AN INTRASPECIFIC INVESTIGATION OF SPATIAL COGNITION, ELEVATION AND SEX IN BIRDS

ETHAN HERMER

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Doctorate in Philosophy in Biology

Department of Biology
Faculty of Science
University of Ottawa

Abstract

What drives variance in wild animals' ability to acquire, process, store, and utilize information is the key question of cognitive ecology. The intraspecific approach is used to examine this question as it can measure the effects of the environment and sex, while directly measuring potential correlations or trade-offs in cognitive abilities. This thesis seeks to study how cognition varies across environments, among-individuals, and between sexes. The 'harsh environment' hypothesis states that food variability and scarcity could drive changes in cognitive abilities that that may aid in foraging. However, this hypothesis is primarily studied in food hoarders, or animals that cache food and retrieve it over winter. In *Chapter 1*, I compare the serial reversal learning performance of high and low elevation non-food hoarding great tits (*Parus major*). I report that low elevation great tits performed more accurately on the task than high elevation birds, showing that elevation drove differences in cognitive ability, in a similar way to scatter hoarders but opposite to what was predicted by the 'harsh environment' hypothesis. In *Chapter 2*, I compare the spatial memory accuracy and proactive interference performance of high and low elevation great tits and find no relationship between these two measures and elevation, but I did find the first significant among-individual trade-off between spatial memory accuracy and proactive interference performance. This trade-off suggests that if the environment does increase or decrease spatial memory, we should see a concurrent evolutionary change in proactive interference. My elevation results differ from those observed in food hoarders which exhibited a positive correlation between elevation and spatial memory. In *Chapter 3*, I performed the second comparison of female and male differences in cognitive variance in a wild animal, by comparing variance in great tits for serial reversal learning and spatial memory. I find that there is no variance difference for serial reversal learning, but greater

female variability in spatial memory. These results differ from those found in food hoarders, which found greater male variability in reversal learning, but not spatial memory. This also is the first time greater female variance has been found for a cognitive ability and provides preliminary evidence that spatial memory accuracy may be coded for on the sex chromosome pair in great tits. In *Chapter 4*, I tested a method to measure motor training and other cognitive abilities and found that black-legged kittiwakes (*Rissa tridactyla*) could complete the motor training task. With some modifications, this task could measure associative learning performance and be the first to investigate its fitness consequences in a seabird. This thesis shows that the relationships between cognition, elevation, and sex is species specific or at least differs between hoarders and non-hoarders. and future studies should continue to utilize variance partitioning approaches to investigate the inter-relatedness of these relationships within species.

Résumé

Ce qui détermine la variation dans la capacité des animaux sauvages à acquérir, traiter, stocker et utiliser l'information est une question clé en l'écologie cognitive. Une approche intraspécifique est utilisée pour examiner cette question puisqu'elle nous aide à mesurer les effets de l'environnement et du sexe, tout en mesurant directement les corrélations ou les compromis potentiels dans les capacités cognitives. Cette thèse étudie les variations cognitives selon l'environnement, entre les individus et entre les sexes. L'hypothèse « harsh environment » stipule que la variabilité et la rareté de nourriture pourraient entraîner des changements dans les capacités cognitives aidant à la recherche de nourriture. Cependant, cette hypothèse est principalement étudiée chez les espèces qui cachent de la nourriture pour la récupérer plus tard en hiver. Dans *le chapitre 1*, je compare les performances d'apprentissage inversé en série chez une espèce qui ne cache pas de nourriture, la mésange charbonnière (*Parus major*), provenant de sites à haute et basse altitude. Je rapporte que les mésanges charbonnières de basse altitude ont exécuté la tâche avec plus de précision que les oiseaux de haute altitude, montrant que l'altitude a entraîné des différences dans les capacités cognitives. Ce résultat est à l'opposé de ce qui était prédit par l'hypothèse « harsh environment ». Dans *le chapitre 2*, je compare la précision de la mémoire spatiale et les performances d'interférence proactive des mésanges charbonnières provenant de multiples sites à haute et basse altitude. Je ne rapporte aucune relation entre ces deux mesures et l'altitude, mais je démontre pour la première un compromis entre la précision de la mémoire spatiale et la performance de la proactivité d'interférence au niveau inter-individuel. Ce compromis suggère que si l'environnement influence la mémoire spatiale, nous devrions voir un changement évolutif correspondant dans l'interférence proactive. Mes résultats en lien avec l'altitude diffèrent des résultats trouvés pour les espèces qui cachent de la nourriture, chez qui les

études antérieures ont rapporté une corrélation positive entre l'altitude et la mémoire spatiale. Dans *le chapitre 3*, j'ai effectué la deuxième comparaison des différences de variance entre mâles et femelles dans la variance cognitive chez un animal sauvage, en comparant la variance pour l'apprentissage inversé en série et la mémoire spatiale chez les mésanges charbonnières. Je note qu'il n'y a pas de différence de variance pour l'apprentissage inversé en série, mais je remarque que les femelles démontrent une plus grande variabilité que les mâles dans leur mémoire spatiale. Ces résultats diffèrent de ceux trouvés chez les espèces qui cachent de la nourriture. Ils ont trouvé une plus grande variabilité masculine dans l'apprentissage par inversion, mais pas dans la mémoire spatiale. C'est également la première fois que nous constatons une plus grande variabilité par rapport aux capacités cognitives chez les femelles. Ceci suggère que la précision de la mémoire spatiale pourrait être codée au moins en partie sur la paire de chromosomes sexuels chez les mésanges charbonnières. Au chapitre 4, j'ai testé une méthode pour mesurer les performances d'apprentissage moteur et d'autres traits cognitifs chez la mouette tridactyle (*Rissa tridactyla*). J'ai trouvé que les mouettes pouvaient accomplir la tâche d'apprentissage moteur. Avec quelques modifications, cette tâche pourrait mesurer les performances d'apprentissage par association et devenir la première tâche qui permet d'étudier les conséquences de l'apprentissage par association sur le succès reproducteur chez un oiseau marin. Cette thèse montre que les relations entre la cognition, l'altitude, et le sexe sont spécifiques à l'espèce, ou du moins diffère chez les espèces qui cachent ou non de la nourriture. Dorénavant, nos études devraient continuer à utiliser des approches pour partitionner la variance statistique afin d'étudier l'interdépendance des variables au sein des espèces.

Acknowledgements

I dedicate this thesis to Ila, Clarence, and Doug Hermer.

I would like to thank the JMF lab, the 4th Floor, ABDG, the BGSA, the EDI book club, and the rest of my UOttawa friends for providing me with knowledge, skills, and a healthy dose of humour.

I specifically would like to thank my collaborators Alexis Chaine, Maxime Cauchoix, Ben Murphy, Shannon Whelan, and Kyle Elliott for allowing me to work with you in your study systems and learn so much from you all. I would like to thank Alice Thiney and Cecile Mahnich for aiding in capture and data collection. I want to thank the volunteers Edoxie Allier, Celia Sineau, Amandine Basset, Gwennan Giraud, Justine Le Vaillant, David Vallecillo, Armand Ziller, Maelle Freteault, Josh Lynton-Jenkins, Alysha Riquier, and James Huynh. Without you, none of this would be possible. I would like to thank Julian Evans, Marie-B. Leduc, Ferran Sayol, Julien Martin, Mathieu Videlier and Cammy Beyts for your immeasurable advice on the statistical and technical aspects of this thesis. Thank you, Florence Jean for your artistic contributions to the second chapter.

I would like to thank my committee for the invaluable advice you provided that drastically improved this thesis and my own development.

I would like to thank my family for the support you gave me during my thesis and throughout life. I share the title of doctor with all of you.

I would like to thank my partner Chloé Montreuil-Spencer. Words cannot describe how important you have been in all of this.

When I had to choose an honours thesis supervisor, you were the best in the department. After searching for a master's thesis supervisor, you were the best in your field. So good in fact, I figured I'd stick around for a Ph.D. After all of this, I know I made the best choice. You once said that we become our supervisors, and I would be lucky if that is even only half true. Thank you for everything Julie.

Table of Contents

Statement of Contributions **xi**

Introduction **1**

References 8

Chapter 1: ***Elevation-related difference in serial reversal learning ability in a nonscatter hoarding passerine*** **18**

Introduction 19

Methods 23

Ethics 29

Results 30

Discussion 37

Supplementary 42

References 52

Chapter 2: ***Great tits who remember more accurately have difficulty forgetting, but variation is not driven by environmental harshness*** **60**

Introduction 61

Methods 68

Ethics 75

Results 76

Discussion 83

Supplementary 89

References 101

Chapter 3: ***Serial reversal learning, spatial memory, and the greater male variability phenomenon*** **110**

Introduction 111

Methods 117

Results 119

Discussion 124

References 129

Chapter 4: ***The relationship between foraging, cognition, and fitness in a vulnerable seabird*** **136**

Introduction 137

Methods 140

Ethics 146

Results 147

Discussion 148

References 151

Conclusion **155**

References 164

Appendix **169**

Appendix 1: *List of co-authored articles not utilized as chapters in the thesis.* 169

Statement of Contributions

Chapter 1 of this thesis is adapted from an article published in *Behavioral Ecology* (Hermer *et al*. 2018). The pronoun "we" is used within the manuscript to include additional authors, however, this chapter was produced by myself. I conceived and designed the study in collaboration with Alexis S. Chaine, Maxime Cauchoix, and Julie Morand-Ferron. I conducted behavioural observations, analyzed the data, prepared the figures, and wrote the chapter. Cecile Mahnich conducted additional behavioural observations and captured birds. Alice Thiney managed bird captures.

