Ecological Implications of Intraspecific Behavioral Variation in a Small Mammal Community

Allison Brehm
University of Maine, allison.brehm@maine.edu

Follow this and additional works at: https://digitalcommons.library.umaine.edu/etd

Part of the Terrestrial and Aquatic Ecology Commons

Recommended Citation
https://digitalcommons.library.umaine.edu/etd/3671

This Open-Access Dissertation is brought to you for free and open access by DigitalCommons@UMaine. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of DigitalCommons@UMaine. For more information, please contact um.library.technical.services@maine.edu.
ECOLOGICAL IMPLICATIONS OF INTRASPECIFIC BEHAVIORAL VARIATION IN A SMALL MAMMAL COMMUNITY

By
Allison M. Brehm
B.S. University of New Hampshire, 2014
M.S. University of Maine, 2018

A DISSERTATION
Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy (in Ecology and Environmental Sciences)

The Graduate School
The University of Maine
August 2022

Advisory Committee:
Alessio Mortelliti, Associate Professor of Wildlife Habitat Conservation, Advisor
Erik J. Blomberg, Associate Professor of Wildlife Population Ecology
Shawn Fraver, Associate Professor of Forest Ecology
Malcolm Hunter Jr., Emeritus Professor of Wildlife Ecology
Brian McGill, Professor of Biological Science
Variation among individuals is at the root of all evolution by means of natural selection. However, only in recent years has intraspecific behavioral variation been embraced as a potential driver of community and ecosystem processes, rather than considered statistical noise. Animal personalities, or behavioral differences between conspecifics that are consistent across time and contexts, are one such form of variation that has received considerable attention in the last two decades. Investigations of the ecological and ecosystem consequences of personality variation is at the current forefront of the field, but much work on this topic remains conceptual. Here, I apply large-scale field experiments and provide empirical evidence for three mechanisms by which personality variation can scale up to influence processes at the population and ecosystem level.

Using the small mammal community in Maine's temperate mixed forests as a model system (specifically, deer mice, *Peromyscus maniculatus*, southern red-backed voles, *Myodes gapperi*, northern short-tailed shrews, *Blarina brevicauda*, and North American red squirrels (*Tamiasciurus hudsonicus*), in Chapter 1 I use a field experiment to confirm that methods of live-trapping do not confound behavioral measurements obtained from standardized assays. Chapter 2 examines four years of detailed trapping data and fine-scale habitat measures to identify population-level correlations between personality traits and habitat selection. Chapter 3
investigates the understudied role of intraspecific behavioral variation in mutualisms by examining the propensity for personality traits of scatter-hoarders to generate context-dependence in the seed dispersal mutualism. Finally, Chapter 4 examines the effects of personality differences on detection, movement, and survival of small mammals.

Using an empirical approach, this dissertation highlights three mechanisms through which animal personalities can influence animal populations and ecosystem function. Ultimately, this dissertation provides important empirical evidence of the ecological consequences of animal personalities and should be a catalyst for prospective work on this topic.
DEDICATION

To my parents,

for teaching me to love and respect the natural world, work hard, and think with an open mind.

Your love and support have given me the courage to chase after my dreams.
ACKNOWLEDGEMENTS

I would first like to thank my advisor, Alessio Mortelliti, for having the utmost faith in me even when I found it hard to believe in myself. You bring such enthusiasm and passion into your work and I am so thankful to have had six years to learn from you and call you a mentor and friend. Thank you for welcoming me into your home, introducing me to real Italian food, and always being honest and understanding. I owe a whole lot of my successes to you and your limitless commitment to my growth as a human and as an ecologist. I will surely miss being able to “pop by” your office for a chat whenever I please, but I look forward to many more years of collaboration with you. Our first field season in the PEF (although full of challenges) will forever hold some of my fondest memories - I can’t wait to teach my own grad students to catch mice!

Thank you to everyone in the WFCB department for making this a friendly and welcoming place to come to each day. To Rena, Katherine, and Molly - I would have been lost without you on numerous occasions and I’ll miss working with you! To Lindsay and Erik, I’m so grateful for the stellar memories at May term(s), and for your help gaining valuable teaching experiences. And of course, Lindsay, thank you for bringing me out to see baby bears and occasionally fueling my sugar addiction with gummy worms!

A sincere thanks to my dissertation committee: Shawn, your office became a little place of sanctuary for me over the years - thanks for all of the meetings (and chocolate)! Erik, thank you for your kindness, calming personality, and support over the years. You have a way of making even something I dreaded (like comps) a little less scary. Brian, thank you so much for thoughtfully and thoroughly answering every one my statistical questions! Mac, your breadth of knowledge and worldview are inspiring. Thank you for your honest and prompt feedback (and the occasional photo of Aram with a squirrel). I’ll always have fond memories of the times I visited your beautiful home!
This work was made possible through the hard work of numerous undergraduate technicians and volunteers. I sincerely enjoyed working with you over the years and I hope you know how instrumental you were to the success of this project! Thank you to my funders: The USDA National Institute of Food and Agriculture McIntire-Stennis projects through the Maine Agriculture & Forest Experiment Station, a National Science Foundation CAREER Award, Penobscot Experimental Research Funds, the University of Maine Graduate Student Government, and the University of Maine's Chase Distinguished Research Assistantship and Janet Waldron Doctoral Research Fellowship. Thank you to Laura Kenefic, and all those who have maintained and facilitated research in the Penobscot Experimental Forest over the years, and thank you to Keith Kanoti for facilitating my pilot studies in the campus forests.

Thank you so much to my family: my parents, who worked so hard to always make sure I had everything I needed to succeed in life, and my siblings, whose artistic talents inspire me to think and work creatively. I’m so lucky to have had your support and love every step of the way.

Finally, I am forever thankful to the lab-mates, office-mates, friends and colleagues I’ve made during my journey here at UMaine. To the strong, smart, wonderful group of women I’ve been lucky to work with over the years: Sara, Bryn, Ivy, Gabi, Maisie, and Brigit - never stop being the unique queens that you are. I will miss you more than you know and I hope we continue to work together in the future! To my “wife”, Gabi, you’ll forever be part of my family - thank you for talking me off my ledge occasionally and putting up with my inability to do dishes. And to everyone who has made memories with me during hikes and fishing trips, camping and bonfires, bird walks, potlucks, dumpster-dives, drinks at Marsh, dumpling nights, dance parties, 5ks, and Friendsgiving’s... I wish I could give you all individually the thanks you deserve for making my time here so memorable. I love you more than I love mice - and that’s saying something.
# TABLE OF CONTENTS

**DEDICATION** .......................................................................................................................... vi

**ACKNOWLEDGEMENTS** ........................................................................................................ vii

**LIST OF TABLES** .................................................................................................................. xiii

**LIST OF FIGURES** .................................................................................................................. xvi

**Chapters**

1. **EFFECTS OF TRAP CONFINEMENT ON PERSONALITY MEASUREMENTS IN TWO TERRESTRIAL RODENTS** ........................................................................................................... 1
   1.1. Introduction .......................................................................................................................... 1
   1.2. Materials and methods ........................................................................................................ 4
      1.2.1. Study site and small mammal trapping ....................................................................... 4
      1.2.2. Behavioral tests ............................................................................................................ 7
      1.2.3. Monitoring capture events .......................................................................................... 8
      1.2.4. Data analysis ............................................................................................................... 10
      1.2.5. Ethical note ............................................................................................................... 11
   1.3. Results ............................................................................................................................... 12
      1.3.1. Repeatability analysis ................................................................................................. 12
      1.3.2. Trap confinement analysis ......................................................................................... 12
   1.4. Discussion .......................................................................................................................... 16
   1.5. Acknowledgments .............................................................................................................. 20
      1.5.1. Author contributions ................................................................................................. 20

2. **LAND-USE CHANGE ALTERS ASSOCIATIONS BETWEEN PERSONALITY AND MICROHABITAT SELECTION** ................................................................................................. 21
   2.1. Introduction ....................................................................................................................... 21
   2.2. Materials and methods ...................................................................................................... 27
      2.2.1. Study area and small mammal trapping ................................................................. 27
2.2.2. Animal tagging and behavioral tests ................................................................. 28
2.2.3. Repeatability of behavior in standardized tests ............................................. 32
2.2.4. Microhabitat selection ...................................................................................... 33
2.2.5. Third-order selection ...................................................................................... 35
2.2.6. Second-order selection ................................................................................... 36
2.2.7. Assessing correlations between personality traits and microhabitat
selection....................................................................................................................... 37

2.3. Results .................................................................................................................. 39

2.3.1. Microhabitat structure ...................................................................................... 39
2.3.2. Trapping and behavioral data .......................................................................... 40
2.3.3. Effects of land-use change on personality-associated microhabitat
selection....................................................................................................................... 40

2.4. Discussion ............................................................................................................ 45

2.4.1. Personality traits and the selection for key resources ..................................... 47
2.4.2. Management implications ............................................................................... 51
2.4.3. Conclusions ..................................................................................................... 52

2.5. Acknowledgments ................................................................................................ 53

2.5.1. Author contributions ....................................................................................... 53

3. SMALL MAMMAL PERSONALITIES GENERATE CONTEXT-DEPENDENCE IN
THE SEED DISPERSAL MUTUALISM ....................................................................... 54

3.1. Introduction .......................................................................................................... 54

3.2. Materials and Methods ....................................................................................... 59

3.2.1. Study site and small mammal trapping .......................................................... 59
3.2.2. Animal processing and behavioral assays ....................................................... 61
3.2.3. Seed experiments ............................................................................................ 61
3.2.4. Statistical analyses .......................................................................................... 63
4.5. Acknowledgments ............................................................................................................. 97

BIBLIOGRAPHY .................................................................................................................... 98

APPENDIX A - Chapter 1 Supplemental Materials ................................................................. 115

APPENDIX B - Chapter 2 Supplemental Materials ................................................................. 116

APPENDIX C - Chapter 3 Supplemental Materials ................................................................. 128

APPENDIX D - Chapter 4 Supplemental Materials ................................................................. 144

BIOGRAPHY OF THE AUTHOR ......................................................................................... 165
LIST OF TABLES

Table 1.1. Model output of top-ranked linear models predicting behaviors performed during standardized tests ................................................................. 14

Table 2.1. Mechanisms with the ability to generate behavioral correlations between personality and habitat selection across landscapes ......................... 23

Table 2.2. Repeatable behavioral traits from three different behavioral tests performed on deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*) ............................................................. 30

Table 2.3. Correlations between repeatable behavioral variables and selection for two components of microhabitat structure at the second and third-orders of selection in deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*) ........................................................................ 42

Table A.1. Repeatability estimates for target behaviors measured in three behavioral tests (handling bag, emergence, and open-field) in deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*) ........................................................................................................ 115

Table B.1. Glossary of the silvicultural treatments used in this study .................. 121

Table B.2. Pairwise correlations between all behavioral variables used in this study for *Peromyscus maniculatus* and *Myodes gapperi* .......................... 122

Table B.3. A glossary of terms associated with hierarchical habitat selection ....... 123

Table B.4. Microhabitat variables measured in this study ..................................... 124

Table B.5. Loadings from the first two components extracted from a principal component’s analysis (PCA) on microhabitat data recorded at each trap location in the study sites ................................................................. 125

Table B.6. Repeatability estimates for target behaviors measured in three behavioral tests (handling bag, emergence, and open-field) in deer
mice (*Peromyscus maniculatus*) and southern red-backed voles
(*Myodes gapper*)

Table B.7. Results of repeatability analysis on variables of microhabitat selection
at the third-order in deer mice (*Peromyscus maniculatus*) and southern
red-backed voles (*Myodes gapper*)

Table C.1. Interactions used to calculate an individual’s score along the
antagonist/mutualist continuum

Table C.2. Repeatability estimates for target behaviors measured in three
behavioral assays (open-field, handling bag, and emergence tests) in
deer mice (*Peromyscus maniculatus*)

Table C.3. Repeatable behavioral traits from three standard behavioral assays
performed on deer mice (*Peromyscus maniculatus*)

Table C.4. Results from linear regression predicting scores along the
predator-mutualist continuum in the deer mouse
(*Peromyscus maniculatus*)

Table C.5. Results for logistic mixed-models predicting the probability of a positive
seed interaction in the deer mouse (*Peromyscus maniculatus*)

Table C.6. Pairwise correlations between all behavioral variables used in model
selection and an individual’s mean body condition index (averaged
scaled-mass index), and mean body mass

Table D.1. Repeatable behavioral traits from three behavioral tests (handling bag,
emergence, and open field) in deer mice (*Peromyscus maniculatus*),
southern red-backed voles (*Myodes gapper*), northern short-tailed
shrews (*Blarina brevicauda*), and North American red squirrels
(*Tamiasciurus hudsonicus*)

Table D.2. Pairwise correlations between repeatable behavioral traits used in
model selection. Behavioral variables shown are the mean BLUP (best linear unbiased predictor) averaged over 1000 simulations.................. 150

Table D.3. Repeatability estimates for target behaviors measured in three behavioral tests........................................................................................................ 153

Table D.4. Best supported models from survival analyses using Robust Design models with the Huggin’s estimator for the deer mouse (Peromyscus maniculatus), southern red-backed vole (Myodes gapper), northern short-tailed shrew (Blarina brevicauda), and North American red squirrel (Tamiasciurus hudsonicus)........................................................................... 154
LIST OF FIGURES

Figure 1.1. Map of the study area at the Penobscot Experimental Forest, Maine U.S.A .......................................................... 6

Figure 1.2. A camera trap (Bushnell NatureView HD) monitors a Longworth trap in the field (circled in red) .............................................................. 9

Figure 1.3. Prior trapping experience influences the behavioral response to trap confinement in deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*) ...................................................... 15

Figure 2.1. A conceptual diagram identifying the two main objectives of this study ..... 25

Figure 2.2. Habitat selection by small mammals occurs at multiple spatial scales....... 34

Figure 2.3. Personality-microhabitat associations are altered by silvicultural practices in deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*) ...................................................... 41

Figure 2.4. Correlations between docility (in a handling bag test) and selection for PC2 at the second and third-order in southern red-backed voles (*Myodes gapperi*) ............................................................................ 45

Figure 3.1. Overview of how individuals are scored along the antagonism-mutualism continuum ................................................................. 58

Figure 3.2. Summary of study area and experimental design............................... 60

Figure 3.3. Individual scores for *Peromyscus maniculatus* along the antagonism-mutualism continuum for red oak (*Quercus rubra*) .............. 67

Figure 3.4. Predicted relationships (and 95% CIs) between a key behavioral variable indicating the degree of boldness/timidness and the probability of interacting positively with seeds ........................................... 69

Figure 4.1. Overview of the experimental design: photos of study areas, variability in small mammal density, and stand structural differences ................. 80
Figure 4.2. Relationship between apparent survival and docility as well as apparent survival and boldness in deer mice (*Peromyscus maniculatus*) and northern short-tailed shrews (*Blarina brevicauda*) .......... 88

Figure 4.3. Relationship between apparent survival and stress de-arousal behavior in deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*) ............................................................................. 89

Figure B.1. Figures showing the total number of captures per individual used in third-order analyses for *Peromyscus maniculatus* and *Myodes gapperi* .................................................................................................. 117

Figure B.2. Scree plot showing the Eigen values of principal components (PC) from a principal component’s analysis (PCA) on microhabitat data recorded at each trap location in our study sites ........................................ 118

Figure B.3. Kernel density plots of the first two microhabitat components (PC1 and PC2) for the reference and treatment areas used in this study ........ 119

Figure B.4. Plot of the first two principal components from a principal component’s analysis (PCA) on microhabitat data recorded at each trap location in our study sites ........................................................................... 120

Figure C.1. White pine (*Pinus strobus*) interaction diagram ................................................. 132

Figure C.2. Red oak (*Quercus rubra*) interaction diagram .................................................. 133

Figure C.3. American beech (*Fagus grandifolia*) interaction diagram ............................. 134

Figure C.4. Individual scores for *Peromyscus maniculatus* along the antagonism-mutualism continuum for eastern white pine (*Pinus strobus*) and American beech (*Fagus grandifolia*) ........................................ 135

Figure C.5. Predicted relationships (and 95% CIs) between an index of body condition and an individual’s score along the antagonism-mutualism continuum ................................................................. 136
Figure D.1. Map of study areas at the Penobscot Experimental Forest.......................... 155
Figure D.2. Positive associations between body mass and apparent survival
obtained from the best supported Robust Design Models. ...................... 156
Figure D.3. Model structure for detection ($p$) and recapture ($c$) probabilities of
North American red squirrels ($Tamiasciurus hudsonicus$) obtained
from the best supported Robust Design Model................................. 157
Figure D.4. The probability of temporary emigration ($\gamma''$) varied by sex in North
American red squirrels ($Tamiasciurus hudsonicus$)............................. 158
Figure D.5. Parameters affecting detection ($p$) probabilities of deer mice
($Peromyscus maniculatus$) from the best supported Robust Design
Model.................................................................................................. 159
Figure D.6. Abundance of small mammals (number of different individual's
captured) over time in the Penobscot Experimental Forest, ME USA. ..... 160
Figure D.7. Parameters affecting recapture ($c$) probabilities of deer mice
($Peromyscus maniculatus$) from the best supported Robust Design
Model.................................................................................................. 161
Figure D.8. Parameters affecting in detection ($p$) probabilities of southern
red-backed voles ($Myodes gapperi$) obtained from the best
supported Robust Design Model...................................................... 162
Figure D.9. Parameter affecting detection ($p$) probabilities of northern
short-tailed shrews ($Blarina brevicauda$) obtained from model
averaging the two best supported Robust Design Models.................... 163
Figure D.10. Parameter affecting recapture ($c$) probabilities of northern short-tailed
shrews ($Blarina brevicauda$) obtained from model averaging the two
best supported Robust Design Models............................................ 164
CHAPTER 1

EFFECTS OF TRAP CONFINEMENT ON PERSONALITY MEASUREMENTS IN TWO TERRESTRIAL RODENTS*


1.1. Introduction

Over the past few decades, the acknowledgement that many species of animals display consistent individual differences in behavior, or personalities, has become widespread (Gosling 2001; Sih et al. 2004b; Carere and Maestripieri 2013; Pennisi 2016). Personalities are heritable (van Oers et al. 2005), have consequences for fitness (Dingemanse and Réale 2005; Biro and Stamps 2008; Smith and Blumstein 2008; Haage et al. 2017), and can limit the ability of individuals to exhibit behavioral plasticity (Dingemanse et al. 2009). This can result in trade-offs where certain personality types perform well in some ecological contexts but not in others (Sih et al. 2003). Because individual personalities can determine the response of individuals to changing environments (Tuomainen and Candolin 2011; Wong and Candolin 2015) and have important ecological implications (Sih et al. 2012; Wolf and Weissing 2012; Brehm et al. 2019a), personality studies in wild populations will likely continue to increase as researchers strive to understand and predict the responses of individuals and populations to anthropogenic changes (Miranda et al. 2013; Berger-Tal et al. 2016; Merrick and Koprowski 2017; Lapiedra et al. 2018a).

Studies of personality in wild populations usually require that wild animals are live-trapped so that one or more standardized behavioral tests can be undertaken, but see (Carter et al. 2010, 2014; Bergvall et al. 2011; Dammhahn and Almeling 2012) for methods of personality observation in non-captured animals. Because live trapping may induce stress (Kenagy 2000; Ortiz and Worthy 2000; Reeder et al. 2004; Fletcher and Boonstra 2006; Boonstra et al. 2009;
Bosson et al. 2012), the process of capturing animals and subsequently measuring their personality offers additional challenges. Specifically, the stress of being trapped might influence the behaviors exhibited by wild animals. When trap-induced stress is unequal among individuals or among capture events and cannot be controlled for during analyses, this could confound the very behaviors at the core of the research.

Several studies have explored the relationship between live trapping and the stress response of animals (Kenagy 2000; Fletcher and Boonstra 2006; Bosson et al. 2012). It is generally accepted that the stress of being captured activates the sympathetic nervous system (secreting catecholamines) as well as the hypothalamic-pituitary-adrenocortical (HPA) axis (releasing glucocorticoids into the bloodstream) (Sapolsky et al. 2000; Fletcher and Boonstra 2006). The hormones secreted from the sympathetic nervous system during the stress response can elevate breathing rate, heart rate, and blood pressure (Fletcher and Boonstra 2006) which, following exposure to a threat (such as a predator attack), stimulates the mobilization of energy to facilitate an escape. Alternatively, the glucocorticoids released from the HPA axis can suppress digestion, inflammation, and reproduction (Sapolsky et al. 2000). When an animal is confined to a trap, this prolonged stressor may result in higher concentrations of stress hormones like catecholamines and glucocorticoids after longer durations spent in a trap (Kenagy 2000), perhaps impacting behaviors exhibited during routine behavioral tests such as grooming, time spent moving, etc. (Carter et al. 2013; Gracceva et al. 2014; Kalueff et al. 2016).

Thus far, studies looking to assess this phenomenon have focused on the hormonal/physiological response to trap-induced stress and results have been mixed (Harper and Austad 2001; Fletcher and Boonstra 2006; Bosson et al. 2012). For example, live trapping does induce an initial stress response (measured using fecal glucocorticoid levels and corticosterone concentrations) in southern red-backed voles (Myodes gapperi) and meadow voles (Microtus pennsylvanicus), but longer times spent in traps do not correlate with increased stress levels (Harper and Austad 2001; Fletcher and Boonstra 2006). By contrast, studies found
that in deer mice (*Peromyscus maniculatus*) and North American red squirrels (*Tamiasciurus hudsonicus*) prolonged time spent in traps was positively correlated with stress hormone levels (Harper and Austad 2001; Bosson et al. 2012). In either scenario, it is unknown whether the time spent in traps may affect behavioral responses, since a change in stress hormones does not necessarily precede a change in behavior.

If confinement duration affects behaviors exhibited during routine testing, this could result in misinterpretation of results and may mask the presence of repeatable behavioral traits in populations of interest. For example, if an individual is captured twice and its behavior assessed each time, but the individual spends one hour confined to a trap on the first capture and eight hours confined to a trap on the second capture, the difference in confinement duration may obscure any consistency in this individual’s observed behaviors. Alternatively, if an individual’s personality influences how quickly it enters a trap, meaning that the boldest individuals enter traps earlier (experiencing longer durations of confinement) this could lead to increased stress levels in only the boldest individuals. If the heightened stress levels caused a behavioral change, for example by causing individuals to behave in a shyer manner, truly bold individuals would appear to act similarly to truly shy individuals, but only because they have been confined to traps longer. This type of confoundment would require studies using behavioral data from trapped animals to further investigate the minimum duration of confinement that alters the behavioral response, and then control for confinement duration. This could be done by: checking traps more frequently, recording the time of capture (obtained using videos from camera traps placed on live traps) then controlling for the duration using imposed covariates in analysis, or using devices that signal when a capture has been made so that animals can be removed promptly (Larkin et al. 2003; Benevides et al. 2008). Empirical evidence is needed to explore the relationship between the time spent in a trap and behavioral response.
The objective of this study was to assess whether personality measurements obtained from live-trapped individuals are being confounded by the amount of time spent inside of a trap. Specifically, we sought to determine whether confinement duration affects the behaviors exhibited in routine behavioral tests. To meet this objective, we conducted a field experiment focused on the deer mouse and the southern red-backed vole in study populations that have been the subject of previous personality research by the authors (Brehm and Mortelliti 2018; Brehm et al. 2019a). Using high-definition trail cameras positioned facing Longworth small mammal traps in the field, we quantified the duration of time that individuals had spent inside a trap before standardized behavioral tests were performed the following morning. Using these data, we evaluated whether behaviors exhibited in behavioral tests varied with the time spent inside the trap.

Results from this study will have implications for researchers who measure behavioral traits following the live-capture of an animal. These results will highlight whether we should take additional steps to ensure that our behavioral measurements are accurate and not unduly influenced by the trapping.

1.2. Materials and Methods

1.2.1. Study site and small mammal trapping

This study was conducted in the Penobscot Experimental Forest (PEF, 44° 51' N, 68° 37' W) at the southern edge of the Acadian forest in east-central Maine, USA. This experimental forest consists of forest units chosen at random and logged separately with varying silvicultural treatments (minimum of two replicates per treatment). Management units average 8.5 ha in area (range 8.1-16.2 ha) and nearly 25 ha of forest (retained in two separate units) serves as reference and has remained unmanaged since the late 1800s (Brissette and Kenefic 2014; Brehm and Mortelliti 2018).
We implemented a large-scale mark-recapture study on six trapping grids (Figure 1.1.): two control (located in reference forest) and four experimental (two replicates in even-aged forest units and two in units treated with a two-stage shelterwood with retention). Trapping grids were 0.81 ha in area and consisted of 100-flagged points spaced 10 m apart. We positioned trapping grids close to the center of the management unit to minimize edge effects (mean distance between grids was 1.44 km; greater than the movements of our study species). We positioned one Longworth trap at each flagged point. Traps were bedded with cotton and baited with a mixture of sunflower seeds, oats, and freeze-dried mealworms. We trapped at each trapping grid for three consecutive days and nights and checked traps each morning and evening. Trapping occurred once per month for five consecutive months each year (June-October 2016, 2017, 2018).
Figure 1.1. Map of the study area at the Penobscot Experimental Forest, Maine U.S.A. (PEF, 44° 51’ N, 68° 37’ W)
1.2.2. Behavioral tests

We used three standard behavioral tests to measure personality of trapped individuals: an emergence test to assess boldness (Brown and Braithwaite 2005; Carter et al. 2013), an open-field test to measure activity and exploration in a novel environment (Walsh and Cummins 1976; Perals et al. 2017), and a handling bag test to measure docility and the response to handling by an observer (Carere et al. 2005; Boon et al. 2007; Montiglio et al. 2012; Taylor et al. 2014). We performed behavioral tests in the order above prior to handling or marking. All tests and processing occurred at a base area in the home grid of the focal individual. For detailed information about how behavioral tests were performed in the field, see methods described by Brehm et al. (2019).

After the completion of the behavioral tests, we recorded sex, body mass (measured using a 100 g Pesola Lightline spring scale), reproductive status (classed as either reproductively active or not based on the presence of an enlarged scrotum, perforated vagina, or signs of pregnancy or lactation), and age class (juvenile or adult). New individuals were anesthetized using isoflurane and tagged with PIT tags (Biomark MiniHPT8) subcutaneously at the mid back. Animals were also marked with a small animal ear tag (National Band, Style 1005-1) and a distinctive haircut. Haircuts were given using one or a combination of small cuts on the following locations: left shoulder, right shoulder, left mid-back, right mid-back, left rear, right rear. These cuts allow for visual identification in camera traps and are superior to methods using dye because they will show up in black and white photo and video. Once per month, we measured the body and tail length of captured individuals (while under anesthesia), and we released all individuals at the exact site of capture post-processing.

To quantify behavior from videotaped emergence and open-field tests, recordings were played back in the laboratory. For emergence tests, an observer recorded the following: the latency for the animal to emerge (defined as all four feet having left the trap), and the total time
spent at the end of the tunnel before emerging. It was determined that an animal was at the end of the tunnel if its nose protruded from the tunnel opening. Open-field tests were analyzed using the behavioral tracking software ANY-maze © (version 5.1; Stoelting CO, USA) to record each individual’s mean speed, distance traveled, relative location in the arena, and supplemental behaviors were recorded like grooming, rearing, and jumping. For the remainder of analyses, we focused on a reduced number of non-redundant and repeatable behavioral variables. See Brehm et al. (2019) for a complete list and biological interpretation of the behaviors used in this study system.

1.2.3. Monitoring capture events

To observe the event of an individual’s capture and calculate the time spent inside the trap before behavioral testing, we positioned camera traps (Bushnell NatureView HD 119740) facing the door of the Longworth trap and its surroundings. We monitored Longworth capture events using camera traps from July–October 2018 (936 total camera trap nights). Cameras were positioned ~50–100 cm from the trap at a height of ~50 cm. Thirteen camera traps were used in total and were positioned on a subset of the 100 available trap locations (Figure 1.2.). We chose camera locations to optimize the chance of observing capture events (hence, we chose trap locations that had successful captures during the previous month). Cameras were positioned simultaneously with Longworth traps and were kept active for the same duration as the traps (three consecutive days and nights at each study grid). We programmed cameras to record a one-minute video whenever movement was perceived (with a one-second delay between videos). Because camera traps occasionally fail to detect movement, we also programmed them to take a one-minute video once per hour (the “field scan” setting). This allowed us to approximate the hour of capture in an instance where the camera failed to trigger at the capture event.
Figure 1.2. A camera trap (Bushnell NatureView HD) monitors a Longworth trap in the field (circled in red).
Videos of capture events were played back in the laboratory, and an observer identified the individual by pairing the information of the date and trap with available capture data. The observer then recorded the time that the individual entered the trap and calculated the total time (in minutes) spent inside the trap before behavioral testing (taken from the time stamp of the open-field video for consistency). This variable will be referred to hereafter as “time in trap”.

1.2.4. Data analysis

To determine which behaviors could be considered personality, we first performed a repeatability analysis on the behavioral variables obtained from the emergence, open-field, and handling bag tests (Dingemanse and Dochtermann 2013; Cleasby et al. 2015). For this analysis, we used data from our study population collected during the 2016, 2017, and 2018 field seasons and used methods described in detail by (Brehm and Mortelliti 2018; Brehm et al. 2019a).

