereotypic (abnormal, repetitive) behaviour (Bildsøe et al., 1991; Svendsen et al., 2007a). Previous research has sometimes found that very stereotypic females have greater reproductive success, manifest as lower rates of barrenness, increased litter size, and decreased kit mortality (e.g. Jeppesen et al., 2004). Complementary findings have come from our own research (Mason, 1992; Meagher et al., submitted): inactive females give birth to fewer kits (i.e. have decreased litter sizes). However, not all differences in their reproductive styles are negative: they are more attentive to their male kits in retrieval tests, and their kits (especially males) grow more quickly.

Active and inactive mink may also differ in several other relevant ways, the first of which is that inactive individuals typically have higher body weights and, likely, more body fat (Hansen and Møller, 2008; Jeppesen et al., 2004). This is significant since high levels of fat are known to interfere with reproduction in many mammalian species, primarily through adipokines, signalling chemicals released by fat cells. These chemicals inhibit gonadotropin-releasing hormone (GnRH), which plays a major role in controlling gametogenesis, implantation of fertilized eggs, and embryo development (see e.g. Gosman et al., 2006; Moschos et al., 2002), as well as copulatory behaviour (Wingfield and Sapolsky, 2003). Such reproductive impairment in relatively fat individuals (i.e. those with higher body condition scores) has been shown in mink (Baekgaard et al., 2007; Joergensen, 1985; Malmkvist and Palme, 2008), and reduced body
weight was found to statistically explain the increased reproductive success in female mink exhibiting higher levels of stereotypic behaviour (Jeppesen et al., 2004).

Alternatively, there are welfare-relevant hypotheses that might explain both profound inactivity and associated reproductive impairments. Since most inactivity occurs in the nest box, which is the only shelter available, it could conceivably be a form of hiding, reflecting chronic fear. Although evidence of such fear is not always detected in inactive, non-stereotypic mink (e.g. Svendsen et al., 2007a in lines differentially selected for SB), where activity and temperament do co-vary, inactive individuals are more likely to be fearful than are their stereotypic counterparts (Hansen and Jeppesen, 2006). Relatedly, some studies have found that inactive phenotypes in mink are associated with increased baseline cortisol levels, indicative of chronic stress (Bildsøe et al., 1991; but cf. Svendsen et al., 2007). Like excess fat, chronic stress can impair reproduction at various stages, from gametogenesis and pregnancy (reviewed in Hemsworth, 2003; Wingfield and Sapolsky, 2003) through lactation (Lau and Simpson, 2004; Rushen et al., 1995). Finally, sickness (Hart, 1988) or injury (Rutherford, 2002) can elicit inactivity, and could interfere with reproduction (see e.g. Aubert, 1999). Although the increased maternal care and kit growth observed in our second experimental group (Chapter 3) render the possibilities of widespread health problems or severe physiological stress unlikely, we wished to rule them out directly.

Our aim here was to determine the practical consequences of the observed differences in reproductive output, and to identify which of the associated traits should be manipulated in order to best increase productivity. In Experiment 1, we collected pilot data on one farm to confirm that inactive mothers in our population do raise fewer kits to weaning successfully. We then conducted a larger-scale experiment, Experiment 2, in which we assessed overall differences in
productivity between active and inactive dams, as indicated by both the total number of kits raised and the kit weights at day 21. We then investigated the various hypothesized mechanisms for a decrease in litter size. If excess fat is wholly responsible, farmers could focus their efforts on manipulating it directly to increase the number of kits produced. If, on the other hand, productivity is mediated by some aspect of temperament or by chronic stress, feed deprivation or similar strategies would not be useful. Finally, given the inverse relationship between inactivity and stereotypic behaviour, and the fact that most research has focused on the latter, we aimed to determine which behaviour was most tightly linked with the observed effects and is therefore the best target for selection to improve productivity.

EXPERIMENT 1

Materials and Methods

Animals and Housing

The subjects were 247 female mink on a commercial farm in Ontario, Canada, of two common colour-types, Black and Pastel (N=141 and 98, respectively). Of these mink, 108 were approximately ten months of age (i.e. in their first breeding season); the others ranged from two to four years of age. The mink were individually housed in conventional wire mesh cages (61 cm x 23 cm x 46 cm or greater) with an elevated wooden nest box at the back. Blacks and Pastels were housed in separate sheds. Then, following standard practice on North American farms, the mink were relocated to larger whelping cages in other sheds, where nest boxes were attached to the front of the cage. Straw and aspen shavings were provided for nest building in these cages.
Pre-breeding Behaviour and Reproductive Success

Twenty behavioural scans were conducted over a four-day period in February, immediately prior to breeding. All observations took place in the mornings, from approximately 08:30 until feeding time in the early afternoon. Behaviour was recorded as stereotypic behaviour, non-stereotypic (normal) activity, or inactivity; definitions are given in Table 4.1.

After weaning, data were collected for the number of live and dead kits on the day after birth, and the number of kits that died by the time of weaning. Litter size at weaning was then calculated for use in analysis. As in Chapter 3, litters in which farmers cross-fostered (added or removed, including kits that were rescued after becoming chilled) kits were excluded from the analysis because such manipulation could obscure biological effects on litter size.

Statistical Analysis

All statistical analyses were conducted in JMP 8 (SAS Institute Inc., NC, USA, 2009). General linear models (GLMs) were used to analyse relationships between age, colour-type (which included the confounding effect of shed) and inactivity, normal activity or stereotypic behaviour (SB), as well as the predictive value of these behaviour patterns for litter size at weaning, for only those individuals that gave birth. Group and age were included as blocking factors, as well as their interactions, which were removed if they had a p-value greater than 0.25. The relationship between behaviour and risk of being barren was assessed by comparing levels of the behaviour between barren and non-barren individuals using a Welch ANOVA to accommodate unequal variances. The conventional probability value of P<0.05 was used to establish statistical
significance; however, exact values are reported for any $P<0.10$ since these effects may still be worth further investigation.

**Results**

Inactivity levels ranged from 0 to 94.7% of observations, with an overall mean of $40.3 \pm 22.1\%$. Blacks were significantly more inactive than Pastels ($F_{1,167}=13.6$, $P=0.0003$). Because these groups were also housed in different sheds, however, the results cannot be used to make inferences about the colour-types in general. There was also a non-significant trend for older mink to be more inactive ($F_{1,167}=3.56$, $P=0.061$). A Welch ANOVA was used to compare SB levels between the top and bottom quartiles for inactivity, since the data did not meet the assumptions of parametric tests; very inactive females had significantly lower mean SB ($F_{1,69.8}=157.1$, $P<0.0001$). Like inactivity, SB ranged from 0 to 94.7% of observations; its overall mean was $32.2 \pm 23.1\%$. Similarly, very inactive females had significantly lower normal activity ($F_{1,66.3}=6.94$, $P=0.011$). The mean for this variable was $23.8 \pm 15.8\%$, with a range of 0 to 89.5%.

There were only six barren individuals; however, inactivity was significantly higher in this group than in those that bred successfully (Welch $F_{1.5,3}=6.62$, $P=0.047$). Among individuals that successfully produced a litter, inactivity was significantly inversely related to litter size at weaning ($F_{1,157}=5.61$, $P=0.019$; Figure 1a). No significant effects of age or colour-type on litter size were detected ($P>0.10$).

SB was significantly lower among barren than non-barren individuals (Welch $F_{1,6.0}=10.9$, $P=0.016$). However, SB did not have a significant main effect on litter size at weaning; instead, it
interacted with both colour-type and age (SB*colour*age F_{1,157}=3.99, P=0.048). Analyses split by age showed no significant relationships between SB and litter size; among the very small group of four-year-olds, there was a trend for a negative relationship (F_{1,9}=4.00, P=0.076). Normal activity levels did not differ between barren and non-barren individuals (P>0.10). For litter size, normal activity interacted with age (F_{1,158}=4.42, P=0.037); split by age, activity was positively related to litter size in two- and three-year-olds only (F_{1,41}=3.64, P=0.064 and F_{1,29}=6.55, P=0.016, respectively).

**EXPERIMENT 2**

**Materials and Methods**

*Animals and Housing*

A total of 549 female mink were observed across three commercial farms: 149 Blacks and 144 Pastels at Farm 1 (the site of Experiment 1), 143 Blacks at Farm 2, and 113 Pastels at Farm 3. Only mink entering their second or third breeding season were used, as these mink are more likely to be very inactive than one-year-old mink are, and focussing on them allowed us to omit the control for age, since this group is relatively homogenous in terms of reproductive output. Housing conditions were very similar to those in Experiment 1. Mink were kept in individual cages, the size of which varied slightly between farms, with a minimum of 61 cm x 19 cm x 46 cm. All cages included a nest box and a single nipple drinker. At Farms 2 and 3, the nesting material provided in the larger cages where the mink gave birth was aspen shavings, while Farm 1 provided straw.
Pre-breeding Behaviour

Nine to ten days of behavioural observations were conducted at each farm in February, before mating began. The protocol was the same as in Experiment 1.

Fear and Physiological Stress

In April, once mink were expected to be pregnant, a ‘glove test’ was applied to determine whether some females were likely experiencing chronic fear. In this temperament test, mink are classified as fearful, curious, aggressive or unresponsive depending on their response to a human extending a gloved hand against the cage (for detailed methods and validation, see Meagher et al., 2011). For a subset of 30 of the least inactive (“active”) and 31 of the most inactive (“inactive”) females on Farm 1, 24-hour faecal samples were also collected prior to the glove test to assess faecal cortisol metabolites (FCM) as an indicator of chronic physiological stress. These samples were stored frozen. After thawing, they were homogenized, and the metabolites were extracted using 80% methanol and analyzed using an 11β-hydroxyaetiocholanolone enzyme immunoassay (Frigerio et al., 2004), previously validated in mink (Malmkvist et al., 2011).
**Body Condition**

The farmers scored animals for body condition in late February, the week before mating began. Body condition scores (BCS) were given on a 5-point scale modified from one developed and validated by Hynes and colleagues (2004), where one represented a very thin mink and five was very heavy. These scores correlate with amount of body fat (Clausen, 2005, cited by Baekgaard et al., 2007). At most times of year, three is considered an ideal score, although during late gestation, somewhat higher scores are expected (Hynes et al., 2004).

Dam body condition was scored again, by the researchers (RM and AB), when the kits were counted on the day after birth. To ensure consistency of use across farms and scorers, in addition to verbal descriptions of the categories, scorers at the other farms used photos of each category from Farm 1 as a reference point. Inter-observer reliability was assessed after the first scores were assigned, and again at the end of the study in comparison with scores given to a subset of animals by a veterinarian (BH) using to the original, similar but more in-depth protocol of Hynes and colleagues (2004) that involved handling the animals to feel for subcutaneous fat.

**Reproductive Output and Kit Performance**

Kits were counted, sexed, and weighed 21 days after birth. Weights were averaged by sex. Day 21 weights were used because this encompasses the period in which kits are fully dependent on their mothers for milk (Dunstone, 1993; Mason, 1994). Litter sizes from this time were used in place of weaning values in this experiment, primarily for pragmatic reasons; however, because kit mortality is extremely low after this age (less than 0.1% of kits: European Commission, 2001;
also Joergensen, 1985), these data are representative of litter sizes at weaning. Litters where cross-fostering had occurred were again removed from all litter size analysis; however, presented kit weights are for all kits raised, whether or not they were biological offspring.

**Health**

During all observation periods, it was noted if any animals did not appear to be eating normally or showed obvious signs of injury. At the conclusion of the study in June, a subset of animals at each extreme for inactivity levels were examined by a veterinarian (BH) to ensure that their previous inactivity was not caused by a physical illness (N=78 total on two farms). This included checking for old physical injuries or foot pad disease that would cause discomfort when standing or walking, checking that the faeces appeared normal and that there were no signs of respiratory illness or infection such as discharge from the nose. Where possible (N=21), the urine was also tested for pH and the presence of leukocytes, protein, glucose or blood/haemoglobin. Leukocytes would indicate infection, while more than trace amounts of protein would suggest kidney pathology (Zellen, 1996), and glucose, a metabolic disorder (Hynes et al., 2004). Blood in the urine can be caused by kidney infections or other conditions (Hunter, 1996).

**Statistical Analysis**

As in Experiment 1, analyses were conducted in JMP. The relationship between behaviour (inactivity or SB) and risk of being barren was assessed by comparing levels of the behaviour between barren and non-barren individuals using Welch ANOVAs to accommodate unequal variances, split by farm and colour-type. Risk of barrenness was also compared between fat and
normal or thin females using Fisher’s exact test, split by colour and farm. All other relationships with inactivity were assessed using mixed model GLMs, because farm was treated as a random effect. Thus, the restricted maximum likelihood (REML) method of estimation was employed (Searle et al., 1992). These GLMs also controlled for colour-type, and its interactions with the other variables; interactions were removed if they had no effect (P>0.25). Adjusted (Type III) sums of squares were used unless otherwise specified. Normality and homogeneity of variance were assessed by inspection of the residual plots (Grafen and Hails, 2002); Bartlett’s test for equal variances was used where the effects of interest were categorical.

Litter size effects were again examined for those individuals that gave birth only, blocking by colour-type and farm. Two outliers were excluded for male kit weights, because they were more than two standard deviations above the mean in absolute value, as were their residuals, and they were likely measurement or recording errors. Their exclusion corrected the non-normality of the original data, without altering the conclusions drawn from the analysis. In the kit weight model, litter size was also included as a covariate.