Chapter 2 of this thesis is adapted from an article published in *Scientific Reports* (Hermer *et al.* 2021). The pronoun "we" is used within the manuscript to include additional authors, however, this chapter was produced by myself. I conceived and designed the study in collaboration with Alexis S. Chaine, and Julie Morand-Ferron. I conducted behavioural observations, analyzed the data, prepared the figures, and wrote the chapter. Ben Murphy conducted additional behavioural observations and captured birds. Alysha Riquier and James Huynh collected video data. Florence Jean provided artwork for the figures. Chloé Montreuil-Spencer managed bird captures.

Chapter 3 of this thesis will be adapted for submission as a research article. This chapter was produced by myself. I designed the study, compiled the data, analyzed the data, and wrote the chapter in collaboration with Julie Morand-Ferron.

Chapter 4 of this thesis will be adapted for submission as a methods article. This chapter was produced by myself. I designed the study, compiled the data, analyzed the data, and wrote the chapter in collaboration with Julie Morand-Ferron. Kyle Elliott and Shannon Whelen helped design the experiment. Shannon Whelan collected acclimatization data.

Introduction

What drives differences in wild animals' ability to acquire, process, store and utilize information is the key question of cognitive ecology (Boogert, Madden, Morand-Ferron, & Thornton, 2018; Cauchoix & Chaine, 2016; Shettleworth, 1998). The intraspecific approach, wherein individuals from the same species are compared, has been used to examine this question as it can measure the associated fitness costs and benefits of cognitive variation (reviewed in: Boogert, Madden, Morand-Ferron, & Thornton, 2018). Increased learning and memory performance has been linked to fitness costs such as decreased survival (Plaçais & Preat, 2013), decreased larval competitive ability (Mery & Kawecki, 2003), decreased offspring production (Kotrschal *et al*., 2013), energetic stress (Jaumann, Scudelari, & Naug, 2013), decreased survival (Mery & Kawecki, 2005; Plaçais & Preat, 2013), and decreased egg laying (Mery & Kawecki, 2004; reviewed in: Burns, Foucaud, & Mery, 2011; but see: Foucaud *et al*., 2016; Kolss, Kraaijeveld, Mery, & Kawecki, 2006). Also, greater learning and memory performance have been linked to fitness benefits such as increased survival (Branch *et al*., 2019; Sonnenberg, Branch, Pitera, Bridge, & Pravosudov, 2019), foraging efficiency (Dukas & Bernays, 2000; Huebner, Fichtel, & Kappeler, 2018; Raine & Chittka, 2008) and reproductive success in different species (Ashton, Ridley, Edwards, & Thornton, 2018).

This cognitive variation can also be linked to environmental differences (Roth II, LaDage, & Pravosudov, 2010; Tebbich & Teschke, 2014; Tello-Ramos, Branch, Pitera, *et al.*, 2018), age (e.g., Bartus, Dean III, & Fleming, 1979), and sex (e.g., Brodin & Urhan, 2015; Guigueno, Macdougall-Shackleton, & Sherry, 2015; Guigueno, Snow, MacDougall-Shackleton, & Sherry, 2014; Lucon-Xiccato & Bisazza, 2017; Zidar *et al.*, 2018). Performance of various wild animals on cognitive tasks is repeatable (Cauchoix *et al.*, 2018) and certain cognitive

abilities are heritable in the lab (Mery & Kawecki, 2002; reviewed in: Croston, Branch, Kozlovsky, Dukas, & Pravosudov, 2015). The intraspecific approach allows for examining the costs, benefits, and heritability of cognition. However, our understanding of which selection pressures lead to differences in cognitive ability, and how these cognitive abilities correlate among-individuals has been studied only in few species and ecological contexts.

It is possible to test hypotheses on which pressures act on cognition by comparing the cognitive ability of individuals from different environments. The social intelligence hypothesis states that cognitive abilities may aid animals in navigating the pressures of interacting with other individuals. Interspecific studies have shown that species that live in groups generally perform better on cognitive tasks than solitary species (e.g., Bond, Kamil, & Balda, 2007), and species in more complex fission-fusion flocks perform better than species in stable flocks (e.g., Amici, Aureli, & Call 2008; reviewed in: Ashton, Thornton, and Ridley, 2018). Intraspecific studies have shown a similar trend, with Australian magpies (*Cracticus tibicen dorsalis*) from larger groups having greater general cognitive ability than those from smaller groups (Ashton *et al*., 2018), for example. However, whether these intraspecific results are due to development, selection, or non-random assortment is yet to be clarified (Lambert, Sewall, & Guillette, 2019). Another candidate selective pressure is environmental harshness, or environments with high food variability and low food availability (Pravosudov & Clayton, 2002; Roth II *et al.,* 2010; Tello-Ramos, Branch, Kozlovsky, Pitera, & Pravosudov, 2018). The 'harsh environment' hypothesis states that cognitive abilities, like learning or memory, may allow animals to cope with harsh environments by increasing their foraging efficiency or circumventing the low food supply entirely (Dukas, 2009; Milton, 1981; Roth II *et al.*, 2010; Sol, Lefebvre, & Rodríguez-Teijeiro, 2005). Both latitude and elevation have been used to investigate the 'harsh environment'

hypothesis in the context of foraging as increases in latitude and elevation can correlate to more snow cover, lower temperatures, and greater seasonality, which leads to greater food scarcity and variability, as well as higher metabolic costs during winter (Boyle, Sandercock, & Martin, 2016; Roth II & Pravosudov, 2009; but see: Körner, 2007). Along these two gradients, the 'harsh environment' hypothesis has been tested in winter resident, scatter hoarding birds (reviewed in: Roth II *et al.*, 2010; Sonnenberg *et al.*, 2019) that store food in multiple areas or caches and utilize spatial memory to remember and return to these food locations during the winter when food is depleted (Krebs, 1990; Sherry, 1998). Therefore, spatial memory is predicted to increase along a gradient of harshness in scatter hoarders (Pravosudov & Clayton, 2002). Empirical support for the 'harsh environment' hypothesis has been found in this context with spatial memory accuracy, and the size of the main brain structure responsible for spatial memory (i.e., the hippocampus) increasing along harshness gradients of elevation and latitude (latitude: Freas, Roth II, LaDage, & Pravosudov, 2013; Pravosudov & Roth II, 2013; Roth II, Chevalier, LaDage, & Pravosudov, 2013; elevation: Freas *et al.*, 2013; reviewed in: Pravosudov, Roth, LaDage, & Freas, 2015). However, very few studies have extended this hypothesis beyond scatter hoarders (e.g., Tebbich & Teschke, 2014).

To better understand the relationship between environmental harshness and cognition outside of the scatter hoarding paradigm, I undertook a series of studies with non-scatter hoarding great tits (*Parus major*). Great tits are small (13-21 g) facultative generalist, winter residents who live in temperate regions and feed on continuously depleting patches of seed and insects (Gosler, 1993; Pagani-Nunez, Barnett, Gu, & Goodale, 2016). These abundant birds' reversal learning and spatial memory performance can be measured using a range of cognitive assays (Amy, van Oers, & Naguib, 2012; Cauchoix, Hermer, Chaine, & Morand-Ferron, 2017;

Morand-Ferron, Hamblin, Cole, Aplin, & Quinn, 2015). In *Chapter 1*, I compared great tits sampled from high and low elevations with a reversal learning task that measures an animal's ability to flexibly switch their learned associations as the environment changes. I then built upon previous studies use of a single reversal (e.g., Tebbich & Teschke, 2014), and added a serial component which measures their ability to not just change behaviour once but track multiple changes in the rewarded option across time (Shettleworth, 2010; Hermer, Cauchoix, Chaine, & Morand-Ferron, 2018). In *Chapter 2*, I compared the performance of great tits sampled from low and high elevations on a spatial memory task, which measures their ability to remember and store the location of an object in space. I also tested their proactive interference performance, or how previous information interferes with their ability to remember new information and causes them to repeat previously correct, and currently incorrect actions (Hermer, Murphy, Chaine, & Ferron, 2021). These studies furthered our understanding of whether harsh environments are potential agents of selection acting on different aspects of cognition in non-scatter hoarding birds.