Once it was determined which behaviors were repeatable and could be considered personality, we sought to determine whether these behaviors would be influenced by the time spent inside the Longworth trap before behavioral testing. We used a nested hypothesis testing approach (Burnham and Anderson 2002a) using linear models and generalized linear models with the repeatable behaviors as response variables. We used only the individual’s behavioral measurement on the specific occasion when its capture was recorded on a camera trap. In the instances where we had repeated measures from the same individual (because we recorded their capture on a camera trap in subsequent trapping sessions - 18 out of 92 individuals), we used only the first event. This allowed us to avoid using mixed-effects models for only a few cases where measures were repeated (Zuur et al. 2009). Again, proportional response variables were logit-transformed to meet the assumptions of normality, and count variables were examined using generalized linear models with a poisson or negative binomial family (depending on dispersion).
We introduced predictor variables one by one to build a base model to control for most of the variability in the data. Predictor variables included sex, body condition, silvicultural treatment, trapping session, body mass, and a variable termed “naïve” which controlled for whether the animal had been captured previously or was naïve to trapping. Models containing each of these variables alone were compared to the null model using the Akaike information criterion corrected for small sample size (AICc) (Buckland et al. 1997; Burnham and Anderson 2002a) and models within 2.0 ∆AICc of the top model were considered to have equal support. If more than one model scored better than the null, we tested a model including multiple additive effects. Once this base model was built, we compared this model to the same model with the addition of the variable “time in trap” to see whether this addition improved the model by AICc. Previous research has shown that males and females may respond differently to trap-induced stress (Bosson et al. 2012), so we subsequently tested for an interaction between the time spent in the trap and sex. Last, to determine whether individuals who are naïve to trapping may be impacted by the time spent inside the trap differently than individuals who have been captured previously, we tested for an interaction between time spent in the trap and the variable “naïve”.

1.2.5. Ethical note

Animal trapping, handling, and marking procedures were approved by the University of Maine’s Institutional Animal Care and Use Committee (IACUC number A2015_11_02). Animals were anaesthetized with isoflurane prior to tagging, and tagging equipment was sanitized with 70% isopropyl alcohol in between animals. All small mammal handling was performed by trained researchers, and all efforts were made to minimize animal stress during the procedure.
1.3. Results

1.3.1. Repeatability analysis

We examined behavioral data collected over three trapping years in our study population from standardized tests for 1791 observations from 603 individual deer mice and 1558 observations from 529 individual red-backed voles. The mean number of repeated observations per individual was approximately three for both deer mice and red-backed voles, with a range of one to six observations. We selected seven significantly repeatable, non-redundant behavioral variables, with a mean repeatability value of 0.81 for deer mice and 0.78 for voles (Table A.1.). The mean 95% confidence intervals for these values were (0.79, 0.84) and (0.74, 0.81), respectively (Table A.1.). These highly repeatable behaviors can be considered personality in our study populations (Bates et al. 2015; R Core Team 2017). The number of observations and individuals shown in Table A.1. differ for behavioral variables obtained from the emergence and handling bag tests since these tests were not performed in 2016.

1.3.2. Trap confinement analysis

The mean time confined to a trap (in minutes) was 611, and the range was 74 minutes to 1085 minutes. This dataset included the capture events from 46 individual deer mice and 43 individual red backed voles for which we performed behavioral tests on the same occasion that a capture was recorded. In 12 out of 14 top models (~86%) predicting behaviors exhibited in standardized tests, the top model did not include “time in trap”. Instead, out of the predictor variables considered (sex, body condition, silvicultural treatment, trapping session, body mass, and the variable “naïve”) behaviors in deer mice were predicted by trapping session and body mass (Table 1.1.). Deer mice with greater body mass showed longer latencies to emerge from the emergence test and the proportion of time spent grooming in the open-field test correlated positively with trapping session (β = 0.26, SE = 0.08, R² = 0.20 and β = 0.58, SE = 0.16, R² =
0.23, respectively). In two cases, (once for deer mice and once for voles) the top model included an interaction between “time in trap” and whether or not the individual was naìve to trapping (Figure 1.3. a-b). Model fit was relatively low for top models (excluding those where the top model included only an intercept), with an average multiple R-squared value (R²) of 0.23 (Table 1.1.).
Table 1.1. Model output of top-ranked linear models* predicting behaviors performed during standardized tests in deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*).

### *P. maniculatus*

<table>
<thead>
<tr>
<th>Latency to emerge</th>
<th>β</th>
<th>St.Error</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>1.21</td>
<td>0.08</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Body mass</td>
<td>0.26</td>
<td>0.08</td>
<td>0.003</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prop. time grooming</th>
<th>β</th>
<th>St.Error</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-3.88</td>
<td>0.51</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Session</td>
<td>0.58</td>
<td>0.16</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

R-squared 0.20
Observations 41

### Prop. time center

<table>
<thead>
<tr>
<th>β</th>
<th>St.Error</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-3.52</td>
<td>0.123</td>
</tr>
<tr>
<td>Time in trap</td>
<td>0.17</td>
<td>0.12</td>
</tr>
<tr>
<td>Naïve</td>
<td>0.04</td>
<td>0.17</td>
</tr>
<tr>
<td>Time in trap*Naïve</td>
<td>-0.53</td>
<td>0.17</td>
</tr>
</tbody>
</table>

R² 0.19
Observations 46

### *M. gapperi*

<table>
<thead>
<tr>
<th>Handling time</th>
<th>β</th>
<th>St.Error</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>45.37</td>
<td>3.68</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Time in trap</td>
<td>-12.4</td>
<td>3.71</td>
<td>0.002</td>
</tr>
<tr>
<td>Naïve</td>
<td>6.04</td>
<td>4.53</td>
<td>0.19</td>
</tr>
<tr>
<td>Time in trap*Naïve</td>
<td>11.3</td>
<td>4.71</td>
<td>0.02</td>
</tr>
</tbody>
</table>

R² 0.28
Observations 43

* Only results from the top model (based on AICc scores) are shown. We have omitted occasions where the null model was the top model, and where the top model did not include the variable “Time in trap”. See materials and methods for more information.
Figure 1.3. Prior trapping experience influences the behavioral response to trap confinement in deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*). (a) Deer mice who were naïve to trapping showed a negative relationship between time in the trap and the proportion of time spent in the center portion of the open-field test. Non-naïve mice showed the reverse relationship. (b) Voles who were not naïve to trapping showed a negative relationship between time in the trap and handling time. Plotted are the relationships predicted from the top linear models and raw data points are shown. The variable “time in trap” has been z-standardized, and the variable “prop. time center” is on a logit scale.
1.4. Discussion

Though previous research has suggested that live trapping may produce a stress response in small mammals, in two small mammal species, prolonged confinement to a live trap does not seem to alter behavior. In an experiment wherein we studied the effects of live trapping on repeatable behavioral variables, our major findings were that for these species, 12 out of 14 behaviors exhibited during routine behavioral tests were not affected by the amount of time that individuals had spent confined in traps. In the two instances where the time spent confined in traps did predict behavior, effect sizes were relatively small, and the direction of the relationship was different for individuals who were naïve to trapping than those who had been trapped previously, indicating that an individual’s previous experience with a trap can influence whether or not trap confinement impacts behavior. Overall, these results suggest that personality data collected from wild, trapped small mammals is not confounded by the trapping process and, where an effect might be present, the predictive power of the time spent confined to traps is relatively weak and possibly not affecting the overall interpretation of results.

Although previous research has not explored the effects of live trapping on personality measurements specifically, studies have investigated the impacts of live trapping on hormonal stress responses and the findings have been mixed. It has been shown in southern red-backed voles and meadow voles that live trapping induces an initial stress response, but that this response is not heightened following prolonged confinement inside traps (Harper and Austad 2001; Fletcher and Boonstra 2006). In our study, the observed behavior of red-backed voles in behavioral tests was consistent with these findings and 6 out of 7 behaviors showed no correlation with the time that the animal had spent previously confined inside of a trap. Previous studies investigating the correlation between stress response and duration of trap confinement in deer mice saw that after prolonged time spent in traps, stress hormone levels were significantly higher than after a short duration of trap confinement (Harper and Austad 2001).
contrast, our results show no correlation between 6 out of 7 behavioral measurements and trap duration in the deer mouse. Although a hormonal change does not necessarily precede a change in behavior, we would expect to see an observable behavioral change in individual deer mice experiencing elevated stress levels (for example, by affecting behaviors that indicate activity level such as speed of locomotion and rearing). Instead, the one behavior in deer mice for which “time in trap” occurred in the top model was the proportion of time spent in the center of the open-field test, a behavior that is most commonly interpreted as indicating the degree of boldness (Choleris et al. n.d.; Barnett 1976; Treit and Fundytus 1989; Ramos et al. 1997; Eccard and Herde 2013; Gracceva et al. 2014). Interestingly, our results show that individuals who had never been trapped previously behaved more boldly in the open-field test (spending more time in the center portion) when their confinement duration was short rather than long. Individuals who had been trapped at least once previously showed the opposite effect; bolder behavior was seen in animals who had spent longer durations in the trap than those who had spent shorter durations (Figure 3a.). This finding suggests that deer mice show some degree of habituation to trapping, and that their experience during trap confinement is different on their first instance of capture than it is during subsequent captures. Those who have experienced trap confinement in the past behave in a more risky manner the longer their confinement duration lasts, which could be caused by increasing stress over time; however, without measuring stress hormones we cannot speculate on the mechanism driving this response. We did not observe any interaction between trap confinement duration and the amount of grooming that mice performed in the open-field test. This would be expected since grooming is commonly used to assess anxiety and stress in both a lab and field setting (Choleris et al. n.d.; Fernández-Teruel and Estanislau 2016; Kalueff et al. 2016). Further, this would not necessarily explain why mice experiencing their first capture behave in less risky manner the longer they spend confined to the trap, unless a naïve individual is more stressed by the initial trapping event than by the prolonged confinement in the trap. This is possible since these individuals do not have any prior experience with the tagging
and processing procedures that follow the trapping. Thus, the naïve individual may have less reason to remain stressed in an environment with ample dry bedding and bait. Again, we would expect to see some relationship with grooming in this scenario, but we cannot rule out that this may be a mechanism playing into this complex behavioral response.

In voles, the one behavior that was affected by the “time in trap” was handling time, or the amount of time spent immobile during a one-minute handling bag test. This behavior is commonly used to assess docility (Boon et al. 2007; Martin and Réale 2008; Montiglio et al. 2012; Taylor et al. 2014). Our results showed that for non-naïve individuals only (i.e., only those who had been trapped at least once previously), shorter durations in the trap correlated with increased docility (Figure 1.3. b.), and docility decreased with increasing duration. Similar to our results for deer mice, this response may be driven by the fact that for an experienced individual, the initial trapping event is less stressful than the period of trap confinement. Again, we saw no effect of confinement duration on grooming behavior, but cannot rule out that increasing stress levels may lead individuals to behave in a more docile manner.

Since 86% of observed behaviors by deer mice and voles showed no correlation with the variable “time in trap”, and all variables commonly used to indicate activity and anxiety showed no correlations, we suspect that the duration of trap confinement is not providing a prolonged stressor for small mammals. It may be noteworthy that the previous trap response studies of deer mice and voles (Harper and Austad 2001; Fletcher and Boonstra 2006) used Sherman traps instead of the Longworth traps used in this study. Longworth traps differ from Sherman traps in that they have a separate nest chamber (providing additional warmth and protection) which may help to limit stress. Additionally, we took further steps to minimize stress by ensure that bedding remained dry (i.e., limiting trapping in adverse weather and replacing damp bedding immediately), and providing ample bait inside the traps. Further, we checked traps twice within a 24-hr period to limit confinement durations (once in the morning, and once just
before dark). We cannot speculate about whether these precautions were adequate in our study to stop an increased stress response after the initial stressor of the trapping event, but regardless, prolonged confinement in a Longworth trap does not seem to result in an observable change for the majority of behaviors in either study species.

Future research examining this relationship in other species and other study populations will help to assess and confirm the generalizability of our findings. We suggest future studies quantifying the effects of trap confinement also include data on the physiological stress response, and consider non-repeatable behavioral traits along with personality traits. Furthermore, we suggest that other studies investigating personality in small mammals consider in analyses whether or not animals have been captured previously. Finally, the response to stressful situations (as in confinement during live-trapping), or an individual’s coping style, may itself represent an aspect of an animal’s personality (Koolhaas et al. 1999, 2010; Dammhahn and Almeling 2012). Within the coping styles framework, it would be interesting to explore to what extent an individual’s behavioral response to trap-induced stress might be plastic vs. relatively fixed over time. It is possible that with a repeated measures design, we could tease apart how much variability exists in the effects of trap confinement on observed behavior and what percent of this variability might be attributable to between individual differences.

Personality studies on wild populations will likely continue to become more common as further research demonstrates the cascade-effects that individual behavioral traits can have on populations and communities (Sih et al. 2012; Lapiedra et al. 2017, 2018a; Merrick and Koprowski 2017; Brehm et al. 2019a). Hence, it is critical to ensure that the very process we seek to illuminate is not being confounded by our methods of obtaining data. Our findings provide evidence that time spent inside of Longworth traps does not determine behaviors performed during standardized tests in two different small mammal species. Therefore, our
results suggest that personality measurements on wild, trapped small mammals are not regulated by trapping procedures.

1.5. Acknowledgments

This work was supported by the USDA-National Institute of Food and Agriculture McIntire-Stennis projects (ME041620 and ME041913) through the Maine Agricultural and Forest Experiment Station, and through a Grant-in-Aid of research from Sigma-Xi, the scientific research society. ST was supported by a grant from the University of Milano - Bicocca. We also thank a number of dedicated field and lab technicians for helping with data collection and video analysis, Bryn Evans for help with map preparation, Malcolm Hunter for comments on a previous version of this manuscript, Laura Kenefic (U.S. Forest Service), and Keith Kanoti for maintaining and facilitating research at the Penobscot Experimental Forest. We are especially thankful to Sara Boone for support in the field.

1.5.1. Author contributions

AM, ST, and AMB conceived and designed the experiment. ST and AMB performed the experiment. AMB analyzed the data and wrote the first draft of the manuscript. ST provided photographs. All authors contributed to the final version of the manuscript.
CHAPTER 2

LAND-USE CHANGE ALTERS ASSOCIATIONS BETWEEN PERSONALITY AND
MICROHABITAT SELECTION*


2.1. Introduction

Land-use change alters the structure and connectivity of landscapes (Lindenmayer et al. 2007; Turner et al. 2007), resulting in substantial changes to the functioning of ecosystems worldwide (Steffen et al. 2004; Dirzo et al. 2014). Managed forests represent one of the most widespread land-uses, with roughly 71% of the earth’s forests actively managed for timber and other commodities (IPCC 2019). The management of forested land for timber, biofuels, and recreation involves changing forest stand structure and composition to meet management goals, often diminishing or degrading key habitat features for wildlife (Fisher and Wilkinson 2005; Zwolak 2009; Gasperini et al. 2016). The way wildlife populations respond to the loss or degradation of key habitat features is, in part, a culmination of the response of each individual in the population. Consequently, we must understand how land-use change alters individual-level resource use and behavior if we are to predict how populations and communities will respond to land-use change (Miranda et al. 2013; Sutherland et al. 2013).

For years, ecologists assumed that all individuals within a species moved throughout the landscape and utilized resources similarly, following the traditional ideas of resource partitioning (Hutchinson 1957; Schoener 1974). The niche variation hypothesis (Van Valen 1965; Roughgarden 1972), however, posits that populations can exhibit among-individual variation in resource use, and this topic has gained traction in recent years (Bolnick et al. 2003; Costa-Pereira et al. 2018). The ability of conspecifics to use different resources and occupy functionally different niches likely facilitates the response of populations under changing conditions (Moran et al. 2016). Therefore, identifying the effects of a changing environment on
intraspecific variation in habitat selection should be of great interest to ecologists and
conservationists alike (Sih et al. 2011; Moran et al. 2016).

Several studies have highlighted the ecological importance of inter-individual differences
in resource use and spatial distribution (Araújo et al. 2011; Bolnick et al. 2011; Schreiber et al.
2016). Such differences can be driven by sex, age classes, morphometric differences, and
individual specialization (Bolnick et al. 2003; Zwolak 2018). Notably, one driver of individual
specialization shown to affect resource use and spatial distribution is animal personality (Boon
et al. 2008; Kobler et al. 2009; Boyer et al. 2010; Pearish et al. 2013; Ghanit et al. 2020; Wat et
al. 2020). *Personality* refers to consistent individual differences in the behavior of conspecifics
(Gosling 2001; Sih et al. 2004b; Stamps and Groothuis 2010b). A growing body of research
highlights the existence of associations between personality types and habitat selection, in
which individuals with certain personality types occur disproportionately in certain physical or
social environments (for example, because certain personality types select specific
microhabitats or select areas with high densities of similar personality types) (Wilson et al. 1993;
Bergmüller and Taborsky 2010; Réale et al. 2010a; Stamps and Groothuis 2010b; Schirmer et
al. 2019, 2020). This has been referred to as *spatial niche specialization* (Schirmer et al. 2019,
2020). A range of mechanisms may generate correlations between personality type and habitat
selection across landscapes (Plomin et al. 1977; Stamps and Groothuis 2010a); eight are
described in Table 2.1.
Table 2.1. Mechanisms with the ability to generate behavioral correlations between personality and habitat selection across landscapes. Provided are the mechanism, a brief description, and a non-exhaustive list of references. Note that these mechanisms are not all mutually exclusive

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>Description</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Niche-picking</td>
<td>individuals of a certain personality type seek out certain habitats or features of the environment</td>
<td>(Stamps and Groothuis 2010a; b; Sih et al. 2018; Schirmer et al. 2019)</td>
</tr>
<tr>
<td>Niche-construction</td>
<td>individuals of a certain personality type modify selective environments</td>
<td>(Odling-Smee et al. 1996; Laland et al. 2016)</td>
</tr>
<tr>
<td>Matching habitat choice</td>
<td>the habitat chosen via niche-picking is that which best “matches” an individual’s abilities to use this environment</td>
<td>(Edelaar et al. 2008)</td>
</tr>
<tr>
<td>The environment influences behavior</td>
<td>the environment itself can influence the behavior of individuals; for example, safe environments can encourage individuals to behave more boldly</td>
<td>(Tuttle and Ryan 1982; Sharpe et al. 1998; López et al. 2005; Webster et al. 2007; Peluc et al. 2008)</td>
</tr>
<tr>
<td>Habitat-specific mortality</td>
<td>certain personality types are more or less likely to survive in specific environments; may be a result of matching habitat choice</td>
<td>(Jaenike and Holt 1991; Edelaar et al. 2008)</td>
</tr>
<tr>
<td>Density-driven assortment</td>
<td>certain personality types may seek out areas of high or low density</td>
<td>(Pearish et al. 2013; Sih et al. 2018; Schirmer et al. 2019)</td>
</tr>
<tr>
<td>Congregating with similar individuals</td>
<td>individuals may congregate in areas with similar personality types</td>
<td>(Johnson et al. 2017)</td>
</tr>
<tr>
<td>Social-networks</td>
<td>conspecifics may distribute within social networks according to personality traits</td>
<td>(Croft et al. 2009; Snijders et al. 2014; Best et al. 2015; Carter et al. 2015)</td>
</tr>
</tbody>
</table>
When anthropogenic changes alter structural characteristics of the environment, the capacity for individuals in a population to adapt will determine the ability of that population to persist. Ultimately, shifts in personality-associated habitat selection caused by land-use change may be an important mechanism affecting the response of populations and communities to global change. However, although a few studies have recognized the capacity for personality traits to influence habitat selection and resource use (Wilson et al. 1993; Bergmüller and Taborsky 2010; Réale et al. 2010a; Stamps and Groothuis 2010a; Schirmer et al. 2019, 2020), previous research has not assessed the extent to which land-use change might affect this process. This knowledge would provide novel justification for the importance of landscape complexity and heterogeneity. That is, it would show that by maintaining landscape complexity and heterogeneity we may maintain suitable habitat for varying personality types and contribute to fostering intraspecific trait diversity.

To address this knowledge gap, we tested the hypothesis that land-use change alters individual patterns of microhabitat selection by small mammals. Specifically, we investigated two main questions (Figure 2.1.): 1) to what extent are personality type and microhabitat selection correlated among conspecifics? and 2) does land-use change (specifically, silvicultural practices) alter these individual patterns of microhabitat selection? We predicted to see correlations between personality traits (such as boldness and activity rates) and selection for major structural components in the forest like the amount of vegetative cover or coarse woody debris (i.e., a mechanism referred to in the literature as niche-picking; Table 2.1.). Specifically, we expected that timid small mammals would select for areas with more vegetative cover than bolder individuals (Carrete and Tella 2010; Holtmann et al. 2017) since bold individuals often take more risks and suffer higher predation (Réale et al. 2010b; Dammhahn et al. 2018) and that more active individuals would select areas with larger, cone-bearing trees due to increased metabolic needs (Biro and Stamps 2010).
Figure 2.1. A conceptual diagram identifying the two main objectives of this study. Deer mice (Peromyscus maniculatus) and southern red-backed voles (Myodes gapperi) were trapped, individually marked, and personality traits were measured using three standardized tests (i.e., an emergence test, an open field test, and a handling test. The open field test is pictured here). We next identified correlations between personality traits (i.e., activity level, exploration, docility, and boldness) and selection for key microhabitat features. We predicted that personality traits such as activity level would correlate with the tendency to select for different microhabitat features (such as tree size - pictured here). Among-individual correlations between personality traits and microhabitat selection were compared across three forest types to assess whether silvicultural practices shift patterns of microhabitat selection.
Further, since a primary mechanism thought to maintain personality variation within animal populations is fluctuating selection associated with environmental heterogeneity (Sih et al. 2004a; Réale et al. 2010b; Le Cœur et al. 2015), we predicted that in forests treated with different silvicultural manipulations, there would be observable differences in the correlation between personality traits and selection for major structural components. We expected that these differences would be associated with the activity rates, docility, or boldness/risk-taking capacities of individuals and explained by variation in key resource availability between forest treatments. For example, we expected this selection to be more prominent where important microhabitats are less abundant (i.e., that correlations between boldness/timidness and coarse woody debris would be stronger in managed forest types where these features are scarcer and only the more bold or aggressive individuals gain access). Additionally, since typically decisions made on broader scales have greater impacts on fitness (Rettie and Messier 2000; Dupke et al. 2017), we expected that patterns of selection would be stronger at coarser ecological scales (i.e., the selection of the home range vs. the selection of daily capture locations).

To achieve our objectives, we conducted a large-scale field experiment over four years wherein we contrasted two replicates of unmanaged forest (control) with four areas of managed forest (two silvicultural treatments - two replicates each). The study was conducted in Maine (USA); target species were the deer mouse (Peromyscus maniculatus) and the southern red-backed vole (Myodes gapperi). We used small mammals as model species because standard methods exist to assess their personalities (Carter et al. 2013; Brehm and Mortelliti 2018; Mazzamuto et al. 2018; Brehm et al. 2020) and microhabitat selection (Dueser and Shugart 1978; Longland and Price 1991; Mortelliti and Boitani 2007; Kellner et al. 2013).
2.2. Materials and methods

2.2.1. Study area and small mammal trapping

This study was conducted at the Penobscot Experimental Forest (44° 51’ N, 68° 37’ W) in Maine, USA. Here, different compartments have been logged separately, managed with contrasting silvicultural treatments, and replicated twice in a randomized experimental design. Approximately 25 ha of forest has been retained in two separate units and left unmanaged since the late 1800s to serve as reference. The Penobscot Experimental Forest is a mixed conifer-deciduous forest (Brissette and Kenefic 2014) and is dominated by shade-tolerant conifers including balsam fir (*Abies balsamea*), red spruce (*Picea rubens*), and Eastern hemlock (*Tsuga canadensis*) (Kimball 2014).

From June through October during four consecutive years (2016–2019) we implemented a large-scale capture-mark-recapture experiment in four areas (representing two contrasting silvicultural treatments) and used two areas of unmanaged forest as reference sites. The treatment areas used were as follows: treatment 1 (even-aged cut) and treatment 2 (shelterwood cut with reserves). Due to the contrasting silvicultural systems, these forest stands differ greatly in the understory density, diameter of trees present, light levels, and quantity of downed woody material and snags. These varying structural characteristics have generated highly contrasting habitat types for small mammals (for further information about the forestry treatments used, a brief description of each site, and photographs see Appendix B: Table B.1.).

We trapped small mammals in six separate study grids (two control grids in reference areas and two grids in each of two replicated treatment areas). The area of the treatments used in this study was 12.8 ha on average (range: 9.49 - 19.39 ha). Each trapping grid was 0.81 ha in area and was positioned at or close to the center of the treatment area to minimize edge effects. The mean home range for *Peromyscus* in this study system was 0.34 ha (range: 0.08-1.01 ha),
and for *Myodes* was 0.33 ha (range: 0.05 - 0.87 ha). Home range calculations were estimated using the `getverticeshr` command from the ‘adehabitatHR’ package in program R to extract the kernel home range contours with a 75% home range estimation (Calenge 2006). Trapping grids consisted of 100 flagged points spaced 10 m apart. The mean distance between grids was approximately 1.42 km, and the mean distance between duplicate grids of the same treatment was approximately 1.44 km. We placed one Longworth small mammal trap at each flagged point. We baited traps with a mixture of sunflower seeds, oats, and freeze-dried mealworms, and bedded traps with cotton stuffing. We checked Longworth traps twice per day (just after sunrise and in the late afternoon). We trapped at each grid for three consecutive days and nights once per month for five consecutive months each year totaling over 35,000 trap nights (trap nights = number of active traps * number of nights).

2.2.2. Animal tagging and behavioral tests

Before animals were handled or tagged, we used three standardized tests to measure behaviors that would later be used to assess personality. An *emergence test* was used to assess boldness (Brown and Braithwaite 2005; Carter et al. 2013), an *open-field test* to measure activity and exploration in a novel environment (Walsh and Cummins 1976; Perals et al. 2017), and a *handling bag test* to measure docility and the response to handling by an observer (Martin and Réale 2008; Montiglio et al. 2012; Taylor et al. 2014). We performed behavioral tests once monthly to ensure that animals would not habituate to the tests. See (Brehm et al. 2019a, 2020a) and Appendix B for detailed behavioral test procedures.

After the behavioral tests were complete, we anesthetized animals with isoflurane and inserted PIT tags (Biomark MiniHPT8) subcutaneously at the midback. Animals were also marked with a small animal ear tag (National Band, Style 1005-1). We recorded sex, body mass (measured using a 100 g Pesola Lightline spring scale), body length, tail length, and age class
(juvenile, subadult, or adult; based on body size and pelage coloration). Animals were released at the site of capture post-processing.

To quantify behavior from videotaped emergence and open-field tests, recordings were played back in the laboratory. For emergence tests, an observer recorded whether or not the animal emerged (defined as all four feet having left the Longworth trap), the latency to emerge, and the total time spent at the end of the Longworth tunnel before emerging. When an individual did not emerge from the test after the three-minute cutoff, the latency to emerge was set to 1.25x the maximum test length. Open-field tests were analyzed using the behavioral tracking software ANY-maze © (version 5.1; Stoelting CO, USA). To assure the independence of response variables in our analyses, all behavioral variables were screened for correlation before analysis (using R<0.7 as threshold) (Dormann et al. 2013). See Appendix B: Table B.2. for pairwise correlations between all behavioral variables used for each study species, and Table 2.2. (modified from Brehm et al. 2019b) for a complete list of the behaviors used, their description and interpretation, and supporting sources. Briefly, we used the following variables: handling time (the number of seconds immobile in a handling test; note that the term handling time should not be confused with a term sometimes used in ecological literature to indicate the time spent handling, processing, and consuming food items), latency to emerge and time at tunnel end (from the emergence test), mean speed (in Peromyscus only), rear rate, proportion of time grooming, and proportion of time in the center (from the open-field test).
Table 2.2. Repeatable behavioral traits from three different behavioral tests performed on deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapper*).

These variables are considered personality traits, and this table provides a guide for their interpretation, a non-exhaustive list of citations supporting these interpretations, the behavioral test each trait was measured using, and a description of the behavior. This table was interpreted from Brehm *et al.* (2019), and a more thorough behavioral interpretation is provided there.

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Personality trait</th>
<th>Behavioral test</th>
<th>Description</th>
<th>Notes about interpretation</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handling time</td>
<td>Docility</td>
<td>Handling bag</td>
<td>Total number of seconds of inactivity during a 1-minute handling bag test</td>
<td>Interpreted as a measure of docility or as a response to stressful confinement.</td>
<td>(Boon <em>et al.</em> 2007, Martin and Réale 2008, Montiglio <em>et al.</em> 2012, Taylor <em>et al.</em> 2014)</td>
</tr>
<tr>
<td>Latency to emerge</td>
<td>Bold/timid</td>
<td>Emergence</td>
<td>Latency (in seconds) to emerge from trap in the emergence test. An animal was considered to have emerged when all four feet left the trap</td>
<td>The latency to emerge from a shelter and into a novel or open environment is commonly assessed on a timid/bold continuum where increased latency signals increased timidness.</td>
<td>(Brown and Braithwaite 2004, Carter <em>et al.</em> 2013, Gracceva <em>et al.</em> 2014)</td>
</tr>
<tr>
<td>Time at end of tunnel</td>
<td>Bold/timid</td>
<td>Emergence</td>
<td>Total number of seconds spent at the end of the tunnel before emerging</td>
<td>We interpret increased time at the end of the tunnel as a sign of timidness. These individuals had a more timid/fearful behavioral tendency and required time to survey the arena before emergence.</td>
<td></td>
</tr>
<tr>
<td>Parameter</td>
<td>Activity</td>
<td>Description</td>
<td>Source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>----------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean speed</td>
<td>Open field</td>
<td>Mean speed in the open field test in (m/s). Calculated by dividing the total distance traveled in the test by the test duration. This is a direct measure of locomotion and activity in the open field test.</td>
<td>(Carter et al. 2013, Gracceva et al. 2014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion time grooming</td>
<td>Open field</td>
<td>Proportion of test duration spent grooming. Grooming in small mammals is an indicator of anxiety and stress. In the deer mouse, a nocturnal species, the open-field test represents an environment of high aversiveness and increased grooming suggests lower anxiety and better coping. In contrast, for the vole (a more diurnal species) low to moderate grooming seems to signal coping, whereas high amounts of grooming indicate high anxiety.</td>
<td>(Choleris et al. 2001, Kalueff et al. 2016, Fernández-Teruel and Estanislau 2016)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rear rate</td>
<td>Open field</td>
<td>Rate of rearing (rears/s). Rearing is defined as forelegs leaving the arena floor. Commonly assessed as correlating positively with activity, and represents part of the animal's <em>diversive</em> exploration, or the <em>search phase</em> of exploratory behavior.</td>
<td>(Choleris et al. 2001, Prut and Belzung 2003, Martin and Réale 2008, Tanaka et al. 2012)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.2 Continued

<table>
<thead>
<tr>
<th>Proportion of test duration spent in the center portion of the arena</th>
<th>Entering into the center, “unsafe” areas in the open-field arena signifies boldness and avoidance of these areas indicates fearfulness.</th>
</tr>
</thead>
</table>

2.2.3. Repeatability of behavior in standardized tests

To determine which behaviors could be considered personality, we calculated the adjusted repeatability and associated 95% confidence intervals for key behaviors performed in the standardized tests using the ‘rptR’ package in R (Stoffel et al. 2017). Statistically, repeatability refers to the proportion of the total phenotypic variation that can be attributed to individual differences (Dingemanse et al. 2009; Dingemanse and Dochtermann 2013; Schielzeth et al. 2020). Practically, this means that repeatability equals the between-individual variance divided by the total phenotypic variance (between-individual variance + within-individual variance). The random effect of individual ID was included in the models and fixed effects included sex, body condition (calculated using the scaled-mass index) (Peig and Green 2009), silvicultural treatment, and trapping session. In all models, we used 1000 parametric bootstraps and 100 permutations. We used Box-Cox transformations on the response variable to approach normality when necessary (Box and Cox 1964; Yang et al. 2011) and we assessed normality by visually inspecting Q-Q plots and by plotting the fitted values against the residual values. Note that we use the term “repeatability” hereafter, but that these estimates are “adjusted repeatabilities” because they have been estimated from mixed models fit with fixed effects (Wilson 2018).
We considered any behavioral trait with a 95% CI for repeatability that excluded zero to be a personality trait (Nakagawa and Schielzeth 2010), but we emphasize that this classification as significantly repeatable does not say anything about the strength of repeatability. We used the raw behavioral variables themselves as measures of personality (Dingemanse et al. 2004, 2012), instead of using a principal components analysis (PCA) or similar dimensionality reduction technique (Boon et al. 2007; Martin and Réale 2008). We made this choice to maximize the variation in behavioral measurements, to remain consistent with our previous work (Brehm and Mortelliti 2018; Brehm et al. 2019a, 2020a), and because components retained from a PCA can sometimes become difficult to interpret biologically (Lever et al. 2017).

2.2.4. Microhabitat selection

Microhabitat selection refers to the disproportionate use of certain microhabitat features in relation to what is available in the environment (Johnson 1980; Jones 2001). The theory of habitat selection is well established, and it has been long recognized that selection occurs at multiple hierarchical orders (i.e., animals select for different habitat features at different spatial and temporal scales) (Johnson 1980; Krausman 1999; Manly et al. 2002; Mayor et al. 2009; McGarigal et al. 2016). This includes broad-scale selection of the geographic range to the fine-scale selection of microhabitat features. In this study, we were interested in selection occurring at the second-order, or the selection of a home range (i.e., where an individual positions its home range), and selection at the third-order, or the selection of microhabitat features within the home range (Figure 2.2.). For a glossary of terms associated with habitat selection and hierarchical orders, see Appendix B: Table B.3.
Figure 2.2. Habitat selection by small mammals occurs at multiple spatial scales. **Second-order selection (A):** within the trapping grid, individuals select a home range. This home range may contain more or less of a specific habitat feature than is available on average in the trapping grid. **Third-order selection (B):** Within the home range, individuals may more frequently select sites that contain more of a specific habitat feature than is available on average in the home range.

In July 2017, we recorded detailed microhabitat measurements within a 5m radius of each trap site (100 sites per trapping grid, 600 sites total). Microhabitats include structural components of the environment such as fallen logs and woody vegetation cover. Structural components such as these can shift drastically following silvicultural disturbance and result in changes to microhabitat use by small mammals (Kellner and Swihart 2014). In line with previous literature on small mammal microhabitat use (Dueser and Shugart 1978; Price and Kramer 1984; Mortelliti and Boitani 2007; Hille and Mortelliti 2010; Fukasawa 2012), variables measured included: percent cover of mosses, herbs and forbs, percent cover of shrubs and saplings at three height categories, meters of coarse woody debris present (in two size categories), percent canopy cover, and the diameter at breast height (DBH) of the largest tree. See Appendix B: Table B.4. for a brief description of each microhabitat variable and more detail about methods of measurement.
We used a principal components analysis (PCA) to reduce the number of measured microhabitat variables and to create meaningful “components” which would explain the structural variation in our study areas. We used a simplimax rotation (Kiers 1998), which is a form of oblique rotation that optimizes for simplicity in the structure of the pattern matrix while minimizing the correlation between components. The Kaiser-Guttman criterion was used to select the number of principal components to retain (Kaiser 1991).

### 2.2.5. Third-order selection

For each microhabitat component, we calculated a variable that would encompass an individual’s tendency to select for that component at the third-order of selection (or selection for specific components of habitat within the home range) (Johnson 1980). To calculate the true selection of a resource, it is necessary to take the availability of that resource into account; since selection is defined as use that is disproportionate to a resource’s availability (Johnson 1980; Mayor et al. 2009). To do this, we subtracted the mean microhabitat component value in the neighborhood of available traps from the component value at an individual’s capture site. Through this subtraction, we were able to quantify whether an individual selected for higher or lower than average microhabitat features. Therefore, positive selection values indicate that the trap site of choice had a greater component value than was available on average, and negative values indicate the opposite. We z-standardized this value of selection by dividing by the standard deviation (meaning that the value would be given in units of standard deviation from the mean).

\[
Select\ PC1 = \frac{PC1_{\text{capture site}} - PC1_{\text{neighborhood}}}{sd(\text{neighborhood})}
\]

The neighborhood radius for each species was calculated by estimating the mean distance between consecutive capture locations, pooled over individuals (calculated using the ‘secr’ package in R [Efford 2021]). This rendered a species-specific estimate for daily activity per
trapping grid, and we used the overall mean per species as the neighborhood radius. Within this neighborhood radius, we included only trap sites that were available (had no capture on the specific trap occasion). Neighborhood radii were as follows: *Peromyscus* (mean: 28.05 m, range: 21.56 m to 46.11 m); *Myodes* (mean: 24.39 m, range: 10 m to 47.89 m).

If daily selection for microhabitat is repeatable among individuals, this suggests limited behavioral plasticity. Therefore, using previously described methods for repeatability analysis, we assessed the extent to which individuals differed consistently in their daily microhabitat selection. We calculated the adjusted repeatability of the microhabitat selection variables using silvicultural treatment, sex, body condition, body mass, and trapping year as fixed effects, and including the random effect of individual ID.

Data used in third-order analyses included 1716 observations from 570 individual deer mice and 1383 observations from 475 individual southern red-backed voles (mean number of observations per individual deer mouse was 3.01 ± 2.89 [range 1 to 29] and per southern red-backed vole was 2.91 ± 2.61 [range 1 to 18]). See Appendix B: Figure B.1. for plots showing the number of captures per individual used in these analyses.

**2.2.6. Second-order selection**

For each microhabitat component, we calculated a variable that would encompass an individual’s tendency to select for that component at the second-order of selection (or selection of the home range). To do this, we removed all individuals who had fewer than four captures. Then, for the remaining individuals, we calculated the mean value of each microhabitat component across all of their capture locations. We then subtracted the mean value available in the trapping grid from each individual’s mean component value following the same procedures as above (and z-standardizing once more). Positive values indicate that an individual’s home
range had a greater component value than was available on average in the trapping grid, and negative values indicate the opposite.

\[
\text{Select PC1 home range} = \frac{PC1 \text{ home range} - PC1 \text{ grid}}{sd(grid)}
\]

Because this analysis rendered one unique value per individual, no repeatability analysis was performed on the second-order selection values.

Data used in second-order analyses included 1071 observations from 162 individual deer mice and 823 observations from 129 individual voles (mean number of observations per individual deer mouse was \(6.61 \pm 3.14\) (range 4 to 29) and per southern red-backed vole was \(6.38 \pm 2.65\) (range 4 to 18)).

2.2.7. Assessing correlations between personality traits and microhabitat selection

Finally, we determined whether personality traits and microhabitat selection at the third-order were correlated using multi-response Bayesian models fit with Markov chain Monte Carlo methods using the ‘MCMCgllmm’ statistical package for R (Hadfield et al. 2010). We estimated the correlation between the habitat selection variables described above and the behavioral variables as measured on the day of this habitat selection (i.e., on the day that each trap site was chosen). When an individual was captured on an occasion when a behavioral test was not performed (such as when weather conditions were too poor to perform behavioral tests) we used the individual’s most recent behavioral measurement. This gave us a repeated measures design and allowed us to measure the among individual-level covariance between traits which can then be used to derive an among individual-level correlation between these traits. We chose these methods as opposed to using the mean behavior or a best linear unbiased predictor (BLUP) (Houslay and Wilson 2017b).
We used non-informative (parameter expanded) priors for both the individual and residual variances of the personality variables and the microhabitat selection variables (prior specifications: R structure degree of belief (nu) = 0.002; G structure degree of belief = 2.0, alpha.mu = rep(0, 2), alphaV = diag(25^2, 2, 2) (Hadfield 2015; Ólafsdóttir et al. 2016; Houslay and Wilson 2017a; Brehm and Mortelliti 2018). We scaled the response variables using the scale function in Base R to help with model fit and ease of interpretation. We fit an unstructured covariance matrix for the grouping variable of individual ID, which allowed us to calculate the variance in each response variable due to differences among individuals as well as the covariance between these variances. For the residual variation (or the within-individual variation), we constrained these to equal zero since, although both traits have repeated measures, they were not measured at the same time (Dingemanse and Dochtermann 2013; Hadfield 2015). Models were run with a burn-in of 25,000 and subsequent 1,000,000 iterations (thinning interval of 500). This resulted in MCMC sample sizes of 1950. We inspected plots of traces and posterior distributions to confirm convergence (Hadfield et al. 2010) and confirmed that autocorrelation between samples was low using the R package ‘coda’ (Monnahan et al. 2017). We performed analyses on each species separately, and we inspected the correlation between the response variables for each silvicultural treatment separately and compared these correlations among treatments. In all, we have run 14 models per treatment for Peromyscus (7 personality traits x 2 microhabitat components) and 12 models per treatment for Myodes (6 personality traits x 2 microhabitat components).

Fixed effects included sex, body condition (scaled and centered), body mass (scaled and centered - used as a proxy for age), and year. To obtain a posterior distribution of the correlation between the dependent variables, we standardized the bivariate model covariances on a scale from -1 to 1 by dividing the corresponding covariance between the two variables by the product of the square root of their variances (Houslay and Wilson 2017a). We assessed the 95%
confidence interval of the correlation and considered intervals that excluded zero to signal significance.

We assessed whether the repeatable behavioral variables correlated with microhabitat selection at the second-order following similar methods as above, however, as outlined by (Houslay and Wilson 2017a), we fixed the within-individual variance in the non-repeated measures to equal a very small positive number since a trait measured only once has no within-individual variance (Houslay and Wilson 2017a). We included the same fixed effects as in the univariate models.

All research was conducted in accordance with and approved by the University of Maine’s Institutional Animal Care and Use Committee (IACUC number A2015-11-02 and A2018-11-01).

2.3. Results

2.3.1. Microhabitat structure

Following dimension reduction of the trap-level microhabitat characteristics with PCA, we retained two components (Table B.5.). The third component was very close to the cutoff using the Kaiser-Guttman rule and so we decided to focus on the first two components (Figure B.2.). The first component (PC1) characterized a gradient ranging from areas with a dense ground cover (mosses, herbs, and forbs) and a more open canopy towards areas with less ground cover and a closed canopy and the second component (PC2) had the highest loadings for the abundance of coarse woody debris (CWD) present and the diameter at breast height (DBH) of the largest tree. We, therefore, interpreted the first component to be a measure of ground cover and light levels and the second to represent key resource availability (seed-bearing trees and CWD).
2.3.2. Trapping and behavioral data

Over four trapping years in our study population (June through October 2016-2019) we obtained behavioral data (via standardized tests) from 577 individual deer mice and 473 individual southern red-backed voles. Both species showed repeatability for all behavioral variables measured. Mean repeatability was 0.346 for deer mice (mean CI = 0.26 to 0.45) and 0.221 for voles (mean CI = 0.12 to 0.37). See Table B.6. for all repeatability estimates and sample sizes.

Both behavioral data and microhabitat selection data (via capture site) were available for 3099 observations from 1045 total individuals from the two study species. We found selection at the third-order for microhabitat characteristics to have low but significant repeatability in all instances (Table B.7.).

2.3.3. Effects of land-use change on personality-associated microhabitat selection

Individual selection for certain components of microhabitat structure and one or more personality traits were significantly correlated across all individuals at one or both orders of selection in both study species. Moreover, the strength and direction of the correlations differed among the unmanaged forest and silvicultural treatments (Table 2.3.). For example, we commonly saw that relationships that were present in the unmanaged forest were lost in one or both silvicultural treatments. For instance, we observed a positive relationship between activity and exploration in an open field test (based on the rate of rearing) and selection for areas within the home range (third-order selection) with increased ground cover and light levels (PC1) in both deer mice and southern red-backed voles (deer mice: mean correlation = 0.38, CI = 0.04 to 0.68; voles: mean correlation = 0.40, CI = 0.11 to 0.68), but this relationship was not significant in either the even-aged or shelterwood treatments (Figure 2.3.). In addition, out of seven relationships in the even-aged forest, five were unique to this treatment and all four relationships seen in shelterwood forest were only present in that treatment.
Figure 2.3. Personality-microhabitat associations are altered by silvicultural practices in deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*). In unmanaged forests, more active and exploratory mice and voles select areas within their home range (third-order selection) with more ground cover and higher light levels (PC1) than less active and exploratory individuals, but this relationship was absent in all of the forest treatments. Activity/exploration was interpreted from the rate of rearing in an open-field test. Correlations were estimated as the posterior mean for a correlation based on Markov chain Monte Carlo (MCMC) Bayesian models controlling body mass, body condition, trapping year, and sex. Estimates in the unmanaged forest were as follows: (*Peromyscus maniculatus*) Posterior mean = 0.38, 95% credible interval (0.04, 0.68). (*Myodes gapperi*) Posterior mean = 0.40, 95% credible interval (0.11, 0.68).
Table 2.3. Correlations between repeatable behavioral variables and selection for two components of microhabitat structure at the second and third-orders of selection in deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*).

<table>
<thead>
<tr>
<th>Species</th>
<th>Order of selection</th>
<th>Reference</th>
<th>Even-aged</th>
<th>Shelterwood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Corr(x,y)</td>
<td>Correlation</td>
<td>95% CI</td>
</tr>
<tr>
<td><em>Peromyscus maniculatus</em></td>
<td>Third-order</td>
<td>(PC1, rear rate)</td>
<td>0.38**</td>
<td>(0.04, 0.68)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(PC1, mean speed)</td>
<td>0</td>
<td>(-0.30, 0.28)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(PC1, prop. groom)</td>
<td>0.2</td>
<td>(-0.70, 0.47)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(PC1, handling time)</td>
<td>-0.35**</td>
<td>(-0.70, 0.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(PC1, time tunnel)</td>
<td>-0.35</td>
<td>(-0.56, 0.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(PC2, mean speed)</td>
<td>-0.08</td>
<td>(-0.43, 0.44)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(PC2, handling time)</td>
<td>-0.03</td>
<td>(-0.34, 0.34)</td>
</tr>
<tr>
<td><em>Myodes gapperi</em></td>
<td>Third-order</td>
<td>(PC1, rear rate)</td>
<td>0.4**</td>
<td>(0.11, 0.68)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(PC1, prop. groom)</td>
<td>-0.05</td>
<td>(-0.69, 0.27)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(PC1, latency emerge)</td>
<td>-0.37**</td>
<td>(-0.58, 0.02)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(PC2, prop. groom)</td>
<td>-0.07</td>
<td>(-0.98, 0.44)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(PC2, handling time)</td>
<td>-0.67**</td>
<td>(-0.98, 0.38)</td>
</tr>
<tr>
<td></td>
<td>Second-order</td>
<td>(PC2, prop. groom)</td>
<td>-0.06</td>
<td>(-0.38, 0.24)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(PC2, handling time)</td>
<td>-0.49**</td>
<td>(-0.77, 0.18)</td>
</tr>
</tbody>
</table>
Table 2.3 Continued

| (PC2, time tunnel) | -0.29 | -0.62 | -0.51** | -0.83 | -0.09 | -0.40, 0.21 |

Correlations were estimated as the posterior mean for a correlation based on Markov chain Monte Carlo (MCMC) Bayesian models controlling body mass, body condition, trapping year, and sex. Behavioral variables involved in significant correlations include mean speed in the open-field, rears per second in the open field (rear rate), proportion of time grooming in the open field (prop. groom), latency to emerge from the emergence test (latency emerge), time spent at the end of the emergence tunnel (time tunnel), and seconds immobile in a handling test (handling time). Only results that are significant in at least one treatment are shown.

Data used in third-order analyses included 1716 observations from 570 individual deer mice and 1383 observations from 475 individual southern red-backed voles (mean number of observations per individual deer mouse was 3.01 ± 2.89 (range 1 to 29) and per southern red-backed vole was 2.91 ± 2.61 (range 1 to 18)). Data used in second-order analyses included 1071 observations from 162 individual deer mice and 823 observations from 129 individual voles (mean number of observations per individual deer mouse was 6.61 ± 3.14 (range 4 to 29) and per southern red-backed vole was 6.38 ± 2.65 (range 4 to 18)).

* 95% CI includes, but does not cross, zero; ** 95% CI does not include or cross zero
In red-backed voles, we observed a correlation between docility level (based on handling time) and microhabitat selection at the second-order (home range selection), and our results showed that this effect was also present in one of the forest treatments (Figure 2.4.). In this instance, we found a negative relationship between docility level and selection for PC2 in the reference forest (mean correlation = -0.49, CI = -0.77 to -0.18). Higher values for this second component are representative of larger trees and hence more abundant food resources (i.e. seeds) and increased structural attributes important for small mammals (runways via coarse woody debris). In the even-aged forest, this negative relationship was less strong, but still significant (mean correlation = -0.40, CI = -0.77 to -0.01). Additionally, our results showed that docile red-backed voles also selected areas with less coarse woody debris and smaller trees than was available on average within their home ranges (third-order selection) compared to less docile, more aggressive individuals (mean correlation = -0.67, CI = -0.98 to -0.38). However, at this third-order, the correlation was only significant in the unmanaged forest (Figure 2.4.).
2.4. Correlations between docility (in a handling bag test) and selection for PC2 at the second and third-order in southern red-backed voles (*Myodes gapperi*). Selection for tree size and coarse woody debris (PC2) decreases with increasing docility in the unmanaged forest at both the second and third-orders of selection, and in the even-aged treatment at the second-order only. Correlations were estimated as the posterior mean for a correlation based on Markov chain Monte Carlo (MCMC) Bayesian models controlling body mass, body condition, trapping year and sex. (Second-order) Unmanaged: Posterior mean = -0.49, 95% credible interval (-0.77, -0.18); even-aged: Posterior mean = -0.40, 95% credible interval (-0.77, -0.01). (Third-order) Posterior mean = -0.67, 95% credible interval (-0.98, -0.38).

2.4. Discussion

Within a fully controlled and replicated field experiment performed over four years, we investigated the relationships between microhabitat selection and personality traits of individual
deer mice and southern red-backed voles. By using MCMC correlation analyses we found that personality traits were associated with patterns of microhabitat selection at the second and third-orders of selection. Furthermore, at both orders these patterns of selection varied depending on land-use differences; resulting in either the loss of personality-associated microhabitat selection or in novel patterns of selection in managed forests. These findings underpin a novel mechanism that may contribute to affecting the response of species to global change and reinforce the need for understanding individual-level resource use to comprehend patterns within populations and communities.

The relationship between personality traits and fitness is well established (Smith and Blumstein 2008; Moiron et al. 2020), and personality-associated habitat selection is a key mechanism through which personality traits can affect fitness. In this study we show that, depending on the environment and the availability of preferred and non-preferred resources, selection patterns of individuals differ; probably allowing for favorable home range placement and daily movement patterns. Specifically, contrasting silvicultural treatments result in dissimilar relationships between personality traits and habitat selection, which could drive fitness differences between habitats if, for example, certain personality types are more or less likely to survive in specific environments. This is a phenomenon that warrants further investigation in this study system since this could in turn influence the behavioral diversity of populations experiencing anthropogenic disturbance (Miranda et al. 2013; Brehm et al. 2019a). Understanding how altering resource availability and habitat structure will influence future resource use by animal populations is fundamental if we are to predict how populations will be affected by land-use change and minimize our impact on ecosystem structure and function. We suggest that this study highlights one overlooked factor, personality-associated habitat selection, which may influence population and community responses to global change.
2.4.1. Personality traits and the selection for key resources

The second microhabitat component, characterized by the diameter of trees and the amount of coarse woody debris present, was a prominent variable correlated with the personality traits of both species at the second and third-orders of selection. This component describes two forest features that are well known to influence small mammal habitat selection and are characteristic of older, less disturbed (especially logged) forests. As primary seed predators and dispersers in forest ecosystems (Vander Wall 1990, 2010; Jansen et al. 2004; Vander Wall et al. 2005b), small mammals rely on mature, seed-bearing trees for a substantial portion of their diet. Although many tree species begin bearing seeds between 5 and 25 years of age, optimal seed production does not occur until much later; age 20 to 30 years in Eastern white pines (*Pinus strobus*), 50+ years for red oaks (*Quercus rubra*), and over 100 years in others (Krugman and Jenkinson 1974; Perkey and Wilkins 2001; Bonner and Karrfalt 2008). As a result, mature trees are the primary contributors to a forest’s seed bank and a central resource for granivorous rodents such as deer mice and red-backed voles.

The observed relationships between the selection of PC2 and personality traits may be influenced heavily by tree species composition and whether the dominant tree type of a forest stand is of a preferred or non-preferred food source. In fact, 4 out of 6 trends found between personality and selection for PC2 in voles may be explained by the selection or avoidance of the dominant tree species by individuals with docile or timid personality types. For example, we found that more docile or timid voles avoided areas with large seed-bearing trees, while bolder or more aggressive individuals selected these areas in the forest (as seen in Figure 2.4. and Table 2.3.). These trends occurred in either the unmanaged forest or in the even-aged forest treatment where Eastern white pine is a dominant tree type. It is well established that Eastern white pine is an important and preferred food source for our study species (Abbott 1962; Abbott and Quink 1970; Duchesne et al. 2000; Boone and Mortelliti 2019). On the other hand, Eastern
hemlock and balsam fir are non-preferred food sources for both small mammal species (Abbott and Hart 1960; Duchesne et al. 2000; Boone and Mortelliti 2019), and these two tree species dominate the shelterwood forest treatment. It is possible that in the forest stands where the white pine is dominant, docile and timid voles are avoiding the areas of the forest with large, cone-bearing white pine trees because of foraging competition with deer mice and the American red squirrel (Tamiasciurus hudsonicus). This finding may support the idea that there is a difference in competitive ability between docile or timid individuals and their aggressive and bolder counterparts, where the timid and docile are actively displaced by the better competitors and instead are found in areas of lower competition (Schirmer et al. 2019).

Coarse woody debris volume and tree diameter varied significantly across our reference sites and the treatments used in this study which could further explain some of the differences in personality-associated habitat selection across treatments. See Figures B.3. and B.4. for kernel density plots showing the distribution of the microhabitat components across the study areas and a PCA plot showing PC1 vs PC2 by treatment, respectively. The reference sites had more coarse woody debris available and trees were larger than in the two managed forest types. Just as mature, seed-bearing trees are critical for small mammals, woody debris in the form of downed trees and stumps is widely recognized as an important resource for these species. Our results are in line with previous research in which associations between woody debris and home range are well documented (Kaminski et al. 2007; Vanderwel et al. 2010; Fauteux et al. 2012). Indeed, coarse woody debris provides key structural complexity, nutrients, and substrate for a variety of plant and animal species and is crucial to the overall health of a forest ecosystem (Maser and Trappe 1984; Harmon et al. 1986; Fukasawa 2012) making it essential for the conservation of biodiversity (Johnson and O'Neil 2001; Lindenmayer and Franklin 2002). In many managed forests, the quantity of large living trees and dead wood has been greatly depleted since large trees are of important value for timber and wood products, and standing
dead wood is often removed before it can cause damage to living trees (Hansen et al. 1991; Berg et al. 1994; Linder and Östlund 1998; Fraver et al. 2002; Gibb et al. 2005).

We also observed several relationships between personality traits and selection for the first microhabitat component (PC1) which described the quantity of herbaceous ground cover, mosses, and grasses that were present at a site, as well as the amount of light present (via canopy openness). This component shows a clear difference between the even-aged forest (which has less ground cover and less open canopy) and the reference/shelterwood forests (both of which have more ground cover and canopy openness) (Figure B.4.). Our results show a positive relationship between the level of activity and exploration and third-order selection for PC1 in the unmanaged forest by both deer mice and red-backed voles (Figure 2.3.). Sites with ample ground cover and a more open canopy were characterized by high scores for this microhabitat component. In other words, in both study species, when compared to less active and exploratory individuals, highly active and exploratory individuals were found at sites that had more ground cover and light than was available on average in the home range. Since more active individuals have higher metabolic requirements (Biro and Stamps 2008; Careau et al. 2008; Réale et al. 2010b), our results may represent an example of personality matching habitat choice (Table 2.1.) (Edelaar et al. 2008). Here, highly active and exploratory individuals may preferentially select for areas with highly caloric, seasonal food resources. In our study sites, these include wild raspberry (*Rubus occidentalis*), wild strawberry (*Fragaria vesca*), bunchberry (*Cornus canadensis*), and Canada mayflower (*Maianthemum canadense*), and these species are common to newly created forest gaps. PC1 was also characterized by the amount of mossy ground cover where moisture-loving insects (another important food source) may be abundant.

Another reason why relatively more active individuals may seek out areas with denser ground cover is for added protection from aerial predators (Longland and Price 1991). Several studies have linked personality traits like activity and exploration to predation-associated
mortality rates (Smith and Blumstein 2008; Rödel et al. 2015; Moiron et al. 2020) therefore it is likely that these differences in predation risk may drive individuals to select more or less strongly for sites of refuge. Further, this could mean that individuals who take more risks offset increased predation-induced mortality through habitat selection, but further work is required to assess this phenomenon. Regardless of the mechanism driving the relationship between activity/exploration and selection for ground cover and light levels, the presence of this trend in both study species is noteworthy.

We found several instances where the personality traits of deer mice and southern red-backed voles correlated significantly with selection for key habitat features at the population level. Interestingly, there were trends that we had expected to see that were not supported by our results. For example, we predicted that we may see an overarching trend between the degree of boldness and selection for vegetative ground cover (Carrete and Tella 2010; Holtmann et al. 2017; Schirmer et al. 2019). Instead, the majority of observed relationships with this habitat component involved activity level in both species. In voles, there were two significant correlations between boldness/timidness (latency to emerge from an emergence test) and selection for ground cover, but this relationship was not uniform among treatments (Table 2.3.). Instead, bolder voles selected home ranges with less ground cover in the unmanaged forest (consistent with the findings of (Carrete and Tella 2010) and (Holtmann et al. 2017), but in the even-aged treatment this relationship was reversed, and bolder individuals selected home ranges with more vegetative cover (Table 2.3.); the finding in this treatment was consistent with those of (Schirmer et al. 2019). The apparent discrepancies between existing empirical studies may reflect differences in the history of land-use or forest management in a study area.

Additionally, we had expected to see stronger patterns of selection at coarser ecological scales (i.e., second-order: home range selection vs third-order: daily activity patterns), since decisions made on broader scales can have greater impacts on fitness (Rettie and Messier
We did see more patterns at the second-order vs the third when looking at deer mice (six significant correlations vs one), but in voles we saw six significant correlations at the third-order compared with four at the second-order and the strength of these correlations was comparable at both orders. The idea that more limiting factors drive habitat selection at coarse scales but are less influential at finer scales (Rettie and Messier 2000) has ambiguous support in the literature (as discussed by (Dupke et al. 2017)), and our study adds to those that do not show clear support for this hypothesis.

2.4.2. Management implications

In this study, we identify a system where land-use change has altered patterns of personality-associated habitat selection and influenced the distribution of individuals across the population. The congregation of like individuals in a population can potentially drive divergent selection (Rice 1987, Via 1999), influence patterns of transmission for diseases and parasites (Boyer et al. 2010; Dizney and Dearing 2013; VanderWaal and Ezenwa 2016; Sih et al. 2018), and could ultimately limit behavioral trait diversity on the landscape. Additionally, within-population differences in resource use or habitat selection could reduce intraspecific competition, perhaps having ecological consequences, for example by impacting patterns of species coexistence (Chesson 2000). Ultimately, these phenomena may influence which individuals persist in response to anthropogenic disturbance and knowledge of these processes should inform management actions (Merrick and Koprowski 2017).

It is well-known that maintaining heterogeneity at multiple ecological scales is vital for fostering biodiversity (MacArthur and MacArthur 1961; Stein et al. 2014), however, practitioners typically promote heterogeneity to maximize inter-specific (rather than intraspecific) richness and diversity. A recent study by the authors shows that behavioral diversity is affected by heterogeneity in vegetation (Mortelliti and Brehm 2020), and we add to this existing knowledge by showing that implementing silvicultural practices (such as even-aged or 2-stage shelterwood)
across a landscape may create the conditions favored by only certain personality types, and likely plays a role in shaping the personality composition of populations (with possible loss of intraspecific diversity). Indeed, previous research on the study system investigated here has shown that contrasting methods for managing forests have shifted the distribution of personalities present in different areas of the landscape (Brehm et al. 2019a); something that can have both ecological and fitness consequences (Smith and Blumstein 2008; Sih et al. 2012; Wolf and Weissing 2012; Carere and Maestripieri 2013; Lapiedra et al. 2018a). To support behaviorally diverse populations and their associated functions, habitat heterogeneity should be maintained at multiple ecological scales. Maintaining landscapes where mature forest stands are interspersed with managed stands will likely create suitable habitat for different behavioral types and thus encourage intraspecific behavioral diversity (and its important ecological and evolutionary consequences).

2.4.3. Conclusions

The adoption of forest management techniques that maintain healthy, functioning ecosystems is a crucial step towards limiting anthropogenic effects on biodiversity. Landscape ecologists and conservation biologists must identify silvicultural methods that promote the maintenance of suitable habitat at multiple spatial scales by promoting connectivity, habitat heterogeneity, and stand structural complexity (Hunter 1990; Lindenmayer and Franklin 2002); this is vital to maintain not only interspecific diversity and richness, but to promote intraspecific behavioral diversity and its associated functions. In this study, we distinguished key patterns of microhabitat selection that are correlated with animal personalities at multiple scales. Our results show that, depending on the type of silvicultural system implemented, forest management can alter patterns of intraspecific habitat selection. This study provides a novel addition to the mechanistic underpinnings of species’ responses in the face of global change.
2.5. Acknowledgments

This project was supported by the USDA-National Institute of Food and Agriculture McIntire-Stennis projects (ME041620 and ME041913) through the Maine Agricultural and Forest Experiment Station. Maine Agricultural and Forest Experiment Station Publication Number 3816. Additional funding was provided by an NSF Career Award (IOS#1940525), Penobscot Experimental Forest Research Funds, the University of Maine RRF Graduate Assistantship and RRF Undergraduate Assistantship. We thank several dedicated field and lab technicians for helping with data collection and video analysis as well as Malcolm Hunter and three anonymous reviewers for comments on an earlier version of this manuscript. We also thank Laura Kenefic (U.S. Forest Service) and Keith Kanoti (UMaine) for maintaining and facilitating research at the Penobscot Experimental Forest. A special thanks to Sara Boone for help with project management and data collection. Illustrations are by Keegan Currier. The authors declare no conflicts of interest.

2.5.1. Author Contributions

A.M.B. and A.M. conceived and designed the study; A.M.B. and A.M. collected the data; A.M.B. analyzed the data; all authors contributed to manuscript preparation.
CHAPTER 3
SMALL MAMMAL PERSONALITIES GENERATE CONTEXT-DEPENDENCE IN THE SEED DISPERSAL MUTUALISM*


3.1. Introduction

From the ants defending an acacia in exchange for nectar to the plover plucking leftovers from the jaws of a crocodile, it is believed that nearly every organism on earth is involved in at least one mutualistic interaction (Bronstein 2015). Mutualisms are relationships between species that are mutually beneficial and are foundational components of ecosystems - even termed “architectures of biodiversity” given their capacity to generate biodiversity through adaptive evolution and coevolution (Vander Wall 2010; Bascompte and Jordano 2013; Bascompte 2019) and direct gene flow within populations (Browne et al. 2018). When a mutualism falls apart, however, there are cascade effects that reach far beyond the players in the interaction (Bascompte and Jordano 2013). Indeed, ecosystem services that we all depend on (such as pollination, seed dispersal, and major biogeochemical cycles) are the products of mutualisms.

The outcome of a mutualism depends on countless interactions between the individuals involved. By extension, there is ample opportunity for variation among individuals to drive these processes (Chamberlain et al. 2014; Snell et al. 2019). Intraspecific variation such as sexual dimorphism, ontogenetic differences, or resource polymorphism can foster individual differences in diet, microhabitat preference, foraging behavior, or other forms of resource use (reviewed by 8), and the implications for community dynamics, competition, predation, demographic rates, and evolution has been reviewed comprehensively (Bolnick et al. 2011; Snell et al. 2019). Empirical investigations examining whether different individuals can affect the outcome of mutualistic interactions are rare, but recent studies have aimed to assess the effects of intraspecific differences related to age, sex, and genetic or morphological variation on
mutualistic interactions (Rodríguez-Rodríguez et al. 2013; Tur et al. 2014; Correa et al. 2015; Pérez-Méndez et al. 2018; Zwolak 2018; Pérez-Izquierdo et al. 2019; Smith et al. 2019; Snell et al. 2019; Fuster and Traveset 2020). However, we lack empirical studies examining sources of variation that do not fit neatly into these categories, such as consistent intraspecific behavioral differences, or personalities (Sih et al. 2004b).

Intraspecific behavioral differences likely play a key role in the seed dispersal mutualism via synzoochorous interactions - the intentional transportation of seeds by animals followed by the hoarding of a portion of those seeds (Gómez et al. 2019), which is often referred to as scatter-hoarding. For approximately 100 million years (Gómez et al. 2019) the behavior of individual insects, rodents, and birds has essentially determined which plants will reproduce and where since dispersers simultaneously consume and kill some seeds while providing an essential service to others (that is, dispersal away from the parent plant where chances of successful recruitment are low) (Howe and Smallwood 1982). Most synzoochorous species are not wholly mutualistic or antagonistic (Gómez et al. 2019) but instead generate an intermediate outcome, as negative and positive effects are exerted simultaneously by all individuals in a disperser population. However, if we placed each individual in a population of scatter-hoarders on a continuum from purely antagonistic to purely mutualistic (Schupp et al. 2017; Gómez et al. 2019), different individuals would likely vary along this continuum with some behaving more cooperatively towards plant species than others (such as by caching a greater number of seeds intact). If so, identifying the forces driving this variation would be imperative.

To understand how mutualistic interactions shape communities and ecosystems, we must identify the mechanisms that underlie their functioning. Synzoochory, for example, has been termed a “conditional mutualism” (Bronstein 1994; Theimer 2005) since outcomes vary depending on several biotic and abiotic factors such as the ratio of seeds to scatter-hoarders, the composition of seeds available, soil moisture, fire events, tree density, and seasonal temperature and precipitation (Vander Wall 2000; Theimer 2005; Lichti et al. 2014; Greenler et
al. 2019; Moore and Dittel 2020). If individual dispersers differ in their propensities to contribute in mutualistic or antagonistic manners, this would mean that certain individuals may be ultimately more important to the seed dispersal mutualism [i.e., keystone individuals (Modlmeier et al. 2014)]. It is the culmination of all decisions made by a seed-dispersing animal (e.g., which seed to choose, whether to consume immediately or cache for later, how far to transport, and where to cache) that determine how mutualistic it is on the whole. Interestingly, among all the factors that shape these decisions (Wang et al. 2013; Lichti et al. 2015), including intraspecific variation attributable to sexual, ontogenetic, or morphological differences (Zwolak 2018), one that is rarely studied is the unique behavioral type of the disperser. A recent conceptual review discusses the capacity for personality traits to influence these processes (Zwolak and Sih 2020), and a few empirical studies (Brehm et al. 2019a; Feldman et al. 2019; Boone et al. 2021) show that personality traits (for example, the boldness or risk-taking capacity of an individual) influence decisions made during seed dispersal. Since a single individual makes thousands of these decisions throughout one season, personality type likely affects how mutualistic each disperser behaves on the whole and drives an individual’s position on the antagonism-mutualism continuum.

The goal of this study is to examine intraspecific variation in the seed dispersal mutualism and understand the role that personality plays. Our objectives were to observe seed dispersal by known individuals, calculate where they fall along a continuum from antagonist to mutualist (Gómez et al. 2019) and finally, assess whether personality traits affect an individual’s position along the continuum. We implemented a large-scale study wherein we trapped small mammals and measured their personality traits using three standard behavioral assays. We targeted one of the most abundant scatter-hoarding rodents in temperate forests of North America, the deer mouse (Peromyscus maniculatus). We performed a seed experiment; offering seeds of three species and using fluorescent powder (Lemke et al. 2009) to track 792 dispersal routes and assess whether individual interactions were positive or negative (Figure 3.1. A) and,
thus, calculate a score along the antagonism-mutualism continuum (Figure 3.1. B). Negative interactions included those that immediately precluded a seed from future germination opportunities (i.e., seeds that were consumed at the experiment site, consumed post removal, or taken down a hole where germination is unlikely [Vander Wall et al. 2005a]). Positive interactions included seeds that were removed from the site and cached intact or left intact at the site after an interaction (Table C.1.). Last, we assessed whether personality traits affected each individual’s score along the continuum.
Figure 3.1. How are individuals placed along an antagonism-mutualism continuum? (A) Each interaction between a scatter-hoarder and a seed can be classified as either positive (seed is dispersed alive and deposited intact) or negative (seed is consumed or taken below ground where germination is unlikely). (B) Using the framework described by (Gómez et al. 2019) all interactions made by an individual can produce an individual’s score along the antagonism-mutualism continuum by subtracting the proportion of negative interactions from the proportion of positive interactions.

\[
\text{Score} = \frac{\#(+)\text{ }}{\#(+) + \#(-)} - \frac{\#(-)\text{ }}{\#(+) + \#(-)}
\]
Only by identifying the mechanisms underlying this conditional mutualism can we accurately predict outcomes in a changing world. If inter-individual differences in behavior is a previously unidentified factor generating context-dependence, the loss of certain individuals may drive relationships to the tipping point, poised to shift from mutualism to antagonism. Anthropogenic changes can modify the distribution of personalities within populations (Miranda et al. 2013; Lapiedra et al. 2017; Brehm et al. 2019a; Mortelliti and Brehm 2020), thus, if personality traits drive an individual’s ecological role (Hunter et al. 2022), altering habitat may impose unexpected consequences on the mutualisms we all depend on. Further, implications of this work could reach far beyond the seed dispersal mutualism; similar mechanisms may shape pollination and plant protection mutualisms where deliberate animal behavior is the driving force.

3.2. Materials and Methods

3.2.1. Study site and small mammal trapping

This study was conducted at the Penobscot Experimental Forest (44° 51' N, 68° 37' W) in Maine, USA (Figure 3.2. A). The Penobscot Experimental Forest is a mixed conifer-deciduous forest (Brissette and Kenefic 2014) and is dominated by shade-tolerant conifers (Kimball 2014). For more detail about the study area see Appendix C.

From June through October (2016 - 2020) we implemented a large-scale capture-mark-recapture experiment in six separate areas of forest. We trapped small mammals in six trapping grids; each grid was 0.81 ha in area and consisted of 100 flagged points spaced 10 m apart. The mean distance between grids was approximately 1.42 km. We placed one Longworth small mammal trap at each flagged point and baited traps with a mixture of sunflower seeds, oats, and freeze-dried mealworms. We bedded traps with cotton stuffing and checked traps twice per day (just after sunrise and in the late afternoon). We trapped at each grid for three consecutive days and nights once per month for five months each year totaling ~ 45,000 trap nights (number of active traps * number of nights).
Figure 3.2. Summary of study area and experimental design. (A) Map of the study area at the Penobscot Experimental Forest (44°51’ N, 68°37’ W) in Maine, USA. (B) Small mammal trapping was performed at six separate trapping grids, and nine seed experiment sites were spaced evenly throughout each grid. (C) Experiment sites contained a set of paired stations monitored by trail cameras. (D) At each station, seeds trays were surrounded by antennas attached to permanent radio frequency identification readers to scan and identify tagged individuals. (E) Fluorescent tracking powder allowed cached seeds to be recovered (F, G), and uniquely marked seeds allowed caches to be attributed to individual dispersers.
3.2.2. Animal processing and behavioral assays

All captures were processed at a base area in the grid. Animals were transferred directly from the trap into three standardized behavioral assays to measure behaviors that would later be used to assess personality. An emergence test was used to assess boldness, an open-field test to measure activity and exploration in a novel environment, and a handling bag test to measure docility and the response to handling by an observer (Carter et al. 2013; Brehm et al. 2019a). We performed behavioral assays once monthly to ensure that animals would not become habituated. Emergence and open-field tests were videotaped and behaviors were quantified from videos in the laboratory. See Brehm et al. (2019, 2020) and Appendix C for detailed field procedures, as well as software and methods used to quantify behavior. See Table C.3. (modified from Brehm et al. 2019a) for a complete list of the behaviors measured, their description and interpretation, and supporting sources.

After the behavioral assays, we anesthetized animals with isoflurane and inserted PIT tags (Biomark MiniHPT8) subcutaneously at the midback. Animals were marked with a small animal ear tag (National Band, Style 1005-1) and a unique haircut. We recorded sex, body mass (using a 100g Pesola Lightline spring scale), body length, tail length, and age class (juvenile, subadult, or adult; based on body size and pelage coloration). Animals were released at the capture site post-processing. Previous research in this study system confirms that sampling methods are not biased towards certain personalities and that trapping methods do not impact behavioral measurements in standard assays (Brehm and Mortelliti 2018; Brehm et al. 2020).

3.2.3. Seed experiments

In September and October 2020, during natural seed drop, we performed a field experiment to observe individuals with known personality types predating and dispersing seeds in their natural environment. We offered seeds of Northern red oak (*Quercus rubra*), Eastern white pine (*Pinus strobus*), and American beech (*Fagus grandifolia*) and prior to the experiment,
seeds were visually inspected for cracks, rot, or weevil holes, and acorns were float-tested. See Appendix C for more detail.

After trapping was completed in September and October, we positioned seed sites spaced evenly throughout each grid (Figure 3.2. B). At each site we placed two seed presentation stations (Figure 3.2. C). One station consisted of a tray with eight individual wells where we offered four acorns and four beech seeds (Figure 3.2. D) and the other station had a tray with six wells to each hold five white pine seeds (30 seeds total). Each presentation tray was mounted onto a 30 x 30 cm piece of vinyl. Each seed station was monitored with a trail camera (Reconyx XR6 Ultrafire) and a permanent radio frequency identification (RFID) reader to scan and identify individuals marked with PIT tags. Stations were dusted with UV fluorescent tracking powder (TechnoGlow) and seeds were painted a unique color using non-toxic, UV fluorescent paint (Neon Glow, ASTMD-4236 certified) and placed in a known location on the presentation tray. See Appendix C for further information about the seed station setups and for further details of methods to assure paint did not influence seed selection, cache recovery or pilferage.

Seed stations were set at dusk and visited before dawn so all cache searching could be done in darkness. The observer (AMB) used a UV flashlight (uvBeast) to follow all fluorescent trails from each station (Figure 3.2. E). The observer recorded whether each trail: 1) ended at a seed, 2) ended down a hole, or 3) ended up a tree or faded out. If a seed was found the observer recorded the identity of the seed and whether the seed was consumed or intact after removal. When trails ended up trees or faded away these trails could not be linked to a known interaction and were, therefore, not used in further analyses. See Figure 3.2. F, G for images of caches located using these methods. All videos from trail cameras were played back in the laboratory and individual identification of each individual visiting the stations was confirmed using the time-stamped RFID reads and the unique haircut seen in the videos. Any instances where a located cache could not be confidently paired with an observed interaction made by a
marked individual was not used in further analyses (27 cases for oak, 18 cases for beech, and 6 cases for white pine). For more details on video analysis see Appendix C.

To minimize visits by non-target species, stations were removed at dawn and re-set at dusk if seeds remained. Seed stations were left active at a site for an average of two nights but were removed after the first night if all seeds were consumed/removed. In total, we offered 412 acorns, 412 beech seeds, and 3,090 white pine seeds at 103 seed sites (206 paired stations).

3.2.4. Statistical analyses
First, we performed a repeatability analysis using the ‘rptR’ package in R (Stoffel et al. 2017) to determine which behavioral variables observed during standardized assays could be considered personality traits. We then calculated each individual’s mean BLUP (best linear unbiased predictor) for each behavioral variable over 1000 simulations with package ‘arm’ and subsequent mentions of personality refer to the mean BLUP value (Gelman and Su 2018). See Appendix C for further details about calculating repeatability and simulated BLUPs.

For all *Peromyscus* who interacted with seeds at presentation stations, we quantified the proportion of interactions that were positive events (the seed was dispersed and cached intact) and the proportion of interactions that were negative events (the seed was consumed at the station, consumed post-removal, or could be confirmed to be taken down a hole). In line with (Gómez et al. 2019), we then calculated each individual’s location on the predator-mutualist continuum by subtracting the proportion of negative interactions from the proportion of those that were positive. Individual locations along the continuum, hereafter referred to as *scores*, were calculated for each seed species separately. For a breakdown of each interaction type and its classification as positive or negative, see Table C.1. We acknowledge that since individuals varied in the number of interactions that went into their calculated score, individuals with a greater number of interactions likely had more precise estimates. While we cannot specifically account for this in the score estimate, we imposed the variable “number of seed interactions” into models predicting the score as detailed below.
3.2.5. Models predicting individual scores along the antagonism-mutualism continuum

We used a nested hypothesis testing approach (Burnham and Anderson 2002b) to assess whether personality type affects an individual’s location along the continuum (one value per individual). We ran linear models using score as the dependent variable. Throughout our analyses, models within 2.0 ΔAICc of the top model were considered to have equal support (Buckland et al. 1997; Burnham and Anderson 2002b). We assured non-collinearity between continuous predictor variables prior to inclusion in models (Table C.6.). We first tested covariates suspected to influence individual scores: the number of seed interactions, forest type, trapping grid, sex, an index of body condition (calculated using the scaled-mass index (Peig and Green 2009)), and body mass (a proxy for age). We compared models to the null, and when more than one model scored higher than the null model, we tested for an additive effect of these variables. We retained the top model from this model set and tested it against seven new models, adding one personality trait measurement to each (Table C.3.). We removed one individual *Peromyscus* from the dataset due to an extreme value (high leverage) in the behavioral variable “Proportion of time in the center” so that this trait would not inflate the strength of a regression in which it is included. All continuous predictor variables were scaled (z-standardized).

3.2.6. Models predicting discrete interactions with seeds

Although both comprehensive and consistent with recent literature, the approach of calculating an individual’s position along the predator-mutualist continuum condenses all interactions by an individual into one single score. To instead assess how personality traits may impact each interaction individually, we performed an additional analysis using mixed-effects models in R package *lme4* (Bates et al. 2015). We ran logistic models on each seed species separately, using each interaction as a separate observation (a repeated-measures design). The dependent variable was a binomial variable with the value 1 if the interaction was classified as positive and 0 if the interaction was classified as negative. We used individual identity as a random intercept.
in these models and used the same fixed effects and model selection procedures as above (with the addition of seed mass as a covariate for oak and beech). Running these mixed models predicting the discrete interactions allowed us to incorporate uncertainty due to differing number of interactions among individuals.

3.3. Results

3.3.1 Repeatability of behavioral traits

From three standardized behavioral assays (an emergence test, an open-field test, and a handling bag test) we examined 819 behavioral observations from 301 individual deer mice with two or more observations and found all behavioral variables to be significantly repeatable (Table C.2.). Mean repeatability was 0.330 (range: 0.191 to 0.447) falling in line with similar field studies on deer mice (Underhill et al. 2021) and near the average previously reported for a variety of field and laboratory studies (Bell et al. 2009). Repeatable traits included the following (described in detail in Table C.3.): *mean speed* (an indicator of activity), *rear rate* (activity and exploration), *proportion time grooming* (anxiety), *proportion time center* (boldness), *handling time* (docility), *latency to emerge* (boldness), and *time at end of tunnel* (boldness).

3.3.2 Seed removal experiments

At 206 paired seed stations, we observed 1,813 visits by small mammals (1,179 of which were deer mice). Other species to visit included the southern red backed vole (*Myodes gapperi*), the American red squirrel (*Tamiasciurus hudsonicus*), and flying squirrels (*Glaucomys spp.*). We observed few visits from other species. In total, 1,110 white pine seeds, 0 acorns, and 17 beech seeds were consumed at the sites (or 40%, 0%, and 4.4%, respectively), whereas 1,215 (44.5%), 261 (68%), and 313 (81.5%) were removed from the sites. We located (or could confidently assume the seed was taken down a hole for 994 (82%), 122 (47%), and 201 (64%) of removed white pine, acorns, and beech seeds.

Of the 1,179 visits made by deer mice, 955 could be counted as seed interactions (where deliberate contact was made with a seed), and 934 were by tagged deer mice. 532 of these 934
interactions were in instances where the cache was located and the fate of the seed was known, or could be confidently inferred (thus, the interaction could be used in further analyses). See Supplementary figures C.1.-C.3. for diagrams showing the number of each interaction type made by tagged individual deer mice.

3.3.3. Calculating scores along the antagonism-mutualism continuum

In total, we calculated scores along the antagonism-mutualism continuum for deer mice using 349 observations from 26 individuals interacting with white pine seeds, 135 observations from 21 individuals interacting with beech seeds, and 48 observations from 19 individuals interacting with acorns. On average, mice had negative scores for all three seeds; mean for mice interacting with white pine was -0.36, with beech was -0.37, and with acorns was -0.07 (See Figure C.4. for the distribution of white pine and beech scores and Figure 3.3. A for the distribution of oak scores).

3.3.4. Effects of personality on scores along the antagonism-mutualism continuum

For red oak scores (n = 19 individuals), we found that an individual’s degree of boldness/timidness (Time at end of tunnel) affected the position along the antagonism-mutualism continuum. Specifically, timid individuals tended to be more antagonistic (β = -0.44 ± 0.15; Table C.4.; Figure 3.2. B). We did not find an effect of personality on scores for white pine seeds (n = 26 individuals); instead the best predictors were body condition index and forest type. Individuals with higher body condition indices fell at a more antagonistic position along the continuum (β = -0.31 ± 0.15; Table C.4.; Figure C.5.). The top model predicting the score for American beech (n = 21 individuals) was the null model.
Figure 3.3. (A) Individual scores for *Peromyscus maniculatus* along the antagonism-mutualism continuum for red oak (*Quercus rubra*; mean score = -0.07). (B) Predicted relationship (and 95% CIs) between a personality trait and an individual’s score along the continuum. Timid individuals have lower (more antagonistic) scores for red oak ($\beta = -0.44 \pm 0.15$ SE) than bolder individuals. Data points represent observed values (one point per individual). Timidness is measured as the time spent at the end of the emergence tunnel before emerging (z-standardized), where high values indicate a longer duration at the tunnel end before emerging (high timidness), and low values indicate short durations at the tunnel end before emergence (low timidness, i.e., boldness).
3.3.5. Effects of personality on discrete interactions

Models predicting each interaction independently showed that repeatable behavioral variables indicating boldness/timidness predicted the probability of interacting positively with white pine and beech seeds (Table C.5.). Bolder individuals (key variable was Proportion of time in the center) were more likely to interact negatively with both white pine (n = 349 observations from 26 individuals) and beech seeds (n = 135 observations from 21 individuals) (β = -1.24 ± 0.41 and β = -0.89 ± 0.33, respectively; Figure 3.4. A-B). Other predictors in the top model for white pine interactions were body condition index (β = -0.43 ± 0.20) and forest type.
Figure 3.4. Predicted relationships (and 95% CIs) between a key behavioral variable indicating the degree of boldness/timidness and the probability of interacting positively with seeds. Bold *Peromyscus maniculatus* are more likely to have a negative interaction with (A) white pine (*Pinus strobus*) and (B) beech (*Fagus grandifolia*) than timid individuals (β = -1.24 ± 0.41 SE and β = -0.89 ± 0.33 SE, respectively). Data points represent observed values (one point per individual interaction). Here boldness is measured as the proportion of time in the center portion of the open-field arena (z-standardized), where high values indicate more time in the center (high boldness), and low values indicate less time in the center (low boldness, i.e., timidness). Relationship for white pine is shown for the treatment 2 forest type.
3.4. Discussion

The seed dispersal mutualism has been termed a “conditional mutualism” due to the array of biotic and abiotic factors that influence the overall outcome (Theimer 2005). When components such as the ratio of seeds to dispersers shift over time or through space, the balance between mutualism and antagonism can be tipped. Until now, knowledge of how individuals may contribute differently to this crucial plant-animal interaction has been limited. We show that personality traits of individual dispersers affect their position along a continuum from antagonist to mutualist. Specifically, the boldness of an individual influenced its tendency to interact positively with seeds. These findings indicate that inter-individual differences in behavior, or personalities, are a previously overlooked mechanism driving context-dependence in the seed dispersal mutualism. To garner a better mechanistic understanding of the mutualisms that provide essential services like seed dispersal and pollination, our findings suggest future research should prioritize understanding the underlying role of intraspecific behavioral diversity in disperser populations.

In this study deer mice were antagonistic on average towards red oak, American beech, and white pine seeds. However, we found that within a population, there is variability in the contributions made by individuals to the seed dispersal mutualism. Specifically, some individuals behave in a far more mutualistic manner than others. In practice, a balance between mutualism and antagonism does not require an equal contribution of predation and successful caching events to break even since just one cache that ends in recruitment has successfully passed on the genetic material of the mother plant. Even if the majority of individuals act as seed predators, a few individuals who cache seeds intact will likely enable the germination rate to exceed 0% - which may translate into a large number of seedlings if seed abundance is high (Hulme and Kollmann 2009). As is true for long distance dispersal events, the frequency of an event is not necessarily positively correlated with its importance (Nathan 2006). Therefore, positive interactions are disproportionately important to the seed dispersal mutualism since their
consequences outweigh those of negative interactions. Here we show that individuals who provide these positive interactions have personality traits in common. Our previous work has shown that our samples are representative of the greater population (i.e., we are not preferentially sampling certain personality types) (Brehm and Mortelliti 2018). Further, the current study was performed among six separate stands within the same experimental forest and individuals were subject to the same predator communities and weather patterns. By extension, these trends, scaled up to an entire scatter-hoarder population over the course of a season, would reflect thousands of seeds cached intact and thousands of others consumed or taken down into underground burrows. Seen in this context, we can begin to understand how heterogeneity in individual contributions can upscale to have ecosystem-level consequences.

When interacting with acorns, bolder mice (i.e., those who emerged from an enclosed space before taking time to assess the safety of surroundings) had more mutualistic scores than timid mice. Bold individuals were more likely to remove acorns from the seed station and cache them intact on or just below the surface (i.e., concealed by a thin layer of detritus, moss, needles, etc.). Timid individuals, instead, were more likely to perform negative interactions such as taking acorns down into an underground burrow (effectively eliminating future chances of germination; 50% of negative interactions observed) or consuming the seed either entirely (8%) or partially (42%) after removal (Figure C.2.). This led to an overall negative relationship between the timidness of an individual and its position on the antagonism-mutualism continuum. This trend may reflect a greater overall perceived risk associated with caching a large seed (a risk that bolder individuals are more willing to take). Larger seeds are generally more conspicuous, and studies have shown that larger seeds are preferentially cached in risky areas (Steele et al. 2014) to avoid being discovered by competitors. Further, whereas beech and white pine seeds are small enough to be transported in the cheek pouch of a deer mouse, an acorn must be carried externally in the jaws and requires the individual to lift its head to maneuver. Studies have shown that species with cheek pouches tend to disperse smaller seeds (and are
likely adapted to do so), whereas pouchless rodents are instead more likely to disperse larger seeds (Vander Wall and Longland 1999; Gómez et al. 2019). Likely, the mass of the acorn slows the disperser down and the inability to store the seed in the cheek pouch inhibits the animal’s ability to see optimally and, therefore, remain vigilant. Individuals who are more timid may perform behaviors deemed “safer” such as moving directly below ground or taking the acorn to a covered area to consume, instead of taking the time (and associated risk) to cache the seed on the surface. Examining the relationship between boldness and the tendency to cache large seeds intact in pouchless scatter-hoarders, such as red-backed voles, would allow us to understand this relationship more fully, as we would expect this trend to be weaker in a rodent better adapted to disperse large seeds.

Though we observed negative relationships between timidness and antagonistic interactions with both white pine and beech seeds, these are likely driven by different factors. Bold individuals (those who spent more time exploring in a risky area of the open-field arena - in the center, away from the walls) were far less likely to cache seeds intact after removal. Specifically, bold mice were up to 82% less likely to cache white pine seeds intact and up to 69% less likely to cache beech seeds intact (Figure 3.4.). The negative interactions observed for white pine seeds were most often instances where individuals remained at the seed site and consumed the seeds immediately (59% of negative interactions observed). Instead, 20% of interactions were seeds consumed after removal, and 21% were seeds taken down into holes. Therefore, the fact that bolder individuals behaved more antagonistically towards white pine is likely because bold individuals were more likely to consume seeds immediately rather than transporting seeds to a “safe” place prior to consumption. Alternatively, beech seeds were rarely consumed at the seed site (this behavior occurred only approximately 4% of the time) but were almost always removed from the site. This finding is consistent with previous research showing the preferential dispersal of larger seeds (Vander Wall 2010). Negative interactions observed for beech were typically consumption directly after removal (approximately 54% of negative
interactions) or transportation of the seed below ground (39%). The propensity for timid individuals to scatter-hoard a beech seed intact may reflect different survival strategies between bold and timid personality types (i.e., prioritizing future energetic reserves over present resource use) and warrants further investigation.

Collectively, deer mice were antagonistic towards red oak, American beech, and white pine seeds. This corroborates previous studies showing that the lower the ratio of seeds to scatter-hoarders, the greater the proportion of the seed crop that is harvested and predated (Theimer 2005), as this study took place in a non-mast year. It is possible, however, that the trends seen here would be exaggerated in years of seed mast, since as the scatter-hoarder population becomes satiated, the greater the number of cached seeds that would escape predation by competitors (pilferage) and the cache owner (recovery) - i.e., the predator satiation hypothesis (Vander Wall 2001). This assumes that the effects of personality on mutualistic behavior are not context-dependent (i.e., it assumes that individuals who cache more seeds intact in a non-mast year would also cache more seeds intact in a mast year). Surplus food items typically initiate caching behavior even if the number of cached seeds far outweighs what the animal would need to survive the winter (Vander Wall 2001), so it is also possible that mast years would dampen the effects of personality type on mutualistic behavior since all individuals would cache more. If the effects of personality on caching probability are consistent among years of high and low resource availability, it is likely in years of seed mast once the population becomes satiated, the surplus caches made by disproportionately mutualistic individuals would contribute more to recruitment in the plant population. Future studies may focus on years of high resource availability in order to make more general statements about the effect of individuality of the animal mutualist on recruitment rates in the plant partner (Schupp et al. 2017).

Future work should also address other situations under which the effect of personality traits on mutualistic behaviors may be conditional (such as the effects of body condition, ontogenetic effects, or environmental cues like predator densities or the availability of refuge). It
is conceivable that, for example, the degree to which boldness affects an individual’s probability of caching a seed intact depends on whether the individual has a higher or lower body condition index. We note, however, that body condition is not always a good indicator of an individual’s energetic reserves in species that hoard resources (discussed by (Underhill et al. 2021)), as these resources are not all stored on the body in the form of fat reserves. In fact, our finding that individuals with lower body condition indices were more likely to cache white pine seeds intact (Figure C.5.) could support the idea that body condition index may be lower in individuals who are caching a greater portion of the resources they find. Last, one potentially important aspect of an individual’s role as seed predator vs mutualist that was not addressed in this study is the tendency to recover or pilfer caches. In order to fill in the remaining black boxes, studies that allow tracking of a seed through subsequent recaching events (such as using telemetric thread tags (Hirsch et al. 2012; Jansen et al. 2012)) while identifying the individual at each recovery/pilferage are needed. This study should act as a catalyst for future work on this topic.

3.4.1. Conclusions

Mutualistic interactions support much of the earth’s primary production and are, therefore, central pillars of healthy, functioning ecosystems. When one partner in a mutualism is compromised, the other becomes threatened indirectly (Markl et al. 2012), and maintaining sufficient population sizes of both mutualistic partners is critical to avoid Allee effects (Winfree et al. 2015). Research has shown that bolder and more active individuals are often more likely to be removed from populations via hunting or fishing and may experience differential fitness in response to urbanization (Miranda et al. 2013; Sih 2013; Merrick and Koprowski 2017). In species that act as dispersers in synzoochorous, mycorrhizal, or pollination mutualisms, the removal of certain individuals from populations may, therefore, have unintended consequences if individuals vary in terms of their functional contributions (McConkey and O’Farrill 2016). We observed individual deer mice (Peromyscus maniculatus) with known personality traits predating and dispersing seeds in a natural environment. We found that within a population of scatter-
hoarders, there is variability in contributions made to the seed dispersal mutualism. Some individuals are far more mutualistic than others, and one driver of this distinction is animal personality. These findings demonstrate the potential for animal personality to act as a mechanism generating context-dependence in the seed dispersal mutualism and provide a conceptual advancement to the study of mutualism - bringing mutualism and intraspecific behavioral variation together in an empirical study.

3.5. Acknowledgments

We would like to thank several field assistants and volunteers for their help with data collection, Sara Boone, Ivy Yen, and Margaux Duparcq for support with data collection and project management, Dr. Laura Kenefic and Keith Kanoti for assistance facilitating research in our study area (Penobscot Experimental Forest), Keegan Currier for illustrations, Malcolm Hunter, Christopher Moore, and two anonymous reviewers for valuable feedback on an earlier version of this manuscript. Rodent silhouettes were obtained from www.phylopic.org. This research was supported by an NSF Career Award to A.M. (IOS#1940525), a Janet Waldron Doctoral Research Fellowship and a Chase Distinguished Research Assistantship to A.M.B, Penobscot Experimental Forest Research Funds, and the USDA National Institute of Food and Agriculture McIntire-Stennis projects ME041620 and ME041913 through the Maine Agricultural & Forest Experiment Station. Maine Agricultural and Forest Experiment Station Publication Number 3883.
Across taxa, individuals of the same species behave consistently differently from one another and these behavioral differences are referred to as personalities (Sih et al. 2004b; Wolf and Weissing 2010). Several studies have documented associations between animal personalities and fitness components such as survival and fecundity (reviewed by [Moiron, Laskowski, & Niemelä, 2020; Smith & Blumstein, 2008]), and evolutionary hypotheses aimed at explaining how such variation in behavioral phenotypes is maintained in populations are often rooted in life history theory (i.e., concerning the allocation of limited resources to survival, maintenance, and current vs. future reproduction to optimize fitness [Stearns, 1989]). Investment in one fitness component inevitably requires a trade-off elsewhere. The ‘pace-of-life syndrome’ hypothesis posits that individuals operate under different strategies to resolve these trade-offs in investment and that behaviors such as activity, aggression, or boldness mediate trade-offs at the individual level (Réale et al. 2010b; Dammhahn et al. 2018). Under this framework, more active, risk-prone, or aggressive individuals are predicted to experience faster growth rates and higher reproductive output but incur life-history trade-offs such as increased mortality and/or exposure to parasites (Biro and Stamps 2008; Barber and Dingemanse 2010; Réale et al. 2010b).

Intraspecific variation in behavior, however, is also hypothesized to be maintained in part because selection on personality traits is context-dependent. That is, individuals with different personality traits experience differential fitness under fluctuating selection pressures (reviewed by [Dingemanse & Réale, 2013]), such as heterogeneity in resource availability, predator density, or competition for resources over space and/or time (Penke et al. 2007; Wolf and Weissing 2010; Le Cœur et al. 2015; Nicolaus et al. 2016). In fact, two meta-analyses show
equivocal associations between risky behaviors and fitness (Smith and Blumstein 2008; Moiron et al. 2020). In the more recent and, thus, more comprehensive paper (Moiron et al. 2020) reviewed empirical studies conducted in both the lab and on wild populations and found no universal directional relationship between risky behaviors and survival or longevity (Moiron et al. 2020). Instead, in the wild, risky individuals lived longer; in contradiction to predictions under the pace-of-life syndrome hypothesis. Explanations for this disparity include differing selection pressures in the lab vs the wild, potentially imperfect measurements of truly risky behaviors, possibly biased estimates of survival, and discrepancies between the level of variation at which theory has been laid out (the among-individual level) and at which it is typically being tested (the phenotypic level) (Niemelä and Dingemanse 2018; Moiron et al. 2020).

Since different personality traits may be more or less advantageous depending on context, without sampling individuals across varying contexts, such as across environments with contrasting habitat structure and resource availability, our ability to make predictions concerning the directional relationship between personality traits and fitness components is limited. Further, the pace-of-life syndrome hypothesis and hypotheses concerning fluctuating selection pressures need not be mutually exclusive. Density fluctuations that act to regulate populations are thought to be nearly universal among taxa (Brook and Bradshaw 2006), and a recently proposed framework suggests that pace-of-life syndromes emerge as a product of fluctuations in population density (Wright et al. 2019). Recent modeling approaches support key predictions made under this framework (Milles et al. 2022). However, empirical studies are necessary to further elucidate the conditions under which density dependent selection may drive covariation between risky behavioral traits and life history traits. Specifically, as Milles et al., (2022) found differences in the slope of the pace of life syndromes that emerged at different hierarchical levels (among vs. between populations), more research is required to understand which factors drive such processes in real ecological systems. Unraveling these mechanisms will further our
understanding of how covariation among behavioral and life-history traits such as survival and reproduction can emerge at the population level.

We tested for the relative support of three non-mutually exclusive hypotheses concerning the potential for spatial and temporal variability to drive misalignment between empirical results and theory: (1) **Spatial heterogeneity in habitat structure alters the relationship between risky personality traits and survival.** Under this hypothesis, we predicted that the relationship between risky traits and survival would differ among forests with varying availability of refuge habitat. Studies have shown that small mammals often use indirect cues (such as cover) rather than direct cues (such as predator scents) to assess risk (Orrock et al. 2004), and risky environments have been shown to favor bolder individuals (Urban 2007; Ingley et al. 2014) as well as promote associations between boldness and aggressiveness (Dingemanse and Wolf 2010). Therefore, we expected that risky traits may have a selective advantage (increase survival) in environments where perceived risk is higher and taking risks is necessary to obtain resources. We define a higher-risk environment as one with less ground cover and coarse woody debris - two key sources of refuge for small mammals (Dueser and Shugart 1978; Loeb 1999; Fauteux et al. 2012). We expected that risky behavioral traits may not influence survival in environments where ground cover and coarse woody debris are more abundant (Figure 4.1C). (2) **Fluctuations in population density alter the relationship between risky personality traits and survival.** The novel framework regarding fluctuating density-dependent selection (Wright et al. 2019; Milles et al. 2022), predicts that risky (or fast-type) personality traits should have a selective advantage during periods of low density, but a disadvantage during periods of high density. However, it is also conceivable that risky traits should be associated with increased survival during periods of high density if these traits improve an individual’s competitive ability (Dingemanse et al. 2004). High degrees of variability in small mammal abundance over space and time in our study system (Figure 4.1B) give us a unique opportunity to examine the effect of density on relationships between risky traits and survival in four species simultaneously. (3) **Fluctuations in**
density interact with perceived predation risk to alter the strength of associations between risky behavior and survival. Under this hypothesis, we predicted that whether or not risky individuals have an advantage in risky environments will depend on small mammal density. Specifically, at low density when there is less competition for resources and space (including refuge habitat), the advantage of risky traits in risky environments should be less compared to non-risky environments. Findings under this third hypothesis should provide more insight about under which circumstances we might expect personality trait variation to be maintained by density-dependent selection.

As part of a large-scale manipulative experiment, we trapped deer mice (Peromyscus maniculatus), southern red-backed voles (Myodes gapperi), northern short-tailed shrews (Blarina brevicauda), and North American red squirrels (Tamiasciurus hudsonicus) in five subsequent years across six different forest stands treated with varying management practices (Figure 4.1A). Using detailed capture histories and robust design models to estimate apparent survival while controlling for imperfect detection and temporary emigration (Kendall et al. 1997), we obtained estimates of apparent monthly survival probability. This study, in assessing the relationship between personality traits and population-level survival estimates, accounts for potential sources of bias in this fitness component (an area targeted as a potential drawback in the current breadth of research on the topic) (Moiron et al. 2020). Additionally, by using datasets collected on four species simultaneously across forest stands with significant structural differences, we should identify common trends regarding the context-dependent fitness associated with risky personality traits, including the potential for population density to alter directional relationships between personality traits and survival.
Figure 4.1. Overview of the experimental design. (A) Photos of study areas at the Penobscot Experimental Forest (44°51’ N, 68°37’ W) in ME, USA. Shown (from top to bottom) are the reference/unmanaged (REF), uniform shelterwood (USW), and irregular shelterwood (ISW) stands. One 90m x 90m trapping grid was positioned near the center of each stand for a total of 6 grids. (B) Small mammal abundance varies over space (i.e., among grids) and time (over the 5 years) in this study system. (C) Major structural differences between the three stand types. The REF grids contained significantly more shrubs/herbaceous material (<1-meter-tall) and total meters of coarse woody debris in two size classes (10-20cm in diameter and 20+ cm in diameter) when compared to the USW and ISW grids (β and SE shown were estimated using linear models - see methods for details).
4.2. Materials and methods

4.2.1. Study area and experimental design

This study was conducted at the Penobscot Experimental Forest (PEF, 44°51' N, 68°37' W) in central Maine, USA, in the transitional zone between the eastern broadleaf and boreal forests (Figure D.1). The PEF is a 1578-hectare, mixed conifer-deciduous forest and is dominated by shade tolerant conifer species including red spruce (*Picea rubens*), balsam fir (*Abies balsamea*), and eastern hemlock (*Tsuga canadensis*) (Brissette and Kenefic 2014). Climate in this region is cool and humid with a mean temperature of 6.8˚C and annual precipitation averaging 1.02m (Cite US Forest Service page). The most abundant small mammal species on this landscape are deer mice (*Peromyscus maniculatus*), southern red-backed voles (*Myodes gapperi*), northern short-tailed shrews (*Blarina brevicauda*), and North American red squirrels (*Tamiasciurus hudsonicus*). The primary experiment in this forest is a “compartment study” wherein different silvicultural treatments were applied to randomly selected and replicated stand-level management units beginning in 1952 and continuing today (Brissette and Kenefic 2014). Approximately 25 hectares of forest have remained unmanaged since the late 1800s and serve as an important reference.

In the current study, we selected two silvicultural treatments: a uniform shelterwood (USW) cut using a two-stage overstory removal, and an irregular shelterwood (ISW) cut using a two-stage overstory removal and retaining reserves, or trees from the older cohort. These two treatments have generated contrasting habitat types for small mammals; the uniform shelterwood has produced dense stands of shade-tolerant trees that are all within the same age class and diameter and the irregular shelterwood has resulted in stands with enhanced vertical structure from the large, retained residual trees which provide shade, increased seed production, as well as snags and downed logs for refuge. We selected two stands that have remained unmanaged since the late 1800s to serve as reference areas (REF). In total, we utilize
six separate stands in this study, and microhabitat measurements were recorded in each study area. For more detail, see methods described in (Brehm and Mortelliti 2021) and the supplementary materials (Supplementary methods). To assess whether the amount of refuge habitat (shrubby ground cover and coarse woody debris) varied among the stand types, we ran linear models with the microhabitat variable in question as the response variable, and stand type (REF, USW, or ISW) as a predictor variable (Figure 4.1C). As small mammals often respond to indirect cues of risk (such as refuge availability) more than direct cues (such as predator scents) (Orrock et al. 2004), we use the amount of refuge habitat as a proxy for perceived risk in this study.

4.2.2. Small mammal trapping

We positioned one small mammal trapping grid close to the center of each forest stand. Trapping grids were 0.81ha in area and consisted of 100 trapping points spaced 10m apart. The mean distance between trapping grids was approximately 1.42km. At each trapping point, we placed one Longworth small mammal trap baited with a mixture of sunflower seeds, oats, and freeze-dried mealworms. Cotton stuffing was provided for bedding. At every other trapping point, we placed one Tomahawk trap baited with a mixture of peanut butter and sunflower seeds. Rain covers were positioned on the Tomahawk traps to limit exposure. Longworth traps were checked twice per day (just after sunrise and in the late afternoon) and Tomahawk traps were activated just after sunrise, checked in the late morning and afternoon, and closed overnight. We trapped at each trapping grid for 3 consecutive days and nights each month for 5 consecutive months each year (June - October) from 2016-2020 totaling approximately 45,000 Longworth trap nights and over 22,000 Tomahawk trap days (trap night/day = number of active traps x number of nights/days). Target species in this study were the deer mouse, the southern red-backed vole, the northern short-tailed shrew, and the North American red squirrel.
4.2.3. Animal processing and behavioral assays

All captures were taken to a location just outside the trapping grid for processing. Animals were transferred directly from the trap into three standard behavioral assays to measure behaviors that would later be used to assess personality. An emergence test was used to assess boldness (Brown and Braithwaite 2005; Carter et al. 2013), an open-field test to measure activity and exploration in a novel environment (Walsh and Cummins 1976; Perals et al. 2017), and a handling bag test to measure docility and the response to handling by an observer (Martin and Réale 2008; Montiglio et al. 2012; Taylor et al. 2014). We performed behavioral assays on the first capture of the month only to ensure that animals would not become habituated. Emergence and open-field tests were videotaped and behaviors were quantified from videos in the laboratory. See Brehm et al. (2019, 2020) and the supplementary materials (supplementary methods) for detailed field procedures, as well as software and methods used to quantify behavior. See Table D.1. (modified from Brehm et al., [2019]) for a complete list of the behaviors measured, their description and interpretation, and supporting sources.

Briefly, from the emergence test we obtained measures related to an animal's degree of boldness (the latency to emerge from a safe, enclosed space, and the time spent at the entrance of the “safe space” before emerging). From the open-field test we obtained measures relating to activity rates (such as the mean speed of the individual), exploratory activity (such as the rate of exploratory rearing), anxiety and stress de-arousal (proportion of time spent grooming), and boldness during exploration (proportion of time spent in the center portion of the arena). Finally, from the handling test we obtained a measure of docility (the number of seconds spent immobile in a one-minute test).

After the behavioral assays, we anesthetized animals with isoflurane and inserted PIT tags (Biomark MiniHPT8 8, 134.2 kHz) subcutaneously at the midback. With the exception of shrews, animals were also marked with a small animal ear tag (Style 1005-1, National Band and Tag Co., Newport, Kentucky, USA). Squirrel ear tags were threaded with a combination of
colored wire (females) or pipe cleaner (males) for identification at a distance (Boon et al. 2007; Brehm and Mortelliti 2018). We recorded sex, body mass (measured using a 100 or 1000 g Pesola Lightline spring scale), body length and tail length (all species except squirrels), age class (juvenile, subadult, or adult; based on body size and pelage coloration), and reproductive status (classified based on the presence of scrotal testes or signs of pregnancy/lactation). Squirrels were processed without the use of anesthetic and instead using a small mammal handling cone to prevent injury and limit stress. Animals were released at the site of capture post-processing.

4.2.4. Statistical analyses

We first performed a repeatability analysis using the ‘rptR’ package in R (Stoffel, Nakagawa, & Schielzeth, 2017) to determine which behavioral variables observed during standardized assays could be considered personality traits. See the supplementary materials (Supplementary methods) for further details about calculating repeatability. We opted not to use a dimensionality reduction technique on the behavioral variables (such as a principal component analysis, or PCA), to remain consistent with our previous work (Brehm & Mortelliti, 2018, 2021, 2022; Brehm et al., 2019, 2020), and because components retained from a PCA can become difficult to interpret biologically and also lead to a loss of information (Lever, Krzywinski, & Altman, 2017). Instead, we elected to use a restricted number of non-correlated behavioral variables consistent with our previous work. To account for variability in the behavior measurements of each individual we calculated individual’s mean BLUPs (best linear unbiased predictor) for each behavioral variable through 1000 simulations (Dingemanse, Moiron, Araya-Ajoy, Mouchet, & Abbey-Lee, 2020; Gharnit, Bergeron, Garant, & Réale, 2020) with package ‘arm’ (Gelman & Su, 2018). Subsequent mentions of personality refer to the mean BLUP value.

To investigate whether personality traits influence survival rates in our target species, we estimated monthly apparent survival using Robust Design Models with the Huggin’s estimator (Kendall 2011). We used year, trapping grid, forest treatment (reference/REF, uniform
shelterwood/USW, or irregular shelterwood/ISW), and sex as grouping variables. We assessed goodness of fit for each species using program RDSurviv (Kendall 2001), and adjusted c-hat during model selection if over-dispersion was detected in the data (Cooch and White 2019). We followed the information-theoretic approach to model selection (Burnham and Anderson 2002b) using the Akaike Information Criteria to rank competing models and considering models within 2.0Δ AICc to have equal support. We determined the top model for each parameter separately, modeling detection (ρ) and recapture (c) probabilities first, while holding emigration (γ’ and γ”) and survival (S) constant, and working with S last as this was the parameter of most biological interest to our study. For parameters used to estimate the best structure for ρ, c, and γ, see the supplemental materials (supplementary methods).

All variables were first included in candidate models singly, and we then tested additive models (by incorporating the next best supported trait one at a time until the addition of a trait did not result in a model scoring ≥2.0Δ AICc better or did not contain a significant beta for the additional variable). Candidate models for survival (S) included time (session [a fully time-varying model], session [continuous], and season [year]), sex, forest treatment (REF, USW, or ISW), trapping grid, body mass (as a proxy for age), body condition (estimated using the scaled-mass index for mice, voles and shrews [Peig & Green, 2009]), the density of conspecifics, and personality traits. In this study, a count of the total number of tagged conspecifics in the trapping grid and year of interest was used as a proxy for density. As deer mice and red-backed voles fill similar niches, we also tested the combined mouse/vole density in addition to the density of conspecifics. For individuals with more than one body mass or body condition measurement, we used the individual’s mean measurement.

After identifying the top model structure from this set of additive models, we incorporated interaction terms involving personality traits to test the following hypotheses: (1) whether the relationship between personality and survival differed between forest treatments [personality*treatment], (2) whether the relationship between personality and survival varied
depending on the density of conspecifics [personality*density], and (3) whether the relationship between personality and survival was affected differently by the forest treatment depending on conspecific density [personality*density*treatment]. For this portion of the analyses we tested interactions only with personality traits that fell within $2.0 \Delta \text{AICc}$ of the null structure or better. In all models, continuous predictor variables were z-standardized and any missing values were set to zero (equivalent to the mean of a scaled variable). For short-tailed shrews, we replaced one behavioral trait value with zero (the mean of the dataset) due to an extreme value (high leverage) in the behavioral variable “Proportion of time in the center” so that this trait would not inflate the strength of a regression in which it is included. All behavioral variables were scanned for collinearity prior to inclusion in robust design models and where pairwise correlations exceeded 0.7 (Dormann et al. 2013) we chose to move forward with the trait with higher repeatability. See Table D.2 for pairwise correlations between all behavioral variables used in model selection. Robust design analyses were performed using the package RMark in Program R (Laake 2013) and the program Mark (White and Burnham 1999).

4.3. Results

4.3.1. Small mammal trapping

Small mammal trapping from June-October in 2016-2020 resulted in 827 individual capture histories for *P. maniculatus*, 870 for *M. gapperi*, 279 for *B. brevicauda*, and 249 for *T. hudsonicus*. Capture histories included 25 primary sampling periods, each with three secondary sampling periods (except the primary period of October 2016, which had only two secondary sampling periods).

4.3.2. Repeatability

For detailed information about sample sizes used in repeatability analyses for each species, see the Supplementary materials (supplementary results and Table D.3.). Repeatable traits for deer mice and voles included the following (described in detail in Table D.1.): mean speed (an
indicator of activity), rear rate (activity and exploration), proportion time grooming (anxiety and stress de-arousal), proportion time center (boldness), handling time (docility), latency to emerge (boldness), and time at end of tunnel (boldness). Repeatable traits for shrews included handling time, mean speed, rear rate, and proportion time center. Repeatable traits for squirrels included handling time, mean speed, proportion time grooming, rate of jumping, and rate of rearing.

4.3.3. Top model structures for survival

For results of model selection using AICc and QAICc to rank Robust Design models with the Huggin’s estimator see Table D.4. Explanation of top model structures for detection, recapture, and emigration can be found in the Supplementary materials (supplementary results).

The apparent monthly survival probability ($S$) for deer mice was greater with increasing body mass ($\beta = 0.30 \pm 0.07$ SE) (Figure D2 A) but lower with increasing docility (handling time) ($\beta = -0.23 \pm 0.07$ SE) (Figure 4.2 A). Apparent monthly survival was lower in uniform shelterwood stands than in the reference and irregular shelterwood stands (real estimates: $S_{\text{REF,ISW}} = 0.67 \pm 0.02$ SE; $S_{\text{USW}} = 0.60 \pm 0.03$ SE). The best supported model included an interaction between stress de-arousal (the proportion of time spent grooming in the open-field test) and conspecific density ($\beta_{\text{personality*density}} = 0.34 \pm 0.07$ SE), but not between de-arousal and forest treatment (classifying all treatments as separate and with the uniform shelterwood forest grouped against reference and irregular shelterwood stands). Specifically, the $[\text{personality*density}]$ model showed that there was a positive relationship between grooming and survival when small mammal density was high (Figure 4.3B), whereas the relationship at low density was weaker but negative (Figure 4.3A). *Peromyscus* density and the combined density of *Peromyscus* and *Myodes* both had significant interaction terms with the grooming trait, but the model with the combined density was more than 8.0 ΔQAICc better, so we used the combined density instead.
Figure 4.2. Relationship between apparent survival and docility (seconds spent immobile in a handling test) as well as apparent survival and boldness (the proportion of time spent in the center portion of an open-field arena) predicted from best supported Robust Design models. Apparent survival is (A) lower in more docile deer mice (Peromyscus maniculatus; $\beta = -0.23 \pm 0.07$ SE) and (B) higher in bolder northern short-tailed shrews (Blarina brevicauda; $\beta = 0.57 \pm 0.32$ SE. Predictions for short-tailed shrews were obtained by model averaging the two top models. Docility and boldness are z-standardized variables. Shaded areas represent 95% CI.
Figure 4.3. Relationship between apparent survival and stress de-arousal behavior (the proportion of time spent grooming in the open-field test, z-standardized) at low vs. high small mammal densities predicted from the top-ranked Robust Design Model with the Huggin’s estimator for (A, B) deer mice (Peromyscus maniculatus) and (C, D) southern red-backed voles (Myodes gapperi). Increased grooming is associated with lower monthly survival probability at (A, C) low small mammal densities, but with higher monthly survival at (B, D) high small mammal densities (β = 0.34 ± 0.07 SE and β = 0.29 ± 0.07 SE, respectively). For Peromyscus models, a count of the total number of Peromyscus and Myodes in the trapping grid and year of interest was used as a proxy for density. For Myodes models, a count of Myodes only was used as a proxy for density (see methods for more detail). Predicted relationships were made using a range of densities spanning (A, C) the 1st quartile and (B, D) the 3rd to 4th quartile, and the grooming trait values present at those densities. Shaded areas represent 95% CI.
Apparent monthly survival probability for voles was greater with increasing body mass ($\beta = 0.49 \pm 0.08 \text{ SE}$) (Figure D.2 B), lower in males than in females ($\beta_{\text{Female} \ [\text{intercept}]} = -0.15 \pm 0.25 \text{ SE}; \beta_{\text{Male} \ [\text{intercept}]} = -0.36 \pm 0.13 \text{ SE}$), and different among years ($\beta_{16 \ [\text{intercept}]} = -0.15 \pm 0.25 \text{ SE}; \beta_{17} = 0.64 \pm 0.24 \text{ SE}; \beta_{18} = 0.67 \pm 0.24 \text{ SE}; \beta_{19} = 1.19 \pm 0.34 \text{ SE}; \beta_{20} = 1.16 \pm 0.36 \text{ SE}$). As was found for *Peromyscus*, the most supported model included an interaction between stress de-\text{arousal} (the proportion of time spent grooming in the open-field test) and conspecific density ($\beta_{\text{personality}*\text{density}} = 0.29 \pm 0.07 \text{ SE}$), but not between de-arousal and forest treatment. Specifically, the [personality*density] model showed that there was a positive relationship between grooming and survival in high density trapping grids/years only (Figure 4.3D), whereas the relationship at low density was stronger and negative (Figure 4.3C). *Myodes* density and the combined density of *Peromyscus* and *Myodes* both had significant interaction terms with the grooming trait, but the two models had nearly identical QAICc and QDeviance, so we used only *Myodes* density.

Two different model structures for apparent survival in short-tailed shrews had equal support by AICc ranking. Both models showed nearly identical positive relationships between body mass and apparent survival ($\beta = 0.45 \pm 0.11 \text{ SE}$ and $\beta = 0.41 \pm 0.10 \text{ SE}$) (Figure D.2 C) as well as differences in survival between years (real estimates: $S_{16} = 0.66 \pm 0.05 \text{ SE}; S_{17} = 0.79 \pm 0.03 \text{ SE}; S_{181920} = 0.59 \pm 0.04 \text{ SE}$ and real estimates $S_{16} = 0.67 \pm 0.05 \text{ SE}; S_{17} = 0.80 \pm 0.03 \text{ SE}; S_{181920} = 0.60 \pm 0.04 \text{ SE}$). One model showed a slight positive relationship between apparent survival and boldness (the proportion of time spent in the center of the open-field arena) ($\beta = 0.57 \pm 0.32 \text{ SE}$) (Figure 4.2 B). The models including the interactions of [personality*density] and [personality*treatment] were not among the top-ranking models.

Apparent monthly survival for red squirrels was greater with increasing body mass ($\beta = 0.37 \pm 0.14 \text{ SE}$) (Figure D.2 D) and decreased as red squirrel density increased ($\beta = -0.44 \pm 0.13 \text{ SE}$). None of the models including the personality traits individually or interactions of
[personality*density], [personality*treatment], and [personality*density*treatment] were among the top-ranking models.

4.4. Discussion
Using five years of mark-recapture data on four small mammal species, we found that in three of the four species, personality traits influenced apparent monthly survival, but not in the direction predicted by life history theory. Instead, riskier behaviors such as defensive aggression (i.e. handling time) and bold exploration (i.e. proportion of time spent in the center of the open-field arena) were associated with increased survival rates in deer mice and northern short-tailed shrews, respectively (Figure 4.2). Furthermore, we found support for fluctuating selection pressures as a mechanism maintaining personality traits in wild populations. Specifically, in both deer mice and red-backed voles, the effects of personality on survival were density-dependent and shifted direction depending on the abundance of conspecifics (Figure 4.3). These latter results supported our second hypothesis, that density generates context-dependence regarding the selective advantage of certain personality traits. Previous meta-analyses have shown that the overall directional effects of personality traits on survival are weak, but our findings suggest that this may be due to fluctuating pressures that shift the direction of selection.

Life history theory predicts that activity, aggression, and boldness mediate trade-offs between increased investment in reproduction and/or resource acquisition and increased mortality (Réale et al. 2010b). In practice, however, studies aimed at assessing the influence of risky personality traits on fitness have found mixed support. We emphasize that we did not directly test whether pace-of-life syndromes were at play in our study populations, because we were unable to assess whether increased survival rates were associated with delayed or suppressed reproduction. However, our findings show that risky behaviors did not lead to lower rates of apparent survival in small mammal populations. On the contrary, in line with the overall conclusions by (Moiron et al. 2020) on wild populations, we saw that specific risky behaviors
were associated with increased survival rates in wild small mammals (Figure 4.2). Defensive aggression in a handling bag test, that is, increased movement/escape attempts in a one-minute handling bag test, and bold exploration patterns in an open-field test (exploring in the center portion of an open-field arena [Archer, 1973]), were positively related to the rate of apparent survival in deer mice and northern short-tailed shrews, respectively. In these species it is possible that the benefits of risky behavior outweigh the costs associated with predation. Bold, aggressive individuals are shown to take more risks while foraging and are likely more competitive at obtaining resources, utilizing novel resources, or acquiring high quality territories (Both et al. 2005; Dammhahn and Almeling 2012). Consequently, these individuals may be able to offset the costs associated with predation risk.

Under our second hypothesis, we predicted that traits representing boldness, aggressiveness, or exploratory activity would influence survival differently when density was high vs. low due to a better competitive ability of risky individuals. Instead, in both deer mice and southern red-backed voles, a density-dependent relationship was seen with the proportion of time spent grooming in the open field test (Figure 4.3). In many rodents, self-grooming is a form of stress coping and de-arousal (Kalueff and Tuohimaa 2005; Fernández-Teruel and Estanislau 2016) and individuals will repetitively groom in a cephalocaudal progression (Kalueff et al. 2016) in response to stressful stimuli. Under low or mildly adverse stimuli, increased self-grooming seems to indicate increasing stress level, but under moderate to highly adverse stimuli, increased grooming has instead been suggested to indicate lower stress levels, as self-grooming ceases under extremely adverse stimuli (Fernández-Teruel and Estanislau 2016). Here, increased grooming was associated with higher apparent survival, but only during periods of high small mammal density (perhaps reflecting a better ability to moderate stress during high competition). Otherwise, at low density, individuals who groomed more experienced lower survival rates (Figure 4.3), possibly indicating that when social stresses are low, increased stress coping behaviors may be detrimental. These results suggest that fluctuations in
competition may play a role in maintaining personality variation in populations with strong seasonal or yearly variation in density.

Our findings under hypothesis #2 may also inspire future work examining self-regulation of rodent populations. The potential for intrinsic factors to self-regulate small mammal population fluctuations is less understood than that of extrinsic factors like food and predators (Krebs 2013). Population self-regulation occurs because high population density is a stressor which may control population growth through morphological, physiological, or behavioral changes that are stimulated by mutual interactions (Krebs 2013). The “polymorphic behavior hypothesis” (Chitty 1967) posits that at high densities, selection favors genotypes that have a worse effect on conspecifics (limiting population growth) whereas at low densities, selection favors genotypes that have a positive effect on conspecifics. Grooming behavior may be involved in a form of population self-regulation (i.e., de-arousal behavior may be connected to processes that limit population growth), but further research is needed to explore this hypothesis. Examining not only the quantity of grooming, but also the quality (i.e., the analysis of grooming behavioral microstructure) (Kalueff and Tuohimaa 2004, 2005), the physiological changes that may be present in high-grooming individuals (such as higher stress hormone levels which suppress reproduction [Sapolsky, Romero, Munck, Romero, & Munck, 2000]), or other behavioral changes that may associate with increased grooming (i.e., increased spacing behavior [Krebs, 2013]) will provide further insight into these findings.

One recent study on multimammate mice (Mastomys natalensis) observed a density-dependent relationship between a behavioral axis interpreted as “stress sensitivity” (characterized by high grooming and low jumping) and survival, where increased grooming was associated with lower survival rates only during the population decrease phase in an intra annual population cycle (Vanden Broecke et al. 2021). Our study approaches the question of density-dependence differently by examining interannual fluctuations in density rather than
seasonal fluctuations. It is interesting to note that there may be opposite effects of grooming on survival when intraspecific competition is broken down into seasonal vs yearly variation. The study by Vanden Broecke et al. was performed using a semi-wild experiment, monitoring individuals within enclosed fields; further work would be needed to assess this process in wild populations where immigration and emigration play a role. Future work may also combine seasonal increase/decrease phases as well as interannual fluctuations to expand more on these phenomena.

Our findings for red squirrels did not include any personality traits in the top model predicting apparent survival. Instead, apparent survival was higher in larger-bodied individuals (Figure D.2 D), and lower when squirrel density (and, thus, intraspecific competition) was high. Previous work on North American red squirrels has shown that more active females had lower overwinter survival (Boon et al. 2008), and work on Eurasian red squirrels (Sciurus vulgaris) identified changing fitness benefits of boldness in relation to habitat type (Santicchia et al. 2018), where bold squirrels had an advantage in forests with highly fluctuating food availability and shy squirrels had the advantage in forests with stable food supplies. A strength of our approach is that we directly accounted for the detection probability of individual squirrels. Specifically, we found that more docile red squirrels were more likely to be detected initially and also more likely to be recaptured than aggressive squirrels (Figure D.3). These methodological differences may explain the differences between studies, but further research is needed to clarify this.

Contrary to our predictions under our first hypothesis, there was no evidence that the relationship between personality traits and apparent survival differed in response to perceived risk. We expected to see that risky personality traits would positively impact survival rates where risk was perceived to be greater. For example, in forests where ground cover such as shrubs and coarse woody debris are less abundant, bolder individuals should have an advantage
because taking risks is necessary to acquire resources in this type of environment and aggressive individuals should outcompete docile ones for limited refuge habitat. In the three forest types used in our study, (reference/unmanaged - REF, uniform shelterwood - USW, and irregular shelterwood - ISW) there are significant differences in three features used as refuge habitat by small mammals (Figure 4.1 C). Specifically, the REF stands provide more shrub cover and coarse woody debris than either the USW or the ISW (Figure 4.1 C, and see Brehm & Mortelliti (2021) for further information for microhabitat differences between the study sites used). Shrub cover and large coarse woody debris are lowest in the USW grids compared to the others. While deer mouse survival rates were significantly lower in this treatment than in the REF and ISW stands, there was no significant interaction between USW and any of the personality traits tested. Previous work suggests that characteristics of the environment may alter the relationship between behavior and fitness, making certain personality traits more or less advantageous in different environments (Miranda et al. 2013; Lapiedra et al. 2017; Santicchia et al. 2018), and in our previous work on this study system we have shown differences in the composition of personality types present in forests managed with different silvicultural systems (Brehm et al. 2019a). For example, our work has shown that in USW forests, deer mice are bolder and more active on average than in the reference forests. It is possible that the behavioral differences observed between forest types in our study system are due to mechanisms other than differing selection pressures among the treatments. For example: niche picking, where individuals of certain personality types seek out habitats with specific features (Stamps and Groothuis 2010a; b; Sih et al. 2018; Schirmer et al. 2019), or density-driven assortment, where certain personality types seek out areas of high or low conspecific density (Pearish et al. 2013; Sih et al. 2018; Schirmer et al. 2019). In this study, our density measures were the number of individuals captured in the grid of interest during the year that the capture took place. Since density also varied spatially in this system, we may have been unable
to disentangle the effects of habitat and density. As our long-term project continues in future years, we should have more power to tease apart the mechanisms driving these phenomena.

In the last two decades, flourishing research on animal personalities has identified links between consistent behavioral differences and individual responses to habitat modification, which can shift the diversity and composition of personality traits in populations through human-induced changes such as forest management (Brehm et al. 2019a; Mortelliti and Brehm 2020), predator introduction (Lapiedra et al. 2018b), and urbanization (Miranda et al. 2013). Understanding how directional selection on personality traits might act is critical, therefore, to predict future phenomena such as personality filtering in populations experiencing anthropogenic changes or anticipating personality-mediated disease spread into urban habitats (Wat et al. 2020). Additionally, behavioral filtering leads to decreased behavioral diversity over time - which could have consequences for ecosystem services that are mediated through animal behavior like seed dispersal and pollination (Brehm et al. 2019a; Brehm and Mortelliti 2022; Hunter et al. 2022).

Using five years of mark-recapture data collected simultaneously on four species, we found associations between personality traits of wild small mammals and apparent survival, however, there were no observable survival costs of risky behaviors. Instead, in opposition to what life-history theory predicts, risky behaviors increased apparent survival rates in populations of deer mice and northern short-tailed shrews. This suggests that bold and aggressive individuals of these species are able to reduce the costs associated with increased predation. Furthermore, in deer mice and southern red-backed voles, the relationship between personality and survival was density-dependent and shifted direction depending on whether small mammal abundance was high or low. These findings supported our second hypothesis, that fluctuations in population density alter the effects of risky personality traits on survival, masking our ability to obtain universal, directional relationships between intraspecific behavioral variation and fitness.
components. These results emphasize the need to explore sources of context-dependence using empirical studies, and suggest that density may be important to consider when predicting personality-dependent selection in fluctuating systems.

4.5. Acknowledgments

We owe thanks to several field assistants and volunteers for their help with data collection, Sara Boone, Ivy Yen, Margaux Duparcq, and Margaret Merz for support with data collection and project management, and Dr. Laura Kenefic and Keith Kanoti for assistance facilitating research in the Penobscot Experimental Forest. Thank you to Erik Blomberg and Malcolm Hunter for feedback on an earlier version of this manuscript. This research was supported by an NSF Career Award to A.M. (IOS#1940525), a Grant-in-Aid of Research from the American Society of Mammalogists, and a Janet Waldron Doctoral Research Fellowship and Chase Distinguished Research Assistantship from the University of Maine to A.M.B, Penobscot Experimental Forest Research Funds, and the USDA National Institute of Food and Agriculture McIntire-Stennis projects ME041620 and ME041913 through the Maine Agricultural & Forest Experiment Station.


BREHM, A. M., A. MORTELLITI, G. A. MAYNARD, AND J. ZYDLEWSKI. 2019a. Land-use change and


CHOLERIS, E., A. W. THOMAS, M. KAVALIERS, AND F. S. PRATO. no date. A detailed ethological analysis of the mouse open field test: effects of diazepam, chlordiazepoxide and an extremely low frequency pulsed magnetic field.


DUPKE, C. ET AL. 2017. Habitat selection by a large herbivore at multiple spatial and temporal scales is primarily governed by food resources. Ecography 40:1014-1027.


Gosling, S. D. 2001. From mice to men: what can we learn about personality from animal research? Psychological bulletin 127:45-86.


IPCC. 2019. Climate change and land: An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems.


TURNER, B. L., E. F. LAMBIN, AND A. REENBERG. 2007. The emergence of land change science for
global environmental change and sustainability. Proceedings of the National Academy of
Sciences 104:20666-20671.

TUTTLE, M. D., AND M. J. RYAN. 1982. The role of synchronized calling, ambient light, and
ambient noise, in anti-bat-predator behavior of a treefrog. Behavioral Ecology and

UNDERHILL, V. ET AL. 2021. Personality and behavioral syndromes in two Peromyscus species:
presence, lack of state dependence, and lack of association with home range size.
Behavioral Ecology and Sociobiology 75.

Academy of Sciences of the United States of America 104:14377-14382.

VAN VALEN, L. 1965. Morphological variation and width of ecological niche. The American

VANDERWAAL, K. L., AND V. O. EZENWA. 2016. Heterogeneity in pathogen transmission:

habitat associations of red-backed voles in boreal mixedwood stands. The Journal of

Illinois, Illinois.

VANDER WALL, S. B. 2000. The influence of environmental conditions on cache recovery and
cache pilferage by yellow pine chipmunks (Tamias amoenus) and deer mice (Peromyscus

117.

VANDER WALL, S. B. 2010. How plants manipulate the scatter-hoarding behaviour of seed-
dispersing animals. Philosophical Transactions of the Royal Society B: Biological Sciences
365:989-997.

VANDER WALL, S. B., P. FORGET, J. LAMBERT, AND P. HULME. 2005a. Seed Fate: Predation,
Dispersal, and Seedling Establishment. CABI Publishing.

VANDER WALL, S. B., K. M. KUHN, AND M. J. BECK. 2005b. Seed removal, seed predation, and

VANDER WALL, S. B., AND W. S. LONGLAND. 1999. Cheek pouch capacities and loading rates of

densities: a new theoretical framework and an empirical test on an invasive zooplankton.
Ecosphere 9.

bulletin 83:482-504.


APPENDIX A: CHAPTER 1 SUPPLEMENTARY MATERIALS

Table A.1. Repeatability estimates for target behaviors measured in three behavioral tests (handling bag, emergence, and open-field) in deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*).

<table>
<thead>
<tr>
<th>Behavioral Variable</th>
<th>Mean</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P. maniculatus</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handling time</td>
<td>15.41</td>
<td>(0, 60)</td>
</tr>
<tr>
<td>Latency to emerge</td>
<td>27.17</td>
<td>(0, 180)</td>
</tr>
<tr>
<td>Time at end of tunnel</td>
<td>7.15</td>
<td>(0, 180)</td>
</tr>
<tr>
<td>Mean speed (m/sec)</td>
<td>0.10</td>
<td>(0, 0.25)</td>
</tr>
<tr>
<td>Prop. time grooming</td>
<td>0.11</td>
<td>(0, 0.96)</td>
</tr>
<tr>
<td>Rear rate</td>
<td>0.19</td>
<td>(0, 0.68)</td>
</tr>
<tr>
<td>Prop. time center</td>
<td>0.03</td>
<td>(0, 0.73)</td>
</tr>
<tr>
<td><strong>M. gapperi</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handling time</td>
<td>47.77</td>
<td>(0, 60)</td>
</tr>
<tr>
<td>Latency to emerge</td>
<td>34.94</td>
<td>(0, 180)</td>
</tr>
<tr>
<td>Time at end of tunnel</td>
<td>12.05</td>
<td>(0, 180)</td>
</tr>
<tr>
<td>Mean speed (m/sec)</td>
<td>0.05</td>
<td>(0, 0.20)</td>
</tr>
<tr>
<td>Prop. time grooming</td>
<td>0.06</td>
<td>(0, 0.81)</td>
</tr>
<tr>
<td>Rear rate</td>
<td>0.09</td>
<td>(0, 0.56)</td>
</tr>
<tr>
<td>Prop. time center</td>
<td>0.04</td>
<td>(0, 0.99)</td>
</tr>
</tbody>
</table>

Repeatability was calculated from univariate mixed-effect models with identity included as a random effect. Parametric bootstrapping was used to calculate 95% confidence intervals. See Methods for more information. Significant repeatability estimates are shown in bold.
Supplementary methods

Details on the behavioral tests

Following capture an animal was transferred directly from the trap of capture into a clean, empty Longworth trap. This trap was then placed into a box sized 46 x 46 x 50 cm. To create a more natural environment, the inside of the box had been painted light brown with a small amount of debris (dead leaves and pine needles) placed on the floor. The box was placed underneath a tarp to control for light levels and canopy cover. A digital camera (Nikon CoolPix S3700) was mounted facing the opening of the Longworth trap, and the observer locked the trap door open before leaving the test area. After three minutes, the observer returned and ended the test. Individuals were caught in a plastic bag and then released into the center of the open field arena.

A five minute open-field test was performed in an arena of dimensions 46 x 46 x 50 cm, placed on a level platform with light levels controlled (Brehm and Mortelliti 2018). After five minutes, the animal was caught in a 4-liter plastic bag and the observer suspended the bag into the open field test to control the visual surroundings. The observer measured the proportion of time that the individual spent immobile for 1 minute (referred to as handling time hereafter). Traps used for emergence tests and the open-field test box were cleaned thoroughly with 70% isopropyl alcohol and wiped with a dry cloth in-between all tests.
Figure B.1. Figures showing the total number of captures per individual used in third-order analyses for *Peromyscus maniculatus* and *Myodes gapperi*. These included 1716 observations from 570 individual deer mice and 1383 observations from 475 individual southern red-backed voles (mean number of observations per individual deer mouse was $3.01 \pm 2.89$ (range 1 to 29) and per southern red-backed vole was $2.91 \pm 2.61$ [range 1 to 18]). Each point represents one individual.
Figure B.2. Scree plot showing the Eigen values of principal components (PC) from a principal component's analysis (PCA) on microhabitat data recorded at each trap location in our study sites at the Penobscot Experimental Forest. The first two components were retained in this analysis.
Figure B.3. Kernel density plots of the first two microhabitat components (PC1 and PC2) for the reference and treatment areas used in this study. Whether or not the component value in each treatment differed from the value in the reference areas is indicated by a * ($p < 0.05$).
Figure B.4. Plot of the first two principal components from a principal component’s analysis (PCA) on microhabitat data recorded at each trap location in our study sites (unmanaged forest - REF, even-aged forest - EA, and 2-stage shelterwood forest - SH).
Table B.1. Glossary of the silvicultural treatments used in this study. Treatment number, silvicultural system used, and a brief definition of the system are given. Photos of each are shown here.

<table>
<thead>
<tr>
<th>Treatment #</th>
<th>Silvicultural system used</th>
<th>Definition</th>
<th>Site description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference</td>
<td>N/A</td>
<td>Unmanaged forest</td>
<td>Mature forest; many large cone-bearing trees; large quantities of downed wood and mosses; sparse understory</td>
</tr>
<tr>
<td>Treatment 1</td>
<td>Uniform shelterwood</td>
<td>Creates a stand of trees that are of the same age class and size (Ek, Katovich, Kilgore, and Palik 2006). This treatment can result in a dense canopy due to the close proximity of trees.</td>
<td>Dense, low-light forest; trees are of smaller diameter; absent of large cone-bearing trees</td>
</tr>
<tr>
<td>Treatment 2</td>
<td>Irregular shelterwood</td>
<td>Commonly used to increase growth and vertical structure in a stand (Brissette and Kenefic 2014), as well as provide downed wood and snags. Large trees from the older cohort are retained and then an even-aged cut is applied. The result is a stand with two defined vertical stages.</td>
<td>Forest with two age classes; dense in places with many mid-size trees; some large cone-bearing trees are present throughout</td>
</tr>
</tbody>
</table>

Reference: Unmanaged  

Treatment 1  

Treatment 2
Table B.2  **Pairwise correlations** between all behavioral variables used in this study for *Peromyscus maniculatus* and *Myodes gapperi.*

<table>
<thead>
<tr>
<th></th>
<th>Handling time</th>
<th>Latency to emerge</th>
<th>Time at tunnel end</th>
<th>Rear rate</th>
<th>Proportion center</th>
<th>Proportion grooming</th>
<th>Mean speed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peromyscus maniculatus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handling time</td>
<td>1.000</td>
<td>0.319</td>
<td>0.466</td>
<td>0.383</td>
<td>-0.191</td>
<td>0.017</td>
<td>-0.421</td>
</tr>
<tr>
<td>Latency to emerge</td>
<td></td>
<td>1.000</td>
<td>0.638</td>
<td>0.306</td>
<td>-0.112</td>
<td>-0.073</td>
<td>-0.221</td>
</tr>
<tr>
<td>Time at tunnel end</td>
<td>0.466</td>
<td>0.638</td>
<td>1.000</td>
<td>0.436</td>
<td>-0.126</td>
<td>-0.171</td>
<td>-0.233</td>
</tr>
<tr>
<td>Rear rate</td>
<td>-0.383</td>
<td>-0.306</td>
<td>-0.436</td>
<td>1.000</td>
<td>0.289</td>
<td>0.056</td>
<td>0.535</td>
</tr>
<tr>
<td>Proportion center</td>
<td>-0.191</td>
<td>-0.112</td>
<td>-0.126</td>
<td>0.289</td>
<td>1.000</td>
<td>-0.073</td>
<td>0.365</td>
</tr>
<tr>
<td>Proportion grooming</td>
<td>0.017</td>
<td>-0.073</td>
<td>-0.171</td>
<td>0.056</td>
<td>-0.073</td>
<td>1.000</td>
<td>-0.427</td>
</tr>
<tr>
<td>Mean speed</td>
<td>-0.421</td>
<td>-0.221</td>
<td>-0.233</td>
<td>0.535</td>
<td>0.365</td>
<td>-0.427</td>
<td>1.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Handling time</th>
<th>Latency to emerge</th>
<th>Time at tunnel end</th>
<th>Rear rate</th>
<th>Proportion center</th>
<th>Proportion grooming</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Myodes gapperi</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handling time</td>
<td>1.000</td>
<td>0.041</td>
<td>-0.029</td>
<td>0.053</td>
<td>-0.137</td>
<td>0.029</td>
</tr>
<tr>
<td>Latency to emerge</td>
<td></td>
<td>1.000</td>
<td>0.579</td>
<td>0.216</td>
<td>-0.071</td>
<td>-0.138</td>
</tr>
<tr>
<td>Time at tunnel end</td>
<td>-0.029</td>
<td>0.579</td>
<td>1.000</td>
<td>0.312</td>
<td>-0.165</td>
<td>-0.219</td>
</tr>
<tr>
<td>Rear rate</td>
<td>-0.053</td>
<td>-0.216</td>
<td>-0.312</td>
<td>1.000</td>
<td>0.560</td>
<td>0.483</td>
</tr>
<tr>
<td>Proportion center</td>
<td>-0.137</td>
<td>-0.071</td>
<td>-0.165</td>
<td>0.560</td>
<td>1.000</td>
<td>0.222</td>
</tr>
<tr>
<td>Proportion grooming</td>
<td>0.029</td>
<td>-0.138</td>
<td>-0.219</td>
<td>0.483</td>
<td>0.222</td>
<td>1.000</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Habitat selection</td>
<td>The usage of certain habitat elements disproportionately to their availability.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-order selection</td>
<td>The physical or geographical range of a species.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second-order selection</td>
<td>Selection of the home range within the physical or geographical range. Here, we specify second-order selection as the selection of the home range within the trapping grid.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third-order selection</td>
<td>Selection of various habitat components within the home range.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fourth-order selection</td>
<td>Selection of fine-scale components. For example, if third-order selection represents the selection of feeding sites, fourth-order selection represents the actual food items obtained.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table B.4. Microhabitat variables measured in this study. A description of the variable is provided.

<table>
<thead>
<tr>
<th>Microhabitat variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>% cover grasses and herbs</td>
<td>Approximate percent ground cover composed of grasses, sedges, and small herbaceous plants. Also included shrubs and saplings below knee height.</td>
</tr>
<tr>
<td>% cover shrubs/saplings &lt;1 meter</td>
<td>Approximate percent ground cover composed of shrubs and saplings &lt;1m in height. Excluded those below knee height.</td>
</tr>
<tr>
<td>% cover shrubs/saplings 1-2 meters</td>
<td>Approximate percent ground cover composed of shrubs and saplings 1-2m in height.</td>
</tr>
<tr>
<td>% cover shrubs/saplings 2-4 meters</td>
<td>Approximate percent ground cover composed of shrubs and saplings 2-4m in height.</td>
</tr>
<tr>
<td>% cover mosses</td>
<td>Approximate percent of ground covered in mosses. Included rocks, roots, and downed wood that were covered in mosses.</td>
</tr>
<tr>
<td>meters of coarse woody debris (CWD)</td>
<td>Total length (in meters) of coarse woody debris (CWD - defined as dead wood that had fallen onto the ground or was at an angle ≤ 45 from the ground and supported by the ground in at least 3 locations) with a diameter between 10 and 20 cm at its widest point.</td>
</tr>
<tr>
<td>10-20 cm</td>
<td></td>
</tr>
<tr>
<td>meters of coarse woody debris (CWD)</td>
<td>Total length (in meters) of coarse woody debris (CWD - defined as dead wood that had fallen onto the ground or was at an angle ≤ 45 from the ground and supported by the ground in at least 3 locations) with a diameter &gt; 20 cm at its widest point.</td>
</tr>
<tr>
<td>20+ cm</td>
<td></td>
</tr>
<tr>
<td>% canopy cover</td>
<td>Percent canopy cover (calculated using a concave spherical densitometer).</td>
</tr>
<tr>
<td>DBH of largest tree</td>
<td>Diameter at breast height of the largest tree within 5m of the trap site.</td>
</tr>
</tbody>
</table>
Table B.5. Loadings from the first two components extracted from a principal component’s analysis (PCA) on microhabitat data recorded at each trap location in our study site at the Penobscot Experimental Forest.

<table>
<thead>
<tr>
<th>Microhabitat variable</th>
<th>PC1</th>
<th>PC2</th>
</tr>
</thead>
<tbody>
<tr>
<td>% cover grasses and herbs</td>
<td>0.77</td>
<td>0</td>
</tr>
<tr>
<td>% cover shrubs/saplings &lt;1 meter</td>
<td>0.48</td>
<td>0.37</td>
</tr>
<tr>
<td>% cover shrubs/saplings 1-2 meters</td>
<td>0.65</td>
<td>0</td>
</tr>
<tr>
<td>% cover shrubs/saplings 2-4 meters</td>
<td>0.14</td>
<td>-0.46</td>
</tr>
<tr>
<td>% cover mosses</td>
<td>0.65</td>
<td>0</td>
</tr>
<tr>
<td>meters of CWD 10-20 cm</td>
<td>0.18</td>
<td>0.59</td>
</tr>
<tr>
<td>meters of CWD 20+ cm</td>
<td>0.06</td>
<td>0.68</td>
</tr>
<tr>
<td>% canopy cover</td>
<td>-0.45</td>
<td>0.21</td>
</tr>
<tr>
<td>DBH of largest tree</td>
<td>-0.21</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Cumulative % variance explained 0.22 0.40

A simplimax rotation was used for this PCA. Bold type indicates microhabitat variables that contributed importantly to a component (using a threshold of 0.3) (Boon et al. 2007, Mazzamuto et al. 2018). See text for more information.
Table B.6. Repeatability estimates for target behaviors measured in three behavioral tests (handling bag, emergence, and open field) in deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*).

<table>
<thead>
<tr>
<th>Behavioral Variable</th>
<th>Mean</th>
<th>Range</th>
<th>RPT</th>
<th>(95% CI)</th>
<th>Observations</th>
<th>Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P. maniculatus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handling time (sec)</td>
<td>16.06</td>
<td>(0, 60)</td>
<td>0.390</td>
<td>(0.299, 0.498)</td>
<td>700</td>
<td>428</td>
</tr>
<tr>
<td>Latency to emerge (sec)</td>
<td>56.20</td>
<td>(0, 225)</td>
<td>0.353</td>
<td>(0.299, 0.498)</td>
<td>641</td>
<td>394</td>
</tr>
<tr>
<td>Time tunnel (sec)</td>
<td>7.08</td>
<td>(0, 180)</td>
<td>0.234</td>
<td>0.361</td>
<td>578</td>
<td>356</td>
</tr>
<tr>
<td>Mean speed (m/sec)</td>
<td>0.09</td>
<td>(0, 0.27)</td>
<td>0.433</td>
<td>0.525</td>
<td>891</td>
<td>574</td>
</tr>
<tr>
<td>Prop. groom</td>
<td>0.13</td>
<td>(0, 0.96)</td>
<td>0.348</td>
<td>0.457</td>
<td>889</td>
<td>572</td>
</tr>
<tr>
<td>Rear rate (rears/sec)</td>
<td>0.20</td>
<td>(0, 0.68)</td>
<td>0.365</td>
<td>0.459</td>
<td>891</td>
<td>574</td>
</tr>
<tr>
<td>Prop. time center</td>
<td>0.03</td>
<td>(0, 0.73)</td>
<td>0.297</td>
<td>0.403</td>
<td>890</td>
<td>573</td>
</tr>
<tr>
<td><strong>M. gapperi</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handling time (sec)</td>
<td>48.96</td>
<td>(0, 60)</td>
<td>0.145</td>
<td>(0.033, 0.295)</td>
<td>528</td>
<td>343</td>
</tr>
<tr>
<td>Latency to emerge (sec)</td>
<td>99.62</td>
<td>(0, 225)</td>
<td>0.295</td>
<td>(0.180, 0.450)</td>
<td>478</td>
<td>316</td>
</tr>
<tr>
<td>Time tunnel (sec)</td>
<td>12.09</td>
<td>(0, 180)</td>
<td>0.261</td>
<td>0.453</td>
<td>351</td>
<td>244</td>
</tr>
<tr>
<td>Mean speed (m/sec)</td>
<td>0.05</td>
<td>(0, 0.20)</td>
<td>0.228</td>
<td>0.360</td>
<td>683</td>
<td>470</td>
</tr>
<tr>
<td>Prop. groom</td>
<td>0.07</td>
<td>(0, 0.81)</td>
<td>0.174</td>
<td>0.305</td>
<td>683</td>
<td>470</td>
</tr>
<tr>
<td>Rear rate (rears/sec)</td>
<td>0.09</td>
<td>(0, 0.56)</td>
<td>0.231</td>
<td>0.371</td>
<td>683</td>
<td>470</td>
</tr>
<tr>
<td>Prop. time center</td>
<td>0.04</td>
<td>(0, 0.99)</td>
<td>0.216</td>
<td>0.361</td>
<td>683</td>
<td>470</td>
</tr>
</tbody>
</table>

Repeatability and the associated 95% confidence intervals were calculated using the rptR package in software R. The random effect of individual ID was included in the models and fixed effects included sex, body condition, silvicultural treatment, and trapping session. Box-cox transformations were used on the response variables when necessary to approach normality. The mean number of tests per individual *Peromyscus* was 1.65 ± 1.08 (range: 1-10) and per individual *Myodes* was 1.54 ± 0.83 (range: 1-5).
Table B.7. Results of repeatability analysis on variables of microhabitat selection at the third-order in deer mice (*Peromyscus maniculatus*) and southern red-backed voles (*Myodes gapperi*).

<table>
<thead>
<tr>
<th>Microhabitat selection</th>
<th>Repeatability (95% CI)</th>
<th>Observations</th>
<th>Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peromyscus maniculatus</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>selectTC1</td>
<td>0.123 (0.073, 0.173)</td>
<td>1716</td>
<td>570</td>
</tr>
<tr>
<td>selectTC2</td>
<td>0.060 (0.022, 0.103)</td>
<td>1716</td>
<td>570</td>
</tr>
<tr>
<td><strong>Myodes gapperi</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>selectTC1</td>
<td>0.130 (0.080, 0.193)</td>
<td>1383</td>
<td>475</td>
</tr>
<tr>
<td>selectTC2</td>
<td>0.092 (0.041, 0.153)</td>
<td>1383</td>
<td>475</td>
</tr>
</tbody>
</table>

Repeatability and the associated 95% confidence intervals were calculated using the rptR package in software R. The random effect of individual ID was included in the models and fixed effects included sex, body condition, silvicultural treatment, trapping session, and trapping year.
Supplementary Methods

Study site

The Penobscot Experimental Forest is a 1578 ha Forest Service experimental forest. Here, different forest units have been logged separately, managed with contrasting silvicultural treatments, and replicated twice in a randomized experimental design. Due to the contrasting silvicultural systems, these forest units differ greatly in the understory density, diameter of trees present, light levels, and quantity of downed woody material and snags. Approximately 25 ha of forest has been retained in two separate units and left unmanaged since the late 1800s to serve as reference. Common coniferous tree species include balsam fir (*Abies balsamea*), red spruce (*Picea rubens*), eastern white pine (*Pinus strobus*), and eastern hemlock (*Tsuga canadensis*), and deciduous species include red maple (*Acer rubrum*), American beech (*Fagus grandifolia*), Northern red oak (*Quercus rubra*), and birch and aspen (*Betula and Populus spp.*.)

In this study, we performed small mammal trapping in six different areas of the experimental forest. Four areas were located in two contrasting silvicultural treatments (a uniform shelterwood cut using a two-stage overstory removal [treatment 1], and an irregular shelterwood cut using a two-stage overstory removal and retaining reserves, or trees from the older cohort [treatment 2]. We also used the two areas of unmanaged forest as reference sites. The area of the treatments used in this study was 12.8 ha on average (range: 9.49 - 19.39 ha).

Behavioral assays

Following capture an animal was transferred directly from the trap of capture into a clean, empty Longworth trap. This trap was then placed into a box sized 46 x 46 x 50 cm. To create a more natural environment, the inside of the box was painted brown with a small amount of debris (dead leaves and pine needles) placed on the floor. The box was placed underneath a tarp to control for light levels and canopy cover. A digital camera (Nikon CoolPix S3700) was mounted facing the opening of the Longworth trap, and the observer locked the trap door open before leaving the test area. A clear plexiglass lid was placed over the box to prevent escapes. After three minutes, the observer returned and ended the emergence test. Individuals were caught in a 4-liter plastic bag and then immediately released into the center of the open field arena.

A five-minute open-field test was performed in an arena of dimensions 46 x 46 x 50 cm, placed on a level platform with light levels controlled. After five minutes, the animal was caught in a 4-liter plastic bag and the observer suspended the bag into the open field test to control the visual surroundings. The observer then performed the handling bag test and measured the proportion of time that the individual spent immobile for 1 minute (referred to as handling time hereafter). Traps used for emergence tests and the open-field test box were cleaned thoroughly with 70% isopropyl alcohol and wiped with a dry cloth in-between all tests. After all three behavioral assays were complete (on average, this took approximately 10 minutes including time to transfer the animal between tests), the animal was processed (tagged and morphometrics measured) and promptly released at the point of capture.

Quantifying behavior

To quantify behavior from videotaped emergence and open-field tests, recordings were played back in the laboratory. For emergence tests, an observer recorded whether the animal emerged (defined as all four feet having left the Longworth trap), the latency to emerge, and the total time...
spent at the end of the Longworth tunnel before emerging. When an individual did not emerge from the test after the three-minute cutoff, the latency to emerge was set to 1.25x the maximum test length (this occurred in 156 tests out of 1164 total). Open-field tests were analyzed using the behavioral tracking software ANY-maze © (version 5.1; Stoelting CO, USA). The following behavioral variables were obtained from the behavioral assays: handling time (the number of seconds immobile in a handling test; note that the term handling time should not be confused with a term sometimes used in ecological literature to indicate the time spent handling, processing, and consuming food items), latency to emerge and time at tunnel end (from the emergence test), mean speed, rear rate, proportion of time grooming, and proportion of time in the center (from the open-field test). See Table C.3. for a complete list of the behaviors used, their description and interpretation, and supporting sources.

**Seed experiments - further details**

Seeds used in this study included Northern red oak (*Quercus rubra*), Eastern white pine (*Pinus strobus*), and American beech (*Fagus grandifolia*). These three seed species were chosen because they are present at our study sites, they represent a variety of sizes (while being large enough to track using the fluorescent powder method), and they represent seeds whose primary dispersal method includes both wind and animal dispersal. Seeds were purchased from F.W. Schumacher Co., Inc. and cold stratified over the winter.

Presentation stations were placed on the forest floor and we mounted a trail camera (Reconyx XR6 Ultrafire) 1.25 to 1.5 meters above the station. Cameras were formatted to take 30 second 1080P HD videos (at 30 fps) and an 8-megapixel photograph prior to the start of the video and once per hour. The cameras were set to the shortest delay between videos (1 s). To identify individual small mammals, we used a permanent radio frequency identification (RFID) reader to scan and identify individuals marked with PIT tags (RFIDLOG dual animal tag rfid data logger). An antenna (Priority1 rfidcoil - 160a) was mounted to each presentation station and positioned to surround the seed presentation trays (Figure 4D). These antennas were built to operate at a frequency of 134.2 kHz for optimal reading of PIT tags used in this study. Records were stored automatically on an SD memory card along with a time stamp of the detection.

To allow for the relocation of seeds removed from the stations, we dusted the vinyl floor tiles with UV fluorescent tracking powder (TechnoGlow; yellow, firehouse orange, or magenta). These pigments are made of earth friendly materials and shine brightly under UV light. Small mammals would leave distinct fluorescent trails when dispersing seeds from the seed stations (Longland and Clements 1995). To permit located caches to be matched to the individual disperser, each oak and beech seed was painted with non-toxic, UV fluorescent paint (Neon Glow, ASTMD-4236 certified). Each seed at a site was painted a different color, and the location of each color on the presentation tray (randomly chosen and rotated between stations), as well as the mass of beech and oak seeds were recorded when stations were set. When a seed was found, the color allowed us to match the cache to the individual small mammal observed dispersing it in trail camera videos. For white pines, five seeds were placed in each of the six wells and all five seeds in a well were painted the same color. Since several white pine seeds can be dispersed at a time, caches were identified by matching the count of each different color seed found within the cache to the seeds seen taken in videos (Figure 4F, G). Since all
experiments were performed in complete darkness, seed color should not have influenced the initial decision to consume or cache the seed.

On average we ran four or five experiment sites at one time and performed the experiment at a trapping grid for 3-4 days. In total, we placed 52 seed sites in September (eight sites at two grids and nine sites at four grids) and 51 in October (eight sites at three grids and nine sites at three grids).

After completion of the seed experiments, we simulated five caches outside of four of our trapping grids. Half of these caches contained one painted beech seed, and the other half contained one unpainted beech seed. We monitored these caches using trail cameras over 11 days and confirmed using a two-sample t-test that the chance of a cache being pilfered or recovered is not influenced by our methods of tagging seeds with paint (p-value comparing the chance of painted vs. unpainted seeds being pilfered within 24 hours = 0.178; p-value comparing cache longevity of painted vs. unpainted seeds = 0.219). We acknowledge that we did not specifically test whether the different colors of paint would have an effect on an individual’s choice to consume or cache a seed, but given the fact that all experiments were performed in the complete darkness of the forest, we have no reason to believe this should have occurred.

**Playback of seed videos**

From the trail camera videos obtained at the seed presentation stations, we recorded each visit to the seed stations, noting the identity of each individual and the nature of their interaction (i.e., which seed species they chose, whether they chose to consume it at the site or remove it, and the color of the seed they chose). We deemed behavior as an “interaction” with a seed if the individual intentionally made contact with the seed. Individual identification was confirmed using both the time-stamped RFID reads and the unique haircut seen in the trail cam videos.

**Repeatability analyses**

We calculated the adjusted repeatability (Wilson 2018) and associated 95% confidence intervals for seven key behaviors performed in the standardized assays using the ‘rptR’ package in R (Stoffel et al. 2017). For this analysis we used only individuals who had two or more repeated observations. Repeatability can be defined as the proportion of the total phenotypic variation that can be attributed to individual differences (Dingemanse et al. 2009; Nakagawa and Schielzeth 2010; Dingemanse and Dochtermann 2013). Practically, this means that repeatability equals the between-individual variance divided by the total phenotypic variance (between-individual variance + within-individual variance). The random intercept of individual ID was included in the models and fixed effects included sex, body condition index (calculated using the scaled-mass index (Peig and Green 2009)), silvicultural treatment, and trapping session. In all models, we used 1000 parametric bootstraps and 100 permutations. We assessed normality by visually inspecting Q-Q plots and by plotting the fitted values against the residual values and used Box-Cox transformations on the response variable to approach normality when necessary (Box and Cox 1964; Yang et al. 2011). We considered any behavioral trait with a 95% CI for
repeatability that excluded zero to be a personality trait (Nakagawa and Schielzeth 2010), but we emphasize that this classification as significantly repeatable does not say anything about the strength of repeatability.

After we confirmed that our behavioral variables were significantly repeatable, we took steps to account for the within-individual variability that occurs between repeated behavioral observations (consistent with methods used by (Dingemanse et al. 2020; Gharnit et al. 2020)). We calculated an individual's best linear unbiased predictor (BLUP) for each of the behavioral variables after controlling for the variation due to sex, body condition, forestry treatment, and trapping session. We then calculated an individual's mean BLUP for each of the repeatable behavioral variables (estimated over 1000 simulations using the sim function in package arm). We then used the mean BLUPs as the personality covariates in our linear models and mixed-effects models.

Further, before moving forward with model selection, we screened all seven behavioral variables for correlation before analysis (using R<0.7 as a threshold, consistent with (Dormann et al. 2013)). See Table C.6. for pairwise correlations between all behavioral variables used in model selection.
Figure C.1. White pine (*Pinus strobus*) interaction diagram showing the number of interactions made by deer mice (*Peromyscus maniculatus*) classified as positive (outlined in green) or negative (outlined in red) and the nature of these interactions.
Figure C.2. Red oak (*Quercus rubra*) interaction diagram showing the number of interactions made by deer mice (*Peromyscus maniculatus*) classified as positive (outlined in green) or negative (outlined in red) and the nature of these interactions.
Figure C.3. American beech (*Fagus grandifolia*) interaction diagram showing the number of interactions made by deer mice (*Peromyscus maniculatus*) classified as positive (outlined in green) or negative (outlined in red) and the nature of these interactions.
Figure C.4. Individual scores for *Peromyscus maniculatus* along the antagonism-mutualism continuum for (A) eastern white pine (*Pinus strobus*; mean score = -0.36) and (B) American beech (*Fagus grandifolia*; mean score = -0.37).
Figure C.5. Predicted relationships (and 95% CIs) between an index of body condition and an individual’s score along the antagonism-mutualism continuum. Individual *Peromyscus maniculatus* with a higher body condition index have lower (more antagonistic) scores for eastern white pine (*Pinus strobus*; $\beta = -0.31 \pm 0.15$ SE). Data points represent observed values (one point per individual). Body condition index represents a scaled-mass index and values were z-standardized. Relationship between body condition index and white pine score is shown for the treatment 2 forest type.
Table C.1. Interactions used to calculate an individual's score along the antagonist/mutualist continuum.

<table>
<thead>
<tr>
<th>Action</th>
<th>Positive or Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seed consumed at the site</td>
<td>Negative</td>
</tr>
<tr>
<td>Seed removed from site and cached intact</td>
<td>Positive</td>
</tr>
<tr>
<td>Seed removed from site and then consumed</td>
<td>Negative</td>
</tr>
<tr>
<td>Seed removed from site and taken down a hole</td>
<td>Negative</td>
</tr>
<tr>
<td>*Seed left intact at the site after interaction</td>
<td>*Positive</td>
</tr>
</tbody>
</table>

The score can be quantified as the proportion of interactions that are positive minus the proportion of interactions that are negative.

* This behavior was observed regularly for white pine seeds only. Further, white pine seeds are primarily wind dispersed, so for this species a non-consumption/rejection can be considered positive.
Table C.2. Repeatability estimates for target behaviors measured in three behavioral assays (open field, handling bag, and emergence tests) in deer mice (*Peromyscus maniculatus*).

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Mean</th>
<th>Range</th>
<th>RPT</th>
<th>(95% CI)</th>
<th>Observations</th>
<th>Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean speed (m/sec)</td>
<td>0.09</td>
<td>(0, 0.29)</td>
<td><strong>0.447</strong></td>
<td>(0.371, 0.515)</td>
<td>815</td>
<td>300</td>
</tr>
<tr>
<td>Rear rate (rears/sec)</td>
<td>0.19</td>
<td>(0, 0.65)</td>
<td><strong>0.361</strong></td>
<td>(0.285, 0.445)</td>
<td>819</td>
<td>301</td>
</tr>
<tr>
<td>Prop time groom¹</td>
<td>0.20</td>
<td>(0, 0.97)</td>
<td><strong>0.389</strong></td>
<td>(0.312, 0.467)</td>
<td>818</td>
<td>301</td>
</tr>
<tr>
<td>Prop time in center²</td>
<td>0.02</td>
<td>(0, 0.53)</td>
<td><strong>0.280</strong></td>
<td>(0.211, 0.368)</td>
<td>816</td>
<td>301</td>
</tr>
<tr>
<td>Handling time (sec)</td>
<td>12.82</td>
<td>(0, 60)</td>
<td><strong>0.342</strong></td>
<td>(0.268, 0.426)</td>
<td>749</td>
<td>268</td>
</tr>
<tr>
<td>Latency to emerge (sec)</td>
<td>47.44</td>
<td>(0, 225)</td>
<td><strong>0.300</strong></td>
<td>(0.221, 0.386)</td>
<td>689</td>
<td>250</td>
</tr>
<tr>
<td>Time at tunnel end (sec)</td>
<td>4.67</td>
<td>(0, 180)</td>
<td><strong>0.191</strong></td>
<td>(0.104, 0.295)</td>
<td>617</td>
<td>229</td>
</tr>
</tbody>
</table>

¹Proportion of time spent grooming in the open-field test.

²Proportion of time spent in center portion of open-field arena.
Table C.3. Repeatable behavioral traits from three standard behavioral assays performed on deer mice (*Peromyscus maniculatus*). This table provides the description of the behavior, the personality trait that this behavior represents, the assay that each trait was obtained from, and how values were interpreted. Included is a non-exhaustive list of citations supporting trait interpretation. This table was adapted from (Brehm et al. 2019a; Brehm and Mortelliti 2021).

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Personality trait</th>
<th>Behavioral assay</th>
<th>Description</th>
<th>Interpretation</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean speed</td>
<td>Activity</td>
<td>Open field</td>
<td>Mean speed in the open field test in (m/s). Calculated by dividing the total distance traveled in the test by the test duration</td>
<td>Higher values indicate greater activity</td>
<td>(Carter et al. 2013; Gracceva et al. 2014)</td>
</tr>
<tr>
<td>Rear rate</td>
<td>Activity and exploration</td>
<td>Open field</td>
<td>Rate of rearing (rears/s). Rearing is defined as forelegs leaving the arena floor</td>
<td>Higher values indicate greater activity and exploration</td>
<td>(Choleris et al. 2001; Prut and Belzung 2003; Martin and Réale 2008; Tanaka et al. 2012)</td>
</tr>
<tr>
<td>Proportion time grooming</td>
<td>Anxiety/stress</td>
<td>Open field</td>
<td>Proportion of test duration spent grooming</td>
<td>Higher values indicate lower anxiety and better coping</td>
<td>(Choleris et al. 2001; Fernández-Teruel and Estanislau 2016; Kalueff et al. 2016)</td>
</tr>
<tr>
<td>Proportion time center</td>
<td>Bold/timid</td>
<td>Open field</td>
<td>Proportion of test duration spent in the center portion of the arena</td>
<td>Higher values indicate more boldness/less timidness</td>
<td>(Barnett 1976; Treit et al. 1989; Ramos et al. 1997; Choleris et al. 2001; Eccard and Herde 2013; Gracceva et al. 2014)</td>
</tr>
<tr>
<td>Handling time</td>
<td>Docility</td>
<td>Handling bag</td>
<td>Total number of seconds of inactivity during a 1-minute handling bag test</td>
<td>Higher values indicate more docility/less defensive aggression</td>
<td>(Boon et al. 2007; Martin and Réale 2008; Montiglio et al. 2012; Taylor et al. 2014)</td>
</tr>
<tr>
<td>---------------</td>
<td>----------</td>
<td>--------------</td>
<td>-------------------------------------------------</td>
<td>-------------------------------------------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>Latency to emerge</td>
<td>Bold/timid</td>
<td>Emergence</td>
<td>Latency (in seconds) to emerge from trap in the emergence test. An animal was considered to have emerged when all four feet left the trap tunnel</td>
<td>Higher values indicate more timidness/less boldness</td>
<td>(Brown and Braithwaite 2005; Carter et al. 2013; Gracceva et al. 2014)</td>
</tr>
<tr>
<td>Time at end of tunnel</td>
<td>Bold/timid</td>
<td>Emergence</td>
<td>Total number of seconds spent at the end of the tunnel before emerging</td>
<td>Higher values indicate more timidness/less boldness</td>
<td>(Brehm et al. 2019a; b)</td>
</tr>
</tbody>
</table>
Table C.4. Results from linear regression predicting scores along the predator-mutualist continuum in the deer mouse (*Peromyscus maniculatus*) for three target seed species: eastern white pine (*Pinus strobus*), red oak (*Quercus rubra*), and American beech (*Fagus grandifolia*).

Models within 2.0 $\Delta$AICc of the top model are shown.

<table>
<thead>
<tr>
<th>Species</th>
<th>Model</th>
<th>$\Delta$AICc</th>
<th>Adj. R squared</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Pinus strobus</em></td>
<td>Body condition + Forest type + Time at tunnel end</td>
<td>0</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>Body condition$^3$ (-0.31 ± 0.15 SE) + Forest type</td>
<td>0.5</td>
<td>0.34</td>
</tr>
<tr>
<td><em>Quercus rubra</em></td>
<td>Time at tunnel end (-0.44 ± 0.15 SE)</td>
<td>0</td>
<td>0.29</td>
</tr>
<tr>
<td><em>Fagus grandifolia</em></td>
<td>Number of interactions</td>
<td>0</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Proportion of time in center</td>
<td>0.37</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Null</td>
<td>0.67</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Forest type</td>
<td>1.76</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Body condition</td>
<td>1.97</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Models for *P. strobus* include scores from 26 individuals, models for *Q. rubra* include scores from 19 individuals, and models for *F. grandifolia* include scores from 21 individuals. Effect sizes and standard errors shown in parentheses for top models.

Behavioral variables in models above represent an individual's average BLUP (i.e. best linear unbiased predictor) estimated over 1000 simulations.

$^1$ Delta Akaike information criterion corrected for small sample sizes

$^2$ Adjusted coefficient of determination

$^3$ Body condition index calculated using the scaled-mass index (Peig and Green 2009)
Table C.5. Results for logistic mixed-models predicting the probability of a positive seed interaction in the deer mouse (*Peromyscus maniculatus*). Models within 2.0 ΔAICc\(^1\) of the top model are shown.

<table>
<thead>
<tr>
<th>Species</th>
<th>Model*</th>
<th>ΔAICc(^1)</th>
<th>Cond. R squared(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Pinus strobus</em></td>
<td>Forest type + Body condition(^3) (-0.43 ± 0.20 SE) +</td>
<td>0</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>Time in center (-1.24 ± 0.41 SE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Quercus rubra</em></td>
<td>Sex</td>
<td>0</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Null</td>
<td>0.02</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Time at tunnel end</td>
<td>0.04</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Body mass</td>
<td>0.67</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Rear rate</td>
<td>1.81</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Seed mass</td>
<td>1.98</td>
<td>0.01</td>
</tr>
<tr>
<td><em>Fagus grandifolia</em></td>
<td>Time in center (-0.89 ± 0.33 SE)</td>
<td>0</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Models for *P. strobus* include 349 observations from 26 individuals, models for *Q. rubra* include 48 observations from 19 individuals, and models for *F. grandifolia* include 135 observations from 21 individuals. Effect sizes and standard errors shown in parentheses for top models.

\(^1\) Delta Akaike information criterion corrected for small sample sizes

\(^2\) Conditional (theoretical) coefficient of determination calculated using the *r.squaredGLMM* command in the MuMIn package in R

\(^3\) Body condition index calculated using the scaled-mass index (Peig and Green 2009)
Table C.6. Pairwise correlations between all behavioral variables used in model selection and an individual’s mean body condition index (averaged scaled-mass index), and mean body mass. Behavioral variables shown are the mean BLUP (best linear unbiased predictor) averaged over 1000 simulations.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - Mean speed</td>
<td>1</td>
<td>0.64</td>
<td>-0.62</td>
<td>0.18</td>
<td>-0.24</td>
<td>-0.33</td>
<td>-0.13</td>
<td>-0.06</td>
<td>-0.10</td>
</tr>
<tr>
<td>2 - Rear rate</td>
<td>0.64</td>
<td>1</td>
<td>-0.54</td>
<td>0.09</td>
<td>-0.18</td>
<td>-0.27</td>
<td>-0.12</td>
<td>-0.08</td>
<td>-0.13</td>
</tr>
<tr>
<td>3 - Prop. time grooming</td>
<td>-0.62</td>
<td>-0.54</td>
<td>1</td>
<td>-0.16</td>
<td>-0.05</td>
<td>0.11</td>
<td>-0.08</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>4 - Prop. time in center</td>
<td>0.18</td>
<td>0.09</td>
<td>-0.16</td>
<td>1</td>
<td>-0.04</td>
<td>-0.01</td>
<td>0</td>
<td>0.02</td>
<td>0.09</td>
</tr>
<tr>
<td>5 - Handling</td>
<td>-0.24</td>
<td>-0.18</td>
<td>-0.05</td>
<td>-0.04</td>
<td>1</td>
<td>0.31</td>
<td>0.22</td>
<td>0</td>
<td>0.05</td>
</tr>
<tr>
<td>6 - Latency to emerge</td>
<td>-0.33</td>
<td>-0.27</td>
<td>0.11</td>
<td>-0.01</td>
<td>0.31</td>
<td>1</td>
<td>0.41</td>
<td>0.02</td>
<td>0.09</td>
</tr>
<tr>
<td>7 - Time at tunnel end</td>
<td>-0.13</td>
<td>-0.12</td>
<td>-0.08</td>
<td>0</td>
<td>0.22</td>
<td>0.41</td>
<td>1</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>8 - Mean body condition</td>
<td>-0.06</td>
<td>-0.08</td>
<td>0.07</td>
<td>0.02</td>
<td>0</td>
<td>0.02</td>
<td>0.01</td>
<td>1</td>
<td>0.41</td>
</tr>
<tr>
<td>9 - Mean body mass</td>
<td>-0.10</td>
<td>-0.13</td>
<td>0.10</td>
<td>0.09</td>
<td>0.05</td>
<td>0.09</td>
<td>0.01</td>
<td>0.41</td>
<td>1</td>
</tr>
</tbody>
</table>
APPENDIX D: CHAPTER 4 SUPPLEMENTARY MATERIALS

Supplementary methods

Behavioral assays

1. Emergence test (deer mice, southern red-backed voles, and northern short-tailed shrews)

Animals were transferred directly from the trap of capture into a clean, empty Longworth trap. This trap was then placed into a box sized 46 x 46 x 50 cm. To create a more natural environment, the inside of the box was painted brown with a small amount of debris (dead leaves and pine needles) placed on the floor. The box was placed underneath a tarp to control for light levels and canopy cover. A digital camera (Nikon CoolPix S3700) was mounted facing the opening of the Longworth trap, and the observer locked the trap door open before leaving the test area. A clear plexiglass lid was placed over the box to prevent escapes. After three minutes, the observer returned and ended the test. Individuals were caught in a 4-liter plastic bag.

2. Open-field test (all species)

Animals were released immediately from the plastic bag into the center of the open field arena (red squirrels were released directly from the trap of capture into the open-field arena). A digital camera (Nikon CoolPix S3700) was mounted above the arena to record the test. The observer left the area, and a five-minute open-field test was performed in an arena of dimensions 46 x 46 x 50 cm, placed on a level platform with light levels controlled using a tarp overhead. After five minutes, the observer returned to the test area and stopped the video recording. The animal was caught in a 4-liter plastic bag (red squirrels were caught instead in a mesh handling bag).

3. Handling test (all species)

The observer immediately suspended the bag into the open field arena to control the visual surroundings. The observer measured the proportion of time that the individual spent immobile for 1 minute (referred to as handling time hereafter). Traps used for emergence tests and the open-field test box were cleaned thoroughly with 70% isopropyl alcohol and wiped with a dry cloth in-between all tests.

Quantifying behavior

To quantify behavior from videotaped emergence and open-field tests, recordings were played back in the laboratory. For emergence tests, an observer recorded whether the animal emerged (defined as all four feet having left the Longworth trap), the latency to emerge, and the total time spent at the end of the Longworth tunnel before emerging. When an individual did not emerge from the test after the three-minute cutoff, the latency to emerge was set to 1.25x the maximum test length. Open-field tests were analyzed using the behavioral tracking software ANY-maze © (version 5.1; Stoelting CO, USA). The following behavioral variables were obtained from the behavioral assays: handling time (the number of seconds immobile in a handling test; note that the term handling time should not be confused with a term sometimes used in ecological literature to indicate the time spent handling, processing, and consuming food items), latency to
emerge and time at tunnel end (from the emergence test), mean speed, rear rate, jump rate, proportion of time grooming, and proportion of time in the center (from the open-field test). See Table D.1. (modified from [Brehm, Mortelliti, Maynard, & Zydlewski, 2019]) for a complete list of the behaviors used, their description and interpretation, and supporting sources.

**Measuring microhabitat**

In July 2017 and July 2021, we recorded detailed microhabitat measurements within a 5m radius of each trap site (100 sites per trapping grid, 600 sites total). Microhabitats include structural components of the environment such as fallen logs and woody vegetation cover. Variables used in this study represent the mean measurement between the two sampling years. For detailed information about field methods used and variables measured, see (Brehm & Mortelliti, 2021).

**Repeatability analyses**

We calculated the adjusted repeatability (Wilson 2018) and associated 95% confidence intervals for seven key behaviors performed in the standardized assays using the ‘rptR’ package in R (Stoffel et al. 2017). For this analysis we used only individuals who had two or more repeated observations. Repeatability can be defined as the proportion of the total phenotypic variation that can be attributed to individual differences (Dingemanse et al. 2009; Nakagawa and Schielzeth 2010; Dingemanse and Dochtermann 2013). Practically, this means that repeatability equals the between-individual variance divided by the total phenotypic variance (between-individual variance + within-individual variance). The random intercept of individual ID was included in the models and fixed effects included sex, body condition (calculated using the scaled-mass index [Peig & Green, 2009]), silvicultural treatment, and trapping session. In all models, we used 1000 parametric bootstraps and 100 permutations. We assessed normality by visually inspecting Q-Q plots and by plotting the fitted values against the residual values and used Box-Cox transformations on the response variable to approach normality when necessary (Box & Cox, 1964). We considered any behavioral trait with a 95% CI for repeatability that excluded zero to be a personality trait (Nakagawa and Schielzeth 2010), but we emphasize that the classification as significantly repeatable does not say anything about the strength of repeatability.

After we confirmed that our behavioral variables were significantly repeatable, we took steps to account for the within-individual variability that occurs between repeated behavioral observations (consistent with methods used by [Dingemanse, Moiron, Araya-Ajoy, Mouchet, & Abbey-Lee, 2020; Gharit, Bergeron, Garant, & Réale, 2020]). We calculated an individual’s best linear unbiased predictor (BLUP) for each of the behavioral variables after controlling for the variation due to sex, body condition, forestry treatment, and trapping session. We then calculated an individual’s average BLUPs for the repeatable behavioral variables (estimated over 1000 simulations using the sim function in package arm [Gelman & Su, 2018]). We then used the mean BLUPs as the personality covariates in our further analyses (survival analyses and local survival analyses).
Model selection for detection, recapture, and emigration parameters

Candidate models for detection included time (session, season, and trap night), and conspecific density (a count of the total number of tagged conspecifics in the trapping grid and year of interest was used as a proxy for density) and personality traits (Royle and Nichols 2003; Stuber et al. 2013; Niemelä et al. 2015; Walsh et al. 2018). We tested for a behavioral effect ($p \neq c$) as well as equal capture and recapture probabilities ($p = c$). We examined real estimates from supported models and when time could be simplified to save parameters (i.e., grouping years together or allowing one year to differ from the rest) we did so (Dingemanse, Both, Drent, & Tinbergen, 2004). Using the best supported model for detection, we then modeled recapture probability by season, trap night, and conspecific density. For deer mice, red-backed voles, and short-tailed shrews, we fixed emigration parameters to specify a no-movement model ($\gamma' = \gamma'' = 0$), as we trapped during summer and fall months when individuals were unlikely to be temporarily unavailable for capture (i.e., below ground in subterranean burrows or in torpor) (Kendall et al. 1997). Further evidence that temporary emigration should be set to 0 for these species was a lack of fit for both random and markovian movement models (Kendall et al. 1997). For red squirrels, we assessed whether emigration should be modeled as random, markovian, or fixed at zero as it is likely that individuals occasionally became temporarily available for capture (average home range of this species is $\sim0.34$ha but can exceed the size of our trapping grid in some cases [Lamontagne et al., 2013]). We modeled movement by sex, body mass, forest treatment, and personality traits (Table D.1.).

Supplementary results

Repeatability

We examined 819 behavioral observations from standardized behavioral assays of 301 individual deer mice and 879 observations from 344 individual voles with two or more observations and found all behavioral variables to be significantly repeatable (Table D.3). Mean repeatability was 0.330 for deer mice (range: 0.191 to 0.447) and 0.222 for voles (range: 0.139 to 0.302) falling in line with similar field studies on deer mice (Underhill et al. 2021) and near the average previously reported for a variety of field and laboratory studies (Bell et al. 2009). We found four repeatable behavioral traits for northern short-tailed shrews after examining 204 observations from 79 individuals with two or more observations. Mean repeatability was 0.375 (range: 0.247 to 0.470). We found five repeatable behavioral traits for North American red squirrels after examining 303 observations from 109 individuals with two or more observations. Mean repeatability was 0.311 (range: 0.222 to 0.439).

Top model structures for detection, recapture, and emigration

For *P. maniculatus*, *M. gapperi*, and *B. brevicauda*, the temporary emigration probabilities ($\gamma''$ and $\gamma'$) were modeled as equal and fixed at 0 (no movement models), and for *T. hudsonicus* a markovian movement model was specified ($\gamma''$ was higher in males and $\gamma'$ was constant [$\gamma'' = \text{sex}$; $\gamma' = \sim1$], Figure D.4).

For *P. maniculatus*, variables included in the top-ranking model for detection probability are: year (years 2017 and 2019 set as equal and different from the other years), the proportion of time spent grooming in the open-field test (anxiety, stress de-arousal), and the proportion of time spent at the end of the emergence tunnel before emerging (timidness) ($p = \sim \text{Year} + \text{Prop. groom} + \text{Time tunnel}$; Figure D.5). Peaks in abundance were observed following years 2017 and 2019; suggesting potential mast in these years (Figure D.6). Variables included in the top-
ranking model for recapture probability are: *Peromyscus* density, handling time (docility) and year (years 2016 and 2018 set as equal and different from other years ($c = \sim \text{density} + \text{Year} + \text{Handling time}; \text{Figure D.7}$).

For *M. gapperi*, variables included in the top-ranking model for detection probability are: year (2019 and 2020 set as unequal and different from other years), and the latency to emerge from an emergence test (timidness) ($p = \sim \text{Year} + \text{Latency to emerge}; \text{Figure D.8}$). The top-ranking model for recapture probability is a null/constant structure ($c = \sim 1$).

For *B. brevicauda*, the top-ranking model for detection includes the variable handling time (docility) ($p = \sim \text{Handling time}; \text{Figure D.9}$). The top-ranking model for recapture probability includes year (years 2019 and 2020 set as equal and different from other years) ($c = \sim \text{Year}; \text{Figure D.10}$).

For *T. hudsonicus*, detection and recapture probabilities were modeled as equal and variables included in the top-ranking model for detection/recapture probability are: handling time (docility) and year (years 2017 and 2019 set as equal and different from other years) ($p = c = \text{Year} + \text{Handling time}; \text{Figure D.1}$).
Table D.1. Repeatable behavioral traits from three behavioral tests (handling bag, emergence, and open field) in deer mice (*Peromyscus maniculatus*), southern red-backed voles (*Myodes gapperi*), northern short-tailed shrews (*Blarina brevicauda*), and North American red squirrels (*Tamiasciurus hudsonicus*). This table provides the assay that each trait was obtained from, the description of the behavior, the personality trait that this behavior represents, and how values were interpreted. Included is a non-exhaustive list of citations supporting trait interpretation. This table was adapted from (Brehm et al. 2019a; Brehm and Mortelliti 2021). For information about which trait was repeatable in which study species, refer to Table D.3.

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Personality trait</th>
<th>Behavioral assay</th>
<th>Description</th>
<th>Interpretation</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean speed</td>
<td>Activity</td>
<td>Open field</td>
<td>Mean speed in the open field test in (meters/sec). Calculated by dividing the total distance traveled in the test by the test duration</td>
<td>Higher values indicate greater activity</td>
<td>(Carter et al. 2013; Gracceva et al. 2014)</td>
</tr>
<tr>
<td>Rear rate</td>
<td>Activity and exploration</td>
<td>Open field</td>
<td>Rate of rearing (rears/sec). Rearing is defined as forelegs leaving the arena floor</td>
<td>Higher values indicate greater activity and exploration</td>
<td>(Choleris et al. 2001; Prut and Belzung 2003; Martin and Réale 2008; Tanaka et al. 2012)</td>
</tr>
<tr>
<td>Jump rate</td>
<td>Activity</td>
<td>Open field</td>
<td>Rate of jumping (jumps/sec)</td>
<td>Higher values indicate greater activity</td>
<td>(Boon et al. 2007, 2008)</td>
</tr>
<tr>
<td>Proportion time</td>
<td>Anxiety/stress coping</td>
<td>Open field</td>
<td>Proportion of test duration spent grooming</td>
<td>Self-grooming is a form of de-arousal and the highest levels of grooming may indicate a lower anxiety level and better coping</td>
<td>(Choleris et al. 2001; Fernández-Teruel and Estanislau 2016; Kalueff et al. 2016)</td>
</tr>
<tr>
<td>grooming</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion time</td>
<td>Bold/timid</td>
<td>Open field</td>
<td>Proportion of test duration spent in the center portion of the arena</td>
<td>Higher values indicate more boldness/less timidness</td>
<td>(Barnett 1976; Treit et al. 1989; Ramos et al. 1997; Choleris et al. 2001; Eccard and Herde 2013; Gracceva et al. 2014)</td>
</tr>
<tr>
<td>Handling time</td>
<td>Docility</td>
<td>Handling bag</td>
<td>Total number of seconds of inactivity during a 1-minute handling bag test</td>
<td>Higher values indicate more docility/less defensive aggression</td>
<td>(Boon et al. 2007; Martin and Réale 2008; Montiglio et al. 2012; Taylor et al. 2014)</td>
</tr>
<tr>
<td>---------------</td>
<td>----------</td>
<td>--------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Latency to emerge</td>
<td>Bold/timid</td>
<td>Emergence</td>
<td>Latency (in seconds) to emerge from trap in the emergence test. An animal was considered to have emerged when all four feet left the trap tunnel</td>
<td>Higher values indicate more timidness/less boldness</td>
<td>(Brown and Braithwaite 2005; Carter et al. 2013; Gracceva et al. 2014)</td>
</tr>
<tr>
<td>Time at end of tunnel</td>
<td>Bold/timid</td>
<td>Emergence</td>
<td>Total number of seconds spent at the end of the tunnel before emerging</td>
<td>Higher values indicate more timidness/less boldness</td>
<td>(Brehm et al. 2019a; b)</td>
</tr>
</tbody>
</table>
Table D.2. Pairwise correlations between repeatable behavioral traits used in model selection. Behavioral variables shown are the mean BLUP (best linear unbiased predictor) averaged over 1000 simulations.

<table>
<thead>
<tr>
<th></th>
<th>Mean speed</th>
<th>Rear rate</th>
<th>Prop. time grooming</th>
<th>Prop. time in center</th>
<th>Handling</th>
<th>Latency to emerge</th>
<th>Time at tunnel end</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peromyscus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean speed</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rear rate</td>
<td>0.58</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prop. time grooming</td>
<td>-0.57</td>
<td>-0.41</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prop. time in center</td>
<td>0.17</td>
<td>0.11</td>
<td>-0.16</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handling</td>
<td>-0.27</td>
<td>-0.21</td>
<td>-0.08</td>
<td>-0.05</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latency to emerge</td>
<td>-0.33</td>
<td>-0.26</td>
<td>0.06</td>
<td>-0.06</td>
<td>0.41</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Time at tunnel end</td>
<td>-0.11</td>
<td>-0.10</td>
<td>-0.08</td>
<td>0.01</td>
<td>0.22</td>
<td>0.32</td>
<td>1.00</td>
</tr>
</tbody>
</table>

| **Myodes** |            |           |                     |                      |          |                   |                   |
| Mean speed | 1.00       |           |                     |                      |          |                   |                   |
| Prop. time grooming | -0.17  | 1.00      |                     |                      |          |                   |                   |
| Prop. time in center | 0.08    | -0.11     | 1.00                |                      |          |                   |                   |
| Handling  | -0.29      | 0.04      | -0.08               | 1.00                 |          |                   |                   |
| Latency to emerge | -0.31  | -0.11     | -0.04               | 0.21                 | 1.00     |                   |                   |
| Time at tunnel end | -0.10   | -0.10     | 0.00                | 0.07                 | 0.26     | 0.26              | 1.00              |

<p>| <strong>Blarina</strong> |            |           |                     |                      |          |                   |                   |
| Mean speed | 1.00       |           |                     |                      |          |                   |                   |
| Rear rate  | 0.49       | 1.00      |                     |                      |          |                   |                   |
| Prop. time in center | 0.32    | 0.17      | 1.00                |                      |          |                   |                   |
| Handling  | -0.45      | -0.36     | -0.30               | 1.00                 |          |                   |                   |
| Latency to emerge | -0.40  | -0.26     | -0.34               | 0.46                 | 1.00     |                   |                   |
| Time at tunnel end | -0.21   | -0.18     | -0.09               | 0.22                 | 0.26     | 1.00              |                   |</p>
<table>
<thead>
<tr>
<th>Tamiasciurus</th>
<th>Mean speed</th>
<th>Rear rate</th>
<th>Prop. time grooming</th>
<th>Handling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean speed</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rear rate</td>
<td>0.68</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prop. time</td>
<td>0.01</td>
<td>0.10</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>grooming</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handling</td>
<td>-0.18</td>
<td>-0.25</td>
<td>-0.09</td>
<td>1.00</td>
</tr>
</tbody>
</table>
**Table D.3.** Repeatability estimates for target behaviors measured in three behavioral tests (handling bag, emergence, and open field) in deer mice (*Peromyscus maniculatus*), southern red-backed voles (*Myodes gapperi*), northern short-tailed shrews (*Blarina brevicauda*), and North American red squirrels (*Tamiasciurus hudsonicus*).

<table>
<thead>
<tr>
<th>Behavioral Variable</th>
<th>Mean</th>
<th>Range</th>
<th>RPT (95% CI)</th>
<th>Observations</th>
<th>Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P. maniculatus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handling time (sec)</td>
<td>12.82</td>
<td>(0, 60)</td>
<td>0.342 (0.268, 0.426)</td>
<td>749</td>
<td>268</td>
</tr>
<tr>
<td>Latency to emerge (sec)</td>
<td>47.44</td>
<td>(0, 225)</td>
<td>0.300 (0.221, 0.386)</td>
<td>689</td>
<td>250</td>
</tr>
<tr>
<td>Time at tunnel end (sec)</td>
<td>4.67</td>
<td>(0, 180)</td>
<td>0.191 (0.104, 0.295)</td>
<td>617</td>
<td>229</td>
</tr>
<tr>
<td>Mean speed (m/sec)</td>
<td>0.09</td>
<td>(0, 0.29)</td>
<td>0.447 (0.371, 0.515)</td>
<td>815</td>
<td>300</td>
</tr>
<tr>
<td>Prop time groom¹</td>
<td>0.20</td>
<td>(0, 0.97)</td>
<td>0.389 (0.312, 0.467)</td>
<td>818</td>
<td>301</td>
</tr>
<tr>
<td>Rear rate (rears/sec)</td>
<td>0.19</td>
<td>(0, 0.65)</td>
<td>0.361 (0.285, 0.445)</td>
<td>819</td>
<td>301</td>
</tr>
<tr>
<td>Prop time in center²</td>
<td>0.02</td>
<td>(0, 0.53)</td>
<td>0.280 (0.199, 0.363)</td>
<td>816</td>
<td>301</td>
</tr>
<tr>
<td><strong>M. gapperi</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handling time (sec)</td>
<td>51.01</td>
<td>(0, 60)</td>
<td>0.187 (0.115, 0.268)</td>
<td>833</td>
<td>321</td>
</tr>
<tr>
<td>Latency to emerge (sec)</td>
<td>73.45</td>
<td>(0, 225)</td>
<td>0.222 (0.135, 0.313)</td>
<td>755</td>
<td>293</td>
</tr>
<tr>
<td>Time at tunnel end (sec)</td>
<td>5.86</td>
<td>(0, 180)</td>
<td>0.235 (0.137, 0.338)</td>
<td>562</td>
<td>229</td>
</tr>
<tr>
<td>Mean speed (m/sec)</td>
<td>0.04</td>
<td>(0, 0.20)</td>
<td>0.302 (0.228, 0.388)</td>
<td>875</td>
<td>342</td>
</tr>
<tr>
<td>Prop time groom¹</td>
<td>0.13</td>
<td>(0, 0.90)</td>
<td>0.139 (0.065, 0.220)</td>
<td>871</td>
<td>342</td>
</tr>
<tr>
<td>Rear rate (rears/sec)</td>
<td>0.10</td>
<td>(0, 0.56)</td>
<td>0.249 (0.172, 0.328)</td>
<td>879</td>
<td>344</td>
</tr>
<tr>
<td>Prop time in center²</td>
<td>0.03</td>
<td>(0, 0.98)</td>
<td>0.220 (0.142, 0.301)</td>
<td>875</td>
<td>342</td>
</tr>
<tr>
<td><strong>B. brevicauda</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handling time (sec)</td>
<td>31.16</td>
<td>(0, 60)</td>
<td>0.470 (0.314, 0.617)</td>
<td>201</td>
<td>74</td>
</tr>
<tr>
<td>Latency to emerge (sec)</td>
<td>96.46</td>
<td>(0, 225)</td>
<td>0.083 (0, 0.279)</td>
<td>157</td>
<td>59</td>
</tr>
<tr>
<td>Time at tunnel end (sec)</td>
<td>13.32</td>
<td>(0, 80)</td>
<td>0.117 (0, 0.373)</td>
<td>114</td>
<td>46</td>
</tr>
<tr>
<td>Mean speed (m/sec)</td>
<td>0.09</td>
<td>(0, 0.21)</td>
<td>0.441 (0.297, 0.589)</td>
<td>204</td>
<td>79</td>
</tr>
<tr>
<td>Rear rate (rears/sec)</td>
<td>0.12</td>
<td>(0, 0.57)</td>
<td>0.247 (0.082, 0.426)</td>
<td>204</td>
<td>79</td>
</tr>
<tr>
<td>Prop time in center²</td>
<td>0.03</td>
<td>(0, 0.19)</td>
<td>0.340 (0.188, 0.496)</td>
<td>204</td>
<td>79</td>
</tr>
<tr>
<td><strong>T. hudsonicus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handling time (sec)</td>
<td>30.53</td>
<td>(0, 60)</td>
<td>0.439 (0.311, 0.578)</td>
<td>249</td>
<td>91</td>
</tr>
<tr>
<td>Mean speed (m/sec)</td>
<td>0.03</td>
<td>(0, 0.40)</td>
<td>0.309 (0.191, 0.447)</td>
<td>299</td>
<td>108</td>
</tr>
<tr>
<td>Prop time groom¹</td>
<td>0.02</td>
<td>(0, 0.59)</td>
<td>0.222 (0.098, 0.359)</td>
<td>291</td>
<td>107</td>
</tr>
<tr>
<td>Jump rate (jumps/sec)</td>
<td>0.01</td>
<td>(0, 0.22)</td>
<td>0.281 (0.149, 0.413)</td>
<td>303</td>
<td>109</td>
</tr>
<tr>
<td>Rear rate (rears/sec)</td>
<td>0.02</td>
<td>(0, 0.25)</td>
<td>0.302 (0.173, 0.444)</td>
<td>296</td>
<td>107</td>
</tr>
</tbody>
</table>

¹ Proportion of time spent grooming in the open-field test
² Proportion of time spent in center portion of open-field arena
Table D.4. Best supported models from survival analyses using Robust Design models with the Huggin’s estimator for the deer mouse (*Peromyscus maniculatus*), southern red-backed vole (*Myodes gapperi*), northern short-tailed shrew (*Blarina brevicauda*), and North American red squirrel (*Tamiasciurus hudsonicus*). Samples included 827, 870, 279, and 249 individual capture histories, respectively. K = number of parameters, AICc = Akaike’s Information Criterion corrected for small sample sizes, Dev = model deviance, and $\hat{c}$ = c-hat (chi-squared/degrees of freedom). Models within 2.0 ΔAICc¹ of the top model are shown.

<table>
<thead>
<tr>
<th>Species</th>
<th>Model</th>
<th>K</th>
<th>AICc</th>
<th>ΔAICc¹</th>
<th>Dev</th>
<th>$\hat{c}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. maniculatus</em></td>
<td>body mass + USW² + Handling + pmmgdensity³⁴Prop. time grooming</td>
<td>5</td>
<td>5048.9</td>
<td>0</td>
<td>5018.7</td>
<td>1.4</td>
</tr>
<tr>
<td><em>M. gapperi</em></td>
<td>body mass + sex + year + mgdensity⁴*Prop. time grooming</td>
<td>5</td>
<td>5174.8</td>
<td>0</td>
<td>5144.6</td>
<td>1.4</td>
</tr>
<tr>
<td><em>B. brevicauda</em></td>
<td>body mass + year + Prop. time in center</td>
<td>9</td>
<td>2413.0</td>
<td>0</td>
<td>2394.8</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>body mass + year</td>
<td>8</td>
<td>2414.2</td>
<td>1.21</td>
<td>2398</td>
<td>1</td>
</tr>
<tr>
<td><em>T. hudsonicus</em></td>
<td>body mass + thdensity⁶</td>
<td>9</td>
<td>2465.9</td>
<td>0</td>
<td>2447.5</td>
<td>1</td>
</tr>
</tbody>
</table>

¹Where c-hat differs from 1, values represent Quasi-AICc (QAICc), ΔQAICc, and QDev. 
ΔAICc is the difference between model AICc (Akaike’s information criterion corrected for small sample sizes) and AICc value of the best model
²A categorical variable allowing survival to differ in the uniform shelterwood stands (USW) when compared the reference (REF) and irregular shelterwood stands (ISW)
³A proxy for small mammal density (equal to the total number of combined *Peromyscus* and *Myodes* tagged in each trapping grid per year)
⁴A proxy for vole density (equal to the total number of *Myodes* tagged in each trapping grid per year)
⁵Model averaging used to obtain predicted values in figures
⁶A proxy for squirrel density (equal to the total number of *Tamiasciurus* tagged in each trapping grid per year)
Figure D.1. Map of the study areas at the Penobscot Experimental Forest (PEF, 44°51' N, 68°37' W) in central Maine, USA.
Figure D.2. Positive associations between body mass and apparent survival obtained from the best supported Robust Design Models in (A) deer mice (*Peromyscus maniculatus*; $\beta = 0.30 \pm 0.07$ SE), (B) southern red-backed voles (*Myodes gapperi*; $\beta = 0.49 \pm 0.08$ SE), (C) northern short-tailed shrews (*Blarina brevicauda*; $\beta = 0.45 \pm 0.11$ SE and $\beta = 0.41 \pm 0.10$ SE), and (D) North American red squirrels (*Tamiasciurus hudsonicus*; $\beta = 0.37 \pm 0.14$ SE). Predictions for short-tailed shrews were obtained by model averaging the two top models. Shaded areas represent 95% CI.
Figure D.3. Model structure for detection ($p$) and recapture ($c$) probabilities of North American red squirrels (*Tamiasciurus hudsonicus*) obtained from the best supported Robust Design Model. Detection and recapture probabilities were equal and higher in years 2016, 2018, and 2020 and lower in years 2017 and 2019 (real estimates: $p_{161820} = 0.28 \pm 0.02$ SE; $p_{1719} = 0.10 \pm 0.02$ SE). Detection and recapture probabilities were positively associated with docility (seconds spent immobile in a handling bag test) ($\beta = 0.15 \pm 0.07$ SE). Shaded areas represent 95% CI.
Figure D.4. The probability of temporary emigration ($\gamma''$) varied by sex in North American red squirrels (*Tamiasciurus hudsonicus*). Results were obtained from the best supported Robust Design Model, specifying a markovian movement structure. Males had a greater probability of being temporarily absent than females (real estimates: $\gamma''_{\text{Female}} = 0.04 \pm 0.03 \text{ SE}; \gamma''_{\text{Male}} = 0.23 \pm 0.06 \text{ SE}$). Error bars represent 95% CI.
Figure D.5. Parameters affecting detection ($p$) probabilities of deer mice (*Peromyscus maniculatus*) obtained from the best supported Robust Design Model. The probability of detection was higher in years 2016, 2018, and 2020 and lower in years 2017 and 2019 (real estimates: $p_{161820} = 0.37 \pm 0.02$ SE; $p_{1719} = 0.25 \pm 0.03$ SE). Detection probability was positively associated with an indicator of stress-coping or de-arousal (the proportion of time spent grooming in an open-field test) ($\beta = 0.12 \pm 0.05$ SE) and negatively associated with timidity (time at the end of the emergence tunnel before emerging) ($\beta = -0.25 \pm 0.10$ SE). Shaded areas represent 95% CI.
Figure D.6. Abundance of small mammals (number of different individual’s captured) over time in the Penobscot Experimental Forest, ME USA. Peaks in abundance can be seen in years 2018 and 2020.
Figure D.7. Parameters affecting recapture ($c$) probabilities of deer mice (*Peromyscus maniculatus*) obtained from the best supported Robust Design Model. The probability of recapture was significantly higher in years 2016 and 2018 of the study than in years 2017, 2019, and 2020 (real estimates: $c_{1618} = 0.64 \pm 0.02$ SE; $c_{171920} = 0.52 \pm 0.02$ SE). Recapture probability was positively associated with *Peromyscus* density ($\beta = 0.24 \pm 0.05$ SE) and negatively associated with docility (seconds spent immobile in a handling bag test) ($\beta = -0.20 \pm 0.05$ SE). Shaded areas represent 95% CIs.
Figure D.8. Parameters affecting in detection ($\rho$) probabilities for southern red backed voles (*Myodes gapperi*) obtained from the best supported Robust Design Model. The probability of detection was highest in year 2020, lowest in year 2019, and relatively constant across the other three years of the study (2016-2018) (real estimates: $\rho_{1618} = 0.38 \pm 0.02$ SE; $\rho_{19} = 0.19 \pm 0.05$ SE; $\rho_{20} = 0.45 \pm 0.02$ SE). Detection probability was negatively associated with timidness (the latency to emerge from an emergence test) ($\beta = -0.39 \pm 0.06$ SE). Shaded areas represent 95% CI.
Figure D.9. Parameter affecting detection ($p$) probabilities of northern short-tailed shrews (*Blarina brevicauda*) obtained from model averaging the two best supported Robust Design Models. The probability of detection decreased with increasing docility (seconds spent immobile in a handling bag test) ($\beta = -0.30 \pm 0.11$ SE). Shaded areas represent 95% CI.
Figure D.10. Parameter affecting recapture ($c$) probabilities of northern short-tailed shrews (*Blarina brevicauda*) obtained from model averaging the two best supported Robust Design Models. The probability of recapture was lower in 2019 and 2020 than in years 2016-2018 (real estimates: $c_{161718} = 0.53 \pm 0.02$ SE; $c_{1920} = 0.38 \pm 0.05$ SE). Error bars represent 95% CI.
BIOGRAPHY OF THE AUTHOR

Allison Brehm was born in Manchester, New Hampshire on March 17, 1991 to William and Sara Brehm. She was raised in Pembroke, New Hampshire with her two siblings and graduated from Pembroke Academy in 2009. She pursued her bachelor’s degree at the University of New Hampshire, and graduated in 2014 with a degree in Zoology and a Minor in Animal Behavior. She then spent a year living and working in the Kalahari Desert on the Kalahari Meerkat Project, which solidified her love of fieldwork and passion for behavioral ecology. She moved to Maine in June of 2016 to pursue her Master’s degree in Wildlife Ecology at the University of Maine under the mentorship of Dr. Alessio Mortelliti.

After receiving her degree in December of 2018, and with the aid of a Chase Distinguished Research Assistantship, she began pursuing a Ph.D. in Ecology and Environmental Sciences from the University of Maine under the continued mentorship of Dr. Mortelliti. During her graduate studies Allison was honored to receive the Janet Waldron Doctoral Research Fellowship as well as the 2021 College of Natural Sciences, Forestry and Agriculture’s Fred Griffee Memorial Award. Together Allison and Alessio have published 11 peer-reviewed manuscripts and have a few more exciting manuscripts in the works with collaborators in the University of Maine’s Cooperative Forestry Research Unit, the School of Forest resources, and at the University of New Hampshire.

After graduating with her doctoral degree, Allison will join the lab of Dr. John Orrock at the University of Wisconsin - Madison as a Postdoctoral Research Associate. There she is excited to continue pursuing her research interests working with small mammals, plant-granivore interactions, and applied behavioral ecology.

Allison is a candidate for the Doctor of Philosophy degree in Ecology and Environmental Sciences from the University of Maine in August 2022.