To determine whether stereotypic behaviour or normal activity are better predictors of productivity than inactivity, the GLMs for litter size and kit weight were re-run with the alternative behavioural predictor added to the model. Sequential (Type I) sums of squares were used to determine whether each of these two behavioural variables was significant with or without controlling for the other. These models included farm and colour-type, as well as litter size when kit weight was the dependent variable, before the effects of interest.
The reliability of BCS was assessed in two ways. First, inter-observer reliability was calculated at one farm (N=112 mink) using Spearman rank correlations or kappa for a categorical scale. Using a scale with acceptable inter-observer reliability (see Results), an ordinal logistic regression was then used to compare scores in February with those on the day after birth, in order to determine its consistency over time.

Finally, relationships between inactivity and the hypothesized intervening variables of FCM, temperament, and body condition were then analysed using GLMs, as above. A log transformation was applied to FCM to achieve normality of the residuals. Since body condition was not a continuous variable, it was treated as the independent variable, and inactivity as the dependent variable to allow the use of a GLM. An arcsine square-root transformation was applied to inactivity. If the variable of interest was significantly related to both inactivity and the measure of productivity, path analysis was conducted using Amos 5 (Arbuckle, 2003) to determine whether inactivity or the possible intervening variable best explained the litter size and kit weight data. Model fit was assessed using the chi-square test, RMSEA (Root Mean Square Error of Approximation), NFI (Normed Fit Index), and AIC (Akaike Information Criterion) statistics. AIC and chi-square values are reported here. Body condition scores from February were used in the litter size models, and from the day after birth in the kit weight models.

Results

In this study, Pastels were more inactive than Blacks, averaging 40.9 ± 1.5% of observations inactive compared to 32.9 ± 1.1% for Blacks (Welch’s ANOVA, F1,470=18.7, P<0.0001). Blacks,
on the other hand, were more stereotypic: 41.6 ± 1.4% vs. 30.4 ± 1.6% (t=-5.28, d.f.=531, 
P<0.0001).

Reproductive Output and Kit Performance

In this experiment, in contrast to Experiment 1, barren mink were significantly less inactive among Blacks at both Farms 1 and 2 (F_{1,15.6}=6.52, P=0.022, and F_{1,11.9}=14.8, P=0.002, respectively). There were no significant differences among Pastels at either farm (P>0.10). The predicted negative relationship between inactivity and litter size was detected (F_{1,327}=4.32, 
P=0.038; Figure 4.1b).

Inactivity predicted increased male and female kit weight at 21 days, even after statistically controlling for the decreased litter size of inactive dams (Figure 4.2). For both sexes of kit, there was a significant main effect of inactivity (females: F_{1,424}=24.4, P<0.0001; males: F_{1,420}=47.9, P<0.0001), but also an interaction between inactivity and litter size (females: F_{1,423}=4.14, P=0.042; males: F_{1,410}=3.87, P=0.050) such that inactivity had a positive relationship with kit weight only in smaller litters.

Stereotypic Behaviour and Normal Activity as Alternative Behavioural Predictors

Again, parametric tests could not be used to assess the relationship between inactivity and SB, and a Welch ANOVA was therefore used, as in Experiment 1, to compare SB levels between mink categorized as active and inactive. Inactive mink had significantly lower mean SB (F_{1,219}=487.9, P<0.0001). Contrary to expectations, average levels of normal activity did not
differ enough between the least inactive (“active”) and the most inactive mink to be detected by a Welch ANOVA (P>0.10), although there was greater variation in activity in the active group (Levene’s test for equality of variance, F\(_{1,231}=2.64\), P=0.001). Activity was therefore not tested as an alternative explanation for the observed relationships with inactivity.

Among Blacks at Farm 2 only, being barren was associated with increased levels of stereotypic behaviour (F\(_{1,11.9}=15.0\), P=0.002). There was no relationship between SB and barrenness in any other group (P>0.10), and neither was it a statistically significant predictor of litter size, regardless of whether it came before or after inactivity in the model. Inactivity, meanwhile, was always significant; indeed, the effect was actually stronger if a control for SB was included first (F\(_{1,302}=5.98\), P=0.015).

As expected, increasing SB predicted decreased kit weight, i.e. the relationship was the inverse of that with inactivity. For female kit weight, as above, SB had a significant main effect only if before inactivity in the model (F\(_{1,417}=19.7\), P<0.0001; else P>0.10). Inactivity, on the other hand, was significant regardless of its position in the model (F\(_{1,413}=4.40\), P≤0.037). For male kit weight, both SB and inactivity were always significant, but the effect of inactivity was stronger: SB had F\(_{1,415}=4.64\) (P≤0.032), compared to F\(_{1,415}=7.25\) (P≤0.007) for inactivity. Because inactivity was thus the more consistent predictor of the reproductive variables of interest, it was the focus of our remaining analyses.
Potential Mechanisms for Inactive Females’ Reduced LS

Fear and stress. Inactive females did not exhibit physiological signs of elevated chronic stress; indeed, they had significantly lower FCM than active females ($F_{1,58}=4.79$, $P=0.033$). Temperament was significantly related to inactivity ($F_{4,474}=8.75$, $P<0.0001$); however, the only category that proved significantly different (using Tukey’s Honestly Significant Difference) was “unresponsive”, in which inactivity was higher ($F_{1,474}=24.5$, $P<0.0001$). Furthermore, a contrast of fearful versus non-fearful mink showed no difference in inactivity ($P>0.10$). Because neither fear nor FCM was elevated in inactive individuals, these variables could not mediate a negative relationship between inactivity and litter size, and thus they were not used in further analyses.

Post hoc analyses were conducted to determine whether the temperament category “unresponsive”, rather than fearful, predicted barrenness or litter size. However, no significant relationships were found with either of these variables ($P>0.10$).

Health. No signs of lasting illness or injury were evident during the observation periods, nor were serious illness or injuries that would limit locomotion detected in active or inactive females during the veterinary exams. At most, small amounts of protein were found in the urine, but the occurrence of this did not differ significantly between the active and inactive females. Of 21 individuals tested, three had some blood in the urine, all inactive females. This was not considered to indicate a severe or chronic condition, and these individuals had litters of average rather than small size; thus, this did not likely explain any differences in reproduction between the groups of interest.
Body condition. Inter-observer reliability of body condition scores did not meet the conventional threshold of acceptability (a rank correlation of 0.7 or greater), ranging from 0.29 to 0.39. Despite this problem, the scores were not invalid, agreement typically being good for animals at the extremes of the scale. For this reason, for subsequent analysis, scores were pooled into three simpler categories: thin, normal or fat (N.B. mink categorized as ‘fat’ were not necessarily obese, but were merely judged to have more body fat than the other mink). With this modified scale, moderate to substantial agreement between observers was obtained, with kappas ranging from 0.42 to 0.65.

BCS did not remain consistent over time (P>0.10). In February, it was significantly related to inactivity (F2,539=37.4, P<0.0001). However, after the birth of the kits, it was not (P>0.10).

Role of Body Condition in Mediating Relationship with Reproductive Output

Being fat rather than normal or thin did not significantly predict risk of being barren, either overall or in any farm by colour group (P>0.10). It could therefore not explain the tendency for reduced barrenness observed in inactive Black females at Farm 1. However, body condition in February interacted with colour-type to predict litter size when inactivity was not in the model (F2,324=3.25, P=0.040). Among Blacks only, fat mink appeared to perform worse than did normal or thin mink, although the contrast was not significant in either colour-type (Blacks: F1,323=32.43, P=0.120; Pastels: F1,323=0.08, P=0.777; Figure 3).
We therefore used path analysis to determine whether the relationship between inactivity and litter size was direct or was mediated by BCS. Figure 4.4 shows the hypothesized pathways tested in the best-fitting model that included all of these variables; path analysis gives relative weights to each path, with a higher weight indicating that a particular path is more likely than the others to explain the observed effects. The direct relationship between BCS and litter size proved to be non-significant. Although this model was adequate (see Table 4.2), the fit could be improved by removing that path, along with the variable colour-type, which also lacked a significant direct relationship with litter size. These results suggest that the relationship between inactivity and litter size is not a result of greater body condition; rather, inactivity predicts decreased litter size and increased BCS independently. Detailed results for both models are presented in Table 4.2.

**Role of Body Condition in Mediating Relationship with Kit Performance**

BCS also significantly predicted male kit weight at 21 days ($F_{2,382}=3.19$, $P=0.042$). It also predicted female kit weight at this age, but interacted with colour-type ($F_{2,382}=3.58$, $P=0.029$). Using path analysis, both inactivity and BCS proved to have significant direct effects on kit weight for both sexes, although BCS was weighted more heavily. The best-fitting model did not include a correlation between inactivity and BCS, suggesting this path was unimportant in explaining our results. The final path diagram is presented in Figure 4.5, and statistical results in Table 4.3.
Discussion

It appears that inactive mink do have smaller litters. In practical terms, this effect is relatively small on an individual level, particularly in Experiment 2. Thus, as seen in Figure 4.1, the difference between the expected litter sizes for the most and least inactive individuals is just under one kit. Although litter size is the most economically important measure of productivity (Lagerkvist, 1997), the increase that would be obtained by reducing inactivity in some instances might not be worth the cost of the associated decrease in kit weight. This will likely vary between farms depending on which aspect of productivity has the most room for improvement. The results for rates of barrenness are less clear, since it was linked to increased inactivity in Experiment 1 but decreased inactivity in Experiment 2. These findings therefore need replication, particularly given the low incidence of barrenness in our samples. Although stereotypic behaviour continues to receive much attention in mink research, it appears that it is less relevant to productivity than is inactivity. Given the strong inverse correlation between these two aspects of behaviour, previously reported relationships with SB (e.g. Jeppesen et al., 2004; Mason et al., 1995) might actually have been driven by decreased inactivity rather than by this abnormal behaviour per se. However, the relationships may also differ between populations: thus, we detected no apparent relationship between SB and litter size even before controlling for inactivity, while others have (see Introduction). Inactivity was also a more consistent predictor of productivity than was normal activity in Experiment 1, and these two aspects of the time budget were uncorrelated in Experiment 2, indicating that normal activity was not the true explanation for the observed variation in productivity.
None of our hypothesized mechanisms could fully explain the relationships between inactivity and our reproductive variables. Both chronic fear and poor health could be ruled out as causal factors, since neither was elevated in inactive females. This was consistent with our findings that inactive dams in some ways appear to be better mothers than their more active peers (Meagher et al., submitted; Chapter 3). It remains possible that females who were inactive in February were experiencing transient illnesses that were not manifest in changes in feeding and had no lasting effects that could be detected by the time full veterinary examinations were conducted, but it seems unlikely that such short-lived conditions had an impact on litter size or kit weight if a litter was conceived. Our findings that very inactive, and thus non-stereotypic, mink tended to have lower concentrations of faecal cortisol metabolites and were no more fearful than were individuals that are more active contrast with the results of Bildsøe and colleagues (1991) mentioned above, but were consistent with more recent data using lines differentially selected for SB (Malmkvist et al., 2011; Svendsen et al., 2007a). These findings could reflect higher psychological stress in individuals with high SB levels, or merely their intense physical exercise (e.g. West et al., 2004). Moreover, our temperament findings are in line with those of Svendsen and colleagues (2007a), who also demonstrated no difference in fearful responses between lines selected for high and low SB, but rather that the low SB mink were more likely to be “uncertain” (which was called “unresponsive” in this paper). The meaning of this category in welfare terms remains unclear (see Meagher et al., 2011). Nonetheless, there is currently no evidence that the welfare of inactive individuals is compromised in comparison with that of others in the same population.

Inactive females did have higher body condition scores, and thus likely possessed more body fat, at least at the time of mating and conception, and this was negatively correlated with
litter size in at least one colour-type. However, path analysis suggested that the variation in litter size was better explained by a direct relationship with inactivity than by the mediating effects of body condition. This analysis cannot definitively determine whether dam inactivity has a causal effect; however, it does indicate that body condition likely did not, in our populations. Body condition also did not fully explain the positive relationships found between inactivity and other aspects of reproductive performance: although it did increase kit weight, it did not fully explain the higher kit weights in offspring of inactive mothers, and it had no relationship at all with rates of barrenness in this sample.

The finding that body condition was not related to litter size independently of inactivity levels contradicts current beliefs; as discussed in the Introduction, previous studies have found an effect of body condition, and there are known physiological mechanisms to explain that effect. There are several possible explanations for these conflicting results. First, they may be due to prior studies not controlling for behavioural differences between animals of different body conditions, in which case the benefits of feed restriction might be mediated by the increased activity it elicits (e.g. Bildsøe et al., 1991; Damgaard et al., 2004; but see Hansen and Møller, 2008). Second, they may reflect differences between populations, due to differences in the distribution of body conditions observed or the colour-types studied. Finally, our results may perhaps simply reflect lower statistical power for detecting body condition effects than relationships with inactivity in the current analysis, since inactivity was assessed with much greater resolution.

Overall, it appears that inactive females are neither stressed nor ill, but that there is some other fundamental difference between the behavioural phenotypes that results in differences in
reproduction. Reducing inactivity through selection or altered husbandry (e.g. the provision of running wheels or other objects to promote activity), rather than focussing on restricted feed to decrease body fat as is currently standard practice, would seem to hold the most potential for increasing reproductive output. Any reduced inactivity, however, may come at a cost in terms of decreased kit growth and potentially increased incidence of barrenness. The overall effects on profitability may therefore vary depending on which factor most limits production on a given farm.

Acknowledgements

Thank you to the farm owners and staff for their help and cooperation, and to NSERC and the Canada Mink Breeders Association for funding the research.
Table 4.1. Ethogram

<table>
<thead>
<tr>
<th>Activity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stereotypic behaviour</td>
<td>Movement or sequence of movements repeated at least three times consecutively</td>
</tr>
<tr>
<td>Inactivity</td>
<td>Lying still, whether eyes open or closed</td>
</tr>
<tr>
<td>General activity</td>
<td>Animal neither inactive nor engaged in stereotypic behaviour; includes eating, drinking and grooming</td>
</tr>
</tbody>
</table>

Table 4.2. Litter size path analysis results

<table>
<thead>
<tr>
<th>Relationships and fit statistics</th>
<th>Models</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inactivity --&gt; Litter size</td>
<td>-0.108†</td>
<td>-0.132*</td>
<td></td>
</tr>
<tr>
<td>Inactivity --&gt; BCS</td>
<td>0.350***</td>
<td>0.350***</td>
<td></td>
</tr>
<tr>
<td>BCS --&gt; Litter size</td>
<td>-0.032</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colour --&gt; Inactivity</td>
<td>0.193***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colour --&gt; Litter size</td>
<td>-0.067</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\chi^2)</td>
<td>0.012, P=0.912 (d.f. 1)</td>
<td>0.308, P=0.579 (d.f. 1)</td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>18.012</td>
<td>10.308</td>
<td></td>
</tr>
</tbody>
</table>

Values given for relationships are standardized regression weights. *** indicates \(P<0.001\), * \(P<0.05\), and † \(P<0.10\). Negative values indicate an inverse relationship, if the predictor is quantitative. A smaller AIC indicates a better fit.

Table 4.3. Kit weight path analysis results

<table>
<thead>
<tr>
<th>Relationships and fit statistics</th>
<th>DV(^1)</th>
<th>Male weight</th>
<th>Female weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inactivity --&gt; weight</td>
<td>0.221***</td>
<td>0.165***</td>
<td></td>
</tr>
<tr>
<td>BCS --&gt; weight</td>
<td>0.257***</td>
<td>0.288***</td>
<td></td>
</tr>
<tr>
<td>Number of kits --&gt; weight</td>
<td>-0.217***</td>
<td>-0.209***</td>
<td></td>
</tr>
<tr>
<td>Colour --&gt; Inactivity</td>
<td>0.169***</td>
<td>0.179***</td>
<td></td>
</tr>
<tr>
<td>Colour --&gt; BCS</td>
<td>0.169***</td>
<td>0.186***</td>
<td></td>
</tr>
<tr>
<td>(\chi^2)</td>
<td>6.046, P=0.302 (d.f. 5)</td>
<td>9.625, P=0.087 (d.f. 5)</td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>26.046</td>
<td>29.625</td>
<td></td>
</tr>
</tbody>
</table>

See notes for Table 4.2.

\(^1\) DV = dependent variable
Figure 4.1. Partial residual plots of litter size vs. inactivity

a) Experiment 1: litter size at weaning. Y-values are calculated using the residuals and regression equation from the model controlling for age and colour-type. b) Experiment 2: litter size at day 21. Y-values are calculated using the residuals and regression equation from the model controlling for farm and colour-type.
Figure 4.2. Kit weight vs. inactivity

PND = postnatal day.
Figure 4.3. Litter size on PND 21 vs. body condition score

Data are least squares means ± SE.
Figure 4.4. Hypothesized path diagram for litter size, with non-significant paths removed for best fit.

Figure 4.5. Hypothesized path diagram for kit weight.

Male and female weights were analysed separately but with identical paths. Again, non-significant relationships were removed.
CHAPTER FIVE
Psychological states induced by non-enriched housing and their relationship to inactivity

This manuscript is in preparation for submission to PLOS One

Abstract

Captive animals housed in impoverished environments are often thought to be ‘bored’. In some situations, particularly when individuals become profoundly inactive, they have also been called ‘apathetic’, or even ‘depressed’. However, the use of these terms has rarely been validated, even though alleviating boredom is one stated aim of some environmental enrichment programs. Boredom in non-humans is difficult to assess empirically, but since it is a negative state resulting from under-stimulation, we can predict that it would increase responsiveness to stimuli of any kind. Apathy (lack of interest), by contrast, would be expressed as decreased responsiveness to all stimuli, while anhedonia (loss of pleasure, a depressive symptom) would decrease responsiveness to normally rewarding stimuli, but not to aversive ones. We tested the competing hypotheses that mink, a model carnivore, experience boredom or a depression-like apathy and/or anhedonia in standard fur-farm cages. We exposed 29 mink to a series of ten stimuli categorized a priori as aversive (e.g. an air puff or predator scent), rewarding (e.g. a moving brush to chase) or neutral (e.g. a plastic bottle). In these 5-minute tests, we assessed latency to make contact and total time oriented to the stimulus as indicators of responsiveness. Effects of long-term housing were tested using repeated measures GLMs controlling for sex and stimulus order. Mink housed in non-enriched (NE) cages made contact sooner (F_{1,28}= 9.65, P=0.004) than mink housed with environmental enrichment (a double cage with wading water and manipulable objects; E). They also spent longer oriented to the stimulus, although this difference was only significant for neutral stimuli (treatment*type F_{2,220}=6.69, P=0.002). Similarly, time in contact with the stimulus was higher for NE mink, but housing interacted with stimulus type (F_{2,191}=3.25, P=0.041) and the difference was significant only for neutral stimuli. However, in separate tests of treat consumption, NE mink ate more treats than did E mink, suggesting they are more responsive to some rewards (F_{1,21}=4.71, P=0.042). Perhaps unsurprisingly since there was no evidence of apathy, inactivity levels assessed via scans three weeks earlier did not predict responsiveness,
within non-enriched housing (all P>0.05). Overall, NE mink show a heightened responsiveness to stimuli of all types, consistent with boredom.

**Introduction**

Boredom, apathy and depression are often hypothesized to occur in captive animals housed in impoverished cages (e.g. Bolhuis et al., 2006; Stevenson, 1983; Wemelsfelder, 1990). However, in very few cases has the use of any of these terms been validated empirically, and often no precise definitions are given. The three terms have also been used in overlapping ways to describe non-human animals, which can lead to confusion. For example, Wood-Gush and Beilharz (1983) suggested piglets in non-enriched environments were bored, but Wood-Gush and Vestergaard (1989) described them as apathetic, even though both the housing conditions and behaviour appeared very similar: in both cases, piglets were more inactive and showed less behavioural diversity than those given enrichment. However, these states have distinct definitions and different clusters of symptoms in humans, which allow them to be distinguished from one another behaviourally.

Depression is the only one of these three states that has been clearly operationalized and demonstrated, at least in part, in non-humans. Chronic stress is believed to be one of its major causes, and more specifically, uncontrollable stressful events (reviewed by Henn and Vollmayr, 2005). Thus, the chronic stress and lack of control known to exist in many captive environments may induce similar states in non-human animals (see Carlstead, 1996; Morgan and Tromborg, 2007). Depression is a very complex phenomenon, with several subtypes (see American Psychiatric Association, 1994); reviewing and assessing all aspects of the disorder in non-humans is beyond the scope of this paper. However, one of its core symptoms, anhedonia, is of interest here; it is common to all subtypes and regularly assessed in animal models (cf. Anisman and Matheson, 2005). Anhedonia is defined as a reduced capacity to experience pleasure, typically measured in terms of the decreased consumption of rewards (e.g. sucrose for rodents: Willner et al., 1987). Apathy is similarly a common symptom of depression, although it can also be a primary complaint (reviewed in Marin et al., 1993). Like depression, it is sometimes a response to uncontrollable stress (see Glover, 1992). It is thought of as a lack of interest or caring, but in
practice, it is typically operationally defined as a state of generally reduced motivation or participation in activities (Marin, 1990; Marin et al., 1991; Sockeel et al., 2006). Thus, apathy should be expressed as a decreased motivation to obtain or interact with any stimuli, while anhedonia would decrease responsiveness specifically to rewarding stimuli. Depressed people whose symptoms include both anhedonia and apathy-like blunted affect typically exhibit decreased sensation-seeking (e.g. Carton et al., 1992).

Boredom in animals has never been empirically demonstrated, because it is very difficult to define operationally (see Wemelsfelder, 1990). In humans, it can be defined as a negative affective state caused by under-stimulation or monotony (Berlyne, 1960). It has been studied in a range of situations, from complete sensory deprivation to the performance of very monotonous tasks (reviewed by Berlyne, 1960). The situation of prisoners serving life sentences probably best parallels the one faced by captive animals; neither prisoners nor animals are deprived of all stimuli, but they do face a very unchanging, inescapable environment, which induces boredom in the humans (see Cohen and Taylor, 1972). In all cases, however, people report that the situation is aversive (Harris, 2000). Thus, bored people are motivated to seek novel stimulation (e.g. bungee jumping: Michel et al., 1997; novel foods: Pliner and Melo, 1997; recreational drugs: Samuels and Samuels, 1974). Impoverished housing conditions may similarly increase motivation to obtain stimulation in rodents; for example, environmental enrichment can decrease instrumental responses to obtain amphetamine in rats (e.g. Stairs et al., 2006). Since we cannot rely on verbal self-reports for non-humans, this motivation to obtain stimulation must form the basis of any operational measures of boredom (as in Kirkden, 2000). Indeed, the concept of boredom has often been used as an explanation for why animals explore, and thus been considered synonymous with motivation to obtain novel stimulation (see Wemelsfelder and Birke, 1997; Russell, 1983; but cf. Wemelsfelder, 1993, who conceptualized it differently).

In addition to predicting different levels of responsiveness, these three states also differ in their predicted relationships with inactivity levels. In humans, symptoms of boredom vary, including both inactivity and stereotypic behaviour, along with restlessness, irritability, lethargy and fatigue (reviewed by Berlyne, 1960; Harris, 2000). Similarly, while boredom is linked to inactivity by some authors, as in the Wood-Gush and Beilharz (1983) example cited above (see
also Gunn and Morton, 1995; Paquette and Prescott, 1988; Stevenson, 1983), it has also been posited to play a role in the development of stereotypic behaviour (Mitchell and Etches, 1977; Wemelsfelder 1990, 1993). Thus, it may not be specific to one particular behavioural phenotype. By contrast, only animals that are inactive and/or unresponsive are typically labelled “apathetic” or “depressed”, and these states have therefore been suggested to underlie the high levels of inactivity sometimes observed in captive animals (see Chapter 1). While depression does not always lead to increased inactivity in humans, certain forms of it do (see Gold and Chrousos, 2002), and the animal models based on response to chronic stress tend to be associated with inactivity (see e.g. Gilmer and McKinney, 2003; Maier, 1984).

Our aim was to test the alternative hypotheses that mink in standard, non-enriched (NE) cages are bored, or that they are apathetic and/or anhedonic. To do so, we compared their responsiveness to a wide range of stimuli to that of mink housed in enriched (E) cages, which here represent ‘content’ animals (i.e. not expressing any of these negative affective states). We predicted that if bored, non-enriched mink would show increased responsiveness to all stimuli presented; if apathetic, they would show decreased responsiveness to all stimuli; and if anhedonic, but not apathetic, they would show decreased responsiveness to rewards only (see Table 5.1). We also predicted that if the mink were apathetic or anhedonic, then NE mink would be more inactive than E mink, and this inactivity would co-vary with the decreased responsiveness; that is, there would be an inverse relationship between inactivity and responsiveness within the NE treatment.

Previous research on how impoverished housing influences responsiveness to stimuli has produced conflicting results. In some cases, long term housing in under-stimulating environments reduced exploration of novel objects (Dahlborn et al., 1996, cited in Olsson and Dahlborn, 2002; Renner, 1987; Renner and Rozenzweig, 1986), while in others, animals housed in non-enriched environments showed increased exploration (Stolba and Wood-Gush, 1980; Wood-Gush et al., 1989; Bracke and Spoolder, 2008) or increased reactivity (Mackay and Wood-Gush, 1980). The interpretation of the results has often been complicated by the fact that differences in levels of fear could act as a confound. Environmental enrichment is known to decrease fearfulness in many species (e.g. Jones, 1982; Klein et al., 1994), and fear might inhibit investigation of novel stimuli,
making non-enriched animals appear less interested in them when the dependent variable is latency to approach or amount of interaction with the stimulus. We thus controlled for differences in fear in two ways: by using multiple measures of responsiveness which would not be influenced by fear in the same direction, and by scoring individuals on the presence or absence of species-typical indicators of fear during the trials, so that we could test for effects of the housing on fearfulness. These other studies of enrichment effects on responsiveness have also looked at response to only one or a few stimuli (e.g. Bracke and Spoolder, 2008), which were all either novel and neutral or novel and aversive, rather than examining the broader pattern of responsiveness.

**Methods**

**Animals and housing conditions**

The subjects were 29 Black mink (*Neovison vison*), all approximately ten months old, and therefore having just reached sexual maturity. Since the age of approximately three months, they had been housed differentially in either enriched (N=6 male, 7 female) or non-enriched conditions (N=8 male, 8 female). Non-enriched (NE) mink had 75 (L) x 60 (W) x 45 (H) cm wire-mesh cages, with a nest box on the front. Enriched (E) mink had an identical home cage, but also had access via a wire-mesh overhead “bridge” to a cage of double that width (details in Dallaire et al., 2011), with running water in a small trough to allow wading and head-dipping. New objects for the mink to interact with were added to this enriched cage every month. These differential housing conditions have been confirmed as causing differences in long-term physiological stress, with E mink less stressed than NE mink (Díez-León et al. 2010, in prep.).

**Behavioural observations**

Prior to the experiment, baseline behaviour was observed using live scans from January 26 to February 1. Scans began at 08:00 and continued every 15 minutes until 12:00. The proportion of scans spent inactive (lying still, whether awake or asleep), was recorded.
Two types of tests were used to assess responsiveness: tests in which stimuli of various types were placed on or near the cage, and tests of consumption in which rewards were offered in the cage. The first type consisted of a series of ten tests, conducted over eight days in late February. In each test, an auditory cue was given to signal the start of a test, and then a stimulus was placed on top of or in front of the cage. The stimuli were categorized a priori as aversive, rewarding, or neutral. The specific stimuli are listed in Table 5.2, along with the rationale for their categorization. Observations began as soon as the mink oriented to the stimulus, and continued for five minutes. The measures of responsiveness were the total duration of orientation to the stimulus, and duration of contact with the stimulus (a subset of time oriented), as well as latency to make contact. Decreased latencies, along with increased duration of both orientation and contact, would indicate heightened responsiveness consistent with boredom, while the reverse would indicate decreased responsiveness. Animals that were fearful would be expected to show increased latencies, but not decreased orientation, since fear and anxiety is associated with vigilance but immobility or avoidance rather than approach (Gray, 1987; cf. Dwyer, 2004; Malmkvist, 2001). There was a single stimulus, the predator silhouette (see Table 5.2), with which the mink could not make contact since it moved some distance above the cage; thus, only duration of orientation was assessed for that stimulus. Latencies were recorded live using a stopwatch, while video was used to obtain durations. Mink were also assigned a fear score out of four on each test, based on the presence or absence of four common indicators of fear in this species: retreat, alternation between retreat and withdrawal, screaming, and spraying a stress odour (reviewed by Malmkvist, 2001; Malmkvist and Hansen, 2002).

Tests alternated between the three stimulus types (see order in Table 5.2), with an equal number of stimuli from each category being presented in the morning (beginning at 08:30) and afternoon (beginning at 14:00) sessions. When aversive stimuli were presented in the morning, no other test took place in the afternoon, to allow the mink time to calm down if stressed by the test. All tests took place in the home cage; if a mink would not return to the home cage to begin a test, it was not conducted, because competing motivations to use the enrichments tended to distract the mink from attending to the test stimulus when in the enriched cage. If mink were asleep and did not wake up or slept continuously for more than 180 seconds of the test, data for that test were
excluded from the analysis, because they were judged not to accurately reflect the individual’s
typical level of responsiveness.

Responsiveness to reward was also assessed as the number of food “treats” (foods the
mink are motivated to eat even when not hungry) the mink consumed. Three treat types were
presented in separate tests: wet cat food (Fancy Feast™ chicken hearts and liver), ferret treats
(Bandits™ chicken flavour), and hot dog sausages. For each test, 30 pieces were given, and the
number consumed in 15 minutes was recorded. Treats removed from the food dish and cached in
the nest box were counted as consumed, as these were typically eaten shortly after the end of the
test. All mink of each sex were given an equal amount of their regular food on these days, and the
proportion of it that had not been eaten was scored visually to the nearest 5% to control for
differences in appetite.

Statistical analysis

All data were analysed using JMP 8 (SAS Institute Inc., NC, USA, 2009). A general linear model
was applied to each dependent variable (duration oriented, latency to make contact, and duration
of contact, as well as proportion of treats consumed), controlling for individual as a random
factor, nested in sex, treatment, and inactivity, since the tests involved repeated measures. The
effects of stimulus type, stimulus order, and all interactions were analysed. Where stimulus order
had no linear effect that would indicate habituation or sensitization over the repeated tests,
stimulus was instead included as a categorical variable. Stimulus order effects are not presented
unless they interact with the variables of interest. This was the case for the treat consumption
tests. Inactivity was included as a categorical predictor in all models, with mink below the
median inactivity level in each sex and treatment classified as “low inactivity” (i.e. “active”), and
those above the median classified as “high inactivity” (i.e. “inactive”). This was to allow nesting
of individual within inactivity, which was not possible if the latter was a covariate. Initially, time
of day (AM or PM) was included in models to test for an effect, but it was not a significant
predictor of any variable of interest and was therefore removed. The models for treat
consumption also included the proportion of the regular diet not eaten, as a covariate. If, for any
reason, the mink were not visible on the video for a portion of the test, duration of time oriented
and duration in contact were calculated based on proportion of visible time. Transformations were required to meet the assumption of normality in some cases: latencies were log-transformed, while the orientation data were squared. A logit transformation with a bias correction factor of 0.003 was applied to the proportion of treats consumed. A non-parametric Wilcoxon signed rank test for matched pairs was used to determine whether fear scores were higher in tests with stimuli categorized as aversive than in any other tests, which would confirm that the choice of stimuli was appropriate. The prevalence of fear was compared between housing treatments using a chi-square test based on presence or absence of fear in any test. Inactivity levels were compared between housing treatments using Wilcoxon tests split by sex, since the data were non-normal.

**Results**

*Validation of stimulus categories*

Overt signs of fear occurred relatively infrequently. However, as expected, fear scores were higher for aversive than for neutral or rewarding stimuli (d.f.=28, W=-162.5 and W=-175.5, respectively; both P<0.0001). In fact, while 26 of the 29 mink showed fear on at least one trial, only four of those had a fear score above zero on any neutral or rewarding trial. There was no effect of housing treatment on fear scores (P>0.10). In addition, latencies to make contact differed significantly depending on stimulus category (F\_2,207= 18.5, P<0.0001; Figure 5.1), being longest for aversive stimuli (Tukey’s Honestly Significant Difference; see next section for interactions).

*General responsiveness: housing treatment effects*

For duration of orientation to the stimulus, housing treatment interacted with stimulus type (F\_2,220=6.69, P=0.002). NE mink spent more time oriented to neutral stimuli than E mink did; for other stimulus types, the pattern was similar but the differences between housing conditions were not statistically significant according to Tukey’s HSD tests (see Figure 5.2). There was also an interaction between stimulus type and stimulus number (i.e. order of presentation; F\_2,223=40.8, P<0.0001). In analyses of orientation split by stimulus type, stimulus number interacted with
housing for rewarding stimuli ($F_{1,26}=7.8$, $P=0.009$): mink spent less time oriented to the second rewarding stimulus than the first, and this effect was stronger in enriched mink. For latency to make contact, housing treatment had a significant main effect only: for all stimuli, NE mink had shorter latencies than E mink did ($F_{1,28}=9.65$, $P=0.004$; Figure 5.1). Results for total time spent in contact with the stimulus were similar to those for orientation. There was a housing by stimulus type interaction ($F_{2,191}=3.25$, $P=0.041$), and NE mink were found to spend longer than E mink in contact with neutral stimuli (see Figure 5.3). Overall, mink spent more time in contact with rewarding stimuli than with the other two stimulus types (main effect of stimulus type: $F_{2,191}=33.7$, $P<0.0001$; Tukey’s HSD).

The dependent variables in these tests co-varied significantly. Latency to make contact with the stimulus was inversely correlated with duration of contact, as expected since the two are non-independent ($F_{1,194}=35.9$, $P<0.0001$). It was also inversely correlated with duration of orientation ($F_{1,218}=23.2$, $P<0.0001$).

In the treat consumption tests, NE mink ate a higher proportion of the treats offered than E mink did ($F_{1,21}=4.71$, $P=0.042$; Figure 5.4). The amount of the regular diet that remained uneaten was not a significant predictor of treats consumed, nor did it differ between housing treatments (both $P>0.10$).

*General responsiveness: relationships with inactivity*

Inactivity was not increased in NE housing. In fact, for females it was higher in treatment E (chi-square=94.3, $P<0.0001$), while there was no significant difference for males ($P>0.10$). Neither did inactivity have a clear relationship with responsiveness. Its main effect was not significant in any of the above models. For latency, duration of orientation and duration in contact, it was often involved in significant interactions (time oriented:
housing*inactivity*stimulus number $F_{1,211}=4.17$, $P=0.042$; time in contact: housing*stimulus type*inactivity $F_{2,188}=4.10$, $P=0.018$). However, when analyses were split by housing condition, no effects of inactivity were detected within the NE group. For treats consumed, housing
condition did not interact with inactivity, so again there was no correlation specifically in the NE group.²

**Discussion**

The data were not consistent with apathy: compared to E mink, NE mink were no more inactive, nor were they less responsive to stimuli. Nor were the data consistent with anhedonia: NE mink did not show reduced responsiveness to rewards. In contrast, they were broadly consistent with boredom. NE mink spent more time oriented to and more time in contact with neutral stimuli, had shorter latencies to reach all stimuli, habituated in their orientation to repeated rewarding stimuli more slowly, and ate more treats in the 15 minute test period. The greater decrease in time oriented to rewards in E mink compared to NE mink is intriguing, since this faster habituation might be another indicator of lower levels of interest in the stimuli (cf. e.g. Eion and Morgan, 1976). However, this effect did not occur for any other stimulus types, and because orientation could only be measured for two rewards, we cannot absolutely rule out the possibility that this effect resulted from differences in response to one specific putative reward rather than true habituation.

Several possible alternative explanations for the observed effects were considered and eliminated. These results cannot be explained by increased fearfulness in NE mink, since they showed no more signs of fear than E mink and they exhibited shorter rather than longer latencies to make contact. In addition, latency to make contact and duration of orientation were inversely correlated; this pattern suggests that both variables reflected “interest” in stimuli. If, by contrast, behaviour was primarily motivated by fear, mink should have spent a long time oriented to the stimulus but been slow to approach it, producing a positive relationship between the variables. The shorter latencies to make contact could alternatively be explained by higher levels of impulsivity, which has been proposed to underlie stereotypic behaviour (reviewed by Dallaire et

---

² Intriguingly, equivalent models using stereotypic behaviour in place of inactivity showed that it was inversely correlated with time oriented (F₁,₂₆=7.31, P=0.012; split by housing due to an interaction, it remained significant in NE housing: F₁,₂₁=6.98, P=0.015) and time in contact (F₁,₂₅=12.6, P=0.002). SB was not significantly associated with latency to make contact (P>0.10).
al., 2011) and might therefore be expected to be more prevalent in animals raised without enrichment. However, a test of impulsivity in these mink showed no effect of housing (Dallaire et al., 2011). We were also able to demonstrate that hunger was not responsible for the results of these treat consumption tests, since amount of the regular diet eaten did not predict treat consumption, thus validating the use of these tests as a measure of response to reward. Finally, for rewarding stimuli only, previous work has indicated that non-enriched housing can increase responsiveness (e.g. van der Harst et al., 2003), an effect reflecting stress-induced sensitization of a reward pathway in the brain, the mesoaccumbens dopamine system (reviewed by Cabib, 2006; Spruijt et al., 2001). The same processes could perhaps be invoked to explain the increased consumption of food treats by non-enriched mink, but they are not consistent with our demonstration that this is part of a broader pattern of increased responsiveness to all stimuli: neutral and aversive ones too. Thus, boredom in the NE animals is the most parsimonious explanation for the pattern of results seen.

The insights yielded by these data were critically dependent on using a diverse array of stimuli. Indeed this is the first experiment of its kind to systematically investigate how enrichment affects responsiveness across stimulus types. Our a priori categorization of the stimuli as aversive or neutral was validated by the increased latencies to make contact with the aversive stimuli, as well as the increased incidence of fear-related behaviour in those trials. The increased time spent in contact with rewarding stimuli compared to aversive or neutral also supports their categorization (although again, this may have been driven by one of the two rewarding stimuli used in these tests). This methodological approach would thus seem a very useful one that now could be applied to other species and other housing conditions.

It remains unclear why previous work has found conflicting results regarding whether enrichment increases or decreases exploratory behaviour, albeit to a narrower range of stimuli than were tested here. They might reflect differences between species, or in the life stage tested; in the processes affected by enrichment, some of which (e.g. fear; stress-sensitization of reward systems) are likely to be stimulus specific, as discussed above; differences in whether the enrichment results in reduced fear; or instead in differences in the severity or duration of the non-enriched treatment, since chronic boredom might eventually turn to apathy and/or depression
(Wemelsfelder 1990; humans: Cohen and Taylor, 1970). The duration of the experimental observations themselves may also play a role. To illustrate, Einon and Morgan (1976) demonstrated that while socially isolated rats made fewer contacts with novel objects, the socially housed rats habituated in their amount of contact more rapidly, and thus might eventually spend less time in contact. These are all factors that could be used to refine future boredom-oriented work on mink and other species.

A number of future studies of mink per se are suggested by our findings and other data from farmed mink. Future work could, for example, investigate whether duration of housing treatment affects results: does spending many years in non-enriched housing tip mink from boredom to apathy, or even depression? Is this most likely in standard farm cages, which are far smaller than the NE cages used here? Formal tests of motivation (e.g. using weighted doors that allow mink to access stimuli by paying a price [cf. Cooper and Mason, 2000]) would also help to confirm that the measures used here truly reflect an increased motivation to obtain stimuli rather than differences only in time engaged with stimuli that are freely offered. The prediction is that for bored animals, any stimuli would have rewarding properties and so could act as positive reinforcers for which animals are willing to work hard for access. Finally, the potential role of inactivity could be investigated in more detail by looking at subtypes of this behaviour. For example, specific forms of alert inactivity have been suggested to indicate either boredom or apathy by other authors (e.g. Bolhuis et al., 2006; Gunn and Morton, 1995).

Overall, this study provides a first step towards operationalizing boredom in non-human animals. Although we cannot yet determine with certainty whether the subjective experience of the animals is similar to that of humans who self-report feeling bored, their behaviour was consistent with that state. Such means of operationally defining boredom for non-humans so that it can be quantified are very much needed, since reducing boredom is often a stated aim of enrichment (e.g. Newberry, 1995), and yet to date we have had no means of judging its success at achieving that aim.
Acknowledgements

Thank you to the staff at the Michigan State University farm, to Kaitlin Bahlmann for help in recording data from the videos, and to Jamie Dallaire and Megan Jones for feedback on the experimental design.
Table 5.1. Responsiveness to different types of stimuli depending on psychological state.

<table>
<thead>
<tr>
<th>State</th>
<th>Aversive</th>
<th>Neutral</th>
<th>Rewarding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apathy</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Anhedonia</td>
<td>Normal</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Boredom</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

Table 5.2. Stimuli used, in order of presentation.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Type</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Plastic bottle</td>
<td>Neutral</td>
<td>Novel and no apparent biological relevance</td>
</tr>
<tr>
<td>2. Handling glove</td>
<td>Aversive</td>
<td>Shown to elicit fear in some individuals (Meagher et al., 2011)</td>
</tr>
<tr>
<td>3. Moving toothbrush</td>
<td>Rewarding</td>
<td>Known to elicit prolonged chasing; can be used as a reward to elicit operant responses (unpublished data)</td>
</tr>
<tr>
<td>4. Peppermint scent</td>
<td>Neutral</td>
<td>Novel and no apparent biological relevance</td>
</tr>
<tr>
<td>5. Air puff</td>
<td>Aversive</td>
<td>Aversive to many species (e.g. Huot et al., 2001: rats; Lansade and Simon, 2010: horses)</td>
</tr>
<tr>
<td>6. Maraca</td>
<td>Neutral</td>
<td>Novel and no apparent biological relevance</td>
</tr>
<tr>
<td>7. Predator silhouette (eagle)</td>
<td>Aversive</td>
<td>Natural predator (Dunstone, 1993); may be innately frightening (cf. Brown et al., 1992)</td>
</tr>
<tr>
<td>8. Ocean scented candle</td>
<td>Neutral</td>
<td>Novel and no apparent biological relevance</td>
</tr>
<tr>
<td>9. Female faeces*</td>
<td>Rewarding</td>
<td>May be attractive to males during mating season, when test was conducted</td>
</tr>
<tr>
<td>10. Predator odour (bobcat urine)</td>
<td>Aversive</td>
<td>Natural predator (Dunstone, 1993); may be innately frightening (cf. Blanchard et al., 1990)</td>
</tr>
</tbody>
</table>

* Presented to males only.
**Figure 5.1.** Latency to make contact with the stimulus, split by housing treatment and stimulus type.

Data are back-transformed least squares means, with error bars indicating confidence intervals. * indicates a significant difference at the $\alpha = 0.05$ level.
Figure 5.2. Duration oriented to the stimulus, split by housing treatment and stimulus type. Data are back-transformed least squares means, with error bars indicating confidence intervals. * indicates a significant difference at the group $\alpha = 0.05$ level (Tukey’s HSD).

Figure 5.3. Duration in contact with the stimulus, split by housing treatment and stimulus type. Data are means ± standard error. * indicates a significant difference at the group $\alpha = 0.05$ level (Tukey’s HSD).
**Figure 5.4.** Proportion of treats consumed by housing treatment.

Data are back-transformed least squares means across all three treat types, with error bars indicating confidence intervals. * indicates that there is a significant difference between treatments at the $\alpha = 0.05$ level.
CHAPTER SIX
Housing effects on inactivity and its subtypes

This manuscript is in preparation for Appl. Anim. Behav. Sci.

Abstract

The effect of impoverished housing on inactivity levels varies across species and experiments; however, when significant increases in inactivity are observed, this inactivity is likely to reflect poor welfare. We aimed to identify specific subtypes of inactivity that could indicate poor welfare in mink. Thus, we assessed inactivity in enriched and non-enriched conditions (Experiments 1 and 2), and after mink were moved from enriched to non-enriched cages, or vice versa (Experiment 2 only). Inactivity levels were established using live scans. Three additional variables were recorded to see if they could help reveal meaningful subtypes of inactivity: posture, location in the cage, and whether the mink was alert or apparently asleep. Posture analysis was exploratory: directional predictions were not made for all postures. Inactivity in the nest box, and lying alert, were expected to be highest in non-enriched conditions, because they might reflect hiding and vigilance respectively. Both of these hypotheses received some support. Inactivity in the nest box was higher in non-enriched housing for females in Experiment 1 (housing*sex P<0.001), and increased by enrichment removal (P=0.018) for both sexes in Experiment 2, although not decreased by being given enrichment (P>0.10). Furthermore, there was some evidence linking it with fear and stress: in “glove tests”, fearful males spent the most inactive time in the nest box (sex*temperament P=0.054), while females showing the largest decreases in faecal corticosteroid metabolites (FCM) when moved to enriched cages also showed the largest decreases in inactivity in the nest box (sex*FCM change P=0.019). Lying alert was similarly higher in non-enriched-housing in Experiment 1 (P=0.013), and increased by enrichment removal (P=0.021); it was also decreased by being given enrichment (P=0.004). However, it did not co-vary with fearfulness in the glove test, or with FCM (both P>0.10). This suggests that it does not reflect anxiety-induced vigilance; instead, we suggest it may be an indicator of boredom. Enriched mink spent more of their lying time curled up (P<0.001) and less lying belly down (P=0.001) in Experiment 1. A trend for decreased time belly down was likewise
found when mink were moved from non-enriched to enriched cages in Experiment 2 (P=0.054).
However, there was no positive relationship between lying belly down and FCM, as predicted if this posture reflects greater stress. The assumption that enriched cages improved welfare was supported by the decreased FCM (P=0.040) as well as decreased stereotypic behaviour in females (sex*housing: P=0.004) in enriched cages. Thus overall, inactivity in the nest box and alert inactivity are potential indicators of poor welfare in farmed mink, while the relationship between posture and welfare now requires further study.

Introduction

As I described in Chapter 1, animals housed in impoverished environments sometimes exhibit elevated levels of inactivity compared to those seen in enriched environments (e.g. chimpanzees: Baker, 2004; foxes: Koistinen et al., 2009; lemurs: Dishman et al., 2009; parrots: Rozek et al., 2010). This raises concern for their welfare for several reasons. First, similar effects are sometimes seen in situations that are known to induce stress. For example, felids spend much of their time inactive and hidden when exposed to stressors such as unpredictable feeding or being moved to new cages (e.g. Carlstead et al., 1993; Rochlitz et al., 1998; Wielebnowski et al., 2002). Similar effects have been seen in rodents exposed to predation threat (e.g. Dalm et al., 2009). Wild-caught animals brought into captivity, which likely experience more severe stress in the captive environment than captive-born individuals do, may be particularly prone to this behaviour (e.g. Jones et al., 2011). It has also been suggested that environments that induce very high levels of inactivity and/or low levels of behavioural diversity also induce boredom (e.g. Wemelsfelder, 1990). However, inactivity is a broad category of behaviour, also including forms such as sleep, which is typically functional and associated with relaxation (in addition to forms associated with stress, such as hiding; see Chapter 1).

It may be possible to differentiate between forms of inactivity that reflect good welfare and forms that reflect poor welfare phenotypically, based on indicators such as location, whether the animal is asleep or alert, and posture. Location is clearly the key to identifying hiding, since hiding by definition means being concealed or out of sight. Thus, in the studies cited above, the animals were remaining under a shelter or behind a wall or other object that visually separated
them from the humans in the area (e.g. behind a litter pan: Carlstead et al., 1993). Level of alertness or responsiveness to stimuli may also be linked to fear: remaining stationary but alert is often considered a form of vigilance (e.g. Rasa, 1989; Soriano et al., 2006), and would thus be expected to occur more often when animals are anxious (e.g. Carlstead et al., 1993). Sleep or rest, by contrast, is more likely to be seen as a sign of relaxation and therefore good welfare (e.g. Abou-Ismail and Mahboub, 2011; Crockett et al., 1995). Researchers have also often used posture to identify specific sub-types of inactivity. It can vary with level of physical comfort (e.g. Tuyttens et al., 2008), or be used to distinguish between true sleep and other forms of rest (reviewed by Langford and Cockram, 2010). Several have also suggested that certain postures might indicate specific affective states such as lethargy or boredom (dog-sitting in sows: Vestergaard, 1984; Fraser, 1975; hunched posture in rabbits: Gunn and Morton, 1995).

In mink, there is some evidence that high levels of inactivity can sometimes reflect poor welfare. Most inactivity occurs in the nest box - an area in which they are largely hidden from sight, and to which they are known to retreat when startled (see Nimon and Broom, 1999) or anticipating an aversive stimulus (Hansen and Jeppesen, 2006). The importance of an opaque nest box as a hiding place has also been demonstrated in another fur-farmed mammal, the fox (reviewed by Nimon and Broom, 2001). Inactivity could therefore reflect hiding in at least some cases. On an individual level, mink that do not exhibit stereotypic behaviour, or exhibit very low levels, in some studies have an increased risk of being fearful (Hansen and Jeppesen, 2006) and increased levels of corticosteroids (Bildsøe et al., 1991) compared to those that exhibit high levels of stereotypic behaviour. Since these non-stereotypic individuals tend to be the most inactive (Axelsson et al., 2009; Bildsøe et al., 1991; Svendsen et al., 2007b), this suggests that inactivity may be linked to fear or chronic stress. However, these individual differences in indicators of fear or stress are not always found (see Svendsen et al. 2007a; Malmkvist et al., 2011). Turning to effects at the population level, to date, few studies have investigated the effects of environmental enrichment on inactivity in this species. Those that have reported no overall increase or even a decrease (e.g. Hansen et al., 2007; Vinke et al., 2005) did not take into account the fact that results may differ for specific subtypes of inactivity. The loss of enrichment may also induce effects on inactivity that are not present in individuals that have never experienced enrichment (cf. Bolhuis et al., 2006, who found pigs moved from enriched to barren pens were
the most inactive). Accordingly, Hansen and colleagues (2007) did find that when mink had access to an extra cage taken away, they then spent more time in the nest box.

The aims of these two experiments were to determine how enrichment influences inactivity in mink, both in terms of total inactivity, and specific subtypes differentiated by alertness, location and posture. This would allow us to identify forms that could be useful as welfare indicators, assuming that enrichment is associated with improved welfare. Our hypotheses were: 1a) enrichment would increase time spent sleeping rather than lying alert; b) enrichment would reduce time spent inactive in locations that might be perceived as “safe”, such as the nest box, and increase inactive time in exposed areas of the cage; and c) enrichment would influence the postures shown while inactive. This posture analysis was exploratory, since there are no published studies of mustelids’ resting postures; however, we hypothesized that enrichment would increase time spent lying on their backs, with the belly exposed, which is generally considered a vulnerable position for any species (e.g. Petru et al., 2009). Our second major hypothesis, tested in Experiment 2 only, was that these effects of housing would be stronger if the mink had prior experience with the other housing type. Finally, we predicted forms of inactivity that were associated with poor welfare and thus decreased by enrichment would be correlated with fear and stress. These experiments rested on the assumptions that enriched housing would reduce fear and stress, and that removal of enrichment would increase stress, which we aimed to confirm using temperament and corticosteroid levels.

Methods

Housing treatments

All mink were individually housed in one of two cage types, enriched or non-enriched (see Chapter 5 for details). Enriched cages included running water in a small trough to allow wading and head-dipping, as well as objects with which the mink could interact, with new ones added each month. For Experiment 1, there was also a nest box in this cage in addition to the one provided in the home cage in both treatments.


**Experiment 1**

The first cohort of subjects consisted of 32 male and 32 unrelated female Black mink. Half of each group were housed in enriched cages and half in non-enriched cages from the time they were born (in late April or early May). Live scanning observations were conducted over six days in the second week of February, from 08:00 to 12:00 each day, with a scan every 30 minutes. Mink were observed from at least two cage lengths away to minimize disturbance, and behaviour was not recorded until they had ceased looking at the observer (see Bildsøe et al. 1990; Svendsen et al. 2007b for validation of live scans, and Dallaire et al., 2011 for more details on methods). Feeding began within an hour after the last scan. Behaviour, location, and posture if inactive were recorded according to the ethogram in Table 1. The two observers conducted additional scans on the same animals one afternoon, to confirm that inter-observer reliability was acceptable.

**Experiment 2**

The subjects in this experiment were 29 unrelated Black mink: 13 initially housed in E cages, and 16 initially housed in NE cages, with half of each group being male and the other half, female (see Chapter 5). They had been placed in these housing treatments at the age of approximately three months, at the time when they were split from sibling pairs.

*Behaviour in original housing*

This cohort of mink was observed for seven days in April (see Figure 6.1 for timeline). Again, all time budget data were collected using live scans in the mornings, from 08:00 to 12:00. They were, however, more frequent, occurring every 15 to 20 minutes. Temperament was also assessed in the winter using a glove test to determine whether housing affected the prevalence of fearfulness. In this test, mink were classified as fearful, curious, aggressive or unresponsive based on their response to a handling glove placed against the cage (see Chapter 2).
**Housing change**

Following the original observation period, the housing treatments were reversed: E mink were moved to NE cages, and vice versa. This was done in two stages, with half of the mink moved each time, to allow us to control for effects of time that were independent of the housing change. In each case, mink were given one week to habituate after being re-located, at which time another six to seven days of scanning observations were conducted.

At the beginning of the observation periods before and after the mink were moved, faecal samples were collected to assess cortisol metabolite output (FCM). The samples were collected every 24 hours, and pooled over three days to get an average level for each individual. Cortisol metabolites were extracted from the homogenized samples using 80% methanol and analyzed using an 11β-hydroxyaetiocholanolone enzyme immunoassay (Frigerio et al. 2004; see also Chapter 4). The use of this FCM measure has been validated for mink, albeit using a sample of females only (Malmkvist et al., 2011). These data, together with the change in stereotypic behaviour, were used to confirm that losing enrichment induced stress.

**Statistical analysis**

Analyses for Experiment 1 were conducted using Minitab 14 (Minitab Inc., PA, USA, 2003). General linear models (GLMs) were used to test for housing effects on total inactivity and all subtypes of interest, controlling for sex. Where interactions were found between sex and treatment, Tukey’s Honestly Significant Difference (HSD) tests were used to identify groups that differed at the level of α=0.05.

Analyses for Experiment 2 were conducted using JMP 8 (SAS Institute Inc., NC, USA 2009). GLMs were then used to test for housing effects on each subtype of inactivity before the housing change. Again, these models controlled for sex, and its interaction with housing. Housing effects on temperament in the glove test were assessed using a Chi-square test. Change in FCM after the move compared to before the move was analysed using GLMs controlling for sex, treatment and whether the individual was in the first or second group moved. The group term and
any interactions were removed from the model if P>0.25. Change in FCM between observation periods 1 and 2 for only those mink that had not yet been moved was analysed in the same way, to determine whether seasonal effects were present. The response of each subtype of inactivity to that housing change was then assessed, in terms of the absolute increase or decrease in time spent on that behaviour compared to the baseline level before the housing was changed. One-way t-tests, or Wilcoxon signed-rank tests in the case of non-normal data, were used to determine whether the change from baseline was significant within each treatment group; these analyses were split by sex if GLMs indicated that there were interactions between sex and treatment. Finally, for any forms of inactivity that were reduced by enrichment, GLMs were used to determine whether they co-varied with temperament or FCM. Levels of these forms of inactivity were also compared in period 2 between individuals that had remained in NE housing and those that had just been moved from E to NE cages. All tests were two-tailed except in the analysis of FCM and stereotypic behaviour changes, in which it was expected that both FCM stereotypic behaviour would increase when enrichment was removed. Assumptions of parametric tests were checked by visual inspection of the residual plots (Grafen and Hails, 2002), and transformations were applied where necessary.

Results

Experiment 1

All data on housing effects (means and statistics) in Experiment 1 are presented in Table 6.2. For total inactivity, there was an interaction between treatment and sex. Non-enriched males were less inactive than enriched males, while no significant effect was detected in females, although numerically they differed in the opposite direction (Tukey’s HSD, P>0.05). When inactivity was divided into alert or resting, time spent resting showed a similar pattern, being lower in NE males than E males, while there was no difference among females. Resting accounted for an average 92 ± 15% of inactive time. Time spent inactive but alert (log-transformed), by contrast, was higher in NE mink regardless of sex.

When inactivity was categorized based on location, NE females spent more time inactive in their nest boxes than E females did, while there was no such housing treatment effect on males.
This seemed to be compensated by E mink spending inactive time in the bridges rather than in the main area of the cage: when inactivity in the nest box was pooled with that in the bridges, females exhibited no difference between treatments (Tukey’s HSD, P>0.05). Instead, the opposite effect was found for males: E males spent more time inactive in those areas than NE males (treatment*sex $F_{1,60}=4.60$, $P=0.036$). Postures differed between treatments: mink in group E spent more of their inactive time curled up than those in group NE, and less of their time lying belly down (i.e. on the sternum). No significant difference in time spent lying belly up was detected ($P>0.10$), perhaps because relatively few individuals in either treatment were observed in this position. See Table 6.3 for a summary of all enrichment effects.

**Experiment 2**

*Effects of original housing*

For total inactivity in the spring, there was an interaction between housing treatment and sex ($F_{1,25}=11.0$, $P=0.003$). Tukey’s HSD tests showed that NE females were less inactive than E females, while there was no significant effect of housing in males. This effect was driven by time spent resting, which showed the same pattern (treatment*sex $F_{1,25}=7.49$, $P=0.011$) and represented an average $92 \pm 6\%$ of total time inactive. No significant effects of housing treatment on time spent lying alert were found ($P>0.10$). There were also no significant effects of housing on time spent inactive in the nest box, or on the relative frequency of any of the lying postures (all $P>0.10$). All results for effects of housing and housing change on time budgets and postures are again summarized in Table 6.3.

There was also no difference between the number of E and NE mink categorized as fearful in the glove test (31% in both groups; $P>0.10$). However, E mink did have lower levels of FCM (log-transformed: $F_{1,25}=4.67$, $P=0.040$). Stereotypic behaviour was highest in NE females, and did not differ between males in different treatments (logit-transformed: treatment*sex $F_{1,25}=9.91$, $P=0.004$; Tukey’s HSD).
Effects of housing change

There was a change in FCM over time, independent of housing change: FCM tended to decrease overall between observation periods in the group that had not yet been moved (t=-2.04, d.f. = 13, P=0.062). This effect was significant in the NE treatment (S=-10.5, d.f. = 7, P=0.031), but not in the E treatment (P>0.10). Thus, there was a decrease in all groups after they were moved, regardless of whether they went from E to NE or from NE to E. However, this decrease was larger in those moved to E housing (F_{1,23}=3.63, one-tailed P=0.035). The change in stereotypic behaviour differed between sexes (direction of housing change*sex F_{1,25}=9.43, P=0.005). One-tailed Wilcoxon signed-rank tests showed that there was a non-significant trend for an increase in females moved to NE (S=5.00, d.f. = 6, P=0.063), while there was a significant decrease in females moved to E (S=-17.0, d.f. = 7, P=0.008). Males did not change in their level of stereotypic behaviour, when housing was changed in either direction (P>0.10).

There was no overall change in inactivity with time for individuals that were not moved to new cages (Wilcoxon signed-rank test, P>0.1). When the housing treatments were reversed, however, there was an interaction between sex and direction of change for total inactivity (F_{1,25}=5.42, P=0.028). Males moved from NE to E showed a significant decrease in inactivity (t=-2.38, d.f. = 7, P=0.049), while females moved in the same direction tended to increase in inactivity once moved to the E housing (t=2.05, d.f. = 7, P=0.08). Neither sex, in contrast, showed significant changes in total inactivity when moved from E to NE (P>0.1). Similarly, for time spent resting, there was an interaction between sex and direction of treatment change (F_{1,24}=4.93, P=0.036). Again, males moved from NE to E showed a significant decrease (S=-17.0, d.f. = 7, P=0.016) while those moved from E to NE did not change significantly in the amount of resting observed (P>0.1). However, in this case, females showed a significant change only if moved from E to NE, in which case there was a decrease (S=-13.0, d.f. = 6, P=0.031; NE-E: P>0.1). For time spent lying alert, the effects did not differ between sexes (P>0.1). This behaviour increased significantly in mink moved to NE (S=32.5, d.f. = 12, P=0.021; Figure 6.2), and decreased in those moved to E (S=-53.5, d.f. = 12, P=0.004). In period 2, however, time lying alert was not significantly higher (although it was numerically so) in mink just moved to NE than in those that had always been housed in NE cages (P>0.1).
Inactivity in the nest box increased significantly in mink moved to NE (sexes pooled; \( t=2.73, \text{d.f.} = 12, P=0.018 \)) but did not change significantly in those moved to E cages (Figure 6.3). In period 2, similarly to lying alert, it was not significantly higher in those just moved to NE than in those remaining in NE that had always been housed in those cages (\( P>0.10 \)). Figure 6.4 illustrates the chances in posture. The relative frequency of lying belly up tended to decrease in mink moved to E (\( t=-1.93, \text{d.f.} = 15, P=0.07 \)), but was unaffected by a move from E to NE (\( P>0.10 \)). Conversely, lying belly down tended to increase in mink moved to NE (\( S=22.0, \text{d.f.} = 15, P=0.054 \)) and remained unchanged in those moved to E (\( P>0.10 \)). Lying curled up was not significantly affected by being moved to E, but tended to decrease if moved to NE (\( t=-1.92, \text{d.f.} = 12, P=0.079 \)). Again, these findings are summarized in Table 6.3.

**Relationships between subtypes of interest and endocrine stress or fearfulness**

The forms of inactivity that were sometimes decreased by enrichment were lying alert, inactivity in the nest box, and lying belly down. The relationships between these variables and FCM and response in the glove test were therefore assessed. Time spent lying alert in the pre-move period did not correlate with either FCM or temperament in the glove test (both \( P>0.10 \)). There was also no relationship between the change in time lying alert after being moved and the change in FCM (\( P>0.10 \)). For time inactive in the nest box, there was no significant relationship with FCM before the move (\( P>0.10 \)). There was a trend for a relationship with temperament as assessed by the glove test, which interacted with sex (\( F_{2,20}=3.40, P=0.054 \)): fearful males spent the most time inactive in their nest boxes, and curious females spent the least. However, Tukey’s HSD tests did not show any significant differences between fearful and curious mink within either sex (\( P>0.05 \)). Similarly, there was an interaction with sex in the analysis of change in inactivity in the nest box versus change in FCM (\( F_{1,22}=6.41, P=0.019 \)), but the relationship with FCM did not attain statistical significance in either sex when the analysis was split (\( P>0.10 \)). By inspection, however, it appeared that for females, a large decrease in FCM corresponded with a large decrease in inactivity in the nest box. Proportion of inactive time spent belly down did not correlate with FCM before the move (\( P>0.10 \)). For temperament, there was an interaction with treatment (logit-transformed; \( F_{2,17}=3.53, P=0.047 \)): in E, fearful individuals spent the most time
in that position, while in NE, they spent the least. Again, however, none of these differences were significant according to Tukey’s HSD. Unexpectedly, there was a negative relationship between change in the relative frequency of this position after the move and change in FCM, such that decreases in the proportion of inactive time spent belly down corresponded to increases in FCM (log-transformed; \( F_{1,23}=10.5, P=0.004 \)).

**Discussion**

My experimental hypotheses were broadly that enrichment was expected to promote resting, but to decrease types of inactivity associated with negative affective states. Specific predictions were that enrichment would: increase time spent sleeping but decrease time spent lying alert (Hypothesis 1a); reduce the time spent inactive in locations that might be perceived as “safe”, such as the nest box, while increasing inactive time in exposed areas of the cage (Hypothesis 1b); and also influence the postures shown while inactive (Hypothesis 1c). A second general hypothesis, tested in Experiment 2 only, was that these effects of housing on forms of inactivity would be stronger if the mink had prior experience with the other housing type. Testing these hypotheses rested on the assumption that, just as in previous mink studies (see Introduction), environmental enrichment would benefit the subjects’ welfare. This assumption was indeed clearly supported by my data. In Experiment 2, female E mink were less stereotypic than female NE mink; and in Experiment 2, E mink also excreted lower levels of FCM. They were no less likely to be fearful, however; this may reflect the reduced power of tests using qualitative data (although they also showed no less fear of novel stimuli; see Chapter 5), or instead indicate that the stress induced by non-enriched housing is of a type other than fear, such as frustration or boredom. Moving mink between housing types confirmed these welfare benefits of E housing, in that the time-dependent FCM decline across all subjects was greatest in mink moved from NE to E housing, and the same move reduced female (although not male) stereotypic behaviour. However, it was unclear that the loss of EE was worse for welfare than lifelong privation (Hypothesis 2).
The effects of environmental enrichment on the total amount of time spent inactive were somewhat variable. Where there were significant effects of housing, inactivity increased overall when given enrichment; however, in Experiment 2, being given enrichment later in life had the opposite effect in males. This might result from a change in the enrichment given: the increased inactivity in males occurred in Experiment 1, in which the enriched cages included an extra nest box which might promote resting, while the additional nest box was not provided in Experiment 2. This variation was not, however, surprising, given the anticipated heterogeneity of this behaviour category. My expectation was therefore that studying different subtypes of inactivity separately – subtypes based on apparent function, on location, and on posture – should produce less variable, more predictable, results.

Analysing time spent resting did not in practice produce very different results from those obtained using total inactivity, since resting represented such a large proportion of inactive time. Thus, in partial support of Hypothesis 1a, resting or sleeping was increased by enrichment in Experiment 1’s males (but not females), and Experiment 2’s females (but not males); and it decreased in females moved from E to NE housing (but puzzlingly increased in males moved from NE to E housing). However, more convincing as support of Hypothesis 1a was the finding that lying alert did seem to be reliably decreased by enrichment. In Experiment 1, this behaviour occurred less in enriched than non-enriched mink, while in Experiment 2 it decreased in mink moved from NE to E housing, but increased in mink moved from E to NE. Several authors have documented similar effects in other species, with both pigs and mice spending more time inactive but awake when in impoverished versus complex living conditions (e.g. Bolhuis et al., 2006; Tilly et al., 2010). This lying alert was unrelated to FCM or glove tests measures of fear or stress, suggesting that this behaviour does not reflect vigilance in mink. Instead, it may be a response to a lack of behavioural opportunities, such that there are few activities to occupy their waking hours. Bolhuis and colleagues (2006) similarly hypothesized that the elevation in alert inactivity they observed in pigs housed in impoverished environments reflected apathy (see Chapter 5), while several other authors have suggested that inactivity under such conditions might reflect boredom (e.g. Wood-Gush and Beilharz, 1983; van Putten, 1980; see Chapter 5).
The data also suggested that, in support of Hypothesis 1b, categorizing inactivity by location holds promise for assessing mink welfare. Thus inactivity in the nest box often seemed to be reduced by enrichment: it was decreased in females given enrichment in Experiment 1 (even though in this experiment, enrichment provision included doubling the number of available nest boxes), and in Experiment 2 it was increased by enrichment removal, while correspondingly decreased by providing enrichment to the formerly NE subjects. Furthermore, we found some evidence to support the hypothesis that this specific form of inactivity reflects fear, since fearful males spent the most time inactive in their nest boxes, while curious females spent the least time there; and for females, changes in this behaviour also co-varied with changes in FCM after the move between treatments, with a large decrease in FCM corresponding to a large decrease in inactivity in the nest box.

Hypothesis 1c concerned posture, irrespective of location. These analyses were primarily exploratory, and thus directional predictions were not made. The broad hypothesis was supported, in that housing type did indeed affect the posture adopted while inactive. Enriched animals in Experiment 1 were more likely than non-enriched mink to lie curled up, and correspondingly, in Experiment 2 this type of lying tended to decrease in mink from which enrichments were removed. Conversely, enriched animals in Experiment 1 were less likely than non-enriched mink to lie belly down, and again correspondingly, in Experiment 2 this type of lying tended to increase in mink from which enrichments were removed. To some extent, these findings likely mirror those on sleep or rest versus lying awake, since lying on belly in practice was a subset of lying awake: mink almost never slept in this position. Analyses did not, in contrast, detect any treatment effects on the category in which we had been primarily interested, lying belly up. This may simply be due to a lack of statistical power, since it proved to be a relatively rare posture in these mink. However, the trend for a decrease in time lying belly up when moved from NE to E conditions suggests that, in contrast to expectations, it may not actually reflect greater relaxation.

No a priori hypothesis had been made for sleeping curled rather than on the belly or sternum, but the predominance of curled postures in enriched conditions thus indicate that it may, in fact, be the most relaxed position for mink. Lying curled has indeed been suggested to indicate
deeper sleep or increased motivation to rest in other species, such as rats (Tromp et al., 1990) and foxes (Tembrock, 1979). Lying belly down, on the other hand, likely allows the fastest escape if awoken (cf. Aristakesyan, 2009) since it is the closest to a standing position. However, inconsistent with this idea that lying belly down indicates poorer welfare than sleeping curled was the finding that the proportion of inactive time spent belly down increased if FCM decreased when the mink were moved to new housing. More research is clearly needed before this posture can be used as welfare indicator.

There was some weak support for Hypothesis 2: that changed housing conditions lead to more dramatic effects than stable housing conditions. Overall, no effects were induced by a change in housing treatment that had not been seen as simple effects of long-term housing in at least one group. However, the loss of enrichment did significantly increase inactivity in the nest box in both sexes, while being reared without enrichment had resulted in elevated inactivity in the nest box only in females in Experiment 1 (with no significant baseline housing effect in Experiment 2). The housing change also induced significant effects on time lying alert that paralleled the treatment effects seen in Experiment 1, but which had not been statistically significant in the original housing in Experiment 2. Although levels of these two behaviour patterns were not significantly higher in those shifted to non-enriched cages before period 2 than in those that had no experience with enrichment, this is likely due to low statistical power in these analyses. This experiment should therefore be repeated with a larger sample size.

Overall, these results help explain why environmental enrichment in other studies on other species has rather inconsistent effects on overall levels of inactivity or activity. Enrichment may promote resting, but conversely it may decrease types of inactivity associated with negative affective states like boredom, apathy or fear. The balance between these contrasting effects will likely depend in part on the type of enrichment offered. For example, physical structures that have the potential to increase perceived safety, such as shelters for rodents, have previously been found to increase inactivity (e.g. Würbel et al., 1998), as do added substrates that seem to increase lying comfort (see Chapter 1); while in contrast, “toys” or foraging enrichments are expected to have the opposite effect, elevating overall activity, as in other examples cited in the Introduction (see also Chapter 1). The enriched conditions offered to the mink in this experiment
offered both safe resting places and opportunities for active interaction (objects to manipulate and wading water), and thus individuals could clearly use the enrichment in different ways. Our data help shows that different subtypes of inactivity -- identified by location, posture and/or apparent level of arousal -- have quite different welfare implications, being differentially affected by enrichment provision or removal and in some cases differentially associated with elevated FCM output or fearful temperaments. Future studies of mink welfare, and future enrichment studies in general, should thus record different forms of inactivity separately rather than using it as a catch-all category, to allow better understanding of how enrichment influences animals’ behaviour and welfare, and of how different sub-types of inactivity relate to welfare.

Specifically for farmed mink, that stress-reducing environmental enrichment differentially influences different subtypes of inactivity helps identify which subtypes now hold most promise as potentially valid welfare indicators for this species. Sleep and resting with the eyes closed, especially in a curled-up posture, may indicate relaxation, comfort or other forms of good welfare. This idea could now be tested with additional temperament data from a larger sample size, to see if these forms of inactivity are most common in curious individuals, and least common in fearful ones. The hypothesis that the curled posture represents deeper sleep could also be tested by threshold experiments to measure the magnitude of a stimulus required to arouse a subject. Obtaining EEG data to more reliably identify sleep and its subtypes might also help to clarify how housing conditions affect true sleep rather than mere rest (since the main effects of current housing conditions here were rather inconsistent). In terms of forms of inactivity possibly linked with poor welfare in this species, results indicated that remaining inactive in the nest box may reflect chronic fear or stress. Future corroborating tests could include offering mink of different temperaments nest boxes that differ in their degree of enclosure or distance from humans. Furthermore, remaining inactive but alert (which often occurs in a belly-down posture) reliably increased in non-enriched conditions suggesting reduced welfare; this now requires further study to identify the underlying affective state, with boredom or apathy (see Chapter 5) being the most likely candidates. Taken together, our data therefore provisionally suggest that sleeping curled up outside of the nest box is the form of inactivity best indicative of a relaxed, high welfare mink; while in contrast lying awake within the nest box is the form that is most concerning from a welfare perspective.
Table 6.1. Ethogram.

<table>
<thead>
<tr>
<th>Behaviour pattern</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stereotypic behaviour</td>
<td>Movement or sequence of movements repeated at least 3 times consecutively</td>
</tr>
<tr>
<td>Borderline stereotypic behaviour</td>
<td>Apparently stereotypic behaviour repeated fewer than 3 times or switching between elements of common stereotypies without repeating a sequence 3 times</td>
</tr>
<tr>
<td>Scratching</td>
<td>Scratching continuously at a wall or edge of the cage for ≥ 3 sec</td>
</tr>
<tr>
<td>Inactivity</td>
<td></td>
</tr>
<tr>
<td>Resting</td>
<td>Lying relatively still with eyes closed or head tucked in (apparently asleep)</td>
</tr>
<tr>
<td>Alert</td>
<td>Lying relatively still with eyes open</td>
</tr>
<tr>
<td>Unknown</td>
<td>Inactive but cannot tell whether alert</td>
</tr>
<tr>
<td>Normal activity</td>
<td>Animal neither inactive nor engaged in stereotypic behaviour; includes eating, drinking and grooming</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Location of inactivity</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>In nest box</td>
<td>At least half of body in nest box</td>
</tr>
<tr>
<td>Other</td>
<td>Elsewhere in home cage</td>
</tr>
<tr>
<td>Enriched cage</td>
<td>In enriched cage</td>
</tr>
<tr>
<td>Bridge</td>
<td>In the bridges connecting the two cages, including the tower (upright portion of the structure)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Posture of mink while inactive</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Belly up</td>
<td>Lying on back, with belly exposed</td>
</tr>
<tr>
<td>Belly down</td>
<td>Lying on sternum</td>
</tr>
<tr>
<td>Curled</td>
<td>Lying curled up</td>
</tr>
<tr>
<td>On side</td>
<td>Lying on side; only recorded in Experiment 2</td>
</tr>
</tbody>
</table>
Table 6.2. Housing effects on inactivity and its subtypes in Experiment 1.

<table>
<thead>
<tr>
<th>Behaviour Pattern</th>
<th>Males</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Females</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E</td>
<td>NE</td>
<td>E</td>
<td>NE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Housing x sex F_{1,60}=10.5, P=0.002</td>
</tr>
<tr>
<td>Total inactivity</td>
<td>0.73 ±</td>
<td>0.54 ±</td>
<td>0.66 ±</td>
<td>0.75 ±</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resting</td>
<td>0.70 ±</td>
<td>0.44 ±</td>
<td>0.64 ±</td>
<td>0.72 ±</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Housing x sex F_{1,60}=13.8, P&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>0.04</td>
<td>0.06</td>
<td>0.05</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lying alert(^1)</td>
<td>0.18</td>
<td>0.28</td>
<td>0.17</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Housing F_{1,60}=6.60, P=0.013</td>
</tr>
<tr>
<td></td>
<td>(0.15-)</td>
<td>(0.21-)</td>
<td>(0.14-)</td>
<td>(0.21-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.23</td>
<td>0.36</td>
<td>0.20</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inactivity in the nest box</td>
<td>0.61 ±</td>
<td>0.41 ±</td>
<td>0.12 ±</td>
<td>0.56 ±</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Housing x sex F_{1,60}=22.7, P&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lying curled up (proportion of</td>
<td>0.70 ±</td>
<td>0.32 ±</td>
<td>0.99 ±</td>
<td>0.54 ±</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Housing F_{1,45}=15.2, P&lt;0.001</td>
</tr>
<tr>
<td>inactive time)</td>
<td>0.12</td>
<td>0.13</td>
<td>0.003</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lying belly down (proportion of</td>
<td>0.19 ±</td>
<td>0.55 ±</td>
<td>0.002 ±</td>
<td>0.32 ±</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Housing F_{1,45}=12.3, P=0.001</td>
</tr>
<tr>
<td>inactive time)</td>
<td>0.10</td>
<td>0.13</td>
<td>0.002</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lying belly up (proportion of</td>
<td>0.11 ±</td>
<td>0.13 ±</td>
<td>0.006 ±</td>
<td>0.13 ±</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Housing P&gt;0.10</td>
</tr>
<tr>
<td>inactive time)</td>
<td>0.09</td>
<td>0.08</td>
<td>0.003</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data are means (proportion of all scans) ± standard errors unless otherwise specified. Significant differences between housing treatments are highlighted in bold.

\(^1\) Data are back-transformed means with confidence intervals in parentheses.
Table 6.3. Summary of effects of enrichment.

<table>
<thead>
<tr>
<th>Behaviour</th>
<th>Cohort 1</th>
<th>Cohort 2</th>
<th>Housing change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>F</td>
<td>E-NE:</td>
</tr>
<tr>
<td>Total inactivity</td>
<td>+ve</td>
<td>n.s.</td>
<td>+ve</td>
</tr>
<tr>
<td>Resting</td>
<td>+ve</td>
<td>n.s.</td>
<td>+ve</td>
</tr>
<tr>
<td>Lying alert</td>
<td>-ve</td>
<td>-ve</td>
<td>n.s.</td>
</tr>
<tr>
<td>Inactivity in NB</td>
<td>n.s.</td>
<td>-ve</td>
<td>n.s.</td>
</tr>
<tr>
<td>Proportion of inactivity spent</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>belly up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion of inactivity spent</td>
<td>-ve</td>
<td>-ve</td>
<td>n.s.</td>
</tr>
<tr>
<td>lying belly down</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion of inactivity spent</td>
<td>+ve</td>
<td>+ve</td>
<td>n.s.</td>
</tr>
<tr>
<td>curled</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

n.s. means no significant effect (P>0.10). Brackets indicate a trend 0.05<P<0.10.

¹ Signs indicate enriched mink relative to non-enriched. For Cohort 2, these are housing effects before the treatment change.
Figure 6.1. Timeline of events in Experiment 2. Boxes above the line indicate behavioural observation periods; white portions of the boxes indicate faecal collection days.
Figure 6.2. Change in time spent lying alert after a change in housing treatment. Data are medians, with error bars showing the interquartile range.

Figure 6.3. Change in time spent inactive in the nest box after a change in housing treatment. Data are means ± standard errors.
Figure 6.4. Change in lying time spent in each posture after a change in housing treatment. Data are means ± standard errors. † indicates an effect that differed from zero with 0.05<P<0.10.
CHAPTER SEVEN
General Discussion

Inactivity is a broad category of behaviour that tends to receive less research attention than overtly abnormal behaviour patterns in captive animals, although certain forms of it are likely linked to welfare. This project aimed to characterize the subtypes of inactivity in farmed mink, and to determine how inactivity and its subtypes relate to their reproductive success as well as their long-term welfare at both the population and individual levels. This involved testing several alternative hypotheses, drawn from analogies with human behaviour, regarding the psychological states that might underlie very high levels of inactivity. The focus was on forms of inactivity that occupy significant portions of the time budget, such as resting, which would be sensitive to chronic stress, rather than very acute forms, such as freezing. Here, I will review the specific hypotheses under test and the evidence for or against them, discuss inconsistencies in the results across experiments, make suggestions for future research and describe the implications of these findings both for researchers and for farmers.

Summary of objectives and results

This project had four major objectives. The first two of these were to determine whether high levels of inactivity reflect chronic stress or fear, on an individual level, and to test whether that resulted in decreased reproductive success (and therefore, likely decreased profit for the farmer). The large-scale experiment that spanned multiple farms and colour-types found no evidence that the most inactive female mink were more likely to be fearful than the most active females. Nor did they show increased baseline cortisol levels, although this is not an ideal measure of stress when comparing animals that differ in physical activity, since activity also influences cortisol concentrations (e.g. West et al., 2004). Some aspects of reproductive performance, including kit growth, were also enhanced rather than reduced among inactive females. Thus, the hypothesis that high levels of overall inactivity reflect fear or stress on an individual level can be rejected. There was, nonetheless, some evidence of decreased reproductive success: inactive females gave birth to fewer kits in a given year. Although the evidence above indicates that this did not reflect stress-induced reproductive impairment, it was demonstrated to reflect some intrinsic difference
between active and inactive females, rather than being a simple physiological effect of body fat. Under natural conditions, their lifetime fitness would likely be no worse than that of active females who produce large litters, because it would be compensated by increased success of their male offspring. However, this litter size effect makes high levels of inactivity practically undesirable on farms, because most males do not survive to reproduce, and the number of pelts produced is more economically important than their size (Lagerkvist, 1997; see Chapter 4).

The third objective of the project was to develop methods of assessing the different psychological states hypothesized to underlie inactivity in suboptimal environments, in order to test which of them was most likely to exist in our populations. I developed and validated a practical temperament test, the “glove test,” which is more sensitive to fear than the existing equivalent is. However, I found no evidence that the prevalence of fear in this test was greater among inactive than active mink, nor was it elevated overall in non-enriched compared to enriched conditions, which were assumed to foster improved welfare. This absence of a difference in fearfulness between enriched and non-enriched conditions, while contrary to many studies in other species (see Chapter 6), is consistent with evidence from an earlier group of mink housed in these same conditions (Díez-León et al., in prep.). The possibility remains, however, that qualitative tests of fear did not have sufficient power to detect effects of enrichment in a sample of this size. As with fear, there was no evidence to support the hypothesis that very inactive mink in standard farm conditions were apathetic or depressed: they did not show decreased responsiveness to rewards (indicative of anhedonia) or to neutral or aversive stimuli, compared to either more active individuals or to mink housed in enriched environments. There was, in contrast, evidence of boredom in non-enriched conditions, as indicated by increased responsiveness to all types of stimuli. This state did not co-vary with total inactivity levels, and is thus not a complete explanation for the profound inactivity sometimes observed on farms. However, it may underlie specific subtypes of inactivity, as discussed below. The elevated corticosteroid levels (FCM) exhibited by mink housed in non-enriched cages further support the assumption that this housing condition did induce some other form of chronic stress; this may reflect boredom, which involves frustration of the motivation to explore or obtain stimulation, or it may reflect the frustration of other motivations (cf. Mason et al., 2001: cortisol increased when mink were deprived of highly valued resources).
The lack of apparent differences in welfare between active and inactive mink thus does not eliminate the possibility that inactivity can be a stress response. Rather, very high levels of inactivity and stereotypic behaviour may be alternative responses (perhaps part of different behavioural syndromes or “coping styles;” see Koolhaas et al., 1999; Sih et al., 2004) that do not differ, on average, in their success at mitigating the effects of a stressful environment. One behavioural phenotype might fare better than the other under a particular set of environmental conditions, explaining why the relationships between stereotypic behaviour and welfare indicators such as corticosteroids seem somewhat variable in this species (reviewed in Chapter 6); Sih and colleagues (2004), for example, state that active or “proactive” coping styles are probably more successful than others in a stable environment, but less successful in a changing environment.

My final objective was to identify specific subtypes of inactivity that could be used as welfare indicators. Location and level of alertness both proved useful in distinguishing subtypes. Inactivity in the nest box seemed linked to stress, because it is elevated in non-enriched environments, linked to fearfulness in a temperament test, and increased by changes that also increase faecal corticosteroid output. We have also found that juvenile mink spend more time in the nest box immediately after weaning, an acute stressor, and that this response was linked to neophobia during the same period (Meagher et al., in prep.). Lying alert likely also reflects poor welfare, but the specific state involved now needs testing; boredom is a likely candidate, since this behaviour pattern was increased by conditions shown to induce signs of boredom, and did not seem linked to fear. Lying alert specifically in a belly down position may be a stronger welfare indicator, while sleeping curled up outside the nest box may be the strongest indicator of good welfare.

**Contradictory results across experiments**

Although time in the nest box seemed linked to fear in the experiments of Chapter 6, inactivity was not linked to fear or to poor kit growth and survival in the breeding study (Chapters 3 and 4) even though most of it (e.g. an average 86.6 ± 14.5% in Blacks in Experiment 2) occurred in the
nest box. There may be a sex difference in the tendency to express fear in this manner: in Experiment 2 of Chapter 6, fear in the glove test interacted with sex to predict inactive time in the nest box, with fearful males spending the most time there, whereas Chapters 3 and 4 focused only on females. However, both sexes increased their inactive time spent in the nest box in response to having enrichment taken away. It seems likely that for both sexes, some inactivity in the nest box is normal, perhaps motivated by increased warmth or comfort, while only very high levels reflect fear. There may be differences between populations in fearfulness (cf. the temperament differences between our study subjects and Scandinavian mink; Chapter 2), as well as differences between cage designs in the perceived safety of the nest box, and therefore in the proportion of nest box use that reflects fear; this would make total inactive time in the nest box a better indicator in some contexts than in others. However, changes in nest box use after alterations in management, as seen in Chapter 6, should be a useful indicator anywhere.

No specific subjective state has yet been linked to the decreased litter size in inactive females in Chapters 3 and 4. Boredom is the only negative affective state for which we found direct evidence in non-enriched housing conditions like those found on the commercial farms. Although we have hypothesized that it may be linked to specific forms of inactivity (lying alert), it was not linked to total inactivity, nor does it predict decreased reproductive success, unlike depression or fear. The mechanism for the decreased litter sizes thus still remains unidentified.

Suggestions for future research

I found no evidence that extremely inactive mink had poorer welfare than their active, stereotypic counterparts; instead, as stated above, they seem to be exhibiting equal levels of stress that are manifest in a different way. However, this qualitative difference in the behavioural response to a given situation makes comparisons of welfare difficult. It is therefore worth confirming using cognitive bias tests; cognitive biases are very closely linked to affective states, at least in humans (reviewed in Mendl et al., 2009), and all individuals would be expected to show qualitatively similar biases (i.e. biases in the same direction) when experiencing negative affect. There could also be a non-linear relationship between inactivity and welfare, such that the true extremes of the population differ, but animals falling closer to the middle of the spectrum do not, or those
individuals in the middle have better welfare than those at either extreme. Comparing results using different cut-off points for the “active” and “inactive” groups might be informative, along with including a group that falls in the middle of the spectrum of activity levels.

There is still work to be done in refining the method of assessing boredom. One aim would be to develop a ‘boredom index’ that is a composite score of consistent high reactivity to diverse stimuli. The hypothesis that lying alert reflects boredom (see Chapter 6) could be tested by assessing whether it co-varies with responsiveness. Finally, although mink appeared bored and not apathetic after several months of being housed in barren cages, it remains to be seen whether apathy or depression would develop after a period of years, or in very small cages.

It would also be of interest to test whether proportion of the time budget spent lying and alert in mink co-varies with temperament in the glove test: mink categorized as “unresponsive” were awake but largely immobile during the test (often lying alert), and tended to spend more time inactive during scan sampling (Chapter 2). However, bored animals should be responsive in a glove test, and spend a lot of time lying alert only when no novel stimuli are available; thus, this behaviour should negatively co-vary across the different experimental contexts. Another question about the glove test that still needs investigating is whether responses in this test relate to fertility because fearfulness predicts barrenness, or because being pregnant reduces fear.

The type of inactivity shown in non-enriched housing may be useful in predicting how animals will use enrichment. This could be tested by assessing what proportion of their inactivity falls into each category, and then adding a variety of enrichment items. This would allow enrichments to be better targeted at specific individuals or groups in the future; for example, individuals that show very low levels of sleep might require a structure that would increase perceived safety, while individuals that spend a great deal of time lying alert might benefit most from objects to manipulate or other forms of stimulation.

To confirm that inactivity in the nest box is an indicator of fear in mink, animals could be offered a choice of more than one nest box that differed in one or more properties. If fear is the main motivator of nest box use, boxes at a greater distance from humans, or ones with covered
tops and three covered sides should be preferred (cf. mice: Van de Weerd et al., 1998), while if the boxes are used because they are warm or more comfortable than the wire mesh cage floor, those with nesting material should be preferred even if all sides are transparent. Whether or not mink are unresponsive in the glove test if it is conducted while they are in the nest box might also help to distinguish between individuals that are motivated to stay in the nest box due to fear and those that are using it for other reasons.

**Applications to other species and recommendations**

The basic variables of posture, location and level of alertness used here are likely informative in most species. Postures, of course, will have to be studied on a species-by-species basis, while the basic principles of choosing sheltered versus exposed locations and being awake versus asleep may have more universal meaning. In some situations, posture and even level of alertness may not be possible to assess during live observations, depending on cage or enclosure design and the distance at which the observer must stand; for example, postures of mink in the nest box were not visible if observed from another aisle. Using automated devices such as accelerometers to assess posture, and EEG to determine when animals are sleeping, could solve this problem.

The finding that inactivity is linked to decreased litter size independently of increased body fat suggests that this should be investigated in other species. Many studies of the effects of fat have likely not controlled for differences in behaviour that might co-vary with body condition, and may thus have misattributed the cause of some differences in reproduction. For agricultural species, this would mean that producers could improve productivity using genetic selection based on activity levels as well as enrichment programmes to stimulate activity and decrease body fat, like those already being attempted in laboratories (e.g. Johnson et al., 2004) and zoos (e.g. Altman et al., 2005). Not only would this potentially increase reproductive success more than simple feed restriction, but would also do so without compromising animal welfare.

The finding that non-enriched cages likely induce boredom rather than apathy is encouraging in the sense that boredom should be more easily reversible. Apathy is a persistent state of decreased motivation, and apathetic animals would be expected to interact with
enrichment less and therefore benefit from it less than other individuals would (see Tilly et al., 2010). Bored animals, by contrast, would be expected to make use of enrichment. Determining which of these states is present in a given population could again help to target the use of enrichment.

**Conclusions**

Standard farm cages likely do induce some form of poor welfare in mink, with boredom being the most likely subjective state involved, although there may also be other types of frustration or perhaps increased fear that was not detected due to low power of the statistical tests with qualitative data and the small number of animals available in Chapter 6. Within populations under given conditions, however, inactive mink likely have no worse average welfare than their very active, stereotypic counterparts. Regardless of the welfare implications, farmers may still want to discourage inactivity for practical reasons, because inactive females produced fewer kits. For welfare research, inactivity should not be treated as a single category of behaviour: different forms have different meanings, and pooling them together often masks actual changes, as evinced by the lack of significant enrichment effects on total inactivity in Chapter 6. Thus, while not all inactive mink have poor welfare, those individuals that spend a higher than average proportion of inactive time lying awake and/or in the nest box may be suffering.
REFERENCES


Mallapur, A., Chellam, R., 2002. Environmental influences on stereotypy and the activity budget of Indian leopards (Panthera pardus) in four zoos in southern India.

Malmkvist, J., 2001. Fear in Farm Mink (Mustela vison) - Consequences of Behavioural Selection.


McPhee, M.E., Carlstead, K., 2010. The Importance of Maintaining Natural Behaviors in Captive Mammals.


Møller, S.H., Hansen, S.W., 2001. Assessment of mink welfare at farm level. NJF Seminar No. 331, Snekersten, Denmark, October 1-3.


van der Harst, J.E., Baars, A.M., Spruijt, B.M., 2003. Standard housed rats are more sensitive to rewards than enriched housed rats as reflected by their anticipatory behaviour. Behav. Brain Res. 142, 151-156.