In addition to cognition-environment correlations, the intraspecific approach can be used to discern correlations between individual traits. Spatial memory and proactive interference, two cognitive measures I recorded in my study subjects, are not necessarily independent aspects of cognition (Croston *et al.*, 2017; Tello-Ramos, Branch, Kozlovsky, *et al.*, 2018). It has been hypothesized that proactive interference is a function of accurate memory making it difficult to forget past information when learning new information (Gonzalez, Behrend, & Bitterman, 1967; Tello-Ramos, Branch, Kozlovsky, *et al.*, 2018). For example, an animal may learn the location of one food patch when foraging. Accurate spatial memory may aid the animal in returning to this food patch until it is depleted. After the food patch is depleted, and the animal learns the location

of a similar, new food patch, the animal may continue to incorrectly return to the area where the first, depleted food patch was, as it is unable to forget the past food patch and learn the new food patch. Therefore, accurate spatial memory may trade-off with an increased cost of proactive interference. Until now, this trade-off has primarily been at the phenotypic level (e.g., Tello-Ramos *et al.* 2017). By not taking multiple measures and partitioning the variance to among and within-individual levels, what may be measured is the within-individual correlation as opposed to the among-individual (i.e., "the individual gambit", Niemelä & Dingemanse, 2018; e.g., Tello-Ramos, Branch, Pitera, *et al.*, 2018). Finding a positive among-individual correlation between spatial memory and proactive interference would indicate that individuals that have more accurate spatial memory, also express more proactive interference (Careau & Wilson, 2017). This would indicate that if elevation does act on spatial memory a concurrent change in proactive interference is expected. To test for this, I measured the among-individual correlation between spatial memory and proactive interference (*Ch. 2*: Hermer *et al*., 2021).

Additionally, the intraspecific approach can be utilized to measure sex differences in a trait. Studies into sex differences in cognitive ability in wild animals have primarily focused on differences in means between males and females. These differences have been attributed to differences in reproductive behaviour (e.g., Sherry & Guigueno, 2019), dominance differences between males and females (e.g., Brodin & Urhan, 2015), spatial demands (e.g., Carazo, Noble, Chandrasoma, & Whiting 2014; but see Lucon-Xiccato *et al.* 2017), and have been shown to be personality dependent (e.g., Titulaer, van Oers, & Naguib, 2012, but see Carazo, *et al.* 2014). Additionally, sex differences in cognitive ability variance have started to be investigated in wild animals (e.g., Branch *et al.*, 2020). Differences between males and females in variance in cognitive ability were originally studied in humans and termed the Greater Male Variability

Hypothesis (hereafter, the GMV phenomena; Hedges & Nowell, 1995; Johnson, Deary, & Carothers, 2008; Lehre, Lehre, Laake, & Danbolt, 2009; Zechner *et al.*, 2001). These differences in variance are hypothesized to be outcomes of sex differences in sexual selection (Pomiankowski & Møller, 1995; Rowe & Houle, 1996; Tomkins, Radwan, Kotiaho, & Tregenza, 2004) and the sex chromosomes (Reinhold & Engqvist, 2013; both hypothesis are reviewed in: Zajitschek *et al.*, 2020). To my knowledge, these hypotheses are poorly studied in wild animals, with only one taking place in wild animals (e.g., Branch *et al.*, 2020). To assess the generalizability of sex differences in cognitive variance, I investigate how serial reversal learning and spatial memory performance variance may co-vary with sex (*Chapter 3*).

Finally, the intraspecific approach can be used to answer questions of how cognition is related to behaviour and fitness. However, intraspecific studies that investigate the link between behaviour, cognition, and fitness are few (reviewed in: Madden, Langley, Whiteside, Beardsworth, & van Horik, 2018). To increase the breadth of species this question is studied in, I undertook a preliminary test of a cognitive task in black-legged kittiwakes (*Rissa tridactyla*), that may be able to test their cognitive ability while they nest. In concert with ongoing studies on both adult and chick fitness measures (e.g., Collins, Hatch, Elliott, & Jacobs, 2019), as well as adult foraging patterns (e.g., Osborne *et al*., 2020), I had planned to study the link between foraging behaviour, cognitive ability, and nestling success. However, COVID-19 stopped these plans and so I present a novel black-legged kittiwake cognitive protocol that will be used in the future to measure cognitive ability in these kittiwakes (*Chapter 4).*

Overall, this thesis provides answers to the questions of how the environment, and individual phenotypic values such as sex, contribute to variation in cognition in wild animals. I answer whether the harsh environment hypothesis extends beyond scatter hoarders into a

generalist species by measuring serial reversal learning, spatial memory, and proactive interference; whether there are costs to increased spatial memory; and whether sex contributes meaningfully to variance differences in cognitive ability, in an avian generalist species. Finally, I tested a novel cognitive task that will be used to answer the link between behaviour, cognition, and fitness in a long-lived seabird.

References

Amy, M., van Oers, K., & Naguib, M. (2012). Worms under cover: relationships between performance in learning tasks and personality in great tits (*Parus major*). *Animal Cognition*, *15*, 763–770.

Ashton B. J., Thornton A, Ridley A. R. (2018) An intraspecific appraisal of the social intelligence hypothesis. *Phil. Trans. R. Soc. B*. 373: 20170288.

Ashton, B. J., Ridley, A. R., Edwards, E. K., & Thornton, A. (2018). Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature*, *554*, 364–367.

Bartus, R. T., Dean III, R. L., & Fleming, D. L. (1979). Aging in the rhesus monkey: Effects on visual discrimination learning and reversal learning. *Journal of Gerontology*, *34*(2), 209–219.

Boogert, N. J., Madden, J. R., Morand-Ferron, J., & Thornton, A. (2018). Measuring and understanding individual differences in cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756).

Boyle, A. W., Sandercock, B. K., & Martin, K. (2016). Patterns and drivers of intraspecific variation in avian life history along elevational gradients: a meta-analysis. *Biological Reviews*, *91*, 469–482.

Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Sonnenberg, B. R., Benedict, L. M., & Pravosudov, V. V. (2019). Elevation-related differences in the age structure of breeding birds suggest stronger selection at harsher elevations. *Behavioral Ecology and Sociobiology*, *73*(10).

Branch, C. L., Sonnenberg, B. R., Pitera, A. M., Benedict, L. M., Kozlovsky, D. Y., Bridge, E. S., & Pravosudov, V. V. (2020). Testing the greater male variability phenomenon: Male mountain chickadees exhibit larger variation in reversal learning performance compared with females. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1931), 13–16.

Brodin, A., & Urhan, A. U. (2015). Sex differences in learning ability in a common songbird, the great tit-females are better observational learners than males. *Behavioral Ecology and Sociobiology*, *69*, 237–241.

Burns, J. G., Foucaud, J., & Mery, F. (2011). Costs of memory: Lessons from "mini" brains. *Proceedings of the Royal Society B: Biological Sciences*, *278* (1707), 923–929.

Carazo, P., Noble, D. W. A., Chandrasoma, D. & Whiting, M. J. (2014). Sex and boldness explain individual differences in spatial learning in a lizard. *Proc. Biol. Sci. 281*, 20133275.

Careau, V., & Wilson, R. S. (2017). Of uberfleas and krakens: Detecting trade-offs using mixed models. *Integrative and Comparative Biology*, *57*(2), 362–371.

Cauchoix, M., Chow, P. K. Y., van Horik, J. O., Atance, C. M., Barbeau, E. J., Barragan-Jason, G., Bize, P., Boussard, A., Buechel, S. D., Cabirol, A., Cauchard, L., Claidière, N., Dalesman, S., Devaud, J.M., Didic, M., Doligez, B., Fagot, J., Fichtel, C., Henke-von der Malsburg, J., Hermer, E., Huber, L., Huebner, F., Kappeler, P. M., Klein, S., Langbein, J., Langley, E. J. G., Lea, S. E. G., Lihoreau, M., Lovlie, H., Matzel, L. D., Nakagawa, S., Nawroth, C., Oesterwind, S., Sauce, B., Smith, E. A., Sorato, E., Tebbich, S., Wallis, L. J., Whiteside, M. A., Wilkinson, A., Chaine, A. S., Morand-Ferron, J. (2018). The repeatability of cognitive performance: a meta-analysis, *Phil. Trans. R. Soc. B. 373*, 20170281.

Cauchoix, M., Hermer, E., Chaine, A. S., & Morand-Ferron, J. (2017). Cognition in the field: comparison of reversal learning performance in captive and wild passerines. *Scientific Reports*, *7*(12945), 1–10.

Collins, S. M., Hatch, S. A., Elliott, K. H., & Jacobs, S. R. (2019). Boldness, mate choice and reproductive success in *Rissa tridactyla*. *Animal Behaviour*, *154*, 67–74.

Croston, R., Branch, C. L., Kozlovsky, D. Y., Dukas, R., & Pravosudov, V. V. (2015). Heritability and the evolution of cognitive traits. *Behavioral Ecology*, *26*(6), 1447–1459.

Croston, R., Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Bridge, E. S., Parchman, T. L., & Pravosudov, V. V. (2017). Predictably harsh environment is associated with reduced cognitive flexibility in wild food-caching mountain chickadees. *Animal Behaviour*, *123*, 139–149.

Dukas, R. (2009). Learning mechanisms, ecology, and evolution. In R. Dukas & J. M. Ratcliff (Eds.), *Cognitive Ecology II* (pp. 7–26). Chicago: University of Chicago Press.

Dukas, R., & Bernays, E. A. (2000). Learning improves growth rate in grasshoppers. *Proceedings of the National Academy of Sciences*, *97*(6), 2637–2640.

Foucaud, J., Moreno, C., Pascual, M., Rezende, E. L., Castañeda, L. E., Gibert, P., & Mery, F. (2016). Introduced Drosophila subobscura populations perform better than native populations during an oviposition choice task due to increased fecundity but similar learning ability. *Ecology and Evolution*, *6*(6), 1725–1736.

Freas, C. A., Roth II, T. C., LaDage, L. D., & Pravosudov, V. V. (2013). Hippocampal neuron soma size is associated with population differences in winter climate severity in food-caching chickadees. *Functional Ecology*, *27*, 1341–1349.

Gonzalez, R. C., Behrend, E. R., & Bitterman, M. E. (1967). Reversal learning and forgetting in bird and fish. *Science.*, *158*, 519–521.

Gosler, A. G. (1993). *The great tit*. London: Hamlyn.

Guigueno, M. F., Macdougall-Shackleton, S. A., & Sherry, D. F. (2015). Sex differences in spatial memory in brown- headed cowbirds : males outperform females on a touchscreen task. *PloS One*, 1–18.

Guigueno, M. F., Snow, D. A., MacDougall-Shackleton, S. A., & Sherry, D. F. (2014). Female cowbirds have more accurate spatial memory than males. *Biology Letters*, *10*(2), 3–6.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, *269*(5220), 41–45.

Hermer, E., Cauchoix, M., Chaine, A. S., & Morand-Ferron, J. (2018). Elevation-related difference in serial reversal learning ability in a nonscatter hoarding passerine. *Behavioral Ecology*, *29*(4), 840–847.

Hermer, E., Murphy, B., Chaine, A. S., & Ferron, J. M. (2021). Great tits who remember more accurately have difficulty forgetting , but variation is not driven by environmental harshness. *Scientific Reports*, 1–11.

Huebner, F., Fichtel, C., & Kappeler, P. M. (2018). Linking cognition with fitness in a wild primate: Fitness correlates of problem-solving performance and spatial learning ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756).

Jaumann, S., Scudelari, R., & Naug, D. (2013). Energetic cost of learning and memory can cause cognitive impairment in honeybees. *Biology Letters*, *9*(4), 20130149.

Johnson, W., Deary, I. J., & Carothers, A. (2008). Sex differences in variability in general intelligence: a new look at the old question. *Perspectives on Psychological Science*, *3*(6), 518–531.

Kolss, M., Kraaijeveld, A. R., Mery, F., & Kawecki, T. J. (2006). No trade-off between learning ability and parasitoid resistance in Drosophila melanogaster. *Journal of Evolutionary Biology*, *19*(4), 1359–1363.

Körner, C. (2007). The use of "altitude" in ecological research. *TRENDS in Ecology and Evolution*, *22*(11), 569–574.

Kotrschal, A., Rogell, B., Bundsen, A., Svensson, B., Zajitschek, S., Brännström, I., Immler, S., Maklakov, A., & Kolm, N. (2013). Artificial selection on relative brain size in the guppy reveals costs and benefits of evolving a larger brain. *Current Biology*, *23*(2), 168–171.

Krebs, J. R. (1990). Food-storing birds: adaptive specialization in brain and behaviour? *Phil. Trans. R. Soc. Lond. B*, *329*, 153–160.

Lambert, C. T., Sewall, K. B. & Guillette, L. M. (2019) Questioning the developmental effects of group size on cognitive abilities. *Learn. Behav. 47*, 280–283.

Lehre, A. C., Lehre, K. P., Laake, P., & Danbolt, N. C. (2009). Greater intrasex phenotype variability in males than in females is a fundamental aspect of the gender differences in humans. *Developmental Psychobiology*, *51*(2), 198–206.

Lucon-Xiccato, T., & Bisazza, A. (2017). Sex differences in spatial abilities and cognitive flexibility in the guppy. *Animal Behaviour*, *123*, 53–60.

Madden, J. R., Langley, E. J. G., Whiteside, M. A., Beardsworth, C. E., & van Horik, J. O. (2018). The quick are the dead: pheasants that are slow to reverse a learned association survive for longer in the wild. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756), 20170297.

Mery, F., & Kawecki, T. J. (2002). Experimental evolution of learning ability in fruit flies. *PNAS*, *99*, 14274–14279.

Mery, F., & Kawecki, T. J. (2003). A fitness cost of learning ability in *Drosophila melanogaster*. *Proc. R. Soc. Lond B*, *270*, 2465–2469.

Mery, F., & Kawecki, T. J. (2004). An operating cost of learning in *Drosophila melanogaster*. *Animal Behaviour*, *68*, 589–598.

Mery, F., & Kawecki, T. J. (2005). A cost of long term memory in *Drosophila*. *Science*, *308*.

Milton, K. (1981). Distribution patterns of tropical plant foods as an evolutionary stimulus to primate mental development. *American Anthropologist*, *83*(3), 534–548.

Morand-Ferron, J., Hamblin, S., Cole, E. F., Aplin, L. M., & Quinn, J. L. (2015). Taking the operant paradigm into the field: associative learning in wild great tits. *PloS ONE*, *10*(8), e0133821.

Niemelä, P. T., & Dingemanse, N. J. (2018). On the usage of single measurements in behavioural ecology research on individual differences. *Animal Behaviour*, *145*, 99–105.

Osborne, O. E., Hara, P. D. O., Whelan, S., Zandbergen, P., Hatch, S. A., & Elliott, K. H. (2020). Breeding seabirds increase foraging range in response to an extreme marine heatwave. *Marine Ecology Progress Series*, *646*, 161–173.

Pagani-Nunez, E., Barnett, C. A., Gu, H., & Goodale, E. (2016). The need for new categorizations of dietary specialism incorporating spatio-temporal variability of individual diet specialization. *Journal of Zoology*, *300*(1), 1–7.

Plaçais, P.-Y., & Preat, T. (2013). To favor survival under food shortage, the brain disables costly memory. *Science*, *339*(6118), 440–442.

Pomiankowski, A., & Møller, A. P. (1995). A resolution of the lek paradox. *Proceedings: Biological Sciences*, *260*(1357), 21–29.

Pravosudov, V. V., & Clayton, N. S. (2002). A test of the adaptive specialization hypothesis: population differences in caching, memory, and the hippocampus in black-capped chickadees (*Poecile atricapilla*). *Behavioral Neuroscience*, *116*(4), 515–522.

Pravosudov, V. V., & Roth II, T. C. (2013). Cognitive ecology of food hoarding: the evolution of spatial memory and the hippocampus. *Annual Review of Ecology, Evolution, and Systematics*, *44*, 173–193.

Pravosudov, V. V., Roth, T. C., LaDage, L. D., & Freas, C. A. (2015). Environmental influences on spatial memory and the hippocampus in food-caching chickadees. *Comparative Cognition & Behavior Reviews*, *10*, 25–43.

Raine, N. E., & Chittka, L. (2008). The correlation of learning speed and natural foraging success in bumble-bees. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1636), 803–808.

Reinhold, K., & Engqvist, L. (2013). The variability is in the sex chromosomes. *Evolution*, *67*(12), 3662–3668.

Roth II, T. C., Chevalier, D. M., LaDage, L. D., & Pravosudov, V. V. (2013). Variation in hippocampal glial cell numbers in food-caching birds from different climates. *Developmental Neurobiology*, *73*, 480–485.

Roth II, T. C., LaDage, L. D., & Pravosudov, V. V. (2010). Learning capabilities enhanced in harsh environments: a common garden approach. *Proceedings of the Royal Society B*, *277*, 3187–3193.

Roth II, T. C., & Pravosudov, V. V. (2009). Hippocampal volumes and neuron numbers increase along a gradient of environmental harshness: a large-scale comparison. *Proceedings of the Royal Society B*, *276*, 401–405.

Rowe, L., & Houle, D. (1996). The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society B: Biological Sciences*, *263*(1375), 1415–1421.

Sherry, D. F. (1998). The ecology and neurobiology of spatial memory. In R. Dukas (Ed.), *Cognitive Ecology* (1st ed., pp. 261–296). Chicago: University of Chicago Press.

Sherry, D. F. & Guigueno, M. F. (2019) Cognition and the brain of brood parasitic cowbirds. *Integr. Zool.* 145–157.

Shettleworth, S. J. (2010). Discrimination, classification and concepts. in *Cognition, evolution and behaviour* 186–188 (Oxford University Press).

Sol, D., Lefebvre, L., & Rodríguez-Teijeiro, J. D. (2005). Brain size, innovative propensity and migratory behaviour in temperate Palaearctic birds. *Proceedings of the Royal Society B*, *272*, 1433–1441.

Sonnenberg, B. R., Branch, C. L., Pitera, A. M., Bridge, E., & Pravosudov, V. V. (2019). Natural selection and spatial cognition in wild food-caching mountain chickadees. *Current Biology*, *29*, 1–7.

Tebbich, S., & Teschke, I. (2014). Coping with uncertainty: woodpecker finches (*Cactospiza pallida*) from an unpredictable habitat are more flexible than birds from a stable habitat. *PLoS ONE*, *9*(3).

Tello-Ramos, M. C., Branch, C. L., Kozlovsky, D. Y., Pitera, A. M., & Pravosudov, V. V. (2018). Spatial memory and cognitive flexibility trade-offs: to be or not to be flexible, that is the question. *Animal Behaviour*, 1–8.

Tello-Ramos, M. C., Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Bridge, E. S., & Pravosudov, V. V. (2018). Memory in wild mountain chickadees from different elevations: comparing first-year birds with older survivors. *Animal Behaviour*, *137*, 149–160.

Titulaer, M., van Oers, K. & Naguib, M. (2012). Personality affects learning performance in difficult tasks in a sex-dependent way. *Anim. Behav. 83*, 723–730

Tomkins, J. L., Radwan, J., Kotiaho, J. S., & Tregenza, T. (2004). Genic capture and resolving the lek paradox. *Trends in Ecology and Evolution*, *19*(6), 323–328.

Zajitschek, S. R. K., Zajitschek, F., Bonduriansky, R., Brooks, R. C., Cornwell, W., Falster, D. S., Lagisz, M., Mason, J., Senior, A. M., Noble, D. W. A., & Nakagawa, S. (2020). Sexual dimorphism in trait variability and its eco-evolutionary and statistical implications. *ELife*, *9*, 1–17.

Zechner, U., Wilda, M., Kehrer-Sawatzki, H., Vogel, W., Fundele, R., & Hameister, H. (2001). A high density of x-linked genes for general cognitive ability: a run-away process shaping human evolution? *TRENDS in Genetics*, *17*(12), 697–701.

Zidar, J., Balogh, A., Favati, A., Jensen, P., Leimar, O., Sorato, E., & Løvlie, H. (2018). The relationship between learning speed and personality is age- and task-dependent in red junglefowl. *Behavioral Ecology and Sociobiology*, *72*(10).

Chapter 1: *Elevation-related difference in serial reversal learning ability in a nonscatter hoarding passerine*

Adapted From:

Hermer, E., Cauchoix, M., Chaine, A. S. & Morand-Ferron, J. (2018). Elevation-related difference in serial reversal learning ability in a nonscatter hoarding passerine. *Behavioural Ecology.* 29, 840–847.

Introduction

Environments characterized by low food abundance and a variable food supply, termed "harsh environments," could favor cognitive abilities that help individuals cope with these challenges by enhancing their ability to find food, remember where it is located, or predict its availability (Dukas, 2009; Kozlovsky, Weissgerber, & Pravosudov, 2017; Milton, 1981; Pravosudov & Roth II, 2013; Sol, Lefebvre, & Rodríguez-Teijeiro, 2005). This "harsh environment" hypothesis has been tested along latitudinal and elevation gradients, as increases in latitude and elevation can correlate to more snow cover, lower temperatures, greater seasonality, and shorter day length, which lead to greater variability in the availability of food and higher metabolic costs (Roth II & Pravosudov, 2009; reviewed in: Boyle, Sandercock, & Martin, 2016; but see: Körner, 2007). Over these two gradients, the "harsh environment" hypothesis has been studied using scatter-hoarding birds, or birds that store food in multiple locations, as these stored caches allow them to circumvent resource variability (Croston *et al*., 2015). This scatter-hoarding behavior has been shown to require spatial memory to remember and retrieve previous caches (Krebs, 1990; reviewed in: Sherry, 1998). Empirical support for the "harsh environment" hypothesis in scatter-hoarders include positive links between elevation or latitude and the size and density of the neurological structures linked to memory (elevation: Freas, Roth II, LaDage, & Pravosudov, 2013; latitude: Freas, Bingman, LaDage, & Pravosudov, 2013; Roth II, Chevalier, LaDage, & Pravosudov, 2013) and behavioral measures of spatial learning and memory (reviewed in: Pravosudov, Roth, LaDage, & Freas, 2015). Although a positive correlation between residual brain size, as a proxy for general cognitive ability, and resource variability has been found in a large-scale comparative study (Sayol *et al*., 2016), few studies have examined the "harsh environment" hypothesis in nonscatter hoarders.

Compared to the amount of knowledge that has accumulated over the past decades on the cognitive ecology of avian scatter hoarders, little is known about the cognitive solutions that may be employed by nonhoarders to survive harsh environments. However, models predict that if the environment contains information that can be used to acquire food efficiently or predict its availability, associative learning, or the ability to learn about the environment, would be advantageous (Mery & Burns, 2010; reviewed in: Dunlap & Stephens, 2016). In support of this, Kotrschal and Taborsky (2010) found that a cichlid fish (*Simochromis pleurospilus*) that experienced a change in food abundance during development performed more accurately on an associative learning task than fish from a control treatment These findings suggest that environmental variability may activate an increase in associative learning ability, pointing to its potential usefulness in a changing environment.

In natural environments, the value of temporally and spatially variable resources can change multiple times, and thus may not only require animals to acquire learned associations, but also to track the value of multiple options and behave accordingly. The ability to collect and utilize information to flexibly adjust one's associations can be measured via reversal learning tasks (Shettleworth, 2010). In a reversal learning task, an animal is required to learn to associate a stimulus with a response (e.g., associate a red well with food, and a blue well with no food) and once the animal has learned this association, the originally unrewarded stimulus is now rewarded and the rewarded stimulus becomes unrewarded (e.g., associate a blue well with food, and a red well with no food). Several reversals can also be used, such that animals have to adjust their choices repeatedly (i.e., serial reversal learning: Shettleworth, 2010). Performance on reversal learning tasks is thought to measure the suite of abilities an animal would use, for example, to change its foraging behavior to match a change in food availability and therefore it can be

predicted to covary with environmental harshness (Janmaat *et al*., 2016; Tebbich & Teschke, 2014). To our knowledge, the only study that examined the relationship between "harsh environments" and reversal learning in nonscatter hoarders is Tebbich and Teschke's (2014) study on woodpecker finches (*Cactospiza pallida*). A single reversal was used and returned evidence that supports the "harsh environment" hypothesis, with woodpecker finches from the food-scarce, more variable desert site performing the reversal faster than those from the more stable cloud forest site.

Performance in reversal learning tasks can be measured in one or several reversals, but also as the rate of change, or improvement of performance across reversals (Bebus, Small, Jones, Elderbrock, & Schoech, 2016; Bond, Kamil, & Balda, 2007; Liu, Day, Summers, & Burmeister, 2016; reviewed in: Izquierdo, Brigman, Radke, Rudebeck, & Holmes, 2017). The improvement of performance across reversals is believed to measure an animal's ability to learn a rule to allow them to more efficiently track resource value (e.g., win-stay/lose-shift: Liu *et al*., 2016) or accumulate past information and better estimate a prior probability that reversals will occur (reviewed in: Izquierdo *et al*., 2017). This ability could allow animals to increase the rate at which they can track recurrent changes in their environment (e.g., Chow, Leaver, Wang, & Lea, 2017; Liu *et al*., 2016; Strang & Sherry, 2014), and could thus be advantageous in highly-changing environments. However, it has never been examined in the context of the "harsh environment" hypothesis.

Here, we present the first test of the "harsh environment" hypothesis using a serial reversal learning task. We studied great tits (*Parus major*) from high and low elevation sites along the Pyrénées mountain range. Great tits are nonscatter hoarding passerines that forage for patchily-distributed seeds and invertebrates during the winter (Gosler, 1993). They are also able

to complete associative learning, initial and serial reversal learning tasks (Amy, van Oers, & Naguib, 2012; Cauchoix, Hermer, Chaine, & Morand-Ferron, 2017; Morand-Ferron, Hamblin, Cole, Aplin, & Quinn, 2015; Titulaer, van Oers, & Naguib, 2012). We predict great tits from higher elevations will have higher accuracy in an initial and serial reversal learning task, and will improve faster (steeper slope) over reversals, than great tits from lower elevations.

Methods

Study sites

Study sites were located near Moulis, France, along the north side of the Pyrénées. High (800–900 m; Galey: 42.93644″ N 0.91702″ E, Cap Sour: 42.93203″ N 1.12322″ E) and low (400–500 m; Aubert: 42.96427″ N 1.10302″ E, Ledar: 42.98306″ N 1.11492″ E, Moulis: 42.96551″ N, 1.08791″ E, Lab: 42.95741″ N, 1.08698″ E) elevation sites covered an elevation range similar to that in other studies of harsh environments (e.g., Croston *et al*., 2017) In the Pyrénées, this results in a temperature decrease of 2 °C (personal communication Chaine, A., Bruendl, 2017; López-Moreno & Nogués-Bravo, 2005), a shorter growing period (Ninot, Carrillo, & Ferré, 2017), and increased snow accumulation (López-Moreno & Nogués-Bravo, 2005). Our high elevation sites were also characterized by increased weekly snow cover compared with our low elevation sites (Supplementary S1.1), which is expected to equate to lower food availability and increased variability in food availability (Boyle *et al*., 2016).

Capture and housing

First capture took place between 15 February 2016 and 1 March 2016 (Winter) when 21 high and 25 low elevation birds were captured. From 7 October 2016 to 22 November 2016 (Fall), an additional 8 high and 12 low elevation birds were captured (Supplementary S1.2). Great tits were captured using mist nets, and marked with a CRBPO (Centre de Recherches sur la Biologie des Populations d'Oiseaux) metal band and colored RFID band (passive integrated transponder, IB Technologies, UK). We used plumage to sex and age the birds (Svensson, 1992). We placed the birds in cloth bags and transported them to outdoor aviaries in Moulis where they were housed in groups of two to five per aviary (1 x 4 x 3 m), depending on how many birds

were caught. Each aviary contained foliage for cover in the nontesting area, four roosting boxes, and two horizontal perches between the foliage and the testing area for perching.

Operant devices

Operant devices consisted of a single portable operant box (Cauchoix *et al.*, 2017; Morand-Ferron *et al*., 2015), which was composed of a waterproof Perspex casing that contained: a motor-activated feeding tray, a printed circuit board (PCB, "Darwin board" designed by Stickman Technologies Inc., UK), and two transparent pecking keys lit up by white light-emitting diode (LEDs). The front panel of the operant box was freely available to any individual and displayed two pecking keys that each contained two live mealworms. The great tits were unable to consume the mealworms which instead acted as bait to elicit pecking actions. A feeding hole 1.5 cm under the pecking keys provided a single dried mealworm reward on the feeding tray when the correct key was pecked. A 10 cm x 5 cm x 1 cm PIT tag identification perch was attached outside of the box, under the feeding hole. The perch instructed the PCB to start the learning program at each individual's current position within the program (see below). For visits that led to a bird pecking a key, which we termed a trial, the PCB recorded whether the choice was correct (1) or incorrect (0). If the bird did not peck a key, no trial was recorded.

Acclimatization

Acclimatization to the aviary occurred for the first 7–12 days of captivity. For the first three days, the birds were given ad libitum black oil sunflower seeds, fat balls, meal worms, and water. Additionally, a small heated (25 °C) room inside of the aviary building was left open and contained a second source of ad libitum food and water, as well as a constant source of light, to encourage feeding. During the acclimatization period, a wooden dummy operant box was placed at the end of the outdoor aviary, on the wooden shelf where the operant box would be during the

test phase (Cauchoix *et al.*, 2017). The dummy box was baited with seed and margarine where the pecking keys would be, and sunflower seeds alone where the reward hole would be. This was done to acclimatize the birds to the components and location of the operant device, and learn that it was a source of food. After three days, the adjoining room to the aviary was shut and the birds were kept only in the outdoor portion of the aviary. Dummy operant boxes were removed at the end of acclimatization.

Testing

For two weeks after acclimatization, the great tits were exposed to a serial spatial reversal learning task for 2 h. Great tits were deprived of food overnight and the first hour after sunrise in the morning (08h00), and for 1 h in the afternoon (14h00). Operant boxes were in the cages for a total of 2 h after each deprivation period, where great tits could engage freely in any number of trials by visiting the device. In total, 28 sessions occurred over 14 days for each group. Reward wheels were checked and replenished halfway through each session when depleted. After the 2 h testing intervals, great tits received ad libitum food and water, and the operant boxes were shut off and covered. The first landing an individual great tit made that resulted in a peck to a pecking key, i.e., the first trial, started its learning program (Cauchoix *et al*., 2017). After this trial, the bird received a reward and the motor training stage started. In this stage, an individual great tit was required to peck the same key it had previously pecked to receive a reward. If the opposite key was pecked, both LEDs would shut off for 500 ms and turn back on. Once an individual obtained nine correct trials in total during the motor training stage, it proceeded to the first reversal learning stage. During the first reversal learning stage, individuals had to inhibit pecking the past correct key from the motor training stage, and instead peck the other key. Incorrect pecks led to the LEDs turning off, and the operant box would become unresponsive to that

individual great tit for a set amount of time. Due to technical issues, a slightly different program was used in each season; in the winter, the punishment period lasted 15 s or until another bird visited the device, while for fall birds it lasted 5 s regardless of whether another bird visited the device or not (included as a covariate in analyses as "season"). The first reversal stage was completed when an individual pecked the correct pecking key in 9 out of 10 successive trials. The second reversal stage would then commence and this individual would have to peck the opposing pecking key from the first reversal stage and again peck the correct key in 9 of 10 successive trials before moving to the next reversal. This alternation between right-key and left-key rewarded pecks would continue with each successive stage up to a maximum of 18 serial reversals. This progression of stages was the same for each bird and progress was maintained across each two-week reversal learning period.

Statistical analysis

To discern whether there was a difference in initial motivation, acclimatization or motor proficiency with the task between high and low elevation birds, we tested for differences in the proportion of birds using the device (i.e., returning a first peck), the proportion of birds completing the motor training phase, the first reversal, and the final reversal (18th reversal) with Pearson's chi-squared tests with Yates' continuity correction. We also compared the mean session number (out of 28 sessions) when the birds registered a first peck and completed the motor training stage between elevations using all birds that made it to each of these stages using a nonpaired Wilcoxon signed-rank test, as these data did not meet the assumption of normality. Due to technical difficulties, one high elevation bird did not have a time for first peck and was excluded from the comparison of the mean session number when the birds registered a first peck.

To determine whether there was a difference in the pace at which trials were taken by birds from high and low elevation, which could have impacted their ability to remember learned associations, we compared their intertrial interval (ITI) durations during reversal learning. We could model only 80.7% (88/109) of the fall bird's reversals due to technical difficulties that returned missing timing information for some ITIs. We used a linear mixed-model with Gaussian distribution (LMM). The dependent variable, ITI (number of seconds from the last reversal trial), was log-transformed to better meet the assumptions of the model. The independent variable was high or low elevation and bird ID was included as a random intercept. We compared this model against the null model using a likelihood ratio test.

We examined the differences in both initial and serial reversal learning accuracy between elevations using generalized linear mixed models (GLMM) with a binary response (1: trial correct, 0: trial incorrect). We included fixed effects of elevation (high, low), sex (male, female), age (juvenile, adult), trial number, and group size. We also included a fixed effect of season (fall, winter), to control for differences in the programmed delay after an incorrect peck and potential seasonal effects. Bird ID and capture site were included as random intercepts (Schielzeth & Forstmeier, 2008).

In the initial reversal model, the random intercept of site returned a zero-variance estimate and we were thus unable to control for a site effect with this approach. To check for a site effect, we used two separate high and low elevation models that contained their respective sites as a fixed effect. The high and low elevation models would not converge with group size included and so it was dropped for these analyses. The effect of site was not significant in either high or low elevation models (Supplementary S1.3, S1.4, S1.5, S1.6) and it was thus excluded from further analyses.

The serial reversal model utilized the same variables as those used for initial reversal learning, with the addition of a reversal number x elevation interaction to test for elevation-related differences in reversal learning improvement, and a random slope of ID x reversal number to account for individual variation in improvement in accuracy over reversals (Schielzeth & Forstmeier, 2008). To check for a site effect in the serial reversal learning models we compared a model with and without the random effect of site using a likelihood ratio test; as it was not significant ($\chi 2 = 24073$, $P = 1$), we excluded it from further analysis.

LMMs and GLMMs were fit using the statistical package lme4 (Bates, Maechler, Bolker, & Walker, 2015) in R (version 3.4.4; R Core Team, 2018). We used Akaike's Information Criterion adjusted for small sample size (AICc) for model selection (R package MuMIn; Bartoń, 2018).

We performed model averaging using models within $< \Delta 2$ AICc of the top model (Grueber, Nakagawa, Laws, & Jamieson, 2011). All continuous predictor variables were standardized by grand mean centering and dividing by 1 standard deviation. Means are presented with their standard errors. Three low elevation birds from the fall experiments were exposed to a similar task during pretrials in the previous winter; two did not contact the device and one only passed the motor training stage. We reran all analyses (Participation, Initial Contact and Motor Training, ITI, Initial Reversal and Serial Reversal) while excluding these data points; our conclusions were unchanged and therefore we report results from the full dataset below.

Ethics

This work was conducted under animal care permits from the French bird ringing office (CRBPO; n° 13619) and animal experimentation review from the state of Ariège (Préfecture de l'Ariège, Protection des Populations, n° A09-4) and the Région Midi- Pyrenées (DIREN, n° 2012-07). Birds were housed in the Moulis experimental aviaries (Préfecture de l'Ariège, institutional permit n° SA-12-MC-054) under animal captivity permits (Préfecture de l'Ariège, Certificat de Capacite, n° 09-321). One great tit from winter capture died during acclimatization.

Results

Participation

All birds pecked the device at least once (High: 28/28; low: 24/24). We found no significant difference between high and low elevation birds in the proportion of birds that passed the motor training phase (High: 24/28, 22/24. Contingency table: $\chi2 = 0.06$, df = 1, P = 0.81), passed at least one reversal (High: 18/28, 13/24. Contingency table: $\chi2 = 0.21$, df = 1, P = 0.65), nor in the proportion of birds that completed the final reversal (High: 11/28, low: 10/24. Contingency table: $\chi2 < 0.001$, df = 1, P = 1).

Initial contact and motor training

There was no significant difference between elevations in the number of learning sessions before they first pecked a key (High: 4.00 ± 4.17; low: 2.21 ± 2.02; Wilcoxon rank sum test, high vs. low: W = 416.5, P = 0.066, N = 51), or before they passed the motor training phase (High: 7.48 ± 6.10; low: 7.23 ± 6.89; Wilcoxon rank sum test, high vs. low: W = 261, P = 0.86, N = 46).

Initial reversal learning

Elevation, sex, age, capture season, and group size had no significant effect on reversal learning accuracy in the first reversal. Trial number had a significant positive effect on accuracy which indicates that the birds improved their accuracy within the first reversal (Tables 1.1 and 1.2).

Serial reversal learning

The reversal number x elevation interaction was not a significant predictor of accuracy (GLMM: z = 0. 42, P = 0.67), suggesting that the rate of improvement over successive reversals did not differ between high and low elevation birds (Supplementary S1.7 and S1.8). This interaction was excluded from further analyses. Across all reversals, low elevation birds

performed significantly more accurately than high elevation birds (Figure 1.1). There was no significant difference between capture season, sex, age, and group size in reversal learning accuracy. Both trial and reversal number had significant positive effects on accuracy, which indicates that the birds improved their performance within reversals (trials) and across successive reversals (Tables 1.3 and 1.4). Elevation was not a significant predictor of intertrial interval duration (GLMM: $\chi 2 = 0.50$, $P = 0.48$).

Table 1.1. Models within <Δ2 AICc of the top model for initial reversal learning accuracy (1: correct 0: incorrect), degrees of freedom, AICc, Δi, and Akaike weights.

Candidate Models	df	AICc	Δi	ωi
Sex+Trial+(1|ID)	4	1919.47	0.00	0.14
Trial+(1|ID)	3	1919.56	0.10	0.14
Elevation+Sex+Trial+(1|ID)	5	1919.65	0.18	0.13
Group Size+Trial+(1|ID)	4	1920.78	1.31	0.07
Season+Trial+(1|ID)	4	1920.85	1.39	0.07
Season+Sex+Trial+(1|ID)	5	1920.91	1.44	0.07
Elevation+Season+Sex+Trial+(1|ID)	6	1920.93	1.46	0.07
Age+Sex+Trial+(1|ID)	5	1921.10	1.63	0.06
Season+Group Size+Trial+(1|ID)	5	1921.22	1.75	0.06
Elevation+Group Size+Sex+Trial+(1|ID)	6	1921.23	1.76	0.06
Elevation+Trial+(1|ID)	4	1921.24	1.77	0.06
Group Size+Sex+Trial+(1|ID)	5	1921.26	1.79	0.06

Table 1.2. Summary of results of model-averaged coefficients for initial reversal learning accuracy (1: correct 0: incorrect).

Parameter	Estimate	Unconditional SE	Confidence interval	Relative importance	*P*
Intercept	0.26	0.32	(−0.37, 0.88)		0.42
Sex (Female)	0.36	0.42	(−0.16, 1.36)	0.60	0.39
Trial	3.33	0.33	(2.68, 3.99)	1.00	<0.001
Elevation (Low)	0.15	0.31	(−0.31, 1.24)	0.32	0.63
Group Size	−0.14	0.41	(−1.85, 0.78)	0.25	0.74
Season (Fall)	−0.09	0.24	(−1.03, 0.38)	0.27	0.71
Age (Adult)	0.19	0.14	(−0.66, 1.27)	0.06	0.89

The estimates are from the full average of 12 models <Δ2 AIC of the top model. Full model: Correct/Incorrect = Group Size + Sex + Age + Season + Elevation + Trial + (1ID).

Table 1.3. Models within <Δ2 AICc of the top model for serial reversal learning accuracy (1: correct 0: incorrect), degrees of freedom, AICc, Δi, and Akaike weights.

Candidate Models	df	AICc	Δi	ωi
Elevation+Season+Reversal Number+Trial+(1+Reversal Number|ID)	8	24,091.56	0.00	0.26
Elevation+Reversal Number+Trial+(1+Reversal Number|ID)	7	24,091.75	0.18	0.24
Elevation+Season+Group Size+Reversal Number+Trial+ (1+Reversal Number|ID)	9	24,091.99	0.42	0.21
Elevation+Group Size+Reversal Number+Trial+(1+Reversal Number|ID)	8	24,093.34	1.77	0.11
Elevation+Season+Sex+Reversal Number+Trial+(1+Reversal Number|ID)	9	24,093.54	1.98	0.10
Age+Elevation+Season+Reversal Number+Trial+(1+Reversal Number|ID)	9	24,093.5	1.98	0.10

Table 1.4. Summary of results of model-averaged coefficients for serial reversal learning: effects of each parameter on the accuracy (1: correct, 0: incorrect) of choice during reversals 1–18.

Parameter	Estimate	Unconditional SE	Confidence interval	Relative importance	*P*
Intercept	0.11	0.07	(−0.02, 0.24)		0.09
Elevation (Low)	0.24	0.08	(0.08, 0.41)	1.00	0.004
Season (Fall)	−0.09	0.09	(−0.30, 0.03)	0.66	0.34
Reversal Number	0.53	0.15	(0.24, 0.81)	1.00	<0.001
Trial	3.40	0.11	(3.19, 3.61)	1.00	<0.001
Group Size	−0.06	0.14	(−0.55, 0.18)	0.32	0.66
Sex (Female)	−0.001	0.03	(−0.18, 0.16)	0.10	0.96
Age (Adult)	0.001	0.03	(−0.19, 0.22)	0.10	0.97

The estimates are from the full average of six models <Δ2 AIC of the top model. Full model: Correct/Incorrect = Group Size + Age + Sex + Season + Trial + Elevation + Reversal Number + (1+Reversal Number|ID).

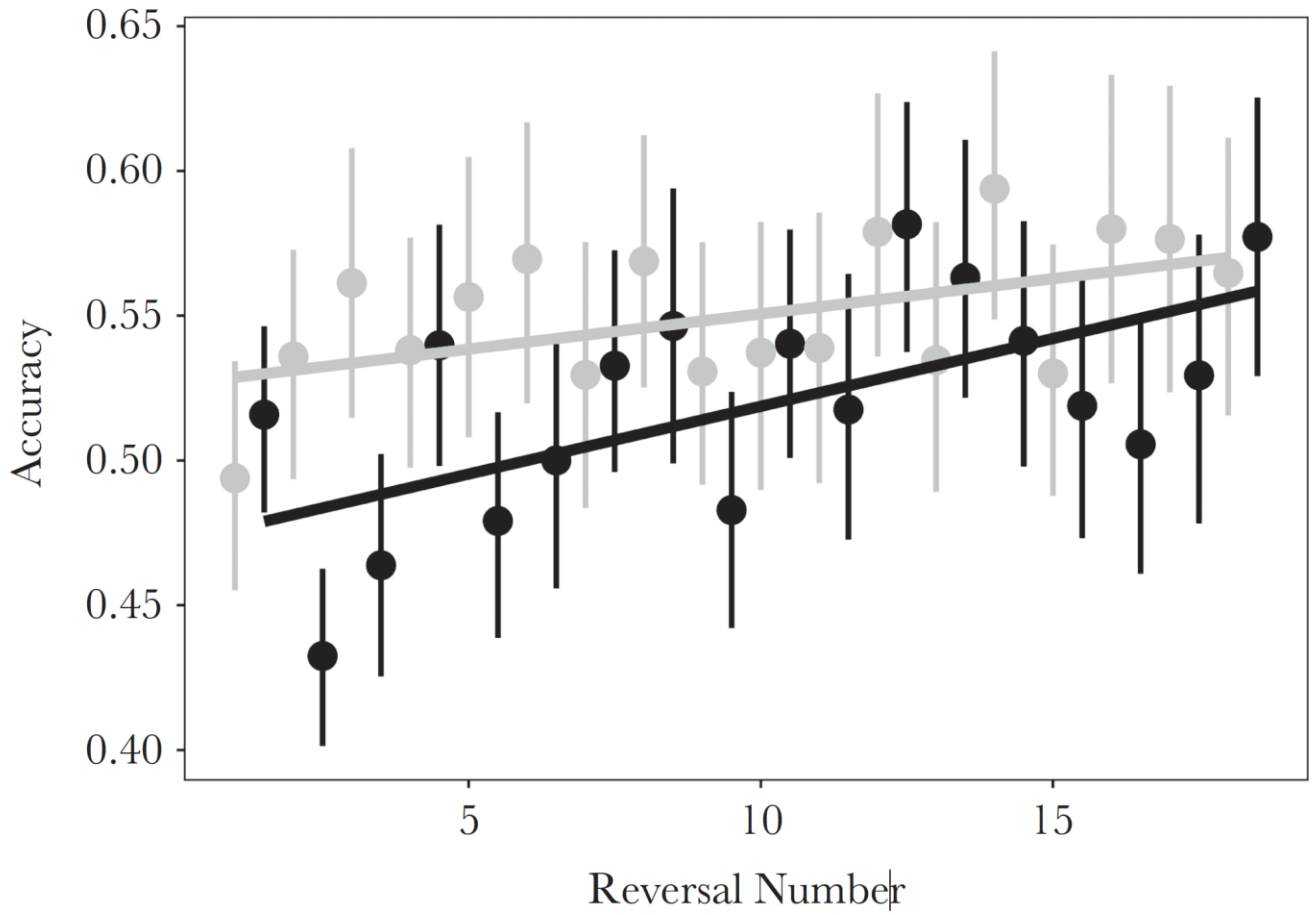


Figure 1.1. Mean accuracy of the individual bird's raw accuracy data, with 95% confidence intervals calculated using a non-parametric bootstrap, as a function of reversal number, for high and low elevations. The trend line is estimated from a general linear model of the relationship of choice accuracy across reversals for high and low elevations. Black points are high elevation, gray points are low elevation.

Discussion

The study of the relationship between cognitive abilities and environmental harshness is still a novel subject and has only been examined in a restricted number of ecological scenarios and species. Here, we performed the first serial reversal learning comparison between individuals from different elevations. We found that there was no difference in the performance of great tits from high and low elevation during the first reversal but when examining performance across all reversals, lower elevation birds were significantly more accurate than high elevation birds. There were no elevation-related differences in the rate of improvement across reversals, although birds from both elevations significantly increased their choice accuracy over successive reversals.

The "harsh environment" hypothesis predicts that reversal learning ability should increase with harshness (Tebbich & Teschke, 2014). In contrast we report: 1) no significant difference in initial reversal accuracy, and 2) higher accuracy over all reversals in birds from low, less harsh environments than in birds from high elevations. Tebbich and Teschke (2014) found support for the "harsh environment" hypothesis with woodpecker finches: individuals from a harsher desert habitat completed a reversal learning task faster than those from a less harsh cloud forest habitat. The opposing results of our study and Tebbich and Teschke's (2014) are surprising, given that our study species, the great tit, and the woodpecker finches in Tebbich and Teschke's (2014) study, are both non-migratory and non-scatter hoarders (Gosler, 1993; Tebbich, Taborsky, Fessl, & Dvorak, 2002). However, the two species differ in other aspects; for instance, woodpecker finches came from populations that differ in tool use, with finches from the harsher environment utilizing tools more frequently than those from the less harsh environment, which may select for greater reversal learning ability in the population from the harsher environment (Tebbich *et al*., 2002; but see: Teschke, Cartmill, Stankewitz, & Tebbich, 2011).

Our results across all reversals concur with those of Croston *et al*. (2017), who utilized a scatter-hoarding species and found that performance on a single spatial reversal of mountain chickadees (*Poecile gambeli*) was inversely related to elevation. The authors suggested that this difference was due to the high elevation chickadees' stronger spatial memory causing proactive interference, i.e., their stronger memories make it difficult to forget past associations and remember new ones (Croston *et al*., 2017; Gonzalez, Behrend, & Bitterman, 1967). This interpretation may also apply to our study population; if food is less abundant at high elevation (reviewed in: Boyle *et al.*, 2016), high elevation great tits may rely more strongly on spatial memory to remember patches of food as they find them, so they can efficiently return when their current patch has been depleted (reviewed in: Janmaat *et al.*, 2016; Milton, 1981). Consequently, they may suffer from greater proactive interference than low elevation great tits, and perform worse at the serial spatial reversal learning task, which required frequently switching spatial associations. It is not currently known whether great tits from the harsh, high elevation sites exhibit increased spatial memory compared with their low elevation counterparts. If this is the case, then high elevation great tits may rely on durable, rather than flexible, spatial memory, to survive in a harsh environment.

Another possible explanation for our results is that environmental harshness itself may not be the main driver of reversal learning ability in our system. Instead, our results may be due to other ecological differences, such as population density or resource breadth. Population density seems to decrease as elevation increases in our system (personal communication, Chaine, A.) and it has been hypothesized (Dunbar, 1998) and shown that living in complex social structures can be associated with increased serial reversal learning performance (Ashton, Ridley, Edwards, & Thornton, 2018; Bond *et al.*, 2007). Larger flocks may require increased reversal

learning ability to compete and interact with more conspecifics at low elevations. Furthermore, if there is a higher density of individuals foraging together at a single patch, food patches may deplete faster than those at higher elevations where only a few birds utilize them. If patches are depleted faster at low elevations, then low elevation birds may require greater reversal learning ability than high elevation birds to more efficiently switch between rapidly depleting resources. Moreover, if lower elevations have a greater diversity of available types of food (i.e., seed from various tree species; Pausas & Austin, 2001; reviewed in: Pausas, Carreras, Ferré, & Font, 2003) this may require greater reversal learning ability to maximize consumption of the best available food. In contrast, high elevation birds may only have the opportunity to consume a few food types and rarely utilize reversal learning to switch between optimal food types.

Although we found a difference in overall reversal learning accuracy between high and low elevation great tits, these differences could potentially be due to differences in these population's reactions to the testing environment (e.g., Calandreau *et al.*, 2011; Thai, Zhang, & Howland, 2013). For instance, high and low elevation birds may have differed in their motivation to start or continue through the task, which could affect reversal learning performance (e.g., Liu *et al.*, 2016). Both high and low altitude birds had at least seven days to acclimate to the testing environment, but to ensure that the effect of elevation on reversal learning performance was not due to differences in acclimatization to the test environment and/or motivation, we compared participation between the high and low elevation great tits. All birds participated in the task. There was no significant difference between elevations in the proportion of birds that completed the motor training stage, at least one reversal, or completed all reversals. Drop-out rates are similar to those found in free-ranging great tits performing the same task (Cauchoix *et al*., 2017). Additionally, there was no significant difference in the session number

the birds started the task nor when they successfully passed the motor training phase. Therefore, we believe the differences in reversal learning between low and high elevation birds were not due to differences in motivation and/or acclimatization. High and low elevation birds could have worked on the task at different rates; for example, high elevation birds could have waited less time between trials and thus retained a stronger memory of the previous trial, increasing the accuracy of their decisions (e.g., Rayburn-Reeves, Laude, & Zentall, 2013). However, this explanation seems unlikely as ITI was also not significantly different between elevations. Finally, differences in social information use between high and low elevations could potentially have influenced my result. To control for increased interference from a larger number of conspecifics, we included group size in each model. Additionally, the birds were asynchronously working on the task, and were at different stages, with different correct sides across testing periods. Due to this, we assume that copying a conspecifics' choice would be correct at only 50% (two options test). As a bird needs to make 9/10 correct pecks to advance through the stages, it is unlikely that high or low elevation birds would be able to complete a reversal by utilizing social information. Given that both high and low elevation were able to complete multiple stages of the serial reversal task, we do not think it was likely that they utilized social information on the serial reversal learning stages of the task.

Birds from both elevations were able to increase their serial reversal learning accuracy across reversals, and did not differ in their rate of performance improvement. The ability to improve performance over successive reversals is conserved across taxa (e.g., Passeriformes: Bond *et al*., 2007; Rodentia: Chow *et al*., 2017; Squamata: Gaalema, 2011; Primates: Hassett & Hampton, 2017; Anura: Liu *et al*., 2016; Hymenoptera: Strang & Sherry, 2014), and has been reported before in free-ranging and captive great tits (Cauchoix *et al*., 2017). To further our

understanding of how great tits were able to increase their serial reversal learning accuracy, future studies could add modifications to the serial reversal learning task. The addition of a second reversal test using a different type of cue (e.g., colour and spatial: Bond *et al*., 2007) or different subsets of cues (e.g., multiple combinations of colours: Parker *et al*., 2012) could provide evidