Welfare benefit from enrichment provision schemes in American mink
(\textit{Neovison vison})

by

Samuel T. Decker

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Master of Science
in
Animal Biosciences

Guelph, Ontario, Canada

© Samuel T. Decker, September 2019
ABSTRACT

WELFARE BENEFIT FROM ENRICHMENT PROVISION SCHEMES IN AMERICAN MINK

*(NEOVISON VISON)*

Samuel T. Decker

University of Guelph, 2019

Advisor: Georgia Mason

Little is known about how animals obtain welfare benefit from environmental enrichment or how to best implement enrichment programs. This thesis aimed to fill these gaps by supplying mink with enrichment provision schemes that differed in their resource number, diversity, and novelty to evaluate the amount of resource use elicited from animals and the welfare benefit that animals subsequently received. Analysis revealed that mink altered resources use over five months, with active use decreasing and passive use increasing. Scheme effects on use were inconsistent, primarily due to unexpected item type effects, but did suggest that use can be increased above levels expressed under minimum enrichment recommendations. However, welfare indicator limitations prevented interpretation of the welfare benefit from schemes or resource use. These results support two considerations: 1) manipulable enrichments are not equivalent in the behavioural opportunity they provide and 2) passive use is a motivational outlet that warrants further study.
ACKNOWLEDGEMENTS

I am first thankful for the guidance and support of my advisor, Georgia Mason, throughout the design, redesign, overhaul, analysis, and writing of the work presented in this thesis. Hopefully my unique research experience truly is a sampling error, but if not, I am confident that I can rely on the acumen and insight you have helped me develop to overcome future challenges. I would also like to thank the members of the Mason Laboratory for their input and inspiration during our weekly meetings. I am especially grateful to Misha, Aileen, Andrea, and Michelle for helping to construct hundreds of rings and to Basma for your tireless work conducting seemingly endless boredom tests.

I also appreciate the comments and suggestions provided by my advisory committee, Lee Niel and Derek Haley, who helped refine the ideas presented here. Similarly, the research community at University of Guelph has expanded my understanding and appreciation for animal welfare far beyond my expectations. I am also thankful for the scholarships, OMAFRA HQP and Joint Mink Research Committee Research Grant, that allowed me to study at University of Guelph and produce this thesis to improve and promote animal welfare consideration in industry.

Finally, I would like to thank the farm staff who accommodated my research. Your care for the animals used in this research is greatly appreciated and our conversations helped me to feel welcome in your industry and community.
# TABLE OF CONTENTS

Abstract ......................................................................................................................................................... ii

Acknowledgements ........................................................................................................................................ iv

Table of Contents .......................................................................................................................................... v

List of Tables ............................................................................................................................................... ix

List of Figures ............................................................................................................................................... x

Chapter 1: Literature Review ...................................................................................................................... 1

1.1 Enrichment and animal welfare ........................................................................................................ 1

1.1.1 Highly motivated behaviour patterns ........................................................................................ 2

1.1.2 General motivation for control, exploration, and stimulation .................................................... 5

1.1.3 How is welfare impact of enrichment assessed? ........................................................................ 7

1.2 Resource use as an indicator of enrichment impact on welfare .................................................... 9

1.2.1 Does active resource use predict welfare impact of enrichment? ............................................. 10

1.3 Habituation to enrichment................................................................................................................ 12

1.3.1 Characteristics of habituation .................................................................................................... 13

1.3.2 When does reduced active resource use impact welfare? ............................................................ 18

1.3.3 How provision schemes could help ............................................................................................ 20

1.4 Enrichment and aggression .............................................................................................................. 22

1.5 Mink: a model carnivore for studying environmental enrichment ............................................. 23

1.5.1 Mink and mink farming .............................................................................................................. 23

1.5.2 Mink and environmental enrichment .......................................................................................... 25
1.5.3 Mink as a model carnivore for research ................................................................. 26

1.6 Conclusions and aims of this thesis .................................................................................. 28

Chapter 2: Provision Scheme Effects On Resource Use ......................................................... 29

2.1 Introduction ....................................................................................................................... 29

2.2 Methods .............................................................................................................................. 31

2.2.1 Data collection ............................................................................................................ 36

2.2.2 Statistical analysis ...................................................................................................... 39

2.3 Results ................................................................................................................................. 42

2.3.1 Relationships between forms of use .......................................................................... 43

2.3.2 Behavioural stability .................................................................................................. 46

2.3.3 Provision scheme effects on active resource use ....................................................... 47

2.3.4 Provision scheme effects on inactive resource use ................................................... 53

2.3.5 Provision scheme effects on resources in the nest box .......................................... 61

2.4 Discussion ......................................................................................................................... 68

2.4.1 Relationships between forms of use .......................................................................... 68

2.4.2 Treatment effects ....................................................................................................... 69

2.4.3 Provision scheme effects .......................................................................................... 71

2.4.4 Limitations ................................................................................................................ 73

2.4.5 Implications ................................................................................................................ 74

Chapter 3: Relationships Between Resource Use and Welfare ........................................... 75

3.1 Introduction ....................................................................................................................... 75
3.2 Methods ......................................................................................................................................... 79
  3.2.1 Statistical analyses.............................................................................................................. 83
3.3 Results ............................................................................................................................................ 85
  3.3.1 Behavioural stability........................................................................................................... 85
  3.3.2 Scheme and treatment effects on stereotypic behaviour..................................................... 86
  3.3.3 Scheme and treatment effects on fearfulness ..................................................................... 92
  3.3.4 Scheme and treatment effects on boredom ........................................................................ 96
  3.3.5 Scheme and treatment effects on bite marks ....................................................................... 98
  3.3.6 Resource use as a welfare covariate ................................................................................... 99
3.4 Discussion .................................................................................................................................... 102
  3.4.1 Treatment and provision scheme effects on welfare ........................................................ 102
  3.4.2 Resource use as a covariate in welfare models ................................................................. 104
  3.4.3 Limitations ...................................................................................................................... 104
  3.4.4 Implications ...................................................................................................................... 107
Chapter 4: Discussion and Future Research ............................................................................................ 108
  4.1 Introduction .................................................................................................................................. 108
  4.2 Summary of objectives and results .............................................................................................. 109
  4.3 Implications ............................................................................................................................. 110
  4.4 Modifications for replication .................................................................................................... 111
  4.5 Future research .......................................................................................................................... 113
  4.5.1 Behavioural opportunity and positive welfare................................................................. 113
4.5.2 Studying novelty and potential complications .......................................................... 115

4.5.3 Active use reduction from habituation, motivational change, or seasonal changes .... 117

4.6 Conclusion ...................................................................................................................... 118

References .......................................................................................................................... 119
LIST OF TABLES

Table 1.1. Characteristics of habituation .................................................................................................... 13
Table 2.1. Hypotheses tested in Chapter 2 .................................................................................................. 31
Table 2.2. Organization of provision schemes and treatment groups ......................................................... 35
Table 2.3. Ethogram of mink behaviours and cage features ....................................................................... 38
Table 2.4. Behavioural stability of active and inactive use ......................................................................... 47
Table 2.5. Evaluation of predictions for active use separated by mink sex and colour .............................. 53
Table 2.6. Evaluation of predictions for inactive use separated by mink sex and colour ......................... 61
Table 2.7. Evaluation of predictions for items in the nest box separated by mink colour .......................... 66
Table 3.1. Hypotheses and predictions for provision scheme and treatment effects on welfare ............ 78
Table 3.2. Hypotheses and predictions for relationships between resource use and welfare .............. 78
Table 3.3. Behavioural stability of stereotypic behaviour and fearful responses ...................................... 86
Table 3.4. Evaluation of predictions for stereotypic behaviour separated by mink sex and colour .......... 92
Table 3.5. Evaluation of predictions for fearfulness separated by mink sex and colour ......................... 96
Table 3.6. Summary of scheme and treatment effect changes when resource use is included in models 101
LIST OF FIGURES

Figure 2.1. Standard mink cage diagram .................................................................................................... 32
Figure 2.2. Shed layout diagram ................................................................................................................ 33
Figure 2.3. Example cage layout with three clusters .................................................................................. 33
Figure 2.4. Resources used in the study ...................................................................................................... 34
Figure 2.5. Sample size estimation for detecting effects with power 0.8 ................................................... 36
Figure 2.6. Experiment timeline for Chapter 2 ........................................................................................... 37
Figure 2.7. Mink whelping cage diagram ................................................................................................... 39
Figure 2.8. General use and its indicator variables across collection periods ............................................. 44
Figure 2.9. Change in relationship between inactive use and presence in the nest box .............................. 46
Figure 2.10. August active use by female mink split by colour type .......................................................... 48
Figure 2.11. August Active use by male mink ............................................................................................ 49
Figure 2.12. October active use by female mink split by colour type ........................................................ 50
Figure 2.13. October active use by male mink split by colour type ............................................................ 51
Figure 2.14. August inactive use by female mink ....................................................................................... 54
Figure 2.15. August inactive use by male mink split by colour type .......................................................... 55
Figure 2.16. Diversity and Number comparisons for October inactive use ................................................ 57
Figure 2.17. Novelty comparisons for October inactive use with pastel males shown separately ............. 58
Figure 2.18. Single housing inactive use by male mink separated by colour type to show differences in Novelty effects ............................................................................................................................................ 59
Figure 2.19. August presence of items in the nest box separated by colour type ........................................ 62
Figure 2.20. Diversity and Number comparisons for October presence of items in the nest box...........64

Figure 2.21. Novelty comparisons for October presence of items in the nest box separated by colour type .........................................................................................................................................................................................65

Figure 2.22. Comparison of items in the nest box between Diversity and Novelty treatments in October 67

Figure 2.23. Mink wearing rings.................................................................................................................................................................................70

Figure 3.1. Experiment timeline for Chapter 3.................................................................................................................................79

Figure 3.2. A catching glove being presented to pair-housed mink.........................................................................................................80

Figure 3.3. A pastel male interacting with an odorant filled tea strainer.................................................................................................82

Figure 3.4. The dorsal side of a female pelt with the leather side out..................................................................................................83

Figure 3.5. October stereotypic behaviour for female mink split by colour type .................................................................88

Figure 3.6. October stereotypic behaviour in male mink split by colour type.........................................................................................89

Figure 3.7. Single housing stereotypic behavior for male mink split by colour type .................................................................91

Figure 3.8. August fearful responses for male mink split by colour type.........................................................................................93

Figure 3.9. October fearful responses for male mink split by colour type.........................................................................................95

Figure 3.10. Time in contact with the stimulus during boredom tests.........................................................................................97

Figure 3.11. Time oriented toward the stimulus during boredom tests.........................................................................................98

Figure 3.12. Number of bite marks present on pelts ..........................................................................................................................99

Figure 3.13. Resource use covariate effects..........................................................................................................................100
Chapter 1: Literature Review

Designing a captive environment that can fulfill an animal’s motivations requires an understanding of what those motivations are and how the animal prefers to resolve them. Environmental enrichment programs have become a cornerstone of animal welfare management and in many cases successfully provide outlets for animal motivations in captivity. Amongst many animal care practitioners and within enrichment literature there is an assumption that the welfare benefit gained from environmental enrichment is linked to direct interaction with enrichments. Attempts to improve the benefit from enrichment have therefore sought to increase and prolong interaction by overcoming limiting factors for enrichment use, such as loss of interest and intraspecific competition. However, the underlying assumption warrants further evaluation to ensure that future work on improving the welfare benefit from environmental enrichment is not misguided. The aim of this thesis is thus to determine whether interaction with enrichments is linked to welfare benefit and to identify optimal methods, in terms of welfare benefit, for provisioning enrichments to captive animals.

This first chapter develops a foundation for the research presented in this thesis. Section 1.1 defines environmental enrichment and how it provides animals opportunities to perform motivated behaviour, followed by a discussion on how welfare impact of enrichment can be assessed. Section 1.2 introduces resource use as a potential welfare indicator and shows that research on active resource use has been inconclusive on the subject. Section 1.3 lists the ten characteristics of habituation, outlines when habituation of active use affects welfare, and proposes methods for mitigating habituation. Section 1.4 examines contradictory effects of environmental enrichment on aggression and possible causes. Section 1.5 describes the subject species for my thesis, American mink, and applies the concepts discussed in previous sections to this model carnivore. Finally, Section 1.6 outlines the hypotheses and predictions tested in later chapters.

1.1 Enrichment and animal welfare

Environmental enrichment, as utilized in animal welfare research and practical animal care, is the provision of resources into a captive environment to meet the needs and preferences of an animal, leading to improvements in physical and mental health (Newberry, 1995; Shepherdson, 1998). These resources fall into one or more of the following categories: physical (e.g. enclosure size or manipulable objects), nutritional (e.g. foraging tasks or increased diet diversity), sensory (e.g. visual or auditory stimuli), social
(e.g. group interaction), and occupational (e.g. computer games) (Bloomsmith et al., 1991). Throughout this thesis, “resources” will signify the objects or features being added to the environment to avoid assuming their efficacy as environmental enrichments.

The earliest development of environmental enrichment programs began in zoos during the early 20th century, and the practice has since expanded to farms and laboratories for nearly every species (Young, 2003). Resources for environmental enrichment are chosen with the intention of improving welfare, whether it be through improving physical health, providing opportunities for species-typical behaviours, or offering animals control over their environment. Most environmental enrichment studies and programs provide resources from multiple categories at the same time to maximize welfare improvement, but this prevents nuanced investigation of what enrichment characteristics drive such improvement (e.g. Latham and Mason, 2010; Swaisgood and Shepherdson, 2006).

When defining animal welfare, it is useful to consider three common conceptions held by those interested in animal care or production. Fraser (2009, 2008) distinguished these conceptions by emphasizing either basic health and functioning, natural living, or affective states. Health and functioning prioritise an animal’s fitness; if they can survive and reproduce, then welfare is not at risk. Natural living focuses on replicating the natural behavioural repertoire, environment, and social structure of an animal to ensure that welfare relative to a wild-type gold standard is maintained. The last conception emphasizes the animal’s affective state—the psychological state or subjective feeling comprising emotion (e.g. fear, joy, and hunger), long-lasting moods (e.g. anxiety and boredom), and clinical conditions (e.g. depression). When an animal experiences prolonged, unpleasant affective states they are considered to have poor welfare and when affective states tend toward joy, satiation, and contentment they are considered to have good welfare. This thesis identifies affective state as the defining feature of welfare, although interaction between fitness, naturalness, and affective state still allows these other features to influence welfare.

1.1.1 Highly motivated behaviour patterns

The interaction between naturalness and affective state is most often targeted when implementing environmental enrichment. The behaviours expressed by wild animals vary in their motivation, with some behaviours having strong motivations because they are essential for maintaining fitness, at least in the wild. Highly motivated behaviours can be split into two categories: behavioural needs and elicited responses (Jensen and Toates, 1993; Mason and Burn, 2018). Dawkins (1988) summarizes behavioural needs by borrowing economic terminology. She argued that behavioural needs have inelastic demand;
these behaviours have a minimum rate of performance, and inability to perform them due to an unsuitable environment leads to poor welfare. A key distinction from elicited responses is that behavioural needs retain their motivation even if the behaviour’s outcome has already been provided within the captive environment. Elicited responses fit within the larger group of elastic behaviours, where their expression changes depending on the presence of appropriate environmental cues. Animals that cannot adequately respond to environmental cues through appropriate behaviour will have poor welfare. Thus, for both behavioural needs and elicited responses welfare detriment stems from an inability to perform evolutionarily necessary behaviour.

Highly motivated behaviours can be identified by measuring the cost paid to access opportunities to perform specific behaviours. For example, demand can be investigated by measuring relative differences in the cost paid to access behavioural opportunities (Dawkins, 1990, 1988). Cost is usually assigned by requiring an operant task to be completed or a weighted door to be pushed open before the environmental stimulus necessary for the behaviour is rewarded. Assessing the maximum cost paid or how many times an animal is willing to pay a cost provides measures to compare the relative motivation for various behaviours. Access to food is usually included as a control in such experiments because feeding is assumed to be one of the most inelastic behaviours, one that must be performed regardless of cost, and therefore measures the absolute maximum cost that an animal will pay.

An issue when using this experimental design is that motivation to perform a behaviour can be altered by the presence of relevant environmental cues used as rewards (Warburton and Mason, 2003). For example, mink preference for access to toys decreased when they were unable to see them prior to paying the cost for access, meaning that the visual presence of toys altered motivational strength (Warburton and Mason, 2003). If reward cues are not accounted for, assessing the cost paid for access cannot differentiate between true inelastic demand seen in behavioural needs and the cue dependent demand seen in elicited responses. In this situation the welfare impact of not performing a behaviour that appears to have high demand cannot be inferred because behavioural needs and elicited responses cannot be differentiated. This does not detract from the idea that performing motivated behaviours, determined by cost paid, promotes positive welfare. Overcoming these caveats to motivational demand testing requires careful experimental and apparatus design, but other behavioural measures that are relatively easier to implement have been developed.

Highly motivated behaviours can also be inferred from behavioural responses to frustration. Frustration refers to the inability to achieve the goal of a behaviour, whether that be performance of the behaviour itself or some other functional consequence (Dollard et al., 1939). The frustration-aggression
hypothesis proposes that frustration increases intraspecific aggression, although aggression must not always follow frustration, and has been documented in many species (Buss, 1963; Dollard et al., 1939; Gallup, 1965; Scott, 1948). Furthermore, providing limited opportunities to perform highly motivated behaviours to group-housed animals may cause aggression due to competition over relevant resources (a topic returned to in Section 1.4).

For a second behavioural response to frustration, Hughes and Duncan (1988) designed a model that suggests stereotypic behaviour, consistently repeated and apparently functionless actions (Mason, 1991), develop when behavioural needs cannot be performed adequately. Their model separates the behavioural need into appetitive and consummatory components, e.g. looking for food and eating it, which include positive feedback to motivation. These feedback loops between the behaviour and motivation continue until the behaviour’s functional consequence has been achieved (e.g. satiation). Highly motivated behaviours likely have strong feedback loops to ensure the behaviours crucial for fitness reach their consequence (e.g. feeding or mating). In this scenario, stereotypic behaviour develops when the consequence of a behaviour cannot be achieved or when the consequence does not directly follow from the appetitive/consummatory behaviour, so these self-reinforcing actions continuously occur. The forms of stereotypic behaviour in such cases usually provide clues regarding the original functional consequence.

Two well-researched examples illustrate how environmental enrichment can provide behavioural opportunity and alleviate welfare concerns when highly motivated behaviours are constrained in captivity. First, non-nutritive sucking occurs in dairy calves that are housed separately from their mothers and fed from a bucket rather than suckling. Non-nutritive sucking is often directed toward pen features or other calves, making it a health concern, and is not usually performed by calves able to suckle from a cow. Efforts to understand and mitigate non-nutritive sucking show that feeding calves from an artificial teat reduces non-nutritive sucking and longer durations of nutritive teat-sucking enhance this effect (Haley et al., 1998). Furthermore, teat-fed calves show lower heart rates during feeding than bucket-fed calves and teat-feeding results in more calm and restful post feeding behaviour (Veissier et al., 2002). Because artificial teat and bucket fed calves have the same nutrient intake (i.e. the outcome of feeding was the same), the benefits of teat sucking suggest this is a behavioural need and that the addition of an artificial teat as environmental enrichment can fulfill this motivation.

Second, stereotypic digging near cage edges occurs in gerbils housed in standard laboratory cages. An initial hypothesis is that digging is a behavioural need and that providing better opportunities to dig would eliminate the stereotypic behaviour, but providing deep sand for digging does not reduce
stereotypic digging (Weidenmayer, 1997). Instead, stereotypic digging is substantially reduced only when a naturalistic burrow is provided as environmental enrichment, consisting of an opaque tunnel and chamber (Waiblinger and König, 2004; Weidenmayer, 1997). Reduced stereotypic digging when the presumed end goal is provided suggests that digging is an elicited response to create a burrow when one is absent rather than a behavioural need.

Both examples show how environmental enrichment, via artificial teats or burrows, can provide animals opportunities to express highly motivated natural behaviours. But creating opportunities for specific natural behaviours is not the only way in which enrichment can enhance welfare.

### 1.1.2 General motivation for control, exploration, and stimulation

A further function of environmental enrichment is to increase an animal’s control over its environment through choice (Carlstead and Shepherdson, 2000; Sambrook and Buchanan-Smith, 1997; Young, 2003). Numerous studies show that gaining control over a negative stimulus reduces stress, losing control is more stressful than not having control at all, and that making choices is inherently rewarding and desirable (reviewed in Leotti et al., 2010). Early work on environmental enrichment in zoos utilized control to improve welfare through operant tasks where animals could choose when and how they interact with their environment (Markowitz, 1982). More recent zoo studies found that offering free choice between exhibit space and off-exhibit holding areas improved welfare, measured via stereotypic behaviour and urinary cortisol (see Mormède et al., 2007 for discussion on cortisol as a welfare measure), in four giant pandas and two polar bears (Owen et al., 2005; Ross, 2006). Providing a single outlet for control may help animals fulfill specific highly motivated behaviours, e.g. the choice to retreat to a burrow or suckle at an artificial teat, but many species also have a recurring need for general stimulation.

The need to explore and experience novel stimulation may be most prominent in generalist and innovative species, meaning those adapted to a variety of environmental circumstances (e.g. raccoons) or those that develop new or modified learned behaviour (e.g. British titmouse; Fisher and Hinde, 1949). These animals typically exhibit cognitive and behavioural traits, such as boldness, neophilia, and increased intelligence, that can predispose them to poor welfare in the monotonous, barren environments often found in captivity (Burn, 2017; Mason et al., 2013; S.M., 2003). Hughes (1997) summarizes the work of Daniel Berlyne (1963, 1960) on the motivation for exploration, which is categorized into extrinsic and intrinsic exploration. Extrinsic exploration is directed toward meeting an external goal. This type of exploration is associated with fulfilling specific highly motivated behaviours, such as retreating
into a burrow to escape a threat. Intrinsic exploration promotes investigation of stimuli for the sake of interest in the stimuli. An early examination of intrinsic exploration showed that rhesus monkeys were motivated to interact with and learn a puzzle without reward and retained interest in the puzzle for thirteen successive days (Harlow, 1950).

Berlyne (1963, 1960) also identified novelty and complexity as key variables governing the initiation and duration of exploration. Novelty is represented by any change in a stimulus or situation between exposures and is the aspect that initiates exploration and diminishes as the behaviour occurs. Complexity is the amount of diversity within a stimulus. It is specified by the number of distinct characteristics and the dissimilarity between those characteristics and impacts the duration of exploration, i.e. the time it takes to become completely familiar with a stimulus. For example, the puzzle given to the rhesus monkeys had five distinct features that were interwoven so that they must be completed in a specific order. This complexity prevented mastery of the puzzle, and subsequent loss of novelty, throughout the study’s duration (Harlow, 1950). Novelty and complexity apply to both extrinsic and intrinsic exploration. Goal based motivation in extrinsic exploration has resulted in most research on these variables being done in intrinsic exploration, where stimulus novelty and complexity presumably have greater impact on expression of the behaviour (Hughes, 1997). Efforts to clarify the underlying motivation or drive behind intrinsic exploration have been unproductive, yet the extensive catalogue of research on intrinsic exploration suggests it may be a behavioural need in some species (Hughes, 1997).

The negative welfare state commonly associated with an inability to satisfy motivation for new experiences through intrinsic exploration is boredom (Burn, 2017). Boredom has been well studied in humans, often resulting in increased sensation seeking or risk taking behaviour (reviewed in Eastwood et al., 2012). More recent work has investigated the potential for boredom-like states (referred to as boredom in the remainder of this thesis) in captive animals measured via hyper-responsiveness to negative, neutral, and rewarding stimuli, which is returned to in Section 1.5.3 (Meagher et al., 2017; Meagher and Mason, 2012). Describing boredom in animals requires argument by analogy, and referring to animal boredom as boredom-like assumes that animals do not share the exact experiences of humans. Designing environmental enrichments that alleviate boredom is difficult because the novelty of a resource is the characteristic that must be maintained. Once a resource has been fully explored, it may no longer satisfy the motivation for new experiences, which is discussed further in Section 1.3 on habituation. Enrichments that provide opportunities for specific highly motivated behaviours do not face this challenge because the functional consequence of the behaviour also drives its performance (e.g. a burrow remains an outlet for retreat even after multiple uses).
Stereotypic behaviour, aggression, and hyper-responsiveness have been presented as measures of affective state, specifically frustration and boredom, but how are researchers able to infer that changes in these behaviours accurately reflect changes in welfare if it is defined by subjective experience? If the goal of environmental enrichment is to improve welfare, accurate measures of welfare are required. Luckily, animal welfare research matured in the latter half of the 20th century to develop methods for validating measures of subjective experience in animals.

1.1.3 How is welfare impact of enrichment assessed?

Before welfare indicators can be discussed, a more thorough description of affective state is required. Russell (2003) proposed affective state as a replacement for the folk concept of emotion with intentionality, i.e. having a feeling about or toward something. Affective states go beyond categorized emotions to include diffuse moods and can be described according to two axes. First, affective states have valence, or a scale from pleasure to displeasure (i.e. positive to negative). Second, affective states have intensity that ranges from activated to deactivated. For example, a positive, deactivated affective state might reflect calm or contentment while a negative, activated affective state might reflect distress or anxiety. These states are often background processes and only consciously attended to when the state changes or when a state is particularly vivid. Welfare is determined by the valence of affective states and can be measured by a momentary value, the general mood over longer periods of time, or the cumulative experiences over an animal’s life (Yeates, 2011). Although affective state is primarily a subjective experience, it also manifests through physiological, behavioural, and cognitive changes, which function as welfare indicators when properly validated.

When validating potential physiological, behavioural, or cognitive indicators of affective state, identifying the subject’s location along the valence axis is crucial to prevent circular reasoning. Altering an animal’s perceived fitness or comparing homologous responses between humans and other species are two methods for determining an animal’s affective state (Mason and Mendl, 1993; Sandøe and Simonsen, 1992). The first method assumes that affective states have an evolutionary connection to fitness and that altering an animal’s perceived fitness causes relative changes in affective state. Attempts to alter fitness must be biologically relevant to elicit changes in affective state (i.e. not evolutionarily new; e.g. a dog does not know that vaccination improves fitness). Once relative change in affective state has been achieved, concurrent changes in physiological, behavioural, and cognitive responses are potential welfare indicators. A related approach is to assume that negative affect promotes avoidance, while positive affect promotes seeking. Exposing animals to stimuli or situations they typically avoid (e.g. vaccinations, even
if these do improve fitness) should induce negative affect, while exposing them to stimuli or situations they find attractive should induce positive affect (Mason and Mendl, 1993). Corresponding responses can then be observed to identify indicators of the affective state change.

The second method relies on an argument by analogy that the subjective experiences of humans and other species are similar, albeit not exactly the same (Dawkins, 1990). Self-report in humans allows for more precise identification of affective state and then concurrent responses to affective state change can be recorded. However, homology between humans and other species in these responses is required to prevent anthropomorphising, thus further validation of human responses in the species of interest is often necessary. Pharmacological validation of responses serves as an intermediate method between altering fitness, providing preferred/avoided resources, and homology with humans (e.g. Keating et al., 2012). Here, varying doses of a drug, such as an analgesic or anxiogenic, are assumed to alter affective state. This assumption is best supported if the mechanism of action for the drug is similar between humans and the species of interest because humans can self-report the associated affective state change. Again, with relative affective state established, responses can be recorded and those that correspond with change in affective state are potential indicators. These methods have developed welfare indicators across many species that can be utilized in evaluating the impact of environmental enrichment on welfare.

Stereotypic behaviour was one of the earliest indicators used in the evaluation of environmental enrichment effects on welfare, primarily because it is exaggerated and detracts from the natural behavioural repertoire of wild animals that zoos aim to present to the public (Hediger, 1950; Meyer-Holzapfel, 1968; Young, 2003). Two meta-analytical reviews found agreement amongst zoo literature that stereotypic behaviour is largely reduced, although not eliminated, by environmental enrichment (Shyne, 2006; Swaisgood and Shepherdson, 2006). However, the limitations of welfare indicators used must be understood in terms of specificity and sensitivity for proper interpretation of effects. Stereotypic behaviour, for example, has good specificity, meaning that expressing the behaviour reliably indicates compromised welfare at some point during the animal’s life. Conversely, stereotypic behaviour is not sensitive because not all animals in negative affective states show stereotypic behaviour, instead potentially becoming inactive and unresponsive (Fureix et al., 2016; Mason and Latham, 2004).

Observing a combination of welfare indicators that account for each other’s limitations and inferring welfare from patterns amongst all indicators is one solution (Mason et al., 2007; Sandøe and Simonsen, 1992). Other candidate measures (focusing on those used in this thesis) include aggression to cage-mates, hyper-responsiveness to diverse stimuli, and fearful responses to a stimulus. As reviewed above, aggression can indicate frustration (see also Section 1.4) and boredom relates to under-stimulation
and subsequent exaggerated responses to new experiences (see also Section 1.5.3). Behavioural signs of fear and anxiety can also indicate poor welfare, as well as being negative affective states themselves, and are often increased when animals are subjected to pain or barren environments (Olsson and Dahlborn, 2002; Parent et al., 2012). Section 1.5.3 provides more detail on how these behavioural measures have been used in welfare and enrichment research with mink. Finally, another behavioural measure frequently utilized when evaluating environmental enrichment is the use of supplied resources by the animal, although inconsistent relationships between resource use and other welfare measures question whether this is a reliable indicator.

### 1.2 Resource use as an indicator of enrichment impact on welfare

Because animals should be highly motivated to interact with successful enrichments, as reviewed in the previous sections, it is tempting to infer a resource’s efficacy for improving welfare from the degree to which it is used. Beyond welfare, resource use may be emphasized by animal caretakers because it makes a zoo exhibit more attractive and interesting to the public (Young, 2003). Resource use can be divided into two main categories: active and passive. Active use involves dynamic contact with resources and includes both extrinsic and intrinsic exploratory behaviours, as described in Section 1.1.2 (Lutz and Novak, 2005). Examples include interaction with balls or toys by mink (Meagher et al., 2014) and primates (Brent and Belik, 1997; Paquette and Prescott, 1988) and other active behaviours previously described (e.g. suckling, retreating, and puzzle solving). Emphasis on active use in environmental enrichment literature is likely due to the focus on exhibiting natural behaviour in zoos (e.g. AZA Tiger Species Survival Plan®, 2016) and a more general “use it or lose it” mentality when designing and funding enrichment programs for farms and laboratories (i.e. it is not worth spending money to provide resources that go apparently unused).

Passive use is a consolidation of multiple types of use that lack the exploratory component of active use. Potential examples include perceived opportunity and inactive use. Perceived opportunity refers to the benefits gained from having a choice to interact with a resource even if the action is not performed (this is examined further in the next section). Inactive use involves any static contact with a resource while the animal is inactive. Inactivity is broadly defined as being relatively motionless and includes being still but awake (Fureix and Meagher, 2015). This category is usually studied when structural resources are designed to alter sleeping or resting behaviours and is often combined with active
use into a general use score when studying multiple types of resources simultaneously (e.g. Dallaire et al., 2012; Díez-León et al., 2016). Mink-specific forms of passive use are also discussed in Section 1.5.2.

The rewarding aspect of inactive use can have multiple components, such as tactile comfort, thermoregulation, or other sensory stimulation. Many species experience a relaxing effect when static pressure is applied to the body (Grandin et al., 1989; Mullen et al., 2008; Valros et al., 2016) and inactive contact with resources could provide similar relaxing benefits. Studies with primates and reptiles have also shown that animals will alter the location of inactive behaviours to make use of thermoregulatory resources (Bashaw et al., 2016; Duncan and Pillay, 2013; Wark et al., 2014). Lastly, when given access to an elevated tunnel, some mink shift inactive behaviour to occur primarily in the tunnel, presumably to provide greater security from humans or to reach a vantage point to monitor their surroundings (Dallaire et al., 2012). The opportunity to choose where inactivity occurs and alter perception during inactive behaviour has potential to improve welfare (see Section 1.1.2 on control), but research into this possibility has been relatively sparse compared to potential benefits of active use.

1.2.1 Does active resource use predict welfare impact of enrichment?

There is conflicting evidence as to whether the welfare benefits of enrichment require or reflect high levels of general resource use, and most of this research utilized farmed mink as a model species. On the one hand, overall use of elevated platforms by mink dams increases as their kits age (Dawson et al., 2013), and is predicted by litter size (Buob et al., 2013). Both papers suggest that platforms were used by nursing female mink to escape their young prior to weaning, which indicates that females provided with platforms exhibited greater control over their environment and therefore may experience better welfare. Díez-León et al. (2016) also found that time spent by mink in an enriched compartment predicted decreased scrabbling stereotypic behaviour. However, general resource use does not negatively correlate with many welfare indicators in mink, such as locomotor stereotypic behaviour (Dallaire et al., 2012; Díez-León et al., 2016), fecal corticoid metabolites (Díez-León et al., 2016), or anatomical stress indicators (Díez-León et al., 2016).

Assessing active use by itself also shows inconsistent effects on welfare indicators. In support of active use reflecting enrichment efficacy, Lumeij and Hommers (2008) found that time spent actively using a foraging resource positively correlated with feather score in parrots, a proxy measure of feather plucking. Therefore, individuals that used the foraging resources had the lowest levels of this specific abnormal behaviour. Further association between active use and enriched treatments that improved
welfare has also been shown in rats (Abou-Ismail, 2011a; Abou-Ismail and Mendl, 2016). Welfare in this case was inferred from increased sleep (validation: Abou-Ismail et al., 2007), a healthier immune system (validation: Corbin et al., 2008; Mitchell et al., 2006), and decreased aggression. However, one study (Abou-Ismail et al., 2010) found that effects of reduced aggression and increased sleep in environmentally enriched rats remained even after statistically controlling for active use. Therefore, these effects were not due solely to time spent actively using resources. Likewise, greater active resource use by mink did not predict reduced stereotypic scrabbling (Díez-León et al., 2016), reduced locomotor stereotypic behaviour, nor improved reproductive success measured via lack of pregnancy, early litter size, and litter size at weaning (Meagher et al., 2014).

The lack of a clear relationship between active use and improved welfare has three potential explanations. One is that the amount of time spent actively using a resource does not accurately represent that resource’s value. Kirkden and Pajor (2006) argue that motivations do not vary consistently with satiation rates or the required amounts of associated resources. For this reason, duration and prevalence of consummatory behaviours, like active use, provide a less accurate assessment of motivation than appetitive behaviours (e.g. pushing heavy weights, as summarized in Section 1.1.1). The second possible explanation is that benefits of active use may continue beyond the behaviour’s performance with a “mere presence” effect of resources improving welfare, such that the perceived opportunity to access the resource is beneficial beyond the time spent actively using it. This is akin to Gibson’s (1977) “Theory of Affordances”, which seeks to explain the attractiveness of certain natural landscapes to humans in terms of the behavioural opportunities available in an environment; passively perceiving these opportunities may boost welfare even if they are not actually exploited (cited in Ross and Mason, 2017). Relatedly, active use may be important for information gain when resources are novel, but once the animal understands the properties of resources in its environment it does not need to continue active exploration to maintain this knowledge. A third option is that the potentially rewarding components of inactive use discussed in the previous section, such as tactile comfort and thermoregulation, contribute to welfare benefit.

Despite the inconsistent results described here, and potential alternative roles of inactive use or passively perceived opportunity, active use remains an intuitive measure of enrichment effectiveness that is widely used by farmers and zookeepers. This has resulted in substantial efforts to prolong active use and understand “habituation”: why many resources receive little interaction shortly after provision.
1.3 Habituation to enrichment

The general concept of habituation is a heterogenous group of response decrement phenomena, and the precise meaning of the term “habituation” is discipline dependent. A series of publications in Frontiers in Integrative Neuroscience on mechanisms of habituation are especially useful when disentangling this topic (e.g. Lloyd et al., 2014; Steiner and Barry, 2014; Typlt et al., 2013). With that said, habituation can provide a useful framework for understanding why active resource use may decrease and in what situations this may be a welfare issue. Rankin et al. (2009) gathered habituation researchers to revisit and clarify the definition of habituation and catalogued ten characteristics that are reproduced in Table 1. The following discussion explains each characteristic with application to environmental enrichment and definitions are repeated before corresponding paragraphs. Whenever possible, references from environmental enrichment literature are provided as examples of characteristics. Some characteristics have not been studied specifically with enrichment, but evidence from other disciplines still influences how animal caretakers mitigate habituation to enrichment.
Table 1.1. Characteristics of habituation

<table>
<thead>
<tr>
<th>#</th>
<th>Habituation characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Repeated presentation of a stimulus results in a progressive decrease in some parameter of a response to an asymptotic level.</td>
</tr>
<tr>
<td>2</td>
<td>If the stimulus is withheld after response decrement, the response recovers at least partially over the observation time (“spontaneous recovery”).</td>
</tr>
<tr>
<td>3</td>
<td>After multiple series of stimulus repetitions and spontaneous recoveries, the response decrement becomes successively more rapid and/or more pronounced (“potentiation of habituation”).</td>
</tr>
<tr>
<td>4</td>
<td>More frequent stimulation results in more rapid and/or more pronounced response decrement and more rapid spontaneous recovery.</td>
</tr>
<tr>
<td>5</td>
<td>Within a stimulus modality, the less intense the stimulus, the more rapid and/or more pronounced the behavioural response decrement. Very intense stimuli may yield no significant observable response decrement.</td>
</tr>
<tr>
<td>6</td>
<td>The effects of repeated presentation of a stimulus may continue to accumulate even after the response has reached an asymptotic level. This effect of stimulation beyond asymptotic levels can alter subsequent behaviour, for example, by delaying the onset of spontaneous recovery.</td>
</tr>
<tr>
<td>7</td>
<td>Within the same stimulus modality, the response decrement shows some stimulus specificity.</td>
</tr>
<tr>
<td>8</td>
<td>Presentation of a different stimulus results in an increase of the decremented response to the original stimulus (“dishabituation”).</td>
</tr>
<tr>
<td>9</td>
<td>Upon repeated application of the dishabituating stimulus, the amount of dishabituation produced decreases.</td>
</tr>
<tr>
<td>10</td>
<td>Some stimulus repetition protocols may result in properties of the response decrement that last hours, days, or weeks. This persistence of habituation is termed long-term habituation.</td>
</tr>
</tbody>
</table>

1Definitions are replicated abbreviations from Lloyd et al. (2014) and more thorough descriptions can be found in the original review by Rankin et al. (2009).

1.3.1 Characteristics of habituation

Characteristic #1 is the general definition of habituation. A key aspect of habituation is that response decrement does not involve sensory adaptation, sensory fatigue, nor motor fatigue, which is supported by characteristics #7-8 and mentioned here to clarify this point early. Instead, decreased response is caused by a change in the perception of the stimulus. The two main responses of interest for habituation studies are reflexes (e.g. startle response) and motivated behaviours (e.g. operant tasks). Most habituation research has focused on reflex responses to study the neurophysiological underpinning of habituation (Typlt et al., 2013). However, active use is categorized as a motivated behaviour because it
involves controlled exploration. Habituation of reflexive and motivated response are likely governed by different mechanisms (Typlt et al., 2013), meaning a large portion of habituation research does not apply to environmental enrichment.

Clarification of the timescale assumed in Table 1, and most habituation research, is necessary because it uses language associated with classical habituation studies and requires translation for proper application to environmental enrichment. “Repeated presentations” and “stimulus repetitions” in classical studies with reflexes are typically repeated every few seconds or minutes. Habituation, as described by #1, refers to a response asymptote within a collection of such repetitions, called a session. Comparatively, presentation of environmental enrichment is essentially continuous because resources are usually supplied for hours or days at a time, with each continuous block of enrichment provision being one session. Thus, depending on the enrichment of interest (see #5), the asymptotic response level referred to in characteristics #1 would occur within the first few minutes or hours of enrichment provision and is called short-term or within-session habituation. Habituation of open-field locomotor activity in mice described by Typlt et al. (2013) provides a nice example of habituation of a motivated behaviour during continuous stimulus presentation. The definition of a session for enrichment studies is complicated by daily fluctuations in active use of resources that are provided continuously for multiple days. In this case, the animal may experience multiple sessions of short-term habituation dictated by changes in attention within the longer continuous session where long-term or between-session habituation may occur. Characteristics #2-9 will first be described in terms of short-term habituation, and long-term habituation will be returned to in #10.

2. **If the stimulus is withheld after response decrement, the response recovers at least partially over the observation time (“spontaneous recovery”).**

3. **After multiple series of stimulus repetitions and spontaneous recoveries, the response decrement becomes successively more rapid and/or more pronounced (“potentiation of habituation”)**

Characteristic #2 should occur if an enrichment is provided on day 1 so that the asymptotic active use level is reached, the enrichment is then removed for the duration of day 2, and then the enrichment is returned on day 3. In this scenario, spontaneous recovery has occurred if there is more active use at the start of day 3 than at the end of day 1. Markowitz and LaForse (1987) and Renner et al. (2000) provide anecdotal reports of spontaneous recovery, the former for pigskin sacks and the latter for puzzle toys inferred from graphs because the original analysis compared reintroduction of a resource (i.e. day 3) to the initial active use (i.e. start of day 1), rather than the level of active use after habituation had occurred (i.e. end of day 1). If the pattern of habituation and recovery is continued, the return to asymptotic active
use will occur more rapidly with each successive recovery, as described in #3. This aspect of habituation has not been directly studied in environmental enrichment (Tarou and Bashaw, 2007), but laboratory research has shown it does occur for motivated responses (Lloyd et al., 2014).

4. **More frequent stimulation results in more rapid and/or more pronounced response decrement and more rapid spontaneous recovery.**

Supplying environmental enrichments continuously risks faster habituation due to characteristic #4. Finding ways to disperse the rewarding consequence of resource use over a longer period may help slow habituation. For example, Lloyd et al. (2012) showed that altering the rate of reward for a motivated behaviour can slow habituation. Rats were trained to snout-poke (i.e. motivated behaviour) to turn on a light (i.e. reward) and then habituation of the behaviour was compared between a fixed rate (FR) 1 schedule, where one snout-poke produced light-onset; 1-minute variable interval (VI 1), where the first snout-poke after approximately one minute produced light; and 6-minute variable interval (VI 6), where the first snout poke after approximately six minutes produced light. The VI 6 schedule rats showed more snout-poking, less within-session habituation of snout poking, and greater recovery of snout-poking between sessions compared to FR 1 and VI 1. For application to extrinsic exploration of enrichment, slowing the rate of reward from a puzzle feeder by making it more difficult to extract food should provide similar mitigating affects on habituation (Tarou and Bashaw, 2007), but making it too difficult can also cause loss of interest (Novak et al., 1998). Application to intrinsic exploration could involve withholding resources in regular intervals, but this may have unintended welfare consequences that are discussed further in characteristic #6.

5. **Within a stimulus modality, the less intense the stimulus, the more rapid and/or more pronounced the behavioural response decrement. Very intense stimuli may yield no significant observable response decrement.**

The original conception of characteristic #5 related to differences in reflex habituation to low and high intensity stimuli (e.g. quiet sounds versus loud sounds) and suggested that highly motivated behaviours in response to biologically relevant stimuli would not habituate (Thorpe, 1966 cited in McSweeney and Murphy, 2009). However, more recent work has shown that the characteristics of habituation also apply to biologically relevant and highly rewarding stimuli, including food (Epstein et al., 2009; Lloyd et al., 2014; McSweeney and Murphy, 2009). Characteristic #5 still applies to motivated behaviours and rewarding stimuli because behaviours that are less motivated/rewarding are likely to habituate faster than those that are highly motivated/rewarding (Lloyd et al., 2014). At the farthest end of this continuum, active use from behavioural needs and highly motivated elicited responses is still susceptible to habituation, albeit potentially at a rate so slow that resources will diminish, or sensory
adaptation/fatigue will occur before habituation. For active use with lesser motivations, the intensity of a resource is determined by the degree of novelty and complexity (e.g. Biondi et al., 2015), as described in Section 1.1.2.

6. **The effects of repeated presentation of a stimulus may continue to accumulate even after the response has reached an asymptotic level. This effect of stimulation beyond asymptotic levels can alter subsequent behaviour, for example, by delaying the onset of spontaneous recovery.**

Providing environmental enrichments continuously for long periods also risks exacerbating habituation due to accumulating effects after the asymptotic rate of active use has been reached (#6). Some environmental enrichment guidelines therefore suggest removing resources before interest in them wanes, primarily concerning long-term habituation (e.g. AZA Tiger Species Survival Plan®, 2016). This may inadvertently result in negative affective states by causing frustration from removal of valued resources (Latham and Mason, 2010). Instead, observing how use of a resource changes over time and at what rate active use asymptotes may reveal whether the resource is still valued or if habituation is a welfare risk (a topic returned to in Section 1.3.2). If a welfare risk is determined, carefully monitoring when use asymptotes and then removing the resource for the minimum amount of time for spontaneous recovery should prevent both accumulated habituation and frustration.

7. **Within the same stimulus modality, the response decrement shows some stimulus specificity.**

As previously mentioned, characteristics #7 and #8 can differentiate habituation from other competing explanations, such as satiation and motor fatigue (Epstein et al., 2009; McSweeney and Murphy, 2009). Epstein et al. (2003) showed stimulus specificity (#7) of habituation to food rewards by training children to complete a computer game to receive a portion of cheeseburger. Once responding for the cheeseburger habituated, the reward was changed to apple pie and responding subsequently increased. This shows that habituation is specific to the reward stimulus (i.e. cheeseburger), that changing the reward characteristics within the same stimulus modality (i.e. food) results in increased responding, and that the habituation could not be explained by satiation.

8. **Presentation of a different stimulus results in an increase of the decremented response to the original stimulus (“dishabituation”).**

9. **Upon repeated application of the dishabituating stimulus, the amount of dishabituation produced decreases.**

Testing for dishabituation (#8) is like testing spontaneous recovery: a reward is given repeatedly until habituation occurs; then, a novel stimulus/reward substantially different from the original is
presented; subsequently, responding for the original reward should increase. The difference between the two tests is that stimulus specificity measures response to a new reward/stimulus while dishabituation measures recovery of response to the original reward/stimulus. For example, Aoyama and McSweeney (2001) trained rats to press a bar to receive a food reward and allowed this response to habituate. Then, one of four dishabituating events occurred: (1) the bar was removed so rats could no longer respond for reward, (2) more bar presses were required to receive a larger reward, (3) more bar presses were required to receive the same reward, or (4) a variable reinforcement schedule was implemented. Once the original reward/schedule was returned, bar press responses increased compared to rats that did not experience a dishabituating event. Increased responding after (2) and (4) suggests that satiation could not be solely responsible for prior habituation because equivalent or larger rewards did not reduce subsequent responding. Similarly, increased responding after (3) shows that prior habituation was not caused by motor fatigue because increased effort for rewards did not reduce subsequent responding. Had these dishabituating events been repeated, they would likely have reduced effectiveness over time consistent with habituation characteristic #9. Stimulus specificity and dishabitation are occasionally confused when attempting to study these characteristics for environmental enrichment (e.g. Pullen et al., 2012), and extensions of them are often suggested as methods to limit habituation of active use (Tarou and Bashaw, 2007).

10. Some stimulus repetition protocols may result in properties of the response decrement that last hours, days, or weeks. This persistence of habituation is termed long-term habituation.

Lastly, characteristic #10 provides a definition for long-term habituation that is more applicable to environmental enrichment research than short-term habituation. Long-term habituation occurs via a combination of characteristics, such as potentiation of habituation and long periods of asymptotic response. It is governed by many of the same characteristics as short-term habituation but likely works through separate mechanisms (Typlt et al., 2013). Most research on long-term habituation focuses on applying these characteristics to mitigate or promote long-term habituation. For example, altering the frequency that a resource is provided (#4) can exploit spontaneous recovery (#2) without causing potentiation of habituation (#3) to maintain active use of a resource over a period of weeks (Kuczaj et al., 2002). Similarly, spreading out repeated presentations of the same food over weeks rather than days can slow long-term habituation and increase consumption in humans (Epstein et al., 2011). Combining reward frequency (#4), stimulus specificity (#7), and dishabituation (#8) also develops a variety paradigm, where groups with multiple food options habituate slower and increase energy intake compared to groups with only one food option over days, weeks, and months (Bouton et al., 2013; Epstein et al., 2015, 2013).
Applying the characteristics of habituation to long-term habituation, as in these examples, is necessary for environmental enrichment programs that seek to maintain sustained interest in resources for weeks, months, or years.

Extensive research has been done on each of these characteristics, but only properties and examples relevant to environmental enrichment and motivated behaviour were covered here. Many of these characteristics relate to aspects of stimulus exploration in Section 1.1.2, such as novelty and complexity of a stimulus altering how often it is explored and how quickly it is habituated to. Habituation can often explain why active use decreases, but this is not always the case. Furthermore, decreases in active use, whether caused by habituation or another mechanism, may not always represent a welfare risk.

1.3.2 When does reduced active resource use impact welfare?

Alternative explanations for declining active use include satiation, motor fatigue, absence of the reinforcer, and developmental changes in behaviour. Extrinsic exploration of resources, such as foraging opportunities, may lead to satiation or motor fatigue prior to habituation, and intrinsic exploration may also lead to motor fatigue. Satiation is unlikely for intrinsic exploration because it does not directly impact relevant physiological processes, such as stomach distension. In situations of satiation or motor fatigue, reduced active use may not represent welfare risk or loss of resource value. Animal caretakers may cease presentation of valued resources if such declines in active use are attributed to habituation because caretakers perceive such resources as ineffective. As previously mentioned, stimulus specificity or dishabituation paradigms are required to test whether habituation is truly responsible for decreased active use. However, these paradigms are rarely implemented in applied studies on environmental enrichment (see Table 1 in Tarou and Bashaw, 2007), which makes many reports of habituation speculative at best.

For resources that elicit extrinsic exploration, reinforcing aspects of the resource (e.g. food or destructibility) may deplete before active use habituates. In this case, a better explanation for decreased responding is extinction: a decrease in responding when the behaviour is no longer reinforced (Bouton and Todd, 2014). Any resource that has a finite amount of reinforcing capability risks extinguishing active use after the reinforcer has been consumed or once the amount of reinforcement for each bout of use has diminished to a certain extent. Still, some studies show gradual decreases in responding over a series of days even when the reinforcing aspect is resupplied. For example, Carlstead et al. (1991) and Anderson et al. (2010) provided similar honey-filled logs to captive bears and measured time spent
interacting with them. Both studies reported decreased time spent using the logs following repeated exposure and suggested habituation as the culprit; however, this change is better explained by extinction. The time it took to remove the supplied honey (i.e. reinforcer) from each log decreased as experience with the logs increased (Carlstead et al., 1991), meaning that the amount of time for extinction to occur (i.e. responding without reinforcement) decreased with experience.

Dishabituation has also been used to explain recovery of responding to empty logs when new honey filled logs were added to the exhibit (Carlstead et al., 1991). Again, this is better explained by disinhibition of extinction; a process similar to dishabituation where responding is restored following introduction of a novel stimulus (Mcsweeney and Swindell, 2002). Smith and Litchfield (2010) and Novak et al. (1998) provide examples where loss of interest in foraging resources was properly attributed to extinction following depletion of food reinforcers. These examples emphasize the need to understand which aspects of a resource are attended to and potentially reinforcing because only those aspects are at risk for habituation. Like confounds with satiation and motor fatigue, misattributing habituation for extinction may overlook resources that can still provide value to animals if their reinforcing components are maintained.

As animals mature, their behaviour and motivations are likely to change and result in decreased expression of some behaviours over weeks and months. For extrinsic exploration of resources, like artificial teat sucking in dairy calves, the motivations for performing the behaviour may decrease as animals progress through different life stages. Similarly, intrinsic exploration has shown age effects in many species. For example, juvenile chimpanzees and giant pandas sustain interest in resources that elicit intrinsic exploration, such as mirrors and balls, while adults seem to prefer extrinsic exploration in foraging resources (Lambath and Bloomsmith, 1992; Swaisgood et al., 2001). Juvenile mink also showed little stability in object play, a form of intrinsic exploration, in early development when controlling for overall activity (Dallaire and Mason, 2016). Once again, reduced performance of behaviours with decreasing motivation due to age cannot be explained by habituation and this situation likely does not pose a welfare risk.

If these competing explanations for decreased active use can be eliminated and true habituation assumed, welfare may still be unaffected. Short-term habituation is not a welfare concern in most cases because behavioural motivation usually does not demand constant resource use, and examples of constant performance of behaviour are often considered abnormal (Tarou and Bashaw, 2007). Some environmental enrichment programs may seek to increase daily performance of behaviours to levels exhibited in wild conspecifics by limiting short-term habituation, but generating constant performance is not the goal. As
mentioned in the previous section, welfare research is more interested in limiting long-term habituation to resources. However, long-term habituation of active use may not represent a welfare risk if the animal still values the resource and is instead using it in other ways, such as the forms of passive use discussed in Sections 1.2 and 1.5.2.

Long-term habituation becomes a welfare risk when an animal no longer values a resource as an outlet for a behaviour (i.e. loss of reinforcer effectiveness) and the behavioural motivation remains unsatisfied. The type of welfare risk depends on whether the behaviour involved extrinsic or intrinsic exploration. Stereotypic behaviour, aggression, or fear may develop if extrinsic exploration habituates and resources are no longer seen as opportunities to perform highly motivated behaviours (Section 1.1.1). Conversely, habituation due to decreased resource novelty could limit opportunities for intrinsic exploration, which could result in boredom (Section 1.1.2). In these situations, developing methods to attenuate long-term habituation are necessary to prevent negative welfare states.

1.3.3 How provision schemes could help

Provision schemes describe how and when environmental enrichments will be supplied in a captive environment. Diversity and novelty have been mentioned as two factors that alter exploration and habituation and likely impact welfare benefits of environmental enrichment. Diversity was first described in Section 1.1.2 within Berlyne’s (1960) definition of complexity: the number of different characteristics for a stimulus. This concept was discussed in the previous section where variety paradigms combined stimuli/rewards to prevent long-term habituation. If each resource included in a provision scheme is considered a characteristic of environmental enrichment, increasing resource diversity could exploit resource specificity and dishabituation to prolong active use. As animals become habituated to one resource they can turn attention to another through resource specificity and thereby facilitate later recovery of active use for the first resource through dishabituation. Increasing diversity of resources also provides greater choice and may improve behavioural opportunity. Abou-Ismail (2011a) investigated resource diversity by comparing rats raised with five different resources to rats raised with each resource alone and found that combining multiple resources in an enriched treatment increased resource contact and active use, as well as improved welfare inferred from increased sleep, healthier immune system, and decreased agonism.

Novelty was also mentioned by Berlyne (1960) as the perceived change in a stimulus and the factor that decreased as a stimulus was explored. Habituation literature further developed this concept and
defined characteristics that can change how quickly novelty is lost. Novelty through perceived change in a provision scheme can be achieved by adding or altering resources over time. Relative novelty is created by altering existing resources or increasing their number, and absolute novelty is created by adding resources or opportunities that have never been experienced (Berlyne, 1960). Provision schemes emphasizing novelty again seek to prevent long-term habituation by leveraging stimulus specificity and dishabituation. Spontaneous recovery could also increase relative novelty by removing and reintroducing resources, but such procedures may inadvertently remove high value resources, as in the discussion of habituation characteristic #6.

Abou-Ismail and Mendl (2016) compared novelty and diversity by designing treatments where five of the same item were supplied to rats and replaced each week with new items (absolute novelty) or five different items were supplied for the study duration (diversity). When the benefits of novelty and diversity were compared, rats enriched through novelty exhibited decreased resource contact and active use and showed signs of poorer welfare: less sleep, lower thymus weights, and total weight gain and increased aggression compared to those in the diversity condition. However, a limitation of this study is how novelty was achieved. Post hoc analysis identified an order effect on resource contact, sleep, and agonism between the resources used in the novelty treatment, potentially meaning that removal of highly valued resources early in the study could reduce efficacy of this provision scheme. Previous research showed that removing resources from enriched mice can increase frustration, evident by elevated corticosterone output and stronger motivation to access resources than barren housed mice (Latham and Mason, 2010). Gross et al. (2011) accounted for this effect by rotating multiple shelter and climbing exemplars so that each cage always had one resource from each category. In this study, no differences in stereotypic behaviour or anxiety, as measured by open-field and elevated zero-maze tests, were found between the novelty, diversity, and baseline nesting material conditions. Conversely, mice provided with two constant resources and one that was changed weekly did show reduced stereotypic behaviour compared to a baseline nesting material condition (Latham and Mason, 2010), so the true relationship between novel and diverse conditions is still uncertain.

Many provision schemes in environmental enrichment research use a combination of novelty and diversity to ensure that active use and welfare are maintained (e.g. Gronqvist et al., 2013; Guy et al., 2013; Nowicki et al., 2015; Schneider et al., 2013; Soriano et al., 2016; Wagman et al., 2018). This method is fantastic for improving welfare but lacks the precision needed to understand how novelty and diversity differ in their ability to sustain use and improve welfare (Swaisgood and Shepherdson, 2006). A fundamental understanding of how different factors alter the efficacy of a provision scheme would
remove much of the guesswork involved in developing environmental enrichment programs. While novelty and diversity show promise for mitigating habituation to resources, the number of enrichments provided may help clarify relationships between environmental enrichment and aggression.

1.4 Enrichment and aggression

A review of enrichment effects on intraspecific aggression reveals a torrent of contradictory results across species, likely because of the two effects outlined in Section 1.1.1: that behavioural frustration can promote aggression (such that enrichment is beneficial) and that competition can restrict access to highly motivating resources (such that enrichment exacerbates aggression). Environmental enrichment has thus been shown to increase intraspecific aggression (Boerrigter et al., 2016 [catfish]; Howerton et al., 2008 [mice]; Marashi et al., 2003 [mice]; Tarou et al., 2004 [orangutans]), decrease intraspecific aggression (Abou-Ismail, 2011b [rats]; Clipperton-Allen et al., 2015 [mice]; Damasceno and Genaro, 2014 [domestic cats]; Gvaryahu et al., 1994 [laying hens]; O’Connell and Beattie, 1999 [pigs]), or have no effect (Brent and Belik, 1997 [baboons]; Dantas-Divers et al., 2011 [domestic cats]; Marashi et al., 2004 [mice]; Nevison et al., 1999 [mice]; O’Connell et al., 2004 [pigs]). Decreasing intraspecific aggression is an additional way that environmental enrichment can improve welfare (Nip, 2018), and conditions leading to increased aggression are often explained by improper resource allocation (sex and species being additional factors).

Improper resource allocation is exemplified by Boerrigter et al. (2016) and Howerton et al. (2008). These studies provided catfish or mice with highly valued resources using a group size to resource ratio of 51:2 and 5:1, respectively, which both resulted in substantial increases in aggression. Providing a limited number of highly valued resources in this way often leads to increased resource defense and competition. Damasceno and Genaro (2014) tested this effect in a colony of domestic cats and found that increasing the number of suspended beef sirloins decreased instances of aggression, prevented monopolization by dominant cats, and allowed opportunity for more individuals to interact with the resource. A similar effect was observed by Paquette and Prescott (1988) where a group of four chimpanzees were provided one manipulable object and the most dominant chimpanzee monopolized use, but when 10 objects were provided this dominant individual showed the least interest in them. This suggests that the other chimpanzees had been prevented from accessing the manipulable object until a surplus was available. In this case, outright aggression did not occur because an established social hierarchy predetermined which chimpanzee would defend resources. Similar monopolization of resources
has been shown for tool-use resources in chimpanzees (Celli et al., 2003) and foraging resources in sea lions (Smith and Litchfield, 2010). These results show that the number of each resource provided can have considerable impact on welfare by altering competition and aggression.

Relationships between aggression and resource number are best explained by resource guarding—the maintenance of resource control in response to real or perceived competition for that resource (Archer, 1988). Maintaining resource control can occur via avoidance (i.e. moving the resource away from a competitor), rapid ingestion (i.e. consuming the resource before a competitor can take it), and/or escalated aggression (Jacobs, 2016). Wild animals typically utilize avoidance to diffuse competition, such as submissive posture or dispersal, but unnatural social structures and limited space in captivity can cause unresolved competition that leads to increased aggression (Koene et al., 2002). Further, many resources are durable and inedible, which prevents rapid ingestion, leaving only escalated aggression as a means of resource guarding. Providing a resource surplus ideally eliminates any real or perceived competition between conspecifics and thereby mitigates enrichment-induced competition. Dispersing resources throughout the environment can also decreased aggression compared to clustered resources, but this method requires ample space and a way to prevent dominant individuals from consolidating resources (Akre et al., 2011).

1.5 Mink: a model carnivore for studying environmental enrichment

1.5.1 Mink and mink farming

The use of model species allows welfare studies to benefit from large sample sizes and generate robust results that have application to source populations. For example, researchers have utilized populations of farmed American mink (Neovison vison) to rigorously assess how various feeding (Malmkvist et al., 2013), housing (Ahola et al., 2011; Díez-León et al., 2017; Hansen et al., 2007), and structural (Dawson et al., 2013) resources impact welfare. American mink are a semi-aquatic mustelid originating from North America but can be found in parts of Europe and Asia due to the spread of mink farming. Over 75 million mink were killed globally in 2016 (Fur Europe, 2017) with approximately 2.1 million killed in Canada (Mink Statistical Briefer, 2017). The true number of mink living on farms at peak production is even higher due to breeding animals. Studies on environmental enrichment with mink thus can have a substantial global impact on animal welfare by influencing research-based management and regulation of farming. Farmed mink also demonstrate many Carnivore-typical behaviours, making
them an excellent model for Carnivores housed in zoos, aquariums, and breeding centers (Dallaire et al., 2012; Diez-León et al., 2016).

Farm mink have been bred into numerous colour-types, each with different behavioural and physiological characteristics, and all follow a similar annual schedule in the northern hemisphere. Mink are born in late April with four to eight kits (infant mink) per litter. Kits are family housed with their mother for approximately six weeks in whelping cages. Whelping cages have a large wooden nest box with wood-shaving bedding and the cage’s wire-mesh floor is fitted with a finer plastic grate to prevent young kits from falling out of the cage (a thorough description of standard and whelping cages can be found in the methods of Chapter 2). Kits are weaned in early June when the mothers are moved to single housing and the plastic floor grate is removed to allow feces to exit the cage more easily. Siblings are housed together in the whelping cages until late-June and then are transferred to male/female pair-housing in standard cages (without access to the “drop-in nest box”—i.e. wooden box hung from the cage ceiling—until vaccinations can be administered one to two weeks later). Identification of mink that will potentially be bred for the next generation occurs during the transfer from sibling to pair-housing and a final screening happens in November. Once given access to the nest box, now without bedding, mink continue to grow through the summer and fall with daily feeding of protein, dairy, and nutritional additives ground into a paste and provided on top of the cage. In mid-October, on some farms, female mink are transferred to whelping cages so that all mink are single-housed. This helps prevent fur damage as mink mature and become much larger, territorial, and potentially aggressive. All mink that are not being kept for breeding in the new year are killed in late November with carbon monoxide, pelted, and prepared for sale. Breeder mink go on to mate in February and the cycle restarts in late-April.

Governments and advisory agencies have developed regulations or codes of practice requiring various forms of enrichment for captive animals. For mink in Canada, the current National Farm Animal Care Council (NFACC) Code of Practice for the Care & Handling of Mink (2013) requires farms to provide a manipulable object to all animals. Manipulable objects are inedible objects that can be easily handled. A review of scientific research that accompanied the NFACC Code of Practice identified gaps in the literature related to manipulable objects and noted potential increased competition and aggression over resources between group-housed mink as an outstanding issue for further study (Finley et al., 2012). Like the common refrain in zoos (see Section 1.2) the review also emphasized the need to find objects that generate long-term interest by mink. This idea again assumes that the welfare benefits of enrichment are closely tied to the active use of resources (a belief also expressed in a recent survey of Canadian farmers: Mason and Dawson, unpubl.). Together, this information generates three questions regarding the
provision and efficacy of manipulable objects: Does the provision of manipulable objects impact intraspecific aggression in mink? Can provision schemes be found for mink that encourage sustained long-term active use? And is active use of resources necessary for welfare benefits to occur?

1.5.2 Mink and environmental enrichment

Access to swimming water usually comes to mind when considering environmental enrichment for a semi-aquatic species and this topic has received the most research over the past few decades amongst all potential enrichments for mink. Consensus within mink literature is that the scientific basis for swimming, or access to a bath, as a behavioural need is inconsistent (Vinke et al., 2008). One possible reason is that swimming is a highly motivated elicited response in the presence of opportunities to swim. In motivational preference tests, as described in Section 1.1.1, mink consistently pay high costs to access a pool compared to other resources (Mason et al., 2001; Warburton and Mason, 2003). Furthermore, when large, naturalistic pools were provided to groups of twenty young mink, they spent over 15% of their time on average in and around water, and habituation to the pools did not occur over the five month study (Schwarzer et al., 2016). The main barriers to providing water basins for farmed mink are the cost to update existing farms and potential health risks caused by poor water management. Unless market trends significantly change (Mink Statistical Briefer, 2017), implementing a water-based enrichment requirement in Canada would be prohibitively expensive for most farms.

Although swimming may be out of reach, other enrichment strategies that also improve mink welfare with simple, economically feasible, on-farm solutions have been identified. First, studies utilizing enriched compartments—secondary cages accessible by an overhead tunnel system and containing manipulable objects and occasionally water basins—have shown that barren housing is detrimental to mink welfare and fitness (Dallaire et al., 2012; Diez-León et al., 2016; Diez-León and Mason, 2016; Meagher et al., 2017; Meagher and Mason, 2012). Meagher et al. (2014) applied this fundamental work to develop affordable manipulable enrichments that successfully improved welfare on commercial farms, such as balls, hanging ropes, or plastic chains. Additionally, as already outlined in Section 1.2.1, two studies focused on the welfare of mothers shortly after giving birth and found that availability of a wire-mesh shelf reduces kit mortality and improves mother health and welfare (Buob et al., 2013; Dawson et al., 2013). More recently, a survey of farmers analyzed what types of manipulable objects farmers had previously tried and which ones were most successful (Mason and Dawson, unpubl.). Again, hanging ropes, plastic balls, and PVC tubes were deemed most beneficial, although sturdy materials were emphasized to circumvent most minks’ destructive tenacity.
Two specific forms of passive resource use by mink related to manipulable objects identified in previous research have yet to be investigated. Farmers have noted that mink occasionally defecate on resources that subsequently go untouched (Mason and Dawson, unpubl.). This active use may reflect scent marking of resources as seen in wild mink (Dunstone, 1993) or may just be an accident. Meagher et al. (2014) and many farmers (Dawson and Mason, unpubl.) also anecdotally observed that resources were often carried into the nest box by mink and subsequently attracted little active use. This may reflect mink’s natural caching behaviour (Burness and Morris, 1993; Gerell, 1967), and resources in the nest box may still be inactively used by sleeping with them. Comparison of welfare indicators between animals that differentially exhibit these forms of resource use would help determine whether they indicate that mink value these resources or instead that the resources are now being ignored.

1.5.3 Mink as a model carnivore for research

Extensive research with mink has validated useful tools for assessing their welfare. As previously mentioned, utilisation of multiple independent welfare indicators helps to account for their individual limitations (Sandøe and Simonsen, 1992). As mentioned earlier, welfare indicators of interest to this thesis are stereotypic behaviour, fear, aggression, and boredom. Stereotypic behaviour—repetitive and unvarying behaviour supposedly caused by mechanisms that include motivational frustration—has been rigorously assessed in mink (Mason, 1993; Polanco et al., 2017). Most mink stereotypic behaviours fit into three categories: (1) head-based, movement of the head with all feet stationary; (2) whole-body, pacing or movement with only rear feet stationary; (3) scrabbling, sustained pawing at walls. Polanco et al. (2017) identified heterogeneity in cause and function for these categories, which suggests that pooling them into a single score may not always be appropriate. Other forms do occur beyond the main categories, such as bar licking and fur chewing. With these caveats in mind, environmental enrichment can reduce stereotypic behaviour, including fur chewing (Dallaire et al., 2012; Díez-León and Mason, 2016).

Turning to fear in mink, Meagher et al. (2011) developed the glove test specifically for Ontario mink because the stick test, which is widely used in Scandinavian research to measure fearfulness, failed to elicit similar fear responses in these mink. The glove test consists of presenting a well-used catching glove to mink and recording their initial response. Catching gloves are used whenever mink need to be transferred or handled (e.g. during vaccination) and therefore quickly gain the scent released when mink are frightened (Dunstone, 1993). Fear of humans or associated objects can have substantial impact on health and affective state in captivity due to increased stress caused by frequent human-animal interaction.
Evidence for such effects were found when gestating mink classified as fearful in the glove test were less likely to reproduce (Meagher et al., 2011). Therefore, finding ways to limit fearfulness in mink can simultaneously improve production and animal welfare. Meagher et al. (2014) found that simple environmental enrichment reduced fearful responses in mink. However, this effect varied between sexes, colour types, and as mink age.

Intraspecific aggression poses a welfare risk to aggressed animals through potential chronic stress and injury (Hurst et al., 1999, 1996). Aggression may also indicate compromised welfare for aggressors due to territorial intrusion or inability to establish appropriate social hierarchies (Boyce and O’Neill-Wagner, 1998; Koene et al., 2002), and/or due to being evidence of sustained frustration (as discussed in Section 1.1.1). Species typical intraspecific aggression in mink develops at approximately 20 weeks (October) and peaks during the spring breeding season (MacLennan and Bailey, 1969). Witnessing aggression during behaviour sampling is rare, yet consequences of aggression between cage-mates are often observed via open wounds or scarring (Dallaire and Mason, 2016; Hänninen et al., 2008b). Bite marks have therefore become a reliable indicator of intraspecific aggression for group-housed mink (Hansen et al., 2014). During the active growth phase of the winter coat and prior to maturation (approximately September-November), bites between mink break hair follicles releasing melanin granules that tattoo the skin. Resulting black marks can be counted to measure the cumulative incidence of aggression during the active growth phase. Furthermore, taking measurements from only the torso and tail removes marks around the neck associated with sexual behaviour (Hansen et al., 2014).

Finally, as mentioned in Section 1.1.2, boredom primarily occurs in suboptimal barren housing and can be identified by hyper-responsiveness to any stimulus (Burn, 2017; Meagher and Mason, 2012). Highlighting their value as a model species for welfare research, mink were the first species in which such states were operationalized. Presenting differentially housed mink with rewarding, neutral, and aversive stimuli and measuring responsiveness allowed discrimination between boredom and other low-arousal, negative affective states (for instance, apathy manifests as minimal interest in all stimuli, while anhedonia should cause normal interest to aversive stimuli but minimal interest to neutral and rewarding stimuli). This tested the hypothesis that aversive, barren conditions induce a generalized motivation to explore (Meagher and Mason, 2012). Comparison of interest scores for numerous aversive, neutral, and rewarding stimuli showed that mink housed without preferred environmental enrichments exhibit higher interest scores than enriched counterparts, suggesting that these responses do validly reflect the negative affective state of boredom (Meagher et al., 2017; Meagher and Mason, 2012).
1.6 Conclusions and aims of this thesis

To further a theoretical approach to environmental enrichment, this thesis aims to identify principles governing the utility of manipulable objects while determining the best provision scheme for improving welfare in socially housed mink. Number, diversity, and novelty of manipulable objects were the provision schemes investigated to explore strategies for improving welfare. Various forms of active and passive use were measured to identify connections with welfare improvement arising from environmental enrichment. The indicators for assessing welfare improvements were stereotypic behaviour, aggression, fear, and boredom because these variables are well described in mink and likely impacted by environmental enrichment. Comparison to a baseline representing current farm practices as recommended by the NFACC codes identified which provision strategies lead to welfare improvement. Distinct provision schemes also likely create variability in resource use, providing an opportunity to determine whether any form of resource use determines welfare improvement.

The following experiment tests two hypotheses: (1) Increasing the number, diversity, and novelty of resources improves resource use and welfare. Supporting evidence would be reduction in indicators of negative affect (stereotypic behaviour, aggression, fear, and boredom) and increase in resource use measures in these provision schemes compared to a baseline condition. It is suspected that competition for resources particularly increases intraspecific aggression, for which supporting evidence would be that increasing resource number reduces aggression compared to a baseline, with reductions being similar between provision schemes with equal resource numbers. (2) Active and/or passive resource use predicts the welfare impact of environmental enrichment. Supporting evidence would be negative correlations between resource use and stereotypy, boredom, fear, and aggression.

The remainder of this thesis accomplishes the following: Chapter 2 discusses provision scheme effects on active use, inactive use, and presence of resources in the nest box. This analysis shows whether the provision schemes described here successfully altered resource use. Chapter 3 begins by examining provision scheme effects on welfare indicators. These results are then compared to the results of Chapter 2 to identify consistent relationship between resource use, provision schemes, and welfare. Last, the potential for resource use as a predictor for welfare or causal factor in provision scheme effects on welfare is covered. Chapter 4 provides a general discussion on the concepts and results of my thesis. Topics for future research are outlined based on key results, a thorough discussion of methods used, and aspects of the experiment that were removed or altered from the initial design.
Chapter 2: Provision Scheme Effects On Resource Use

2.1 Introduction

Active and passive resource use are two ways for animals to obtain welfare benefit from environmental enrichment, as highlighted in Chapter 1. Active use involves extrinsic or intrinsic exploration of resources and is the focus of most environmental enrichment research (see Section 1.1.2). Alternatively, passive use includes inactive contact with resources or indirect use of resources, such as knowing that opportunities for active use are available or storing resources (see Sections 1.2 and 1.5.2). Active and inactive use, their combined measure of general use, and presence of resources in the nest box are the factors of interest for this chapter, and the term “resource use” collectively refers to them all.

As reviewed in Chapter 1, animal care researchers, providers, and governing organizations have emphasized developing environmental enrichment programs that bolster sustained active use of resources. Section 1.3 proposed altering how environmental enrichment is provided to encourage sustained interest in resources by exploiting the characteristics of habituation, such as stimulus specificity and dishabituation. Additionally, Section 1.4 noted that monopolization of resources could limit opportunity for some animals to use them. Habituation and competition are also potential limiting factors for inactive use and the presence of resources in the nest box because both forms of use may represent a response to stimuli (the resources), and monopolization of items by one animal could prevent the other from inactively using them or placing them in the nest box. The goal of this chapter is to understand which provision schemes best promote and sustain resource use, and how the different forms of resource use relate to each other, before exploring potential welfare benefits of resource use in Chapter 3.

To understand how altering the provision of resources impacts resource use, an appropriate baseline for comparison is necessary. The NFACC Code of Practice for Canadian mink farms offers such a baseline through its minimum requirement of one manipulable resource per cage (here, the “Baseline” provision scheme). Increasing the number of resources provided to groups of mink should allow more animals access to resources by limiting monopolization (e.g. Damasceno and Genaro, 2014; Paquette and Prescott, 1988). For pair housed mink, providing four manipulable objects should limit monopolization and competition (here, the “Number” provision scheme). Therefore, I hypothesized that resource use by mink increases with the number of resources provided. Once animals can use resources freely, attention then turns to limiting potential habituation of use.
Resource diversity modulates loss of novelty during use according to Berlyne’s (1963, 1960) conception of exploration: diversity exploits habituation specificity and dishabituation to increase and sustain use (Abou-Ismail, 2011a), and it can be increased by supplying multiple types of resources simultaneously (here, the “Diversity” provision scheme). I thus hypothesized that resource use by mink increases with the diversity of resources provided. Novelty was also identified as a resource characteristic that can decrease through use, leading to cessation of interaction in some instances. Increasing or reestablishing resource novelty relies on habituation specificity and dishabituation, and it can be achieved by adding new resources to an environment, a form of absolute novelty (here, the “Novelty” provision scheme) (Berlyne, 1960). Absolute novelty is preferred here over relative novelty because it offers the greatest change and therefore should have larger effect sizes when compared to other conditions (see Section 1.3.3). Therefore, I hypothesized that adding new items over time increases resource use.

Manipulable objects, resources that can be easily carried or moved with the paws or mouth, are studied here because these resources elicit a variety of uses and questions remain about their efficacy as environmental enrichments for farmed mink (Finley et al., 2012). Recent work with mink has also identified numerous manipulable objects that are viable for mink (Mason and Dawson, 2019 unpubl.; Meagher et al., 2014). Simultaneously studying how number, diversity, and novelty alter resource use requires at least two different resources. Small wiffle balls and plastic rings were chosen here because they are easily manipulable, beneficial for mink in combination with other resources, and can be taken into the nest box (Meagher et al., 2014). Axelsson et al. (2009) also found that manipulable plastic tunnels and balls received similar amounts of interaction. Beyond being required for the experimental design, repeating Baseline and Number conditions with two types of objects will reveal any robust scheme effects that are potentially generalizable to other resources.

The hypotheses investigated in this chapter all refer to differences between provision schemes and are summarized in Table 2.1. The predictions for these hypotheses can be split into practical and fundamental categories. Practical predictions identify provision schemes that increase resource use over current management practices by making comparisons to the Baseline scheme. These predictions are useful for guiding incremental change of environmental enrichment recommendations. Fundamental predictions identify which characteristics of a provision scheme increase resource use by comparing schemes to their relevant control group (see Table 2.1). For the Number scheme, comparison to Baseline fulfills both practical and fundamental predictions. The predictions can be summarized with the following schema for predicted resource use: Baseline < Number < Diversity < Novelty. In this order, each provision scheme adds one new characteristic while holding all previous characteristics constant.
Additionally, post hoc comparison between forms of resource use determined if any were redundant. Results from testing these hypotheses will then form predictions on differences in welfare between provision schemes that will be tested in Chapter 3.

Table 2.1. Hypotheses tested in Chapter 2. “Resource use” is interchangeable with active use, inactive use, general use (composite score of active and inactive), and presence of resources in the nest box. This creates a maximum of twelve hypotheses being tested and twenty predictions. The Baseline condition acts as a control for Number because the only difference in animal care and housing is the number of resources being provided. Comparison between Diversity and Baseline would reveal whether the former has increased resource use over current practices, but it would not provide information specifically about resource diversity because creating diversity requires more than one resource (i.e. there is a number confound). The appropriate control group for Diversity would have the same number of resources but of one type (the Number scheme). Novelty faces similar confounds as Diversity when compared to Baseline. The proper control group for assessing absolute novelty effects would have the same overall number and type of resources as Novelty, except they are held constant rather than being added over time (the Diversity scheme).

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Practical prediction</th>
<th>Fundamental prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource use increases with the number of resources available</td>
<td>Number will induce more resource use than Baseline</td>
<td></td>
</tr>
<tr>
<td>Resource use increases with the diversity of resources available</td>
<td>Diversity will induce more resource use than Baseline</td>
<td>Diversity will induce more resource use than Number</td>
</tr>
<tr>
<td>Adding new items over time increases resource use</td>
<td>Novelty will induce more resource use than Baseline</td>
<td>Novelty will induce more resource use than Diversity</td>
</tr>
</tbody>
</table>

2.2 Methods

The research presented here and in Chapter 3 was approved by the University of Guelph Animal Care Committee (AUP #3903.160788) and meets the most recent recommendations from the NFACC Code of Practice for the Care and Handling of Farmed Mink.

A southern Ontario mink farm supplied and housed the mink for this study. 816 juvenile mink were housed in male/female pairs in 408 cages sized to meet the NFACC Codes of Practice (Figure 2.1). The farm had previously identified these mink for pelting in November 2018. Cages were organized into
two rows separated by a walkway, with one row containing demi colour-type mink and the other pastel (Figure 2.2). Two colour-types were used due to limited numbers of each available on the farm and to increase external validity. Treatments were assigned to clusters of three cages (Figure 2.3). The two outside cages in each cluster were excluded from analysis to ensure that all cages of interest had neighbours with the same treatment as their own. Therefore, the middle cage in each cluster was the experimental unit (N = 136). Cages received their treatments in July 2018, immediately following the shift from family housing to pair housing, which is part of the farm’s standard operating procedure. Randomly assigning treatments within the demi and pastel rows provided blocking for colour-type.

Figure 2.1. Standard mink cage diagram. A) A top-down view of a cage. The walkway separating the two rows of cages is at the bottom of the figure. The drop-in nest box, which is supported by the cage ceiling, is represented by a grey square. All cages included an 8-inch strip of hanging plastic as the farm’s standard environmental enrichment (black rectangle). A water line runs along the walkway side of the cage and food is presented on top of the cage, in front of the nest box. B) A front view of a cage. The nest box entrance was either a 4-inch hole, as shown here, or a larger, rectangular gap made by removing the top 6 inches of the front panel. The hanging enrichment is hidden behind the nest box from this view. Cage sides were opaque plastic, but the rear wall, which backed against another cage, was wire mesh. Cage measurements: (L) 24 in X (W) 15 in X (H) 18 in. Nest box measurements: (L) 7 in X (W) 7 in X (H) 11 in.
Figure 2.2. Shed layout diagram. A diagram showing the shed layout. The shed had six rows of approximately 206 cages with three walkways providing access to cages. Mink for this study were housed in the center two rows. Green refers to the demi mink and blue to pastel. The outer rows in the shed contained whelping cages. Pastel females were moved to the whelping row outlined in red and demi females to the one outlined in blue.

Figure 2.3. Example cage layout with three clusters. Each colour represents a different treatment. Boxes with an ‘X’ were excluded from data collection and analysis.

Small white wiffle balls (PrideSports® Perforated Practice Golf Balls™) and PVC rings (4-inch diameter by 2-inch length) were the resources used in varying combinations for each treatment (Figure 2.4). Wiffle balls cost 0.25CAD each and the material cost of rings was 0.35CAD each (the emotional cost being immeasurable). Rings were constructed by cutting 4-inch diameter by 10-foot PVC pipes into 2-inch rings. A lengthwise cut made the rings flexible to prevent mink from becoming stuck in them and to ease their removal if they did. This cut caused rings to partially collapse, so the final diameter was slightly less than 4 inches. Rings were then sanded to remove any loose material and sharp edges. Approximately 660 rings and wiffle balls each were used in this study.
Four provision schemes created seven groups (Table 2.2). The Baseline scheme provided the minimum NFACC requirements for environmental enrichment; its groups (1 and 2) each comprised a single resource for the study’s duration and were separated into ring or ball treatments. The Number scheme provided four resources of a single type, again with groups (3 and 4) separated by ring or ball treatments. To control for lifetime number of resources, the remaining treatments provided four resources per cage by the end of the study. The Diversity scheme provided two of each resource and had a single group (5) of two balls and two rings. The Novelty scheme increased the number and diversity of resources over time; its groups (6 and 7) comprised two of a single resource to start, with two of the other resource added halfway through the study (two balls later supplemented by two rings, or two rings later supplemented by two balls). Therefore, the Novelty scheme had the order of resources as its treatment.
Table 2.2. Organization of provision schemes and treatment groups. The treatment row describes the difference between groups within the same provision scheme. Y and N refer to yes or no, respectively.

<table>
<thead>
<tr>
<th>Provision Scheme:</th>
<th>BASELINE</th>
<th>NUMBER</th>
<th>DIVERSITY</th>
<th>NOVELTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group Number:</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Group Resources:</td>
<td>1 Ball</td>
<td>1 Ring</td>
<td>4 Balls</td>
<td>4 Rings</td>
</tr>
<tr>
<td></td>
<td>2 Balls</td>
<td>2 Rings</td>
<td>2 Rings</td>
<td>2 Balls</td>
</tr>
<tr>
<td>Treatment:</td>
<td>Item Type</td>
<td>Item Type</td>
<td>None</td>
<td>Item Order</td>
</tr>
<tr>
<td>Experimental units per group:</td>
<td>18</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Resources per cage at end of study:</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Resource Diversity:</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Resources Constant:</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

For clarification, mink were assigned to the Baseline, Number, Diversity, or Novelty scheme; then, those in the Baseline and Number schemes were assigned to an item type treatment, and those in the Novelty scheme were assigned to an item order treatment. This created seven groups with treatment nested within scheme because treatment was not consistent across all schemes. Additionally, the Novelty scheme acted as a Two-item scheme when it only has the first set of resources. Cages were periodically checked to verify that the appropriate resources were present, and any that had fallen out of cages were replaced.

*A priori* sample size estimates indicated that a medium effect size, determined by the $f^2$ effect size index for general linear models (Cohen, 1992), at power 0.8 could be detected with a sample size of 18 experimental units per group (Figure 2.5). This was estimated by creating a sample size estimation plot for linear models based on the maximally fitted model with 11 independent variables (scheme/treatment*sex*colour; see Section 2.2.2). This rough sample size was relevant for all dependent variables with an expected medium effect size. Prior research on the dependent variables chosen here with similarly nuanced treatments and provision schemes was rare. Therefore, ability to detect a medium effect size was chosen because detecting small effects required sample sizes beyond what was available. Additionally, small effect sizes are less likely to be biologically relevant. Number, Diversity, and Novelty
groups each had 20 replicates between the two colour-types and Baseline groups had 18. Reduced replication in some groups was necessary for efficient use of space and allows for maximum replication possible in all groups, while having 18 or more replicates buffered against losses as the study progressed.

Figure 2.5. Sample size estimation for detecting effects with power 0.8. Cohen’s (1992) suggested effects size based on $f^2$ for linear models are 0.02 (small), 0.15 (medium), 0.35 (large). The dashed horizontal line indicates the sample size for groups other than those from the Baseline scheme.

2.2.1 Data collection

The different forms of resource use were measured via time budgets (along with some welfare indicators; see Chapter 3). Figure 2.6 presents an experiment timeline and summarizes when resource use was recorded. Time budgets for all mink were collected twice during the experiment, once in late-August and again in early-October, closely following the second addition of resources to the Novelty groups. This allowed observation of changes in dependent variables as the study progressed.
Time budgets were collected via instantaneous scan sampling with a 40-minute inter-observation window (Martin and Bateson, 2007). This interval is the minimum possible given the sample size, reliably reflects video recording (Svendsen et al., 2007), and is consistent with previous research by our group (e.g. Buob et al., 2013; Dawson et al., 2013). Observations began at 8:30am, approximately 30 minutes after leftover food was spread amongst cages by the farm staff (part of normal farm practice). A 15-minute break occurred around 11:00am while farm staff scraped any remaining food from the tops of cages. Observations then resumed until 1:00pm, at which point a 1.5-hour break occurred while mink were being fed. Whenever possible, one round of observations was made immediately prior to feeding to include this period of relatively high activity. Observations ended at 5:30pm. This schedule produced around 10 samples per mink each day and was repeated for five consecutive days in each collection period. The starting cage for observations was changed each day to prevent cages from being observed at the same time each day.

During pilot trials, if the observer stood directly in front of the cage while collecting data, mink ceased their current behaviour and attended to the observer, which prevented accurate description of their behaviour. The observer (SD) prevented such disturbance by watching the cage of interest from two cages away until the behaviour of both mink could be accurately recorded (< 10 seconds). The cage was then approached to record whether any resources were in the nest box before moving to the next set of cages. Sexual dimorphism, with males being larger and having squarer heads than females, allowed for reliable identification of both mink in a cage. Table 2.2 presents an ethogram of behaviours and cage features recorded. Resource contact in the cage and nest box were recorded as separate behaviours but were combined to form active and inactive use measures for analysis. Internal reliability for each behaviour within collection periods was checked via split-half analysis between odd and even days (Martin and Bateson, 2007).
Table 2.3. Ethogram of mink behaviours and cage features

<table>
<thead>
<tr>
<th>Behaviour</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource use</td>
<td>Active</td>
</tr>
<tr>
<td>Resource use</td>
<td>Inactive</td>
</tr>
<tr>
<td>Stereotypic behaviour (Polanco et al., 2016)</td>
<td>Scrabbling</td>
</tr>
<tr>
<td>Stereotypic behaviour (Polanco et al., 2016)</td>
<td>Head-based</td>
</tr>
<tr>
<td>Stereotypic behaviour (Polanco et al., 2016)</td>
<td>Whole-body</td>
</tr>
<tr>
<td>Aggression (Dallaire and Mason, 2016)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cage Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resources in nest box</td>
</tr>
</tbody>
</table>

In mid-October, females were transferred to the shed’s outer rows of whelping cages as part of the farm’s standard operating procedure (to prevent fur damage as the mink matured). All experimental resources remained with males in the home cage, meaning females no longer had access to experimental resources. Whelping cages provided a larger nest box outside the cage with wood-shaving bedding, a wire-mesh shelf, and a shorter ceiling (Figure 2.7). A third time budget collection occurred shortly after females were moved. The presence of resources in the nest box was not collected at this time due to practical constraints (the doubling of cages being observed). The general method and schedule for observations was consistent with the previous time budgets.
Figure 2.7. Mink whelping cage diagram. A) A top-down view. The nest box is attached to the outside of the cage while a wire-mesh shelf now hangs inside the cage. The standard environmental enrichment is an 8-inch strip of hanging plastic. Cage sides are opaque plastic. B) A front view of a whelping cage. Cage dimensions: (L) 24 in X (W) 15 in X (H) 15 in. Shelf dimensions: (L) 10 in X (W) 4 in X (H) 5 in. Nest box dimensions: 10 in X 10 in X 10 in.

2.2.2 Statistical analysis

Analyses were run in R version 3.5.2 (R Core Team, 2018). Generalized linear mixed models and linear mixed models were run with the “lme4” package (Bates et al., 2015). Three cages were removed from analysis. A pastel cage from the ring-second Novelty group was removed because it contained two males. A demi cage from the ball-second Novelty group was removed because it contained a pastel female and demi male. A demi cage from the four-ring Number group was removed because the male died prior to starting the experiment and the female was singly housed thereafter.

The most complete model structure included sex, colour, treatment nested in provision scheme, interaction between these variables, and a random cage variable \( y = \text{scheme/treatment} * \text{sex} * \text{colour} + (1|\text{cage}) \). Treatment was nested in provision scheme because treatments were not repeated across all schemes. The colour blocking factor is included in interactions because there were multiple independent units for treatment, scheme, and sex within each block. Because males and females within a cage are not independent, cage was included as a random variable whenever both sexes were
included to prevent psuedoreplication. During single housing, only resource use in males was investigated because females no longer had access to resources. These single-housing models therefore did not include a ‘sex’ fixed effect or ‘cage’ random effect.

All forms of use were analyzed as a proportion of observations. Active use was analyzed via logistic regression and modelled with generalized linear mixed models (GLMM) with a binomial distribution. Logistic regression was chosen because these proportion data did not meet the assumptions of normality and homogeneity of variance when using Gaussian models, even after appropriate transformation. The dependent variable for the GLMM was a two-column matrix with the number of observations where active use was present and absent. GLMMs were analyzed with the Type II Wald Chi-square test. The Wald Chi-square test approximates the likelihood ratio test for large sample sizes, but only requires one model and tests whether coefficients are equal to zero (Fox and Weisberg, 2019). Type II sums of squares were chosen because nesting treatment within scheme resulted in aliased coefficients that prevented the use of Type III sums of squares (see Hector et al., 2010).

Generalized linear models (GLM) were used whenever each sex needed to be modelled separately. Resource presence in the nest box was also modelled with GLMs because a single score was taken for each cage and these models provided better homogeneity of variance than Guassian models. Nest box resources were analyzed as the proportion of observations with at least one item in the nest box (e.g. 1/1, 2/4, and 4/4 items in the nest box all received the same score). GLMs were analyzed with likelihood-ratio tests using the `Anova()` function in the “car” package (Fox and Weisberg, 2019). Differences between the likelihood-ratio tests and Wald Chi-square tests for these models suggested that the more reliable likelihood-ratio test should be used. General and inactive use were modelled with linear mixed models following logit transformation to meet the assumptions of normality and homogeneity of variance. A composite variable for general use was created by combining the proportion of observations exhibiting active and inactive use. Linear mixed models were analyzed with Type II Wald F tests with Kenward-Roger method for determining the denominator degrees of freedom (see Spilke et al., 2005).

Models for each collection period were run separately, rather than as a repeated measures design, because the hypotheses being tested do not refer to changes in behaviour over time and differences in husbandry between males and females prevent complete models from being carried through the entire study period. If a model failed to converge due to over-fitting, it was first simplified by modeling each sex separately. This removed the random variable (cage) and reduced the interaction structure. If the models still failed to converge, they were again split by colour. If this did not solve the problem, it was decided that there were too little data for analysis and only descriptive statistics were reported. Any interactions
with p-value between 0.1 and 0.05 were further analyzed, but main effect trends were not reported due to
the number of tests being run (to avoid Type I errors). Models for interpreting interactions with provision
scheme or treatment were chosen by fitting separate models by the levels of the other factors. For
example, an interaction between provision scheme, colour, and sex would first be split by sex, and then
resulting scheme by colour interactions would be split by colour to find differences in scheme main
effects.

Post hoc investigation of the relationships between forms of use were run to reveal any statistical
redundancies between these measures. In these models one form of use was the dependent variable, and
another was included as a covariate \[ y = \text{scheme/treatment*sex*colour*covariate} + (1|\text{cage}) \]. Four comparisons were modelled: general/active, general/inactive, inactive/active,
inactive/nest box items. The first form of use listed was the dependent variable and the second was the
covariate. When added as covariates, inactive use and nest box items were both a continuous proportion
of observations, while active use was a 1/0 score for any expression in the collection period due to the
rarity of this behaviour. Partial R^2 for the covariate was determined for any models with a covariate main
effect with p < 0.05 using the \text{r2beta()} function in the “r2glmm” package (Jaeger, 2017). Interactions
with the covariate were split by the levels of the other factor(s) until the covariate main effect occurred
without interactions. Active use was not included as a dependent variable to prevent the use of pseudo R^2
statistics for logistic regression. The threshold for redundancy between forms of use was R^2 > 0.6 because
such values represent large effect sizes based on the r and \( f^2 \) standardized effect size measures (Cohen,
1992). If two forms of use were consistently redundant across collection periods, further analysis was
only conducted on one of them.

Before hypothesis testing commenced, split-half analysis for stability of each behaviour
compared odd and even days in each collection period to assess consistency over time as an indication of
data quality (Martin and Bateson, 2007). The odd day proportion of observations for the behaviour was
arbitrarily the dependent variable and the even day proportion was included as a covariate \[ \text{odd} = \text{scheme/trmt*sex*color*even} + (1|\text{cage}) \]. GLMMs or GLMs were again used for
modelling active use and nest box resources, while linear mixed models were used for inactive use. The
even day covariate for active use was a binary variable for presence of the behaviour, again due to rarity.
Any interactions or main effects with p > 0.05 were ignored due to the large number of tests being run.

The planned comparisons for interpreting scheme main effects differed between August and the
Fall (October and single housing) but were always designed to test hypotheses while minimising multiple
testing (cf. Table 2.1). In August, the Novelty scheme comprised two resources of one type, so it acted as an intermediary group between Baseline and Number and was better considered a Two-item scheme, as mentioned in Section 2.2. Therefore, for any August data where scheme was significant, a 2x3 factorial design was used to model the ball and ring treatments across Baseline, Number, and Two-item schemes to identify effects of number of resources on resource use. If scheme was significant in this model, additional Baseline/Number and Baseline/Two-item comparisons determined whether either scheme affected usage compare to Baseline. If at least one affected usage, they were then compared to each other. To identify effects of Diversity on usage in August, Diversity was first compared to Baseline using the nested treatment structure, and if Diversity affected usage, a further comparison of Diversity to Number was made (see Section 2.1 for how this comparison controls for the Diversity group also comprising more items, not just diverse items). In the Fall, scheme effects were interpreted by first comparing Number, Diversity, and Novelty schemes individually to Baseline. The Baseline/Number comparison used a 2x2 factorial design while the other Baseline comparisons used the original nested treatment design. If a scheme improved usage from Baseline, relevant comparisons were made to understand the relationship between Number/Diversity and Diversity/Novelty (and so test the hypotheses laid out in Section 2.1).

Post hoc models for interpreting treatment effects were run where treatment effects emerged in the planned scheme comparisons described in the previous paragraph or if only treatment effects (but no scheme effects) occurred in the model. Note that treatment effects were not relevant for the Diversity scheme because it only had one group. To investigate item type effects in August, a 2x3 factorial design was used to compare ball and ring treatments across Baseline, Number, and Two-item schemes. Likewise, in the Fall a 2x2 factorial design was used to compare ball and ring treatments between Baseline and Number schemes. To investigate item order effects in the Fall, a model comparing the two Novelty groups was run to identify any difference between ball-second and ring-second groups.

2.3 Results

This section begins by identifying relationships between forms of use (Section 2.3.1) and reporting the stability of each form of use within collection periods (Section 2.3.2) followed by sections reporting provision scheme effects on each form of use (Sections 2.3.3 - 2.3.5). Least square means (LSM) and associated standard error for the proportion of observations where the behaviour occurred are provided at the beginning of each section. Arithmetic means are presented in some instances, and noted, where model convergence was an issue. Each section also includes a graphical representation of the
results, with asterisks denoting P values (* < 0.05; ** < 0.01; *** < 0.001). After reporting the results within a collection period (e.g. active use in August), a short paragraph summarizes key outcomes. A concluding paragraph in each section summarizes the results for the form of use across collection periods. Additional tables at the end of each section compare the results to the predictions outlined in Table 2.1.

All items received substantial active use when first provided to cages in July (personal observations). Rings initially elicited fearful responses when dropped into cages, such as frantically running around the cage and into the nest box, but this subsided after a few minutes and object directed play commenced (see Dallaire and Mason, 2016 for discussion of object directed play). Some early instances of resource defense between cage-mates were observed, but no quantitative data were collected.

2.3.1 Relationships between forms of use

Active and inactive use

In August, a treatment*sex*active use interaction (F3,208.24 = 3.04, p = 0.03) and subsequent scheme*active use interaction (F3,105 = 2.72, p = 0.048) for female mink occurred. Only females in the Baseline and Number schemes showed a relationship between active and inactive use (p < 0.034). However, the direction of this effect differed between provision schemes. Baseline females showed a positive relationship between the forms of use (R² = 0.008), while Number females showed a negative relationship (R² = 0.07). No effect of active use on inactive use was found in October nor during single housing (October: F1,117.97 = 0.19, p = 0.66; single housing: F1,111 = 0.0005, p = 0.98). These results show that active and inactive use were distinct behaviours.

General and active use

An active use*provision scheme interaction occurred when active use was included as a covariate for general use in August (F3,181.91 = 2.66, p = 0.049). Subsequent models showed a positive relationship between active use and general use for only Baseline mink and males in the Novelty scheme (p < 0.04). The R² range for active use in these models were again very low at 0.02-0.08, suggesting that active use explained a miniscule amount of variation in general use. Furthermore, no effect of active use on general use occurred for any mink in October or during single housing (October: F1,117.97 = 1.96, p = 0.16; single housing: F1,111 = 0.06, p = 0.8). Together, these results show that active use and general use were not statistically redundant measures of resource use.
General and inactive use

All mink showed a positive relationship between inactive and general use in August (p < 0.0002). Interactions with other factors in the model did occur, but these were explained by differences in effect size (rather than the direction of presence of the relationship). The $R^2$ range for inactive use in models where only the main effect occurred in August was very high: 0.67-0.99. This strong positive relationship continued in October and during single housing for all mink (p < 0.0001), and the proportion of variation in general use explained by inactive use remained consistently high (October $R^2$: 0.61-0.98; single housing $R^2$: 0.8-0.99). These results, combined with those from active use, show that the general use composite variable was statistically redundant with inactive use due to the little amounts of active use observed in all collection periods (Figure 2.8). Therefore, general use was not included in further analysis.

**Figure 2.8. General use and its indicator variables across collection periods.** Only males are included in the single housing plots. Only main effects for the full model are shown due to similarity in the relationship between forms of use for all mink. All measurements of use are proportions of observations.
Inactive use and nest box resources

In August, there was a scheme*sex*colour*nest box resource interaction ($F_{3, 105} = 2.62, p = 0.055$) that led to scheme*colour*nest box and treatment*colour*nest box interactions in only female mink (scheme: $F_{3, 105} = 4.30, p = 0.0066$; treatment: $F_{3, 105} = 2.84, p = 0.042$). Subsequent predictive effects of nest box resources on inactive use only occurred in some groups of female mink, but were always in the same direction. A positive relationship between these forms of use occurred for pastel females in the one-ball group, Diverse females, females in the two-ball group, and demi females in the two-ring group ($p < 0.03, R^2: 0.03-0.88$). In October, all mink now showed a fairly strong positive relationship between inactive use and nest box resources ($p < 0.005, R^2: 0.42-0.82$). Interactions with other factors did occur but were explained by differences in effect size between groups of mink, rather than by differences in direction or presence of the relationship. The strengthening of this relationship between August and October seems best explained by an increased proportion of inactive use occurring in the nest box in October ($0.64 \pm 0.03$) compared to August ($0.14 \pm 0.02$). Although the presence of resources in the nest box was not measured during single housing, the continued increase in this proportion of inactive use for males ($0.85 \pm 0.027$) suggests that the relationship between forms of use may have again occurred. However, inconsistency in the relationship between inactive use and nest box resources over time and the moderate amount of variation in inactive use explained by the presence of resources in the nest box indicates that these forms of use were not redundant (Figure 2.9). Both forms of use were therefore included in further analysis to better understand how mink were using resources.
2.3.2 Behavioural stability

The stability of each form of use within collection periods is presented in Table 2.1. Active use had little stability within all collection periods. This likely resulted from the rarity of observed active use (see Section 2.3.2). Inactive use and presence of resources in the nest box, in contrast, were predominantly stable in all collection periods. Scheme*colour*odd day and treatment*odd day interactions occurred in August and October for nest box resources. However, these interactions were explained by differences in effect size (rather than the direction or presence of the relationship) and are not reported here.
Table 2.4. Behavioural stability of active and inactive use

<table>
<thead>
<tr>
<th>Collection Period</th>
<th>Behaviour</th>
<th>Stable</th>
<th>Not Stable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>August</td>
<td>Active use</td>
<td>P Males ($X^2(1) = 5.16$, $p = 0.023$)</td>
<td>D Males ($X^2(1) = 1.25$, $p = 0.26$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P 4-ball Females ($X^2(1) = 4.63$, $p = 0.031$)</td>
<td>D Females ($X^2(1) = 0.6$, $p = 0.44$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other P Females ($X^2(1) = 3.55$, $p &gt; 0.059$)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inactive use</td>
<td>All mink ($F_{1,208.13} = 38.3$, $p &lt; 0.0001$)</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Nest box resources</td>
<td>All mink ($X^2(1) = 360$, $p &lt; 0.0001$)</td>
<td>None</td>
</tr>
<tr>
<td>October</td>
<td>Active use</td>
<td>None</td>
<td>All mink ($X^2(1) = 0.24$, $p &gt; 0.63$)</td>
</tr>
<tr>
<td></td>
<td>Inactive use</td>
<td>All mink ($F_{1,205.95} = 66.8$, $p &lt; 0.0001$)</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Nest box resources</td>
<td>All mink ($X^2(1) = 991$, $p &lt; 0.0001$)</td>
<td>None</td>
</tr>
<tr>
<td>Single housing</td>
<td>Active use</td>
<td>None</td>
<td>Male mink ($X^2(1) = 1.39$, $p &gt; 0.24$)</td>
</tr>
<tr>
<td>(males only)</td>
<td>Inactive use</td>
<td>Male mink ($F_{1,105} = 246.8$, $p &lt; 0.0001$)</td>
<td>None</td>
</tr>
</tbody>
</table>

2.3.3 Provision scheme effects on active resource use

August active use

The arithmetic mean proportion of time spent actively using resources was low: 0.008 ± 0.001 for females and 0.018 ± 0.002 for males. The full model for active use in August did not converge, so each sex was modelled separately (Figure 2.10 and 2.11). Females showed a scheme by colour interaction ($X^2(3) = 13.45$, $p = 0.0038$) and treatment main effect ($X^2(3) = 9.68$, $p = 0.022$). Splitting by colour, demi females showed a scheme effect ($X^2(3) = 15.99$, $p = 0.0011$) and no treatment effect ($X^2(3) = 1.54$, $p = 0.67$). The mink in the Baseline, Number, and Two-items schemes did not differ in their active use ($X^2(2) = 3.57$, $p = 0.17$), but the Diversity scheme did induce more active use than both Baseline ($X^2(1) = 10.37$, $p = 0.0013$) and Number ($X^2(1) = 4.55$, $p = 0.033$). Pastel females showed no difference between schemes ($X^2(3) = 2.92$, $p = 0.40$), but there was a treatment effect ($X^2(3) = 1.54$, $p = 0.013$). When comparing Baseline, two-items, and four-items schemes, mink in ring groups showed more active use than ball.
groups ($X^2(1) = 9.92, p = 0.0016$), but there was no difference between schemes ($X^2(2) = 2.19, p = 0.33$).

Since Baseline could not be considered a homogenous group for pastel females, separate comparisons between Baseline ball and ring groups and Diversity were run but, in contrast to demi females, they showed no scheme effect (ball: $X^2(1) = 0.85, p = 0.36$; ring: $X^2(1) = 2.12, p = 0.15$).

**Figure 2.10. August active use by female mink split by colour type** Active use is measured as a proportion of observations. Error bars show 95% confidence intervals around the LSM.

Males showed both scheme effects ($X^2(3) = 16.96, p = 0.0007$) and treatment effects ($X^2(3) = 59.19, p < 0.0001$). Like in pastel females, rings received more active use than balls when comparing Baseline, two-items, and four-items schemes ($X^2(1) = 56.98, p < 0.0001$), and there was no difference between these schemes ($X^2(2) = 4.56, p = 0.1$). Comparisons between Diversity and other schemes were split by item type. Males in the Diversity scheme showed more active use than those with one or four balls (1 ball: $X^2(1) = 31.68, p < 0.0001$; 4 balls: $X^2(1) = 30.95, p < 0.0001$), but they did not differ from those with one ring ($X^2(1) = 3.56, p = 0.059$).
Figure 2.11. **August Active use by male mink.** Active use is measured as a proportion of observations. Error bars show 95% confidence intervals around the LSM.

In summary, active resource use was rare for all mink in August. Rings consistently received more active use than balls in all mink except demi females. Comparing provision schemes, only the Diversity scheme seemed to increase active use; however, this effect did not occur in pastel females and only occurred in males when comparing Diversity to ball groups.

October active use

The arithmetic mean proportion of time spent actively using resources was even lower now: 0.002 ± 0.001 for females and 0.003 ± 0.001 for males. The full model again did not converge, so each sex was modelled separately. Demi and pastel females were modelled separately because the complete female model also had issues with separation (Figure 2.12). Demi females showed no difference in active use between schemes ($X^2(3) = 4.48$, $p = 0.21$) or treatments ($X^2(3) = 4.49$, $p = 0.21$). Pastel females showed a treatment effect ($X^2(3) = 9.94$, $p = 0.019$), although there was still no difference between schemes ($X^2(3) = 0.52$, $p = 0.91$). Investigating the item type treatment in Baseline and Number schemes showed that rings still received more active use than balls ($X^2(1) = 6.95$, $p = 0.0084$). No order effect was
found between the two Novelty groups ($X^2(1) = 2.99, p = 0.08$). Because the Baseline scheme could no longer be considered homogenous, separate models were run for ring and ball comparisons to Diversity and Novelty. However, there was still no difference between these schemes and groups (ball: $X^2(2) = 2.57, p = 0.28$; ring: $X^2(2) = 0.69, p = 0.71$).

Like females, the full male model had issues with separation, so each colour was modelled separately (Figure 2.13). Demi males showed a scheme effect ($X^2(3) = 11.65, p = 0.0087$) and no treatment effect ($X^2(3) = 7.50, p = 0.057$). Comparison between the Baseline and Number schemes could not be completed because mink in neither scheme showed any active use. Subsequently, mink in the Diversity scheme showed more active use than both Baseline ($X^2(1) = 4.12, p = 0.042$) and Number ($X^2(1) = 4.26, p = 0.039$) schemes. The Novelty scheme also induced more active use than Baseline ($X^2(1) = 6.68, p = 0.0098$), but the ball-second group also showed more active use than the ring-second group ($X^2(1) = 7.50, p = 0.0062$). Comparison between Baseline and Novelty was rerun with separate

![Figure 2.12. October active use by female mink split by colour type. Active use is measured as a proportion of observations. Error bars show 95% confidence intervals around the LSM.](image)
models for each Novelty group. All active use previously shown by Novelty mink occurred in the ball-second group, which still differed from Baseline ($X^2(1) = 11.03, p = 0.0009$). The ball-second group did not differ from the Diversity scheme ($X^2(1) = 1.68, p = 0.20$). In contrast to demi males, pastel males showed a treatment effect ($X^2(3) = 11.25, p = 0.01$) but no scheme effect ($X^2(3) = 2.73, p = 0.44$). Rings received more active use than balls when compared between Baseline and Number schemes ($X^2(1) = 11.14, p < 0.001$), while there was no difference between Novelty groups ($X^2(1) = 0.11, p = 0.74$). Like pastel females, no scheme effects were found when comparing Diversity and Novelty groups to Baseline ring or ball groups separately (ball: $X^2(1) = 3.90, p = 0.14$; ring: $X^2(1) = 3.26, p = 0.19$).

**Figure 2.13. October active use by male mink split by colour type.** Active use is measured as a proportion of observations. Error bars show 95% confidence intervals around the LSM.

By inspection, active use decreased in males and females compared to their August values. Despite this, rings again received more active use than balls, although this effect was only present in pastel mink. The ball-second Novelty group showed more active use than the ring-second group in demi males, which is surprising given that rings general received more active use than balls in other mink and
that rings were given more recently in the ring-second group. Finally, the Diversity scheme was again the only scheme that increased active use, albeit with this effect now only found in demi males.

**Single-housing active use**

Only active use by males was investigated during this observation period because females no longer had access to rings or balls. The arithmetic mean proportion of time spent actively using resources for males was 0.002 ± 0.0005. No scheme or treatment altered active resource use at this time (scheme: \(X^2(1) = 1.41, p = 0.7\); treatment: \(X^2(1) = 5.22, p = 0.16\)).

**Active use summary**

Active use of resources was always rare and appeared to decrease over the course of the study. Still, somewhat consistent patterns were observed across August and October observations (Table 2.5). Rings always received more active use than balls in pastel mink, and this effect also occurred in demi males in August. Diversity was the only scheme that increased active use compared to Baseline and relevant control groups, albeit only in some mink. The ball-second group of demi males also showed increased use compared to the ring-second group and Baseline scheme in October, but this group did not differ from demi males with Diversity. Despite consistently higher active use for rings, which was observed in August but not October for demi males, no other differences in presentation order between Novelty groups were seen. None of the previous effects remained in males after females were removed from cages. Overall, all analyses here were also likely hampered by the rarity and instability of active use observations.
Table 2.5. Evaluation of predictions for active use separated by mink sex and colour. Values within a cell represent whether the prediction was met in a collection period, with collection periods ordered chronologically and separated by bars. Some predictions were not tested in all collection periods (e.g. Two-items comparisons were only possible in August, and Novelty comparisons were only possible in October and single housing). Cell colour indicates consistency of effects across collection periods. The order of consistency from lowest to highest follows darkest to lightest. X = no effect. ✓ = effect found. * = effects that were complicated by treatment interactions.

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Demi Females</th>
<th>Demi Males</th>
<th>Pastel Females</th>
<th>Pastel Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number &gt; Baseline</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2 items &gt; Baseline</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Number &gt; 2 items</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Diversity &gt; Baseline</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Diversity &gt; Number</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Novelty &gt; Baseline</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Novelty &gt; Diversity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Treatment Effects

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Demi Females</th>
<th>Demi Males</th>
<th>Pastel Females</th>
<th>Pastel Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rings &gt; Balls</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ball 2nd &gt; Ring 2nd</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

2.3.4 Provision scheme effects on inactive resource use

August inactive use

The LSM proportion of observations spent in inactive use was 0.13 ± 0.006 for females and 0.21 ± 0.009 for males. The full model for inactive use showed a treatment effect and scheme*sex interaction (treatment: F_{3,119} = 44.5, p < 0.0001; scheme*sex: F_{3,119} = 3.82, p = 0.012). Separate models were run for each sex to further investigate scheme effects. The female model showed scheme and treatment effects (scheme: F_{3,119} = 7.39, p = 0.0001; treatment: F_{3,119} = 17.0, p < 0.0001) (Figure 2.14). Comparing Baseline, Two-items, and Number schemes again showed a scheme effect (F_{2,101} = 8.81, p = 0.0003), and rings received more inactive use than balls (F_{1,101} = 47.9, p < 0.0001). Females in the Number scheme showed more inactive use than those with Baseline or Two-items (Baseline: F_{1,67} = 16.9, p = 0.0001; Two-items: F_{1,66} = 10.19, p = 0.0021), and the latter two schemes did not differ from each other (F_{1,66} = 1.05, p = 0.31). Comparisons between Diversity and other schemes were split by item type due to item type effects in the previous models. Females with the Diversity scheme showed more inactive use than
those with one or four balls (1 ball: $F_{1,34} = 27.9, p < 0.0001$; 4 balls: $F_{1,36} = 4.33, p = 0.045$), but they did not differ from those with one ring ($F_{1,34} = 0.096, p = 0.76$).

The complete male model showed a scheme effect and treatment*colour interaction (scheme: $F_{1,119} = 29.61, p < 0.0001$; treatment*colour: $F_{1,119} = 2.89, p = 0.038$) (Figure 2.15). The treatment*colour interaction was ignored for the moment while investigating scheme effects. The model comparing Baseline, Two-items, and Number schemes showed a scheme*item type*colour interaction ($F_{1,101} = 3.95, p = 0.022$). Subsequent models for this comparison were split by colour type. Both colour types again showed a scheme effect (D: $F_{1,50} = 17.1, p < 0.0001$; P: $F_{1,51} = 6.03, p = 0.0045$), and rings again received more inactive use than balls (D: $F_{1,50} = 63.9, p < 0.0001$; P: $F_{1,51} = 68.1, p < 0.0001$). Both colour types also showed that males with the Number scheme expressed more inactive use than those with Baseline or Two-items (D Baseline: $F_{1,33} = 42.9, p < 0.0001$; P Baseline: $F_{1,34} = 10.0, p = 0.0032$; D Two-items: $F_{1,34} = 5.02, p = 0.032$; P Two-items: $F_{1,35} = 5.82, p = 0.021$). Demi and pastel males differed in the

Figure 2.14. August inactive use by female mink. Inactive use is measured as a proportion of observations. Error bars show 95% confidence intervals around the LSM.

The complete male model showed a scheme effect and treatment*colour interaction (scheme: $F_{1,119} = 29.61, p < 0.0001$; treatment*colour: $F_{1,119} = 2.89, p = 0.038$) (Figure 2.15). The treatment*colour interaction was ignored for the moment while investigating scheme effects. The model comparing Baseline, Two-items, and Number schemes showed a scheme*item type*colour interaction ($F_{1,101} = 3.95, p = 0.022$). Subsequent models for this comparison were split by colour type. Both colour types again showed a scheme effect (D: $F_{1,50} = 17.1, p < 0.0001$; P: $F_{1,51} = 6.03, p = 0.0045$), and rings again received more inactive use than balls (D: $F_{1,50} = 63.9, p < 0.0001$; P: $F_{1,51} = 68.1, p < 0.0001$). Both colour types also showed that males with the Number scheme expressed more inactive use than those with Baseline or Two-items (D Baseline: $F_{1,33} = 42.9, p < 0.0001$; P Baseline: $F_{1,34} = 10.0, p = 0.0032$; D Two-items: $F_{1,34} = 5.02, p = 0.032$; P Two-items: $F_{1,35} = 5.82, p = 0.021$). Demi and pastel males differed in the
relationship between Baseline and Two-items schemes, with demi males in the Two-items scheme showing more inactive use than those in Baseline and pastel males showing a scheme*item type interaction (D: F\textsubscript{1,33} = 10.89, p = 0.0023; P: F\textsubscript{1,33} = 6.30, p = 0.017). Pastel males with two balls showed more inactive use than those with one ball (F\textsubscript{1,16} = 5.99, p = 0.026), but there was no difference between those with one or two rings (F\textsubscript{1,17} = 1.04, p = 0.32). Further comparisons with the Diversity scheme were separated by item type due to differences in inactive use between rings and balls. All males in the Diversity scheme showed more inactive use than those with one or four balls (1 ball: F\textsubscript{1,34} = 126, p < 0.0001; 4 balls: F\textsubscript{1,36} = 36.8, p < 0.0001), but there was a scheme*color interaction for the comparison to males with one ring (F\textsubscript{1,34} = 7.25, p = 0.011). Demi males in the Diversity scheme showed more inactive use than those with one ring (F\textsubscript{1,17} = 20.1, p = 0.0003), but pastel males in these groups did not differ (F\textsubscript{1,17} = 1.25, p = 0.27). There was no difference between demi males in the Diversity scheme and those with four rings (F\textsubscript{1,17} = 0.0004, p = 0.98).

![Figure 2.15. August inactive use by male mink split by colour type. Inactive use is measured as a proportion of observations. Error bars show 95\% confidence intervals around the LSM.](image-url)
In summary, effects on inactive use were mostly consistent across mink in August. Rings always received more use than balls and having four of either item resulted in more inactive use than having only one or two. However, mink with two items occasionally did not differ from those with only one item. Mink in the Diversity scheme also showed more use than those with one and four balls, but rarely showing more use than those with one ring.

October inactive use

The LSM proportion of observations spent in inactive use was 0.16 ± 0.015 for females and 0.16 ± 0.014 for males. The full model showed scheme and treatment effects on inactive use (scheme: $F_{3,119} = 7.35, p = 0.0001$; treatment: $F_{3,119} = 7.85, p < 0.0001$). Comparing Baseline and Number schemes revealed a scheme*item type interaction ($F_{1,67} = 5.87, p = 0.018$) (Figure 2.16). Scheme effects only occurred between ring groups; mink with four rings showed more inactive use than those with one ring (ring: $F_{1,33} = 9.22, p = 0.0046$; ball: $F_{1,34} = 0.55, p = 0.46$). Item type effects only occurred in the Number scheme; mink with four rings showed more inactive use than those with four balls (Number: $F_{1,35} = 24.8, p < 0.0001$; Baseline: $F_{1,32} = 0.43, p = 0.51$). Because item type effects only occurred in the Number scheme, further comparisons to Baseline were not split by item type. Mink with the Diversity scheme showed more inactive use than those in Baseline and the four-ball group but did not differ from the four-ring group (Baseline: $F_{1,50} = 14.1, p = 0.0004$; four-ball: $F_{1,36} = 22.7, p < 0.0001$; four-ring: $F_{1,35} = 0.38, p = 0.54$) (Figure 2.15).
Figure 2.16. Diversity and Number comparisons for October inactive use. Inactive use is measured as a proportion of observations. Error bars show 95% confidence intervals around the LSM.

Mink in the Novelty scheme also showed more inactive use than Baseline ($F_{1,66} = 10.4, p = 0.002$), although there was a treatment*sex*colour interaction ($F_{1,66} = 3.12, p = 0.051$) (Figure 2.17). Only pastel males showed an item order effect, with the ball-second group showing more inactive use than the ring-second group ($F_{1,17} = 5.57, p = 0.028$). The ball-second group still increased inactive use over Baseline for pastel males ($F_{1,25} = 11.7, p = 0.0022$), but the ring-second group did not differ from Baseline ($F_{1,24} = 0.008, p = 0.93$). The original Novelty scheme effect also remained for demi mink and pastel females when pastel males were not included ($F_{1,66} = 10.8, p = 0.0016$). The Novelty scheme did not differ from Diversity in demi mink, pastel females, nor pastel males in the ball-second group (DM and PF: $F_{1,52} = 0.71, p = 0.40$; PM: $F_{1,18} = 1.03, p = 0.32$).
As in August, the Number scheme again increased inactive use over Baseline, although only between groups with rings. Item type effects also remained consistent within the Number scheme, with rings receiving more inactive use than balls. The Diversity scheme increased inactive use over Baseline and the four-balls group, and no difference was found between Diversity and the four-rings group. Lastly, Novelty effects were complicated by item order interactions with sex and colour type. Pastel males that received balls second showed more inactive use than those that received rings second. The former group also showed more inactive use than Baseline while the latter did not differ from Baseline. Demi mink and pastel females with the Novelty scheme also showed more inactive use than Baseline. No Novelty group increased inactive use over the Diversity scheme.

Figure 2.17. Novelty comparisons for October inactive use with pastel males shown separately. Inactive use is measured as a proportion of observations. Arithmetic means and 95% confidence intervals with the modified Wilson method (Brown et al., 2001) are shown for demi mink and pastel females. Error bars show 95% confidence intervals around the LSM for pastel males.
Single-housing inactive use

The arithmetic mean proportion of observations spent in inactive use was $0.41 \pm 0.26$ for males. The full model showed scheme and treatment effects on inactive use during single housing in male mink (scheme: $F_{3,119} = 5.47$, $p = 0.0015$; treatment: $F_{3,119} = 3.36$, $p = 0.021$) (Figure 2.18). Comparing Baseline and Number schemes showed no scheme effect ($F_{1,67} = 2.03$, $p = 0.16$), but rings did receive more inactive use than balls ($F_{1,67} = 7.33$, $p = 0.0086$). Further comparisons between Baseline and other schemes were split by item type. Males with the Diversity scheme showed more inactive use than those with one or four balls (1 ball: $F_{1,34} = 7.68$, $p = 0.009$; 4 balls: $F_{1,36} = 11.72$, $p = 0.0016$), but they did not differ from those with one ring ($F_{1,34} = 0.24$, $p = 0.63$). The Novelty scheme only increased inactive use over the one-ball group (1 ball: $F_{1,50} = 15.6$, $p = 0.0074$; 1 ring: $F_{1,50} = 0.04$, $p = 0.83$), and a scheme*colour interaction occurred when comparing Novelty to Diversity ($F_{1,52} = 6.89$, $p = 0.011$). Surprisingly, demi males with the Novelty scheme showed less inactive use than those with Diversity ($F_{1,26} = 4.49$, $p = 0.044$), and there was no scheme effect for pastel males ($F_{1,26} = 2.59$, $p = 0.12$).

Figure 2.18. Single housing inactive use by male mink separated by colour type to show differences in Novelty effects. Inactive use is measured as a proportion of observations. Error bars show 95% confidence intervals around the LSM.
Unlike previous collection periods, mink in the Number and Baseline scheme did not differ in their amount of inactive use, although rings still received more use than balls. Like previous collection periods, the Diversity scheme only increased inactive use over groups with one or four balls. Novelty effects were sparse, with this scheme only increasing inactive use over the one-ball group and even showing less inactive use than the Diversity scheme.

Inactive use summary

Scheme effects changed slightly as the study progressed (Table 2.6). The Number scheme increased inactive use over Baseline in August and October, but this effect did not occur during single housing. Rings consistently received more inactive use than balls in all collection periods. The Diversity scheme always increased inactive use over groups with only balls and rarely differed from groups with rings. Novelty effects were more inconsistent. In October, the Novelty scheme increased inactive use over Baseline in most mink, while in single housing, males with the Novelty scheme only showed more use than those with one ball. Item order effects were rare, only occurring in pastel males in October. Overall expression of inactive resource use was consistent between August and October for both sexes. However, males in single housing doubled their inactive use compared to levels recorded two weeks prior.
Table 2.6. Evaluation of predictions for inactive use separated by mink sex and colour. Values within a cell represent whether the prediction was met in a collection period, with collection periods ordered chronologically and separated by bars. Some predictions were not tested in all collection periods (e.g., Two-items comparisons were only possible in August, and Novelty comparisons were only possible in October and single housing). Cell colour indicates consistency of effects across collection periods. The order of consistency from lowest to highest follows darkest to lightest. X = no effect found. ✓ = effect found. “<” = opposite effect found. * = effects that were complicated by treatment or scheme interactions.

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Demi Females</th>
<th>Demi Males</th>
<th>Pastel Females</th>
<th>Pastel Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number &gt; Baseline</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2 items &gt; Baseline</td>
<td>X</td>
<td>✓</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>Number &gt; 2 items</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Diversity &gt; Baseline</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Diversity &gt; Number</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Novelty &gt; Baseline</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Novelty &gt; Diversity</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

### Treatment Effects

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Demi Females</th>
<th>Demi Males</th>
<th>Pastel Females</th>
<th>Pastel Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rings &gt; Balls</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ball 2nd &gt; Ring 2nd</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>✓</td>
</tr>
</tbody>
</table>

2.3.5 Provision scheme effects on resources in the nest box

August presence of resource in the nest box

The LSM proportion of observations with resources in the nest box was 0.13 ± 0.007 for demi mink and 0.17 ± 0.007 for pastels. The full model for presence of resource in the nest box in August showed scheme*colour and treatment*colour interactions (scheme: $X^2(3) = 101.6, p < 0.0001$; treatment: $X^2(3) = 23.69, p < 0.0001$). Subsequent models were thus split by colour type (Figure 2.19). Both demi and pastel mink showed scheme effects when comparing Baseline, Two-item, and Number schemes (D: $X^2(2) = 31.72, p < 0.0001$; P: $X^2(2) = 10.32, p < 0.0057$), and rings were in the nest box more often than balls (D: $X^2(2) = 171.6, p < 0.0001$; P: $X^2(2) = 73.64, p < 0.0001$). For demi mink, items were in the nest box less often in the Number scheme than Baseline or Two-items (Baseline: $X^2(1) = 27.02, p < 0.0001$; Two-items: $X^2(2) = 21.50, p < 0.0001$). Conversely, pastel mink with the Number scheme did not differ from Baseline (Baseline: $X^2(1) = 2.09, p = 0.15$), but did have items in the nest box more often than those with Two-items ($X^2(1) = 10.25, p = 0.0014$). The Baseline and Two-items schemes did not differ in either
colour type (D: $X^2(1) = 0.35, p = 0.55$; P: $X^2(1) = 2.91, p = 0.088$). Comparisons between Diversity and other schemes were split by item type due to item type effects in the previous models. Demi mink with the Diversity scheme had items in the nest box more often than those with Baseline or Number (1 ball: $X^2(1) = 136.8, p < 0.0001$; 1 ring: $X^2(1) = 12.81, p = 0.0003$; 4 balls: $X^2(1) = 202.4, p < 0.0001$; 4 rings: $X^2(1) = 66.59, p < 0.0001$). Pastel mink in the Diversity scheme also had items in the nest box more often than those in the one-ball and four-ball group (1 ball: $X^2(1) = 11.22, p = 0.0008$; 4 ball: $X^2(1) = 4.66, p = 0.031$), but less often than those in the one-ring or four-ring group (1 ring: $X^2(1) = 4.10, p = 0.043$; 4 ring: $X^2(1) = 8.75, p = 0.0031$).

The only consistent effect across color types in August was that rings were in the nest box more often than balls. The Number scheme did increase nest box presence over Baseline and Two-items only for demi mink and neither colour showed differences between Baseline and Two-items. Demi mink with

![Figure 2.19. August presence of items in the nest box separated by colour type. Presence of items in the nest box is measured as a proportion of observations. Error bars show 95% confidence intervals around the LSM.](image-url)
the Diverse scheme had items in the nest box more often than those in Baseline and Number, but pastel mink only showed this effect when comparing Diversity to ball groups.

October presence of resource in the nest box

The LSM proportion of observations with items in the nest box was 0.3 ± 0.009 for demi mink and 0.37 ± 0.009 for pastels. The full model for items in the nest box in October showed a scheme effect and treatment*colour interaction (scheme: \(X^2(3) = 405.2, p < 0.0001\); treatment: \(X^2(3) = 12.95, p = 0.0047\)). Scheme effects were investigated first, and further treatment effects were addressed as they occurred. Comparing Baseline and Number schemes revealed a scheme*item type interaction (\(X^2(1) = 73.04, p < 0.0001\)) (Figure 2.20). The interaction occurred because the one-ball group had items in the nest box more often than the four-ball group (\(X^2(1) = 7.51, p < 0.0061\)), and the one-ring group had items in the nest box less often than the four-ring group (\(X^2(1) = 111.7, p < 0.0001\)). Item type effects were consistent across schemes, with rings again being in the nest box more often than balls (Baseline: \(X^2(1) = 9.74, p < 0.0018\); Number: \(X^2(1) = 277.3, p < 0.0001\)). Further comparisons to Baseline or Number were split into separate models for each item type. Mink with the Diversity scheme had items in the nest box more often than both Baseline groups and the four-ball group but did not differ from the four-ring group (1 ball: \(X^2(1) = 215.9, p < 0.0001\); 1 ring: \(X^2(1) = 134.9, p < 0.0001\); 4 ball: \(X^2(1) = 315.6, p < 0.0001\); 4 ring: \(X^2(1) = 0.92, p = 0.34\)) (Figure 2.19).
The models comparing Novelty to Baseline groups showed a potential item order effect that was analyzed separately ($X^2(1) = 113, p < 0.0001$). Comparing ball-second and ring-second groups revealed an item order*colour interaction ($X^2(1) = 10.69, p = 0.0011$) (Figure 2.21). Both demi and pastel mink in the ball-second group had items in the nest box more often than those in the ring-second group (D: $X^2(1) = 26.72, p < 0.0001$; P: $X^2(1) = 97.04, p < 0.0001$), with this difference being more pronounced in pastel mink (D Cohen’s d: 0.52; P Cohen’s d: 0.89). Separate models for each item order were used when investigating Novelty scheme effects. Ball-first and ring-first groups had items in the nest box more often than both Baseline groups (Ball-first 1b: $X^2(1) = 71.62, p < 0.0001$; Ball-first 1r: $X^2(1) = 28.24, p < 0.0001$; Ring-first 1b: $X^2(1) = 353.8, p < 0.0001$; Ring-first 1r: $X^2(1) = 249.1, p < 0.0001$), although there was a scheme*colour effect when comparing ball-second to the one-ring group ($X^2(1) = 8.30, p = 0.004$). This interaction was like the previous colour type interaction, which was driven by pastel mink in the ball-second group having items in the nest box more often than demi mink (D: $0.49 \pm 0.11$; P: $0.67 \pm 0.10$). Only pastel mink in the ball-second group had items in the nest box more often than those in the
Diversity scheme ($X^2(1) = 181.4$, $p < 0.0001$), and all mink in the ring-second group had items in the nest box less often than those in the Diversity scheme ($X^2(1) = 40.95$, $p < 0.0001$).

**Figure 2.21.** Novelty comparisons for October presence of items in the nest box separated by colour type. Presence of items in the nest box is measured as a proportion of observations. Error bars show 95% confidence intervals around the LSM.

Scheme and treatment effects on presence of resources in the nest box in October were mostly consistent across cages. As in August, rings were in the nest box more often than balls. Number effects differed between item type, with the one-ball group having an item in the nest box more often than the four-ball group, while ring groups unexpectedly showed the opposite effect. The Diverse scheme had items in the nest box more often than Baseline and the four-ball group but did not differ from the four-ring group. Lastly, the Novelty scheme had items in the nest box more often than Baseline. Comparisons between Novelty and Diversity were inconsistent, with only pastel cages in the ball-second group having items in the nest box more often than Diversity and all cages in the ring-second group having nest box presence less than Diversity.
The proportion of observations with one or more items in the nest box doubled for all mink between August and October. Given the proportion of observations that males exhibited inactive use in the nest box during single housing (0.41 ± 0.02), the mean proportion of observations where resources were in the nest box must have again increased over October values (D: 0.3 ± 0.03; P: 0.37 ± 0.03). Some effects on items in the nest box remained consistent throughout the study (Table 2.7). Treatment effects occurred in August and October: rings always were in the nest box more often than balls, and the ball-second group had items in the nest box more often than the ring-second group. Mink in the Diversity scheme had items in the nest box more often than balls groups in both observation periods. Other Number and Novelty comparisons were inconsistent, possibly due to strong item type effects.

### Table 2.7. Evaluation of predictions for items in the nest box separated by mink colour.

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Demi Mink</th>
<th>Pastel Mink</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number &gt; Baseline</td>
<td>&lt;</td>
<td>✓</td>
</tr>
<tr>
<td>2 items &gt; Baseline</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Number &gt; 2 items</td>
<td>&lt;</td>
<td>✓</td>
</tr>
<tr>
<td>Diversity &gt; Baseline</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Diversity &gt; Number</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Novelty &gt; Baseline</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Novelty &gt; Diversity</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

**Treatment Effects**

<table>
<thead>
<tr>
<th>Treatment Effects</th>
<th>Demi Mink</th>
<th>Pastel Mink</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rings &gt; Balls</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ball 2nd &gt; Ring 2nd</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Further analysis of the Diversity scheme revealed that the proportion of observations with only rings in the nest box (August: 0.23 ± 0.04; October: 0.47 ± 0.04) was much higher than that for only balls (August: 0.03 ± 0.012; October: 0.002 ± 0.001) and both items (August: 0.006 ± 0.003; October: 0.008 ± 0.002) in both collection periods. Similarly, exploring item type effects for the Novelty scheme in October, when rings and balls were available in both treatments, revealed differences in how items were used based on when they were added to cages (Figure 2.22). Mink that received balls second had a proportion of observations with items in the nest box consistent with the Diversity scheme in October (rings only: 0.58 ± 0.04; balls only: 0.002 ± 0.001; both: 0.02 ± 0.004). Conversely, mink that received rings second had substantially fewer observations with only rings in the nest box and increased their observations with balls and both items present compared to the Diversity scheme in October (rings only: 0.17 ± 0.02; balls only: 0.04 ± 0.01; both: 0.10 ± 0.03).

**Figure 2.22.** Comparison of items in the nest box between Diversity and Novelty treatments in October. Presence in the nest box is a proportion of observations. Arithmetic means are shown with 95% confidence intervals with the modified Wilson method (Brown et al., 2001).
2.4 Discussion

2.4.1 Relationships between forms of use

The rarity of and decline in active resource use during the study likely contributed to the instability of this behaviour within collection periods, although the amount of active use displayed in August was not unusual, being consistent with the level recorded in a previous study with mink (Dallaire and Mason, 2016). Meagher et al (2014) also observed substantial decreases in active resources use through the fall and winter that were similar to those observed here. For active use in October, the few differences observed between provisions schemes may be unreliable because many groups did not perform the behaviour during observations. The decline in active use between observation periods cannot be attributed solely to habituation because maturation and colder weather are also associated with overall decreases in active behaviour for mink, which were also observed here (Dallaire and Mason, 2016). Disentangling which factors are most important would require additional research (see Chapter 4).

Low amounts of active use relate to the positive relationship observed between inactive and general use. For this study, it appears that the composite score of general use was simply a measure of inactive use with added variability caused by small amounts of active use. Because measures of general and inactive use are statistically redundant, only potential welfare benefit from inactive use will be investigated in Chapter 3.

Inactive use was also a composite variable that combined inactive contact in the main cage and nest box. The proportion of total inactive use drawn from each of these original variables changed as the study progressed, with inactive contact in the nest box becoming predominant in October and single housing. This is possibly explained by items accumulating in the nest box over time combined with general increases in inactivity later in the year. This also suggests that mink chose to sleep with/inactively contact items because items were unavoidable in the nest box and could have easily been removed. An alternative explanation is that mink removed items from the main cage space to create more room for movement. However, the proportion of observations that mink were in the nest box, with or without items, increased during the study for both males and females (Aug M: 0.12 ± 0.01; Oct M: 0.60 ± 0.01; Single M: 0.66 ± 0.01; Aug F: 0.15 ± 0.01; Oct F: 0.67 ± 0.01). Rather than place items in the nest box in August to create space for movement or resting, when mink were spending the majority of their time in the main cage, mink instead kept items in the space they themselves predominantly occupied, which continued as mink increased their time in the nest box later in the study.
As inactive use occurred in the nest box more often over time, the positive relationship between inactive use and presence of resources in the nest box became more pronounced. Because some cages had resources in the nest box even when mink were not in the nest box, it is still possible that some non-contact benefit of having resources there occurred. For instance, storing items in the nest box may have provided an outlet for caching behaviour exhibited by wild mink (Burness and Morris, 1993; Gerell, 1967). Therefore, both resources in the nest box and inactive use will be investigated in Chapter 3.

### 2.4.2 Treatment effects

Treatment effects complicated analysis of provision scheme effects. Differential effects of item type (rings versus balls) and order (rings second versus balls second, in the Novelty group) occasionally prevented specific evaluation of provision schemes because treatments were nested within provision schemes. The experimental design employed here aimed to identify similarities between treatment groups and relied on them for broader comparison between provision schemes to test the hypotheses outlined in the introduction. In practice, treatment effects were sufficiently large for one conclusion to be that provision schemes do not have effects that are consistent across different resources: an interesting finding, if somewhat disappointing for the hypotheses being tested. The specific challenges caused by treatment effects when analyzing provision scheme comparisons are discussed in the next section. However, consistent treatment effects may provide another way to assess possible connections between resource use and welfare, as will be shown below.

Item type effects were consistent across all forms of resource use in all collection periods, except for active use during single housing, which may have been a Type II error due to increasing rarity of active use as the study progressed. Surprisingly, rings received more use than balls, contradicting the previous finding that these items receive similar amounts of use that had been part of the basis for their inclusion in this study (Axelsson et al., 2009). For active and inactive use, item type effects only refer to differences in use when only one item type was available, i.e. Baseline and Number schemes, because these forms of use were not separately analyzed for each item type in the Diversity and Novelty schemes. Resource presence in the nest box was measured separately for each item type and revealed that item type effects on this form of use also applied to Diversity and Novelty schemes. This suggests that mink preferred rings in the nest box more than balls when both items were available.

It is possible that rings appeared to receive more inactive use simply because they were larger than balls, which causes a greater probability of randomly contacting them while inactive. However, it is
unlikely that mink would accidentally wear rings, which was frequently observed (Figure 2.23). A better explanation for the difference in use is that the rings used here offered a greater variety of behavioural opportunities than balls due to modifications made compared to those used in previous research. Axelsson et al. (2009) provided female mink with a 25 cm long plastic tunnel and noted that it received active and inactive use and that attempts were made to bring them into the nest box. The modified rings used here were one-fifth that length (being c. 5 cm long) and therefore could be more easily manipulated, worn/slept in, and brought into the nest box.

These item type differences do not, however, support any of the hypotheses defined for this chapter and should be considered *post hoc* results. Recognizing that, item type effects shown here do support the idea that resource use increases with the number of different behavioural opportunities provided by an item (which is developed further in Chapter 4).

*Figure 2.23. Mink wearing rings.* The pastel male is wearing two rings while inactive. The pastel female would also be classified as in a state of inactive use because she is contacting one of the rings the male is wearing. Photo by SD.

Item order effects in the Novelty scheme were less consistent than item type effects and only occurred in October. Whenever item order effects occurred in active use, inactive use, or items in the nest box, the ball-second group displayed more use than the ring-second group. This is especially surprising
because the novelty of rings, the item presumably receiving most use based on item type effects, was greater in the ring-second group, and habituation theory predicts that novelty would enhance usage. It is possible that having the greater behavioural opportunities provided by rings earlier in life led to fundamental changes in the extent of resource use later in life. The difference in which items were found in the nest box between mink that received rings or balls first in the Novelty scheme does show that mink used items differently depending on when items were introduced. Further investigation of such time/order dependent effects is required before any robust conclusions can be made.

Finally, the order effect found here also potentially differs from that found by Abou-Ismail and Mendl (2016), where rats were given five items of the same type for one week that were then swapped for five items of a different type. This schedule was repeated over five weeks for five different items with the order of item introduction randomized between groups of rats. An item order effect on resource use was found that could be attributed to either the removal or late addition of valued objects. The removal of items was avoided in the present study to limit potential order effects, but continued presence of such effects offers further credibility to the order of introduction explanation. Still, the treatment effects shown here and by Abou-Ismail and Mendl (2016) emphasize that, when attempting to promote use, the behavioural opportunities provided by resources should remain constant even if access to all resources is not always available.

2.4.3 Provision scheme effects

The number of resources provided did not affect active use throughout the study, the rarity of this behaviour during observations and poor data quality may be responsible (this is returned to in Section 2.4.3). Effects on inactive use, and resources in the nest box changed as the study progressed. In August, the Number scheme showed more inactive use than the Baseline and Two-item schemes. While this could be explained by random chance, the same effect applied to presence of resources in the nest box, at least for demi mink: a form of passive use preceded by the active movement of items. It is unlikely that items would randomly enter the nest box because the entrance was located approximately seven inches above the cage floor. Similar effects occurred in October for both forms of use, now across all mink: the Number scheme increased use over Baseline, but only for comparisons between ring groups. These results for inactive use and the presence of resources in the nest box support the hypothesis that these forms of resource use increase with the number of resources provided, as if reducing competition between mink allowed more utilization. Consistent with this, the number of resources did not impact resource use of any kind during single housing (a period where competition obviously does not apply). However, the Two-
item scheme in August also did not differ from Baseline in any form of resource use. Although this last result does not support the resource number hypothesis, it does combine with the single housing result to further show that differences in resource use between groups of mink with different numbers of items cannot be solely explained by the random probability of contacting an item or having an item in the nest box.

The diversity of resources provided had somewhat consistent effects on resource use throughout the study. Mink in the Diversity scheme always displayed more inactive use, and had resources in the nest box more often, than the one-ball and four-ball groups. These forms of resource use were also occasionally increased over the one-ring group but were rarely increased over the four-ring group, which is potentially explained by mink in the Diversity scheme directing use preferentially toward rings. Active use showed similar differences between groups in August, but declines in active use (to negligible levels) during later collection periods caused these effects to disappear or become inconsistent. These differences between the Diversity scheme and other groups are thus unable to generate conclusions about the general impact of resource diversity on use because of the item type effects. The differences between ball groups and the Diversity scheme, but lack of difference between ring groups and the Diversity scheme, cannot therefore be explained by item diversity but instead seem more likely to reflect the presence of an object—rings—that inherently received more use. The hypothesis that resource use increases with resource diversity is therefore refuted by these results. Furthermore, similarity between the Diversity scheme and ring groups can be extended to conclude that balls do not reduce resource use; instead, balls simply elicit less resource use than rings.

The novelty of resources had fewer consistent effects than the other provision scheme characteristics. In October, mink in the Novelty scheme showed more inactive use, and had resources in the nest box more often, than mink in Baseline but never more than mink in Diversity. The latter result indicates that differences between Novelty and Baseline are caused by some factor other than the novelty of resources provided (i.e. number or diversity). Similarly, inactive use during single housing was increased in the Novelty scheme over only the one-ball group. Here, specific effects caused by the novelty of resources cannot be inferred due to confounds in the type, number, and diversity of resources provided. Furthermore, only the ball-second group, and not the ring-second group, increased resource use over Baseline and Diversity. This further supports the item order effects discussed in the previous section. Overall, the hypothesis that resource use increases with the novelty of resources is ultimately not supported by these results.
Turning to practical implications, the cumulative interpretation of provision scheme effects on resource use shows that resource use by farmed mink can be increased beyond levels expressed under current management practices (which are to provide a single, unchanging item). Specific characteristics that achieve increased resource use are still unclear, but rings seem more effective than balls. This differential effectiveness can be explained by the greater variety of behavioural opportunities throughout an animal’s life provided by rings. Furthermore, the ability for items to be employed in a variety of uses may allow them to remain viable environmental enrichments as an animal’s motivations and preferences change (an idea explored further in Chapter 4).

2.4.4 Limitations

Beyond the analytical limitations imposed by the presence of treatment effects, some item characteristics and aspects of the sampling method created challenges for this study. First, the wiffle balls used here were occasionally crushed by mink as they were used, and this occurred more frequently as mink grew larger (personal observation). If the beneficial aspect of a ball was depleted after being crushed, extinction of use may occur. This is a possible explanation for why balls received less use than rings, the latter being much sturdier. Also, after being crushed, some balls would fall through the cage floor and need to be replaced. Replacing balls whenever they are damaged or leave the cage could replenish their reinforcing characteristics, but this is not economically sustainable given the cost per ball used here (0.25 CAD). Furthermore, consistent monitoring for item presence and condition is logistically infeasible for the thousands of cages on most farms. Farmers have previously recommended small wiffle balls as environmental enrichments if they are robust enough to withstand use by mink (Meagher et al., 2014). However, this study emphasizes the need for further research to identify manufacturers that are suitable for farm purposes.

Second, multiple prototypes of the rings used here were designed before the final structure was chosen. Early models were less than two inches long and could be manipulated by mink to fall out of the cage through holes in the wire mesh floor. Increasing the ring length so that it was impossible for rings to fit over the cage wiring accounted for the ingenuity of animals to circumvent even the most well designed environmental enrichments (see Markowitz, 1982). Although rings could no longer leave the cage, they occasionally wrapped around each other due to the flexibility provided by the lengthwise cut. This effectively caused the number of rings in some cages to decrease and for subsequent double thickness rings to become inflexible. The presence of such doubled rings was not recorded, but this could alter the probability of contacting a ring if it occurred frequently enough within a provision scheme. Ring
flexibility also became an issue as mink, especially males, grew larger. In October, the farm owner became worried that rings would become stuck around males and damage them or their fur, although this was not observed during the study. Some rings were therefore stretched so that they formed an open U shape, which prevented them from being worn by mink while moving around the cage. Such alteration of rings was not recorded, but it may have caused further changes in how mink used some rings later in the study.

Third, the rate of instantaneous scan sampling may not have been ideal for recording rare behaviours, such as active use. Svendsen et al. (2007) found that stereotypic behaviour by a low stereotyping line of mink was underreported by direct scan sampling with a 30-60 minute interval between observations compared to continuous video recording. In this case, either interruption of the behaviour by the observer, which was observed during pilot observations here (see Section 2.2.1), or inaccuracy due to a longer observation interval, could explain underreporting. Other research on direct sampling methods has noted that 500 observations per individual would be required to accurately estimate behaviours that an animal spends 1% of its time exhibiting (Jauhiainen and Korhonen, 2005). The present study collected approximately 50 observations per animal for each collection period, potentially causing rare behaviours to be underreported. This could explain the potential floor effect in active use and why no active use was observed for many groups of mink in October and during single housing.

2.4.5 Implications

If the forms of resource use assessed in this chapter are relevant measures of the welfare impact of environmental enrichment, the same relationships between treatments and provision schemes that occurred for resource use should also occur for welfare indicators. The most consistent effect across forms of resource use and collection periods was for rings to receive more use than balls. Although this result does not fit within the hypotheses described in Section 2.1, it does follow the principles governing resource use described in Chapter 1. Provision scheme effects were mostly limited to experimental schemes (Number, Diversity, and Novelty) showing more resource use than Baseline, with inconsistent differences between experimental schemes preventing investigation of specific provision scheme characteristics. The inclusion of treatment effects and three experimental provision schemes resulted in a substantial number of hypothesis tests being run, which increases the probability of Type I errors. These results should therefore be replicated in more targeted research assessing each provision scheme or treatment effect individually. Still, the treatment and provision scheme effects found here offer a foundation for analysis of potential welfare effects in the next chapter.
Chapter 3: Relationships Between Resource Use and Welfare

3.1 Introduction

The previous chapter identified differences in resource use between treatments and provision schemes, but further investigation is required to determine whether this also reflects differences in welfare. In this chapter, welfare indicators previously validated for farmed mink, namely stereotypic behaviour, intraspecific aggression, fearfulness, and hyper-responsiveness to diverse stimuli, are measured to assess whether active use, inactive use, and/or the presence of items in the nest box influence welfare. The validated welfare indicators used here cover three negative affective states that are potentially exacerbated or caused by a lack of behavioural opportunity: frustration, fear, and boredom.

Stereotypic behaviour (SB) is one indicator of frustration (see Section 1.1.1), but limitations of this measure must be considered when interpreting effects. Mason and colleagues (Mason et al., 2007; 2004) separated SBs into those caused by frustration or brain dysfunction. The latter form is likely more resistant to environmental enrichment once fully developed because repetitive behaviour has become centrally controlled, and thus separated from any previous frustration, or results from broader perseveration (Mason and Latham, 2004). Because mink typically begin developing SBs at five months old, with the largest increase in performance occurring between 8-12 months (Mason, 1993), any SB exhibited by the young mink studied here is unlikely to be centrally controlled and thus still susceptible to enrichment effects. Within frustration-based SBs, increased SB in relatively barren environments, compared to those with enrichment, predominantly indicates poorer welfare. However, when assessing individual differences in welfare among animals within a barren environment, those that exhibit more SB may have better welfare than those that do not. This may be caused by repetitive motions providing a calming effect, SBs being a functional substitute for the frustrated behaviour, or inactivity being an alternative response to frustration (e.g. Mason and Latham, 2004). The present study relies on comparisons between housing conditions and also controls for overall activity; thus, groups that express more SB should have poorer welfare. Furthermore, agreement between effects on SB and other welfare indicators can help interpret the welfare significance of differences in SB (Mason et al., 2007).

Intraspecific aggression is the other indicator of frustration used here (see Section 1.4). Frustration that induces intraspecific aggression typically stems from competition over limited resources leading to resource guarding (Jacobs et al., 2018, 2016). Behavioural measures of aggression between
young mink are rarely observed and hard to differentiate from rough-and-tumble play (Dallaire and Mason, 2016; Hänninen et al., 2008b), but physiological measures can provide reliable estimates of aggressive interactions. Bite marks, small “tattoos” on the skin side of pelts (see Section 1.5.3), are caused by aggressive interactions between mink that occur during the active growth phase of the winter coat, between 5-8 months old (Bassett and Llewellyn, 1949; Hansen et al., 2014). Currently, bite marks have only been measured for group size comparisons (e.g. Axelsson et al., 2017; Hänninen et al., 2008a, 2008b), leaving potential effects of physical enrichments unexplored.

Aggression is also recorded during glove tests with mink, where a familiar, aversive stimulus (here, a mink catching glove) is presented, but determining the motivation (e.g. self-defense or predation) for this behaviour is challenging because it is directed toward an object rather than conspecific, which limits its usefulness as a welfare indicator (Meagher et al., 2011). Conversely, fearful responses during glove tests are valid indicators of persistent fear and chronic stress, as determined by consistent individual differences and decreased reproduction likelihood for fearful mink (Meagher et al., 2011). Persistent or excessive fear in captive animals can also lead to pathological anxiety (Rosen and Schulkin, 1998), deteriorated health, and lower productivity (Hemsworth, 2003), making reductions in fear beneficial for welfare, health, and economic sustainability. Environmental enrichments have been shown to decrease fearful responses in demi juvenile mink (Meagher et al., 2014), which suggests that fearfulness is a useful measure for comparing the welfare benefit of environmental enrichments in the present study.

The final welfare indicator used here is hyper-responsiveness to stimuli with diverse valence, which compares boredom between differentially housed animals. Environmental monotony has been emphasized as the main cause of boredom in humans and animals (Burn, 2017; Wemelsfelder, 2005), and connects this negative affective state to the exploration and habituation concepts discussed in Chapter 1. Environmental enrichment’s ability to create opportunities for behavioural expression in otherwise barren environments, and to thereby decrease monotony, establishes its potential for mitigating boredom. As mentioned in Section 1.5.3, hyper-responsiveness was first validated as a boredom indicator in farmed mink by showing that standard housing increased interaction time with rewarding, neutral, and aversive stimuli compared to an enriched housing condition that had previously shown improvements in other welfare measures (Meagher et al., 2017; Meagher and Mason, 2012). Applying similar boredom tests across the housing conditions used here should also reveal differences in interaction time caused by variations in monotony between provision schemes and treatments.

Two types of analyses are employed here to understand the relationship between resource use and welfare. First, the predictive value of resource use is assessed by comparing provision scheme and
treatment effects on the welfare indicators described above to the same effects on resource use found in Chapter 2 (Table 3.1). Second, the potential for resource use to explain any scheme and treatment effects on welfare is assessed. This is achieved by statistically controlling for resource use in indicator models that show scheme or treatment main effects (second hypothesis in Table 3.2). Any loss of scheme and treatment main effects in these models when controlling for resources use would suggest that the previous main effects can be explained by the amount of time animals spent using resources. Opportunistically, these models also allowed for testing the second hypothesis in Table 3.2, by revealing any negative relationships between resource use and welfare indicators. Together, the predictive and explanatory analyses described here will reveal any relationships between resource use and welfare to achieve the broader goal of this thesis: determining the efficacy of resource use as an indicator of, and explanation for, welfare improvement from environmental enrichment.
Table 3.1. Hypotheses and predictions for provision scheme and treatment effects on welfare. Hypotheses and predictions are based on the results from Chapter 2.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Predictions Based on Chapter 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schemes and treatments that increase active use also improve welfare</td>
<td>Mink with rings in Baseline, Two-items, and Number will have better welfare than those with balls in the equivalent schemes</td>
</tr>
<tr>
<td></td>
<td>Diversity mink will have better welfare than those from ball groups in Baseline and Number schemes</td>
</tr>
<tr>
<td>Schemes and treatments that increase inactive use also improve welfare</td>
<td>Mink with rings in Baseline, Two-items, and Number will have better welfare than those with balls in the equivalent schemes</td>
</tr>
<tr>
<td></td>
<td>Number mink will have better welfare than those in Baseline and Two-items schemes</td>
</tr>
<tr>
<td></td>
<td>Diversity mink will have better welfare than those from ball groups in Baseline and Number</td>
</tr>
<tr>
<td></td>
<td>Novelty mink will have better welfare than Baseline mink</td>
</tr>
<tr>
<td>Schemes and treatments that increase the presence of resources in the nest box also improve welfare</td>
<td>Mink with rings in Baseline, Two-items, and Number will have better welfare than those with balls in the equivalent schemes</td>
</tr>
<tr>
<td></td>
<td>Number mink will have better welfare than those in the Baseline scheme.</td>
</tr>
<tr>
<td></td>
<td>Novelty mink that received balls second will have better welfare than those that received rings second</td>
</tr>
<tr>
<td></td>
<td>Diversity mink will have better welfare than those from ball groups in Baseline and Number</td>
</tr>
<tr>
<td></td>
<td>Novelty mink will have better welfare than Baseline mink</td>
</tr>
</tbody>
</table>

Table 3.2. Hypotheses and predictions for relationships between resource use and welfare.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Resource use statistically accounts for scheme and treatment effects on welfare</td>
<td>Resource use will eliminate scheme and treatment effects on welfare indicators when included as a covariate, while also having a negative relationship with those indicators.</td>
</tr>
<tr>
<td>2. Resource use predicts individual differences in welfare</td>
<td>Resource use will have a negative relationship with indicators of negative affect within and across schemes and treatments.</td>
</tr>
</tbody>
</table>


### 3.2 Methods

This chapter presents welfare indicators collected from the same population of mink described in Chapter 2, which spanned four provision schemes (Baseline, Number, Diversity, and Novelty) and seven subsequent treatment groups (see Table 2.2). As noted in Section 2.2.2, three cages were removed from analysis, bringing the sample size for this study to 133 cages housing male/female pairs. Further reductions due to logistical constraints on the number of mink that could be assessed are noted in the description of relevant welfare indicators. Figure 3.1 shows an experiment timeline of when welfare indicators were measured.

![Figure 3.1. Experiment timeline for Chapter 3. A) movement to pair-housing and administration of treatments (at age 10 weeks); B) time budgets and glove tests; C) administration of second resource type to the Novelty group; D) separation of females into whelping cages; E) modified time budget assessment during single housing; F) boredom tests; and G) mink are killed as part of the farm’s standard procedure and bite marks are recorded (at age seven months).](image_url)

SB was recorded in all time budgets and aggression was recorded during pair-housing in August and October. Time budgets followed the same method described in Section 2.2.1 (see Table 2.2 for ethogram). SB was separated into three subtypes (head based, whole body, and scrabbling; see Table 2.2), based on heterogeneity of these groups suggested by Polanco et al. (2017). Descriptions of each are consistent with previous work by our group. Aggression was included here to account for instances outside the period included in bite mark analysis, with presence of screaming or sustained pursuit of a fleeing partner differentiating aggression from rough-and-tumble play (Dallaire and Mason, 2016).

Glove tests for assessing fear were conducted after time budget collections were completed in August and October. Methods were adapted from Meagher et al. (2014, 2011). A well-used mink catching
glove was brushed against the cage-front and an empty finger extended through the cage mesh (Figure 3.2). Moving the glove, as opposed to stationary presentation used in prior research, was decided upon after pilot trials showed that most mink were unresponsive to the latter method. The cage front and glove were large enough to allow both mink access to the glove. The initial reaction of each mink in the cage was categorized as: (1) fearful, if the mink retreated to, or attended to the stimulus from, the back of the cage; (2) curious, if they approached and contacted the glove without sustained biting; (3) aggressive, if they exhibited hard, sustained biting on the glove; (4) unresponsive, if they did not respond within 30 seconds of glove presentation; (5) or indifferent, if they briefly oriented toward the glove and then ignored it for the remaining time. Only fearful responses were analyzed as a welfare indicator (see Section 3.1). Unresponsive mink were retested after all other cages, and if unresponsive a second time, they were removed from analysis. Cages were tested twice on consecutive days in each collection period, starting with a different cage each day to prevent mink from being tested at the same time each day.

Figure 3.2. A catching glove being presented to pair-housed mink. Both mink would be classified as “curious” because they approached and contacted the glove without sustained biting. Photo by SD.
Boredom tests began after the third time budget was completed during the single-housing phase and were adapted from Meagher et al. (2017). Only males were tested because females no longer had their experimental resources. The stimuli used, and their associated biological relevance, were mountain lion urine (aversive; PredatorPee®), vanilla (neutral; No Name®), and pheasant scent (rewarding; D.T Systems, Inc.). These stimuli had previously been ranked 3.5, 3.5, and 1 respectively for housing effect size (enriched vs standard housing) in boredom tests, with 1 being the largest effect size of 21 stimuli tested (Meagher et al., 2017). The biological relevance of a stimulus was assumed based on similarity to natural cues, such as predator and prey species for aversive and rewarding, or lack thereof for neutral (Dunstone, 1993). The stimuli were also chosen for their uniform presentation method as odorants. Tests occurred over six days. On each day, one color type was tested with one stimulus, resulting in approximately 68 tests each day: the maximum number possible given daylight availability. The order of stimulus presentation was aversive > neutral > rewarding for all mink (i.e. two days of aversive stimulus testing, two days of neutral stimulus testing, and then two days of rewarding stimulus testing). This non-random order was chosen to reduce noise and to enhance the valence differences between stimuli (because the first stimulus presented in boredom tests can elicit fear even when otherwise rewarding: Meagher et al. 2017).

Two people conducted the boredom tests simultaneously. Stimuli were presented as 4 drops of odorant on a cotton ball, which was then placed inside a stainless steel mesh tea strainer. One person then hung the strainer on the cage front so that the mink could interact with it but not pull it or the cotton ball inside the cage (Figure 3.3). Each person held a stopwatch to measure either time in contact with or time oriented toward the stimulus. Any manipulation of the tea strainer with the snout or front paws was scored as contact. Orientation was recorded any time the mink directed attention toward the stimulus from anywhere in the cage, such as looking at the tea strainer from the back of the cage. These states are not mutually exclusive, so any time contacting the stimulus was included in time orienting toward the stimulus. Latency to contact was not recorded because it is an insensitive measure of environmental enrichment effects (Meagher et al., 2017).

Stimuli were presented for five minutes. Each trial used a new cotton ball and the strainer was sprayed with 70% ethanol solution, left to dry, and wiped between tests. When left to evaporate, ethanol can remove oily odorants deposited on the strainer, which ensures that no mink cues remain between tests (Leo and Pamplona, 2014). Two strainers were alternated for each odorant so that one could dry while the other was in use. Individual strainers were also used exclusively with one stimulus type to prevent cross contamination. If a mink was asleep and unresponsive at the start of a trial, he was skipped and returned
to once awake. This allowed all males to be tested and included in analysis. If all cages of a colour type could not be tested on the same day, the remaining cages were tested first the following morning.

Figure 3.3. A pastel male interacting with an odorant filled tea strainer. A keyring and carabiner securely attached the tea strainer to the top of the cage. Photo by SD.

In late November, all mink were killed using carbon monoxide, as is normal farm practice. Scheduling conflicts between when female pelts were being processed and when males would be killed prevented both sexes from being examined. Female pelts were prioritized because they typically show more bite marks than males (Hansen et al., 2014). Females from each treatment group were thus killed in a single batch and subsequently given matching, treatment-specific nose tags (but individual identification was no longer possible). Carcasses then passed through the normal pelting procedure until they would be attached to boards for drying prior to sale. Here, the pelt was cleaned with a stiff brush to remove any wood shavings (as used to partially dry pelts) and two photographs of the leather side (dorsal and ventral) were taken (Figure 3.4). The number of bite marks, identified as small, black dots on the leather side of the pelt, were later counted using ImageJ software (Schneider et al., 2012). Only bite marks between the
forelimbs and caudal end of the pelt were included in the score because the area between the forelimbs and cranial end was obscured by remnant tissue. Marks in the latter area are also typically associated with sexual behaviour in mink, which was not of interest for this study (Hansen et al., 2014). ID labels were obscured in the pictures so that the counter (SD) was blind to treatment. Any instances of scarring or other pelt damage were also noted.

![Figure 3.4. The dorsal side of a female pelt with the leather side out.](image)

Pelts were placed on a whiteboard with an identification code that was later replaced with a neutral ID number to blind the scorer to treatment. The black bracket indicates the area of the pelt where bite marks were recorded. Bite marks can be seen near the caudal end of the pelt (right) as black dots. Photo by SD.

### 3.2.1 Statistical analyses

SB and fearfulness were measured as proportions of observations. Too little aggression occurred during time budget observations for analysis, which is consistent with previous research with young mink (Dallaire and Mason, 2016; MacLennan and Bailey, 1969). SB is frequently confounded with total active time and thus analysed as a proportion of active observations (Garner and Mason, 2002). Instead, this behaviour was modelled here as a proportion of total observations and active observations was included.
as a covariate \([SB = \text{scheme/treatment} * \text{sex} * \text{colour} + \text{active} + (1|\text{cage})]\). This accounts for the potential confound while maintaining the original structure of the response variable (Lazic, 2016, p. 166). The three forms of SB recorded were combined into one score for analysis due to scarcity of these behaviours. Unresponsiveness was rare during glove tests: two mink were removed from analysis in August (female, ring-second; male, four rings) and one in October (male, diversity). The basic model for resource use in Chapter 2 was used for fear responses \([\text{fear} = \text{scheme/treatment} * \text{sex} * \text{colour} + (1|\text{cage})]\).

Fearfulness and SB were both analysed via generalized linear mixed models (GLMM) with binomial distributions. Generalized linear models were implemented whenever the sex and random cage factors were removed from the model. Split-half analysis for SB and fearfulness was completed prior to hypothesis testing to assess behavioural stability and data quality, again with even day values included as a covariate (Martin and Bateson, 2007). Because only two glove tests were performed for each mink in a collection period, split-half models now had a binary dependent variable and covariate for the presence or absence of a fearful response. As with active use in Chapter 2, rarity of SB required the even day covariate to be included as a binary variable for presence of the behaviour.

Orientation toward the stimulus and stimulus contact during boredom tests were modelled via linear mixed models with repeated measures for each mink \([\text{response} = \text{scheme/treatment} * \text{colour} * \text{stimulus} + (1|\text{mink ID})]\). No sex or cage factors were included because only male mink were tested.

Bite mark data exhibited substantial zero-inflation and were therefore modeled with zero-inflated negative binomial models using the zinbivfl function in the “pscl” package (Zeileis et al., 2008). The negative binomial distribution was chosen because a likelihood ratio test between models fitted with negative binomial or Poisson distributions showed that overdispersion in the model was not caused solely by an over-abundance of zeros \((X^2(1) = 493.16, p < 0.001;\) Zuur et al., 2009). The zero-inflated models would not run when the nested treatment structure of the experimental design was included, so separate planned contrasts for Number, Diversity, and Novelty and a post hoc model for item order effects were run (item type effects were included in the Number model, see Section 2.2.2). These contrasts were analyzed with Type II Wald Chi-square test for the count portion of the model. The binomial portion had issues with separation that could not be resolved and is not reported.

Least square means are reported with the associated standard error. Arithmetic means are presented in some instances where model convergence was an issue. The planned contrasts for scheme
effects and post hoc models for treatment effects described in Section 2.2.2 were again employed here when investigating these effects on welfare. The complete set of planned contrasts was used because the predictions in Table 3.1 included most scheme and treatment comparisons. Further models investigated whether resource use explains scheme and treatment effects on welfare by including resource use covariates in models that previously showed such main effects. If scheme and treatment effects were statistically eliminated ($p > 0.05$) when the resource use covariate was included, and a covariate main effect also occurred, that form of resource use sufficiently explained the previous scheme and treatment effects (see Nip et al., 2019 for similar analysis). Opportunistically, these models also assessed whether the degree of resource use predicted levels of welfare indicators within and between groups. Lack of individual identification for bite mark scores prevented inclusion of resource use covariates. Therefore, this welfare indicator was not included in this analysis. These models were repeated for all forms of resource use (active, inactive, nest box presence).

### 3.3 Results

As in Chapter 2, a short paragraph summarizes key outcomes after reporting the results within a collection period (e.g. stereotypic behaviour in August). Each section also includes a graphical representation of the results, with asterisks denoting P values (* $< 0.05$; ** $< 0.01$; *** $< 0.001$). Additional tables at the end of each section compare the results to the predictions outlined in Table 3.1. A concluding paragraph in each section also summarizes the results for each welfare indicator across collection periods.

#### 3.3.1 Behavioural stability

Stability results for SB in time budgets and fearful responses during glove tests are presented in Table 3.3. SB became stable in more mink as the study progressed and prevalence of SB increased. Few mink in August performed SB, which prevented analysis for some groups. Furthermore, SB stability in August for demi females in the Two-item scheme is likely an artifact of SB rarity (see Section 3.3.2). Fearful responses showed the opposite change over time, becoming less stable in more mink in October compared to August. This is likely associated with the decrease in fearful responses as the study progressed (see Section 3.3.3).
Table 3.3. Behavioural stability of stereotypic behaviour and fearful responses. N/A values in the Not Stable column represent models where the behaviour did not occur on either odd or even days and analysis could not be run.

<table>
<thead>
<tr>
<th>Collection Period</th>
<th>Behaviour</th>
<th>Stable</th>
<th>Not Stable</th>
</tr>
</thead>
</table>
| August            | Stereotypic behaviour | **D Two-item Females**  \(X^2(1) = 4.23, p = 0.04\)              | **Males N/A**  
                      |                    | **Other Females N/A**                                      |                                                                            |
|                   | Fearfulness        | **D Baseline, 4-ball, and Two-item Males**  \(X^2(1) = 6.45, p < 0.011\) | **Other Males**  \(X^2(1) = 1.81, p > 0.18\)                                |
|                   |                    | **P 4-ball, Diversity, and Two-item Males**  \(X^2(1) = 7.27, p = 0.007\) |                                                                            |
|                   |                    | **Females**  \(X^2(1) = 30.2, p < 0.0001\)                            |                                                                            |
| October           | Stereotypic behaviour | **D Ring-2\textsuperscript{nd} Females**  \(X^2(1) = 8.43, p = 0.004\) | **Males**  \(X^2(1) = 2.94, p = 0.086\)                                      |
|                   |                    | **P Females**  \(X^2(1) = 15.0, p = 0.0001\)                          | **Other D Females N/A**                                                    |
|                   | Fearfulness        | **D Novelty Males**  \(X^2(1) = 6.19, p < 0.013\)                    | **Other Males**  \(X^2(1) = 3.51, p > 0.061\)                                |
|                   |                    | **D Females**  \(X^2(1) = 16.3, p < 0.0001\)                         | **P Females**  \(X^2(1) = 3.73, p > 0.053\)                                |
| Single-housing    | Stereotypic behaviour | **D Novelty Males**  \(X^2(1) = 4.18, p = 0.041\)                    | **Other D Males**  \(X^2(1) = 0.89, p > 0.34\)                              |
| (males only)      |                    | **P Males**  \(X^2(1) = 10.1, p = 0.0015\)                           |                                                                            |

3.3.2 Scheme and treatment effects on stereotypic behaviour

**August**

The arithmetic mean proportion of observations performing SB for those mink that exhibited it was 0.025 ± 0.006 for females and 0.024 ± 0.005 for males. Too few mink – only three females and four males – exhibited SB in August for analysis. One female was pastel and the other two were demi. The pastel and one demi female were in the two-ring group, while the remaining demi female was in the four-ring group. Three males were from the pastel colour-type and one was demi. Two pastel males were in the Diverse scheme and the remaining pastel and demi males were in the four-ring group.
October

SB was now seen in 14 females and 15 males. Eight females were pastel and six were demi. Eleven males were pastel and four were demi. The arithmetic mean proportion of time exhibiting SB, of those who exhibited it at all, was $0.04 \pm 0.007$ for females and $0.02 \pm 0.003$ for males. The full model for October SB did not converge, so separate models were run for each sex.

The female model was further separated by colour-type because it also did not converge. Demi females showed scheme and treatment effects (scheme: $X^2(3) = 14.31, p = 0.0025$; treatment: $X^2(3) = 14.34, p = 0.0025$) (Figure 3.5). The Number scheme did not differ from Baseline ($X^2(1) = 0.41, p = 0.52$), and no item type effect occurred ($X^2(1) = 1.51, p = 0.22$). The Diversity scheme also did not differ from Baseline ($X^2(1) = 0.13, p = 0.72$). There was a treatment effect when comparing Novelty to Baseline ($X^2(2) = 13.48, p = 0.0012$), where the ring-second group showed more SB than ball-second ($X^2(1) = 12.44, p = 0.0004$). Further models comparing Novelty and Baseline were therefore split by Novelty treatment. The ring-second Novelty group showed more SB than Baseline ($X^2(1) = 14.01, p = 0.0002$), while there was no difference between the ball-second group and Baseline ($X^2(1) = 0.80, p = 0.37$).

Pastel females showed no scheme effects ($X^2(3) = 5.29, p = 0.15$), but they did show a treatment effect ($X^2(3) = 10.55, p = 0.014$) (Figure 3.5). Ball groups showed more SB than ring groups when investigating item type effects in the Baseline and Number schemes ($X^2(1) = 6.55, p = 0.01$), and there was no item order effect within the Novelty scheme ($X^2(1) = 3.53, p = 0.06$). Because the Baseline scheme could no longer be considered a homogenous group, separate models were run comparing Baseline ball and ring groups to Diversity and Novelty. Neither Diversity nor Novelty mink differed from the one-ball group (Diversity: $X^2(1) = 1.64, p = 0.2$; Novelty: $X^2(1) = 1.06, p = 0.3$). Conversely, both Diversity and Novelty increased SB over the one-ring group (Diversity: $X^2(1) = 5.4, p = 0.02$; Novelty: $X^2(1) = 5.55, p = 0.02$).
The complete male model had a treatment*color interaction ($X^2(3) = 6.69, p = 0.082$), so data were split by colour-type (Figure 3.6). Unlike demi females, there were neither scheme nor treatment effects in demi males (scheme: $X^2(3) = 3.07, p = 0.38$; treatment: $X^2(3) = 1.42, p = 0.70$). However, both scheme and treatment effects occurred for pastel males (scheme: $X^2(3) = 9.22, p = 0.027$; treatment: $X^2(3) = 9.53, p = 0.023$). The following analysis therefore refers to only pastel males.

Number mink did not differ from Baseline mink ($X^2(1) = 0.48, p = 0.49$), but mink in ball treatments showed less SB than mink in ring treatments in both schemes ($X^2(1) = 6.49, p = 0.01$). The Baseline scheme therefore could no longer be considered a homogenous group, so further comparison of Diversity and Novelty schemes to Baseline were split into separate models for each Baseline group. The Diversity scheme mink unexpectedly showed more SB than the one-ball group ($X^2(1) = 6.53, p = 0.011$), but did not differ from one ring ($X^2(1) = 0.02, p = 0.88$). Finally, Novelty mink did not differ from one ball mink ($X^2(1) = 0.67, p = 0.41$), but did show less SB than one ring mink ($X^2(1) = 10.44, p = 0.0012$). There was no difference between the Novelty groups in either of these models ($X^2(1) = 1.21, p = 0.27$).
Because Novelty reduced SB compared to the Baseline ring group, it was then compared to Diversity and subsequently showed less SB than the Diversity ($X^2(1) = 8.07, p = 0.0045$).

Figure 3.6. October stereotypic behaviour in male mink split by colour type. Stereotypic behaviour is measured as a proportion of observations. Error bars show 95% confidence intervals around the LSM.

In summary, the proportion of observations performing SB, amongst those mink that displayed any, was similar to August, but the number of mink exhibiting SB did increase. No scheme reduced SB compared to Baseline in pastel females, but the resource type provided in Baseline and Number schemes did impact SB, with balls being associated with more SB than rings. In pastel males, providing more items again did not reduce SB, and only resource type effects were found. However, balls were now associated with less SB than rings. Providing two balls and rings in the Diverse scheme did not reduce SB in pastel males (even though this included two balls), and even increased it compared to the one-ball group. The only scheme that seemed to benefit pastel males was Novelty, at least when compared to the one-ring group (but not one-ball). Provision scheme effects in demi mink were sparse: only the ring-
second group showed more SB than Baseline in demi females. No effects of scheme or treatment were found in demi males.

Single housing

The arithmetic mean proportion of observations performing SB for those who performed any was 0.03 ± 0.003 for male mink. Twenty-two males now exhibited SB. Males showed a treatment effect \( \chi^2(3) = 11.4, p = 0.0097 \), but no difference between schemes \( \chi^2(3) = 7.45, p = 0.059 \) (Figure 3.7). When investigating item type effects between Baseline and Number, males showed a scheme*colour interaction \( \chi^2(1) = 6.27, p = 0.012 \). In demi males, neither treatment \( \chi^2(1) = 2.40, p = 0.12 \) nor scheme \( \chi^2(1) = 0.022, p = 0.88 \) effects occurred. Conversely, pastel males with the Number scheme showed less SB than those with Baseline \( \chi^2(1) = 11.1, p = 0.0009 \), and there was again no treatment effect \( \chi^2(1) = 0.05, p = 0.82 \). Comparisons between Diversity and Baseline were not run due to the lack of item type effects in the previous models. Comparisons between Novelty groups with all males revealed that the ring-second group showed more SB than ball-second \( \chi^2(1) = 9.40, p = 0.0022 \). The ring-second group did not differ from Baseline \( \chi^2(1) = 0.23, p = 0.63 \), while ball-second did reduce SB compared to Baseline and Diversity schemes (Baseline: \( \chi^2(1) = 15.4, p < 0.0001 \); Diversity: \( \chi^2(1) = 5.62, p = 0.018 \).
The only scheme effect during single housing was thus for pastel males in the Number scheme to show less SB than those in Baseline, but this effect was not seen in demi males. There were no item type effects, but an order effect between Novelty groups did occur. The ring-second group showed more SB than ball-second and subsequent analysis showed that the ball-second group displayed less SB than both Baseline and Diversity schemes.

**Stereotypic Behaviour Summary**

Prevalence of SB was low in all observation periods. Table 3.4 presents a summary of scheme and treatment effects on SB. Provision scheme effects were sparse and inconsistent across observation periods. In October, the only potentially beneficial scheme was Novelty in pastel males, but this is countered by increased SB for demi females in the Novelty scheme. The beneficial Novelty effect was also shown by males in single housing, albeit only for those that received rings second. Treatment effects were similarly rare. In October, pastel males and females showed opposite effects of resource type and the
effect in males was not present after females were moved. Item order effects were more consistent, with ring-second groups performing more SB than ball-second for demi females in October and all males after female removal.

Table 3.4. Evaluation of predictions for stereotypic behaviour separated by mink sex and colour. Values within a cell represent whether the prediction was met in a collection period, with collection periods ordered chronologically and separated by bars. Some predictions were not tested in all collection periods (e.g. Two-items comparisons were only possible in August, and Novelty comparisons were only possible in October and single housing). Cell colour indicates consistency of predicted effects across collection periods. The order of consistency from lowest to highest follows darkest to lightest. X = no effect. ✓ = predicted effect. * = effects that were complicated by treatment interactions. “>” = opposite effect found.

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Demi Females</th>
<th>Demi Males</th>
<th>Pastel Females</th>
<th>Pastel Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number &lt; Baseline</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2 items &lt; Baseline</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Number &lt; 2 items</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Diversity &lt; Baseline</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Diversity &lt; Number</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Novelty &lt; Baseline</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Novelty &lt; Diversity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Treatment Effects

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Demi Females</th>
<th>Demi Males</th>
<th>Pastel Females</th>
<th>Pastel Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rings &lt; Balls</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ball 2nd &lt; Ring 2nd</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

3.3.3 Scheme and treatment effects on fearfulness

August

The LSM proportion of fearful responses was 0.7 ± 0.056 for females and 0.45 ± 0.064 for males. The full model showed a scheme*sex*colour interaction ($X^2(3)=6.76, p = 0.08$), so further models were first split by sex to interpret this interaction. No scheme or treatment effects were found in the female model (scheme: $X^2(3)=1.51, p = 0.68$; treatment: $X^2(3)=5.03, p = 0.17$), but the male model showed
treatment*colour ($X^2(3) = 10.42, p = 0.015$) and scheme*colour interactions ($X^2(3) = 7.07, p = 0.07$) (Figure 3.8).

When split by colour, demi males showed a treatment effect ($X^2(3) = 7.85, p = 0.05$) and no scheme effect ($X^2(3) = 1.12, p = 0.77$). Neither scheme nor item type effects occurred when comparing ball and ring treatments in Baseline, Number and Two-items (scheme $X^2(2) = 0.69, p = 0.71$; treatment $X^2(1) = 2.91, p = 0.088$). Comparisons with Diversity were not run due to the lack of scheme effects in the full model. Conversely, pastel males did show a scheme effect ($X^2(3) = 8.24, p = 0.041$) but had no treatment effect ($X^2(3) = 4.77, p = 0.19$). When comparing Baseline, Number, and Two-items, there was a scheme effect ($X^2(2) = 7.81, p = 0.02$) and again no treatment effect ($X^2(1) = 2.42, p = 0.12$). Number reduced fearfulness compared to Baseline ($X^2(1) = 8.00, p = 0.0047$) but not Two-items ($X^2(1) = 1.42, p = 0.23$), and there was no difference between Two-items and Baseline ($X^2(1) = 2.57, p = 0.11$). Ring treatments also reduced fearful responses compared to balls in the Baseline-Number comparison ($X^2(1) = 4.07, p = 0.044$). However, this effect did not occur in any other model, meaning that two rings did not differ from two balls. Lastly, Diversity did not differ from either Baseline group ($X^2(1) < 0.99, p > 0.32$).

**Figure 3.8.** August fearful responses for male mink split by colour type. Fearful responses are measured as a proportion of observations. Error bars show 95% confidence intervals around the LSM.
In summary, scheme and treatment effects were scarce and inconsistent for fearful responses in August. No scheme or treatment reduced fearful responses in females. The Number scheme may have benefited pastel males and having rings occasionally reduced fearful responses compared to balls for these mink.

October

The arithmetic mean proportion of fearful responses was $0.20 \pm 0.03$ for females and $0.10 \pm 0.02$ for males. The full model failed to converge so males and females were modelled separately. There were again no scheme or treatment effects in the female model (scheme: $X^2(3) = 5.74, p = 0.13$; treatment: $X^2(3) = 2.53, p = 0.47$). The full male model had issues with separation, so it was split again by colour (Figure 3.9).

Demi males again showed no scheme or treatment effects (scheme: $X^2(3) = 2.48, p = 0.48$; treatment: $X^2(3) = 6.40, p = 0.09$). Pastel males similarly showed no treatment effect ($X^2(3) = 4.30, p = 0.23$), but they did show a scheme effect ($X^2(3) = 12.55, p = 0.0057$). The model comparing Number and Baseline schemes had issues with separation and further models for this comparison were not run (only two males in the Baseline ball group showed fearful responses). Pastel males with Diversity unexpectedly showed more fearful responses than Baseline and Number (Baseline: $X^2(1) = 4.26, p = 0.039$; Number: $X^2(1) = 11.9, p = 0.0006$), and those with the Novelty scheme did not differ from Baseline ($X^2(1) = 0.63, p = 0.43$).
Overall, scheme effects were again scarce, and no treatment effects were found. No scheme was able to reduce fearfulness in any group of mink. Surprisingly, the only effect seen was for Diversity to show more fearful responses than Baseline and Number in pastel males. This may reflect mink in the Diverse scheme not reducing their fearfulness over time as much as the other schemes, given that this effect was not seen in August. The relatively low number of fearful responses in October may also impact the presence of this odd effect.

Fearfulness Summary

The prevalence of fearful responses amongst all mink decreased substantially from August to October. Most mink who exhibited fearful responses in August switched to curious responses in October. Table 3.5 presents a summary of scheme and treatment effects on fearfulness. The only predicted scheme and treatment effects on fearfulness occurred in August for pastel males. Providing four items reduced fearfulness, and rings also reduced fearfulness compared to balls, although the latter effect did not apply to mink with two items. The only effects seen in October also occurred in pastel males and were opposite...
of the predicted relationship, but these effects may be explained by a floor effect caused by the reduced fearfulness in all mink.

Table 3.5. Evaluation of predictions for fearfulness separated by mink sex and colour. Values within a cell represent whether the prediction was met in a collection period, with collection periods ordered chronologically and separated by bars. Some predictions were not tested in all collection periods (e.g. Two-items comparisons were only possible in August, and Novelty comparisons were only possible in October and single housing). Cell colour indicates consistency of predicted effects across collection periods. The order of consistency from lowest to highest follows darkest to lightest. X = no effect. ✓ = predicted effect. “>” = opposite effect found. * = effects that were complicated by treatment or scheme interactions.

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Demi Females</th>
<th>Demi Males</th>
<th>Pastel Females</th>
<th>Pastel Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number &lt; Baseline</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>2 items &lt; Baseline</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Number &lt; 2 items</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Diversity &lt; Baseline</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Diversity &lt; Number</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Novelty &lt; Baseline</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Novelty &lt; Diversity</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Treatment Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rings &lt; Balls</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>✓*</td>
</tr>
<tr>
<td>Ball 2nd &lt; Ring 2nd</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

3.3.4 Scheme and treatment effects on boredom

The LSM time spent in contact with the negative, ambiguous, or rewarding stimulus was 170 ± 3.42s, 161 ± 3.42s, and 156 ± 3.42s, respectively. A treatment*colour*stimulus interaction occurred when modelling contact with stimuli ($F_{6, 238} = 2.03, p = 0.06$). Subsequent models were split by stimulus type due to a colour*stimulus interaction ($F_{2, 238} = 20.73, p < 0.0001$) and stimulus main effect ($F_{2, 238} = 5.75, p = 0.0036$). However, no further treatment or scheme effects on contact time could be detected for any stimulus type (scheme: $F_{3, 119} < 1.57, p > 0.2$; treatment: $F_{3, 119} < 0.86, p > 0.46$) (Figure 3.10).
Figure 3.10. Time in contact with the stimulus during boredom tests. Error bars show 95% confidence intervals around the LSM. There were no scheme or treatment effects.

The LSM time spent oriented toward the negative, ambiguous, or rewarding stimulus was 240 ± 3.14s, 206 ± 3.14s, and 200 ± 3.14s, respectively. A scheme*stimulus interaction occurred when modelling orientation time ($F_{6, 238} = 2.01, p = 0.065$). Subsequent models split by stimulus type did not, however, show scheme or treatment effects on orientation time (scheme: $F_{3, 119} < 1.91, p > 0.13$; treatment: $F_{3, 119} < 2.16, p > 0.096$) (Figure 3.11). In summary, none of the predicted scheme or treatment effects occurred during boredom tests.
3.3.5 Scheme and treatment effects on bite marks

The arithmetic mean number of bite marks on pelts across schemes was 6.25 ± 0.82. This number refers to the number of individual marks because marks did not always occur in pairs. Schemes did not differ in their number of bite marks ($X^2(1) < 2.14, p > 0.14$), and neither item type nor order effects occurred (type: $X^2(1) = 2.71, p = 0.1$; order: $X^2(1) = 1.08, p = 0.3$). Figure 3.12 presents the similarities in pelt scores across schemes and treatments. Some pelts also exhibited scarring potentially caused by minor dermatophytosis (ringworm infection) (Overy et al., 2015).
Figure 3.12. Number of bite marks present on pelts. Raw means are presented due to limitations of zero-inflated modeling (See Section 3.2.1). Error bars show 95% confidence intervals. There were no scheme or treatment effects.

3.3.6 Resource use as a welfare covariate

The previous results for welfare indicators revealed 15 models that showed provision scheme or treatment main effects, with 13 for stereotypic behaviour and two for fearfulness. When active use, inactive use, or presence of items in the nest box was included as a covariate, only two models showed resource use effects (Table 3.6). First, the Diversity and Baseline comparison for pastel males in October showed a negative relationship between inactive use and SB ($X^2(1) = 3.89, p = 0.048$) (Figure 3.13A). Second, the Baseline and Number comparison for pastel males during single housing showed an unexpectedly positive relationship between active use and SB ($X^2(1) = 5.45, p = 0.02$) (Figure 3.13B). Provision scheme effects remained in both models, however, meaning these apparent resource use effects did not statistically explain previous scheme effects (see Table 3.6). Four models for fearfulness lost scheme or treatment effects when resources use was included, but these effects were already marginal and were not explained by resource use, due to no relationship between resource use and welfare.
Figure 3.13. Resource use covariate effects. A) Inactive use and SB for Baseline and Diversity pastel males in October. B) Active use and SB for Baseline and Number pastel males during single housing. Stereotypic behaviour is measured as a proportion of observations. Error bars show 95% confidence intervals.
Table 3.6. Summary of scheme and treatment effect changes when resource use is included in models. Resource use columns show relevant scheme and treatment effects when the respective form of use is included in the model. S: scheme effect. T: treatment effect. Bold effects indicate models where resource use predicted individual differences in welfare. Italic effects indicate models where scheme or treatment effects no longer occurred when resource use was included.

<table>
<thead>
<tr>
<th>Original Model</th>
<th>Original Effect</th>
<th>Active Use Added</th>
<th>Inactive Use Added</th>
<th>Resources in Nest Box Added</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>October SB for pastel males with</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline and Number</td>
<td>T: $X^2(1) = 6.49$, $p = 0.011$</td>
<td>T: $X^2(1) = 7.44$, $p = 0.0064$</td>
<td>T: $X^2(1) = 5.57$, $p = 0.018$</td>
<td>T: $X^2(1) = 6.57$, $p = 0.01$</td>
</tr>
<tr>
<td>One ball and Diversity</td>
<td>S: $X^2(1) = 6.53$, $p = 0.011$</td>
<td>S: $X^2(1) = 8.07$, $p = 0.0045$</td>
<td><strong>S: $X^2(1) = 9.67$, $p = 0.0019$</strong></td>
<td>S: $X^2(1) = 9.01$, $p = 0.0027$</td>
</tr>
<tr>
<td>One ring and Novelty</td>
<td>S: $X^2(1) = 10.4$, $p = 0.0012$</td>
<td>S: $X^2(1) = 9.97$, $p = 0.0016$</td>
<td>S: $X^2(1) = 11.5$, $p = 0.0007$</td>
<td>S: $X^2(1) = 7.84$, $p = 0.0051$</td>
</tr>
<tr>
<td>Diversity and Novelty</td>
<td>S: $X^2(1) = 8.07$, $p = 0.0045$</td>
<td>S: $X^2(1) = 7.72$, $p = 0.0055$</td>
<td>S: $X^2(1) = 9.58$, $p = 0.002$</td>
<td>S: $X^2(1) = 9.21$, $p = 0.0024$</td>
</tr>
<tr>
<td><strong>October SB for pastel females with</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline and Number</td>
<td>T: $X^2(1) = 6.54$, $p = 0.01$</td>
<td>T: $X^2(1) = 5.58$, $p = 0.018$</td>
<td>T: Does not converge</td>
<td>T: Does not converge</td>
</tr>
<tr>
<td>One ring and Diversity</td>
<td>S: $X^2(1) = 5.40$, $p = 0.02$</td>
<td>S: Separation</td>
<td>S: $X^2(1) = 5.56$, $p = 0.02$</td>
<td>S: $X^2(1) = 6.04$, $p = 0.014$</td>
</tr>
<tr>
<td>One ring and Novelty</td>
<td>S: $X^2(1) = 5.55$, $p = 0.019$</td>
<td>S: Separation</td>
<td>S: $X^2(1) = 4.20$, $p = 0.04$</td>
<td>S: $X^2(1) = 3.98$, $p = 0.046$</td>
</tr>
<tr>
<td><strong>October SB for demi females with</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ring-2nd and Ball-2nd</td>
<td>T: $X^2(1) = 12.4$, $p = 0.0004$</td>
<td>T: $X^2(1) = 11.9$, $p = 0.0006$</td>
<td>T: $X^2(1) = 12.2$, $p = 0.0005$</td>
<td>T: $X^2(1) = 11.0$, $p = 0.0009$</td>
</tr>
<tr>
<td>Baseline and Ring-2nd</td>
<td>S: $X^2(1) = 14.0$, $p = 0.0002$</td>
<td>S: $X^2(1) = 14.0$, $p = 0.0002$</td>
<td>S: $X^2(1) = 9.69$, $p = 0.0019$</td>
<td>S: $X^2(1) = 11.8$, $p = 0.0006$</td>
</tr>
<tr>
<td><strong>Single housing SB for pastel males with</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline and Number</td>
<td>S: $X^2(1) = 11.1$, $p = 0.0009$</td>
<td><strong>S: $X^2(1) = 13.5$, $p = 0.0002$</strong></td>
<td>S: $X^2(1) = 12.9$, $p = 0.0003$</td>
<td>S: $X^2(1) = 8.00$, $p = 0.0047$</td>
</tr>
<tr>
<td><strong>Single housing SB for males with</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ring-2nd and Ball-2nd</td>
<td>T: $X^2(1) = 9.40$, $p = 0.0022$</td>
<td>T: Separation</td>
<td>T: Separation</td>
<td>T: $X^2(1) = 10.1$, $p = 0.0015$</td>
</tr>
<tr>
<td>Baseline and Ball-2nd</td>
<td>S: $X^2(1) = 15.4$, $p &lt; 0.0001$</td>
<td>S: Separation</td>
<td>S: $X^2(1) = 16.3$, $p &lt; 0.0001$</td>
<td>S: $X^2(1) = 9.20$, $p = 0.0024$</td>
</tr>
<tr>
<td>Diversity and Ball-2nd</td>
<td>S: $X^2(1) = 5.62$, $p = 0.018$</td>
<td>S: Separation</td>
<td>S: $X^2(1) = 4.84$, $p = 0.028$</td>
<td>S: $X^2(1) = 5.11$, $p = 0.024$</td>
</tr>
<tr>
<td><strong>Fearfulness for pastel males with</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline and Number in August</td>
<td>S: $X^2(1) = 8.00$, $p = 0.0047$</td>
<td>S: $X^2(1) = 8.68$, $p = 0.0032$</td>
<td>S: $X^2(1) = 8.61$, $p = 0.0033$</td>
<td>S: $X^2(1) = 7.91$, $p = 0.0049$</td>
</tr>
<tr>
<td>T: $X^2(1) = 4.07$, $p = 0.044$</td>
<td>T: $X^2(1) = 6.01$, $p = 0.014$</td>
<td>T: $X^2(1) = 3.80$, $p = 0.051$</td>
<td>T: $X^2(1) = 3.51$, $p = 0.061$</td>
<td></td>
</tr>
<tr>
<td>Baseline and Diversity in October</td>
<td>S: $X^2(1) = 4.26$, $p = 0.039$</td>
<td>S: $X^2(1) = 4.40$, $p = 0.036$</td>
<td>S: $X^2(1) = 3.29$, $p = 0.07$</td>
<td>S: $X^2(1) = 2.76$, $p = 0.097$</td>
</tr>
</tbody>
</table>
3.4 Discussion

3.4.1 Treatment and provision scheme effects on welfare

Stereotypic behaviour and fearfulness data quality for many mink were limited by instability of these behaviours within collection periods. As with active use in Chapter 2, results from groups of mink with poor data quality should therefore be interpreted hesitantly: there were treatment and scheme effects, as discussed below, but all need replicating. Boredom tests and bite mark scores, in contrast, revealed no differences in welfare between schemes or treatments. This may indicate that these measures were too insensitive to detect any differences between groups, or that schemes and treatments genuinely had little impact on welfare. Potential explanations for the instability of some behaviour data, and for the lack of convincing scheme and treatment effects, are returned to in Section 3.4.3.

Treatment effects on welfare were less frequent than the treatment effects on resource use found in Chapter 2. Fearfulness for pastel males in August and SB for pastel females in October were the only two instances where ring groups showed improved welfare over ball groups. However, the former effect is limited: first, the item type effect did not occur for pastel males in the Two-item group in August, and second, for pastel males in October, it is countered by ring groups possibly having poorer welfare than ball groups in terms of stereotypic behaviour. Further, many mink who showed an item type effect on resource use did not also show that effect for welfare; any effects that did occur were inconsistent across collection periods. These results provide little overall support for the prediction that ring groups would have better welfare than ball groups and applies to all three hypotheses on links between resource use and welfare (Table 3.1).

Item order effects were similarly rare. Only in demi females in October and all males when single housed did ball-second mink have better welfare than ring-second mink in terms of SB. And although no group of mink showed the opposite effect, some groups of mink with an item order effect on resource use again did not show the same effect on welfare. These results provide tentative support for the prediction that mink receiving balls-second have better welfare than those receiving rings-second, which applies to the hypothesis that storing resources in the nest box predicts welfare improvement (Table 3.1).

Provision schemes differed in welfare measures more often than did treatments, albeit still less often than their effects on resource use. The Number scheme improved welfare over Baseline only for
pastel males in terms of fearfulness in October and SB during single housing. However, the Number scheme did not improve welfare over Two-items, nor did Two-items ever improve welfare over Baseline. It is interesting that Number effects on SB occurred during single housing, when competition over resources was no longer a factor. Because SBs were only developing in the young mink observed here, it is possible that reduced competition early in life prevented later SB development. However, the possibility that male mink received greater benefit from having more items cannot be eliminated, and neither can the explanation that this is a Type I error. These results are thus again sparse in comparison to the item number effects shown for inactive use and therefore offer little support for the prediction that Number mink should have better welfare than mink with Baseline and Two-items. These results do, however, provide limited support for the prediction that Number mink would show improved welfare over Baseline based on previous results for resource presence in the nest box.

Diversity effects on welfare only occurred in October and were always opposite to the predictions based on Chapter 2. These effects only involved pastel mink, predominantly males, and were occasionally limited by item type differences in the Baseline scheme. Consistency within collection period, sex, and colour type and between SB and fearfulness may lend credibility to some aspect of item diversity reducing welfare compared to other schemes, but limitations of the welfare indicators used here could also be responsible. Many of the comparisons that showed Diversity decreasing welfare included Baseline or Number groups that had a zero score for the welfare indicator. In these cases, the rarity of relevant behaviours could overestimate the difference between groups, especially considering that males in October had poor data quality for SB and fearfulness. Even with this caveat, the results for the Diversity scheme do not support predictions based on active use, inactive use, or presence of items in the nest box.

Turning to the impact of novelty, sex and colour type interactions generated inconsistent effects when comparing SB between Novelty and Baseline schemes, and no other welfare indicators showed Novelty effects. Pastel males in October and all males during single housing were the only groups of mink where Novelty treatments occasionally shared reduced SB. Conversely, female mink in October in the Novelty scheme sometimes showed more SB than Baseline. The rarity and variability of these effects across collection periods and welfare indicators prevents the Novelty predictions based on inactive use and presence of items in the nest box from being met.

In all cases, provision scheme and treatment effects on welfare were thus less frequent than similar effects on resource use. Hypotheses on similarities between effects on active/inactive use and
welfare were largely unsupported. The hypothesis on similar scheme and treatment effects between resource presence in the nest box and welfare received the most support, based on item order and Number effects, but the overwhelming inconsistency of these results between sex, color, and welfare indicator suggests that this support is tentative at best.

3.4.2 Resource use as a covariate in welfare models

Resource use failed to explain previous scheme and treatment effects on welfare. Scheme and treatment effects dropped in only four of the 45 models after controlling for resource use, and resource use did not also predict individual differences in welfare in these models. Similarly, individual differences in welfare were predicted by resource use in only two models: inactive use predicted less SB for pastel males in the Baseline and Diversity schemes in October, and, unexpectedly, active use predicted more SB for pastel males in the Baseline and Number schemes during single housing. Both of these resource use effects were minimal and potentially caused by an excess of zero scores for SB, which could result in minor variations in SB for a few mink driving these apparent relationships between resource use and welfare.

3.4.3 Limitations

When the previous sections are considered together, they generate two possible conclusions. The first is that resource use is not linked to negative affective states, i.e. that the degree of resource use does not predict the welfare benefits of enrichment. The second alternative explanation is that the welfare indicators used here were insensitive to differences in affective state, such that the lack of effects represent Type II errors. This second conclusion is supported by developmental changes in mink during the study and potential methodological shortcomings. These must be addressed before the merit of the first conclusion can be assessed.

As with active use in Chapter 2, the rarity of the SB observed here may have been escalated by the scan sampling method used. Rare behaviours are more accurately estimated with a shorter inter-observation window and with more observations than could feasibly be achieved here due to the number of cages being recorded (Jauhiainen and Korhonen, 2005; Svendsen et al., 2007). Two factors likely contributed to the rarity of SB. First, SBs exhibited by young mink are in early development and therefore
are less rigid and consume a smaller proportion of time than the more established SBs of adults (Mason, 1993). Personal observation of cage mates successfully interrupting each other’s SBs by blocking cage space supports the flexibility of these SBs and their classification as still developing. When interpreting SB, the presence of developing SBs typically indicates ongoing frustration and suggests that young mink exhibiting this behaviour are currently experiencing poor welfare (Mason and Latham, 2004). However, the apparent differences in developing SBs found here may be artifacts of previously mentioned methodological inaccuracies and should be interpreted cautiously.

A second factor contributing to SB rarity could be population differences in prevalence and severity of SB. Multi farm studies have found substantial differences in SB performance between farms (e.g. Mason, 1993; Meagher et al., 2014). Lines of high/low SB mink have also been established in research populations (e.g. Svendsen et al., 2007), suggesting that there is a genetic basis for individual differences in performance. For low SB populations potentially like the one used here, emphasis on differences in SB may overlook other ways animals are expressing frustration, and so may be subject to floor effects that exaggerate differences between groups (as seen for Diversity comparisons in October). Considering the two limiting factors together, SB may thus be an insensitive measure of welfare for young mink, especially in low SB populations.

Fearfulness measures were primarily limited by changes in the response to the glove tests between collection periods, which led to issues with floor effects like those discussed for SB. The reduction in fearfulness between glove tests follows similar results found during the test’s initial development (Meagher et al., 2011). Habituation to the glove test and developmental changes in the response to aversive stimuli are potential explanations for differences between collection periods, but the current experiment is unable to disentangle these mechanisms. However, the mink tested by Meagher et al. (2011) increased aggressive responses following repeated testing, whereas mink here primarily switched to curious responses. Age differences between the young mink used here and adult mink used in prior research may explain this discrepancy. Comparisons to other studies with young mink are not possible because this was the first time that the glove test has been used with juvenile mink (other temperament studies employing the stick test instead).

This study is also the first time that boredom tests have been given to juvenile mink, which may explain why no scheme or treatment effects were found. A potential explanatory hypothesis is that the monotony that typically leads to boredom takes a long time to develop on environments that frequently
change or have components that are resistant to decreases in novelty, such as social interaction. The period of monotony development would begin when an animal is first introduced to the environment and ends when consistent boredom states cannot be resolved by the animal. The young mink studied here had experienced two major housing changes in the four months prior to boredom testing (family housing to pair housing to single housing) and were experiencing seasonal changes in daylight and weather for the first time, which may have impeded monotony development. In contrast, the mink used to validate the boredom test as a welfare indicator were 10 months old and had been in consistent single housing for at least six months (Meagher et al., 2017; Meagher and Mason, 2012). Therefore, these mink may have experienced greater monotony due to the extended period without social interaction within an unchanging environment.

Another difference between current and prior boredom tests was in the number of unique stimuli presented. Prior research presented a suite of 10-14 stimuli with diverse presentation methods to estimate response times for aversive, ambiguous, and rewarding stimuli (Meagher et al., 2017; Meagher and Mason, 2012). The choice to present only one stimulus from each valence grouping arose from logistical constraints on the number of tests that could be completed, but it may have reduced the accuracy of response times estimates for each grouping. For example, if the odorants presented here differed in their arousing capabilities, perhaps due to odorant concentration, they may not accurately estimate response times for valence groupings. Unexpectedly, mink here were also substantially more uniform in their contact and orientation time between stimuli compared to the same stimuli presented in Meagher et al. (2017), which could suggest that mink were attending to the tea strainer presentation method rather than specific odorants.

Finally, bite mark analysis was limited by the shift to single housing for all mink in late October and the inclusion of light-coloured mink. As mentioned in Section 3.1, bite marks only develop between September and November. Therefore, the shift to single housing in October halved the amount of time that aggressive interactions between mink could result in bite marks. Bite marks are also more difficult to count in light-coloured mink (Hansen et al., 2014), so the inclusion of pastel mink in this study may have limited accuracy of scheme and treatment estimates. No color type main effect on bite marks was found, but this may be due to the zero-inflated dataset and overdispersion potentially caused by the shortened bite mark development window.
These limitations on the welfare indicators measured provide useful suggestions for improving future implementation, but they also emphasize that indicators validated for adult animals may not always apply to juvenile conspecifics. A more general limitation of this study is the large number of hypothesis tests run, which inflates the probability of Type I errors (as mentioned in Section 2.4.5). Results from this study should therefore be replicated in more targeted research before steadfast conclusions are made.

3.4.4 Implications

No form of resource use was able to consistently predict individual differences in, or scheme and treatment effects on, welfare. However, dissimilarity between effects on resource use and welfare does not mean that the former is useless in indicating the benefit gained from enrichment. Only indicators of negative effective state were measure here, which may not be expressed until later in life and are often insensitive to changes in positive welfare (as discussed in Sections 3.1 and 3.4.3). The greater behavioural opportunities provided by the resources used here, e.g. the ability to cache or sleep with items, may be positive experiences that are beneficial for mink but do not address causes of negative welfare (a topic returned to in Chapter 4). Future research should compare enrichment use and welfare in adult mink to identify any changes in their relationship as mink reach maturity and find valid measures of positive affective state to cover a broader spectrum of welfare indicators.
Chapter 4: Discussion and Future Research

4.1 Introduction

The study presented in the preceding chapters was motivated by persistent assumptions and questions about how environmental enrichment improves welfare. First, there was the assumption that physical contact, primarily active handling, was required for the welfare benefits of manipulable objects to occur (see Section 1.2). This idea likely grew from early laboratory psychology research that investigated causes for brain and behaviour changes elicited by environmental enrichment. For example, Ferchmin et al. (1975) placed rats in a wire cage suspended within an enriched cage so that these rats could observe others interacting with enrichments but not do so themselves. When brain weights, home-cage behaviour, and exploratory behaviour were compared between enriched, observer, and impoverished rats, the observers did not exhibit changes associated with enrichment. Observers were, instead, more akin to impoverished rats, leading to the conclusion that direct interaction with enrichments was required for the development of physiological and psychological effects. However, observers in this study were not provided the choice to interact with enrichments, which is possibly an essential component of environmental enrichment (a topic further discussed in Section 4.5.1).

Questions about whether physical contact with enrichments is required for welfare benefits to occur persisted through the subsequent decades (e.g. Jones, 1997; Tarou and Bashaw, 2007), although at least one code of practice regarding environmental enrichment maintained the assumption that sustained active use was necessary (e.g. NFACC, 2013). Recent work primarily with mink and rats sought to clarifying the relationship between contacting enrichments and welfare, but these species have generated conflicting results (as reviewed in Section 1.2.1). The present study aimed to reconcile this discrepancy by thoroughly investigating how mink interact with manipulable objects, including methods beyond active use, and determining whether differences in amounts of interaction accurately reflect differences in welfare.

The second question that motivated this study asked what aspects of enrichment provision cause the largest increase in resource use and welfare improvement. Literature reviewed in Chapter 1 on exploration, habituation, and aggression identified the number, diversity, and novelty of resources as
provision scheme characteristics of interest. Abou-Ismail and colleagues (2011a; 2016) have previously investigated diversity and novelty; however, limitations of these studies prevented definitive conclusions about the importance of these factors (see Section 1.3.3 and 4.5.2). Most other studies on environmental enrichment have incorporated multiple provision characteristics into super-enriched schemes, which increases the probability of welfare benefit at the expense of identifying specific principles for supplying enrichments to captive animals (Swaisgood and Shepherdson, 2006). The study presented here therefore sought to simultaneously assess provision scheme characteristics to develop a fundamental understanding of how each contributed to resource use and welfare benefits of enrichment.

This final chapter begins by summarizing the overarching objectives of this thesis and the extent to which they were met. The results from Chapters 2 and 3 are then discussed, regarding their application to current management practices and contribution toward developing principles of environmental enrichment. Lastly, directions for future research are explored based on the conclusions and limitations of this research.

4.2 Summary of objectives and results

Two overarching objectives guided the research presented in Chapters 2 and 3: first, to determine if resource use, whether it be active, inactive, or the presence of items in the nest box, accurately reflects or affects the animal welfare benefits of enrichments; and second, to identify provision schemes, such as increasing the number, diversity, or novelty of items, that promote use and improve welfare. The inclusion of the Baseline provision scheme, representing current farm practices, also allowed this study to investigate practical improvements to environmental enrichment standards for farmed mink.

The results can be summarized in four main conclusions. The first three relate to the results of Chapter 2 and the last to Chapter 3. 1) The amount of resource use can be increased above levels exhibited under current management recommendations (Baseline) for farmed mink. However, unexpected item type and order effects prevented clear assessment of which provision characteristics reliably cause use to increase and suggest that provision scheme effects can be inconsistent between items (here, balls and rings). 2) Item type effects also showed that the manipulable objects currently recommended, based on prior research and farmer surveys (Mason and Dawson, unpubl.; Meagher et al., 2014), are not equivalent in the amount of use they elicit, with rings being more effective than balls. 3) The presence of
resources in the nest box likely represents caching behaviour due to increased presence of these items in the nest box as the study progressed and as mink increased their own time in the nest box. Caching and inactive use are ways, other than active use, that mink interact with manipulable objects that warrant further study. 4) No form of use consistently predicted negative affective states, but limitations of the welfare indicators used here prevent drawing conclusions on the relationship between resource use and welfare.

Considered together, these conclusions demonstrate that this study was successful in detecting differences in resource use between provision schemes, although these effects were primarily limited to Baseline comparisons. This prevents fundamental interpretation of the relative importance of provision characteristics (i.e. number, diversity, and novelty) but does allow practical application to current farm standards. The unexpected item type results also show differences in resource use despite being unrelated to the original hypotheses being tested. The welfare assessment component of objectives was not met. Potential modifications to this study that would better prepare future work to meet this objective are discussed in Section 4.4.

4.3 Implications

The first conclusion on increasing resource use above current practices suggests that the current environmental enrichment recommendations for farmed mink can be improved. However, exact recommended changes require further research to better understand how various provision schemes relate to each other. Including demi and pastel mink in this study revealed unexpected differences between colour types in their response to provision schemes, in terms of both resource use and welfare measures, albeit at the expense of statistical power when testing the hypotheses of interest. Although not directly related to the hypotheses tested in this thesis, these findings contribute to prior work showing differences between colour types in fearfulness (Meagher et al., 2014, 2011). This emphasizes that colour types may differ in their management preferences and that any assumptions about similarities between colour types should be thoroughly investigated before being acted upon.

Current recommended manipulable objects for farmed mink also require further evaluation to determine the type and extent of use that each elicits. It is likely that these items differ in the behavioural opportunities provided to animals, thus differing in their potential to retain interest, whether that be
through active use, inactive use, or caching. To inform future codes of practice, research should focus on designing and assessing enrichments that maximize behavioural opportunity. For example, rings could be played with, worn, and carried into the nest box, while balls could only be played with or stored in the nest box and can be too fragile to withstand sustained use. The evaluation of different forms of resource use presented here suggests that items being inactively used or cached may still be valued by mink, meaning that their removal from cages could be detrimental to welfare. Therefore, enrichment recommendations that prioritize only active use may overlook items that fulfill more passive motivations.

Lastly, welfare indicators validated for adult animals may not be applicable to juveniles. Fearfulness and intraspecific aggression were the most promising indicators of negative affective states employed here for juvenile mink; however, developmental changes in fearfulness during the study and methodological limitations of bite mark analysis still prevented robust conclusions on welfare effects. Other than statistical limitations caused by including two colour types that differed in response to treatments, issues with detecting biologically relevant effects were primarily methodological (see Section 3.4.3). Additionally, the differences in behavioural opportunities provided by the rings and balls used here may be better suited for generating positive affective states, which were not assessed, rather than addressing specific causes of frustration or fear that produce negative affective states. For example, contacting rings may have provided comfort to mink, based on frequent observation of mink wearing rings while resting/sleeping and bringing them into the nest box when most inactive time occurred there later in the year, but this may not have dealt with frustrations that would lead to stereotypic behaviour or aggression. Valid indicators of positive welfare in mink may provide better measures to understand the relationship between resource use and welfare, especially in juveniles.

4.4 Modifications for replication

Over the course of this study, areas where modifications to methods could improve the sensitivity and accuracy of measurements were noted and are presented here to inform future replication of this work. When studying provision schemes that require multiple resources, such as diversity and novelty, running pilot studies to identify resources that provide similar behavioural opportunities and elicit similar amounts of use is advised. This could prevent item type or order effects from complicating results, as occurred here. Additionally, recording which item type is being actively or inactively contacted when
comparing use between schemes would allow unexpected treatment effects to be investigated in all schemes. This was done for the presence of items in the nest box here but was not completed for the other forms of use.

For active use, finding alternative measurements, beyond direct observation, would help overcome limitations caused by the large number of observations required to accurately estimate this rare behaviour (see Section 2.4.4). One possibility is tracking when items change location within the cage, either between observations or longer periods. This requires being able to identify individual items when multiple of the same type are present in a cage, such as including unique markings on some items. A more technologically savvy solution could be to attach accelerometers to items that can track when they are being handled. However, both solutions create a problem where any incidental contact with items is recorded as active use, which would need to be addressed before they can be considered valid measures. Signs of wear could also be recorded at the end of the study to assess general interaction with items, but this requires items to be equivalent in their ability to be damaged or altered, which was not the case between the rings and balls used here (see Section 2.4.4).

The most beneficial modification to assessing welfare would be to pre-screen the specific study population to identify welfare indicators that are more sensitive to changes in affective state (e.g. those without floor/ceiling effects). Tailoring the suite of welfare indicators employed in a study to a specific population could prevent differences in indicator prevalence between populations (e.g. age groups, farms, laboratories, zoos) from causing relevant welfare effects from being overlooked. For example, fur clipping/chewing is sensitive to benefits gained from simple environmental enrichments but was not recorded here due to logistical restraints (Meagher et al., 2014). Welfare assessment could also be improved by designing a study with matched pairs of littermates between treatments (e.g. Meagher et al., 2014) or only including one colour type to reduce data noise caused by genetic differences between groups. The colour type recommended for replication depends on the question of interest. Resource use effects were more prevalent in pastel mink, but demi mink may offer better aggression measurement through bite mark analysis. Similarly, altering the methods for measuring some welfare indicators used here could improve their validity.

First, the presentation method for stimuli during boredom tests should focus the animal’s attention on valence differences. Tea strainers were potentially too dynamic and rewarding, with interest in batting/biting this new toy overshadowing any perceived differences between odorants. Prior
implementation of boredom tests with mink has shown that the responses to such rewarding stimuli are unable to detect differences in boredom (Meagher et al., 2017). Second, although bite mark analysis can assess rare instances of aggression, this measure requires extensive coordination with farm staff to be successful. Mink must be pair housed through the growth period of the winter coat for meaningful differences in bite marks to become apparent, which is consistent with the validation work for this indicator (Hansen et al., 2014). This measure also becomes more complicated if individual identification of mink is required, such as when investigating behavioural correlates of aggression, because most farms kill mink in batches of 15-20. Individually killing and tagging mink would likely be prohibitively slow during one of the most labour-intensive seasons on mink farms. Finally, including welfare indicators that cover a wider variety of affective states (i.e. ranging from positive to negative, see Section 4.5.1) would increase the probability that environmental enrichment effects on welfare are not miscalculated.

4.5 Future research

While preparing this thesis, additional topics that warrant further research extending from the background and results of the presented study were identified. These include the relationship between environmental enrichment, behavioral opportunity, and positive welfare; complications with designing and studying novelty treatments; and the cause of reductions in active use over time. These topics are discussed in the following sections.

4.5.1 Behavioural opportunity and positive welfare

Animal welfare frameworks that include positive affective states have become more prevalent in recent decades (Boissy et al., 2007; Yeates and Main, 2008) and aim to foster peak welfare: a state where negative affect can be relieved and positive affect established. The ongoing discussion of the relationship between positive affect and welfare is largely supported by a wealth of neuroscience research showing that positive and negative affect both contribute to subjective experience and are equally capable of motivating behaviour (reviewed by Panksepp, 2011). When animals can successfully respond to motivations to perform rewarding behaviours they may generate positive affective engagement: an experience of satisfaction, fulfillment, or comfort (Mellor, 2015). Multiple recent hypotheses have emphasized agency – the capacity for self-reliance and control through choice – as the foundation for
producing positive affective engagement and maintaining peak welfare (Franks, 2019; Špinka, 2019; Veasey, 2017). Within the context of the study presented in this thesis, interactions with manipulable objects may create outlets for agency and be at least partially motivated by positive affect. Therefore, measures of negative affect alone could be insensitive or unrelated to relevant welfare improvements caused by resource use.

Incorporating measures of affect that span the valence spectrum could provide a more accurate assessment of the welfare impact of environmental enrichment. However, there are few valid measures of positive affect for most species due to late research interest in this topic relative to interest in monitoring and reducing negative affect. Behavioural diversity, the variety and frequency of species-typical behaviours (i.e. excluding stereotypic and anticipatory behaviours), is one proposed measure of positive affect that has been employed when studying environmental enrichment, predominantly in zoos (Hirt and Wechsler, 1994; Kistler et al., 2009; Miller et al., 2016, 2011; Part et al., 2014; Shepherdson et al., 1993; Swaisgood et al., 2001; Swaisgood and Shepherdson, 2005). The underlying assumption of this measure is that animals who exhibit relatively greater behavioural diversity are more capable of fulfilling their motivations and thus potentially experience greater positive affective engagement (Miller et al., 2016).

Preliminary validation as a welfare indicator was contributed by Miller et al. (2016), who showed an inverse relationship between individual differences in behavioural diversity and fecal glucocorticoid metabolites in cheetahs, although this result cannot differentiate between relative changes in negative welfare and the presence of positive welfare. Furthermore, the most common method for quantifying behavioural diversity (the Shannon diversity index, i.e. H-index; Shannon and Weaver, 1949) has debilitating mathematical limitations caused by the arbitrary nature of ethogram development (Cronin and Ross, 2019). Identifying reliable methods for quantifying behavioural diversity is a necessary step before further validation work can commence. This would require interdisciplinary collaboration between statisticians and welfare scientists to create a measure that has predetermined methods for ethogram development and is sensitive to behavioural valence. Once such a measure is available, replication of prior validation and further research with more species may allow behavioural diversity to become a useful measure in animal welfare research and management.

One possible method for additional validation of behavioural diversity as an indicator of positive affect involves a within-subject design. The animals’ behavioural diversity in its home environment, which provides enough opportunity to fulfill motivations, would be measured before and after some
treatment that is known, or reasonably inferred, to create positive affective states. Treatments could be pharmacological (more applicable to laboratory research) or the presentation of stimuli assumed to be positive based on evolutionarily adaptive relevance (see Section 1.1.3). Behavioural diversity would be predicted to increase after the treatment has established a positive affective state. Importantly, the capacity for behavioural diversity in the environment, meaning the maximum variety and frequency of behaviour possible, should remain constant between pre- and post-treatment observations. Treatments should also avoid introducing other confounds that may affect behavioural diversity, such as altering general activity levels. These considerations ensure that any differences found between observations are caused by changes in affective state. Miller et al. (2011) used a similar experimental design when investigating the welfare impact of education programs with dolphins (*Tursiops truncatus*). However, this study suffered from circular reasoning, in that education programs had not previously been confirmed to generate positive affect, and behavioural diversity had not yet been validated as a welfare measure.

Once behavioural diversity has been successfully and repeatedly validated as a measure of positive affect, it can then be used to more precisely estimate the welfare impact of environmental enrichment. For example, if the behavioural opportunity available is held constant between Number, Diversity, and Novelty provision schemes (by prescreening resources as in Section 4.4), any differences in behavioural diversity should be related to differences in welfare and suggest that provision characteristics alter the benefit gained from environmental enrichment. As shown by the range of species that behavioural diversity has already been applied to (e.g. foxes, dolphins, and small cats), future work assessing provision characteristics need not focus on mink, although this species does offer substantial sample sizes and uniform housing. Similarly, numerous other potential measures of positive affect have been proposed (see Boissy et al., 2007), but all still require proper validation and understanding of limitations for whichever species is being studied (e.g. Ahloy-Dallaire et al., 2018 on play behaviour).

### 4.5.2 Studying novelty and potential complications

In Section 1.3.3, relative and absolute novelty were defined according to Berlyne’s (1960) framework for exploratory behaviour. Absolute novelty added resources and/or opportunities that were previously unavailable, while relative novelty altered existing resources to modify opportunities. This section further inspects relative and absolute novelty to describe four ways to establish novelty and potential complications associated with each.
The study presented here included one form of absolute novelty: new items were added to the environment, and old items were never removed. Abou-Ismail and Mendl (2016) and Trickett et al. (2009), instead, generated absolute novelty by adding new items and removing old items. Both forms of absolute novelty encountered issues because the behavioural opportunities afforded to animals were not constant. As discussed in Section 2.4.2, unexpected item order effects found in this study show that the order in which items are added to the environment alters the extent of their use. Initially withholding items that presumably afford greater behavioural expression resulted in less resource use once those items were added compared to groups that had constant access to those items (as in the reduced resource use by mink that received rings second). Item order effects were also found by Abou Ismail and Mendl (2016), but these could not be definitively attributed to the addition or removal of a specific item. Similarly, Trickett et al. (2009) found that removing hanging ropes from pig pens and replacing them with wooden blocks caused resource use to plummet until the more preferred or useful items (ropes) were returned. In all three cases, absolute novelty conditions elicited less resource use than diversity conditions that included constant access to all resources. Additionally, this comparison also revealed that removal or late addition of resources may limit welfare, assessed via indicators such as sleep, agonism, and audible vocalizations (Abou-Ismail and Mendl, 2016), which is consistent with prior research showing that removing opportunities for behavioural expression can be more damaging to welfare than never having those opportunities (Latham and Mason, 2010). Reassessing the effects of absolute novelty with a suite of items that all provided similar behavioural opportunities may reveal different relationships between item order treatments and diversity schemes. Gross et al. (2011) accomplished this by using three exemplars of shelters (cardboard house, plastic house, or plastic tunnel) and climbing structures (ladder, rope, wooden bars) to create nine combinations of enrichment to compare diversity and novelty conditions in mice. However, this study did not show differences in stereotypic or anxiety-related behaviour between diversity, novelty, and a baseline condition consisting of only nesting material.

Relative novelty can first be achieved by making slight changes to familiar resources and associated experiences, such as moving them within the environment, changing their structure or function, or replenishing aspects of the resource that are consumed. The second method is to add more items of the same type already present in the environment. For example, increasing the number of building blocks available to a child would create novelty by expanding the number of possible structures that could be made. Trickett et al. (2009) included relative novelty treatments using the former method by providing either hanging ropes or wooden blocks for pigs and replenishing or cleaning them as they became worn or
soiled (for ropes and blocks, receptively). They found that replacing hanging ropes increased their use, but cleaning and replacing wooden blocks did not have the same effect. This highlights that relative novelty in some cases may be too subtle for changes to be perceptible or meaningful for animals.

Future research comparing these types of novelty to each other and to constant diverse conditions may reveal why novelty, a concept frequently emphasized in environmental enrichment research, does not always increase resource use nor improve welfare. Likely causes for such effects that require targeted evaluation are inconsistency of choice and control (primarily related to absolute novelty) and inability to make meaningful increases in novelty (primarily related to subtle changes in relative novelty). Increased human interaction in most novelty conditions, from the need to add, remove, or alter resources, is also a potential confound between novelty and diversity that should be addressed. Furthermore, once consistent resource use and welfare effects from each novelty method are established, comparison between them could identify optimal ways of producing novelty, as measured by improvements in welfare and feasibility of practical implementation.

4.5.3 Active use reduction from habituation, motivational change, or seasonal changes

As a final extension of this thesis, longitudinal changes observed here in all forms of resource use should be explored further. Focusing on active use, reductions in this behaviour over the course of the study could be due to habituation, motivational change during maturation, seasonal changes in general activity levels, or some combination of these factors.

Determining whether habituation is responsible for reductions in active use requires either a stimulus specificity or dishabituation paradigm (as discussed in Section 1.3.1, Epstein et al., 2009; McSweeney and Murphy, 2009), which involves behavioural observations immediately following provision of items. Such observations were not done here because this study was interested in how longer-term measures of use related to welfare. If habituation is identified as a partial cause for reductions in active use, this still may not represent a complete loss of value to the animal (see Section 1.3.2). For instance, animals may have shifted their interest in items towards other forms of use, as was evident here by increases in inactive use and presence of items in the nest box over time.

Understanding whether relative changes between forms of use represent motivational changes caused by maturation or seasonal change is slightly more challenging. This would require long-term
observation of how resources are used, ideally through at least two years to identify any seasonal effects on resource use. A maturation effect would be apparent if active use never returned, even if new items were continuously added to the environment. It may be tempting to avoid long-term observation by doing a direct comparison of resource use between juveniles and adults previously housed without the items of interest (e.g. Lambath and Bloomsmith, 1992; Swaisgood et al., 2001), but such a comparison could ignore potential developmental effects on resource use caused by having expanded behavioural opportunity at a younger age (similar to item order effects previously discussed).

Additionally, observing animals through adulthood would increase the number of validated welfare indicators available for assessment (e.g. stereotypic behaviour without floor effects, readiness to mate, fertility, and litter size in mink). Continuing observations into adulthood would also increase the amount of time mink are single housed, which might provide a better representation of any benefits gained from physical environmental enrichment. Finally, investigating provision scheme characteristics with adult mink would contribute to prior research showing that recent addition of resources to female mink, compared to lifelong enrichment, had inconsistent effects on reproductive measures of barrenness and early litter size (Meagher et al., 2014). The present study did not continue observations into this period because it was more interested in seeing how environmental enrichment impacts welfare for most of the mink on farms, who typically have a seven-month lifespan.

### 4.6 Conclusion

Although the objectives of this thesis were only partially met, new insights into how animals may benefit from environmental enrichment have been identified. Forms of passive use, such as inactive contact and caching items in secure areas, warrant further investigation and consideration when researching and designing environmental enrichments. Furthermore, a focus on offering more opportunities for motivational expression, even for behaviours like wearing a plastic ring, may grant animals more agency within their environment and allow more frequent positive affective engagement. Lastly, the ideas presented in this thesis should be investigated in a variety of species to continue development of general principles for environmental enrichment.
REFERENCES


Buob, M., Meagher, R., Dawson, L., Palme, R., Haley, D., Mason, G., 2013. Providing “get-away bunks” and other enrichments to primiparous adult female mink improves their reproductive productivity.


Confidential United States Government work not protected by U.S. copyright. 2023 by Elsevier Inc. All rights reserved.


Jaeger, B., 2017. r2glmm: computes R squared for mixed (multilevel) models.


125


Nip, E., 2018. The Long-Term Effects of Environmental Enrichment on Agonism in Female C57BL/6, BALB/c, and DBA/2 Mice. University of Guelph.


Shepherdson, D.J., Carlstead, K., Mellen, J.D., Seidensticker, J., 1993. The influence of food presentation...


Swaisgood, R.R., Shepherson, D.J., 2005. Scientific approaches to enrichment and stereotypies in zoo animals: What has been done and where should we go next? Zoo Biol. doi:10.1002/zoo.20066


Veasey, J.S., 2017. In pursuit of peak animal welfare; the need to prioritize the meaningful over the measurable. Zoo Biol. 1–13. doi:10.1002/zoo.21390


doi:10.1016/j.tvjl.2007.05.009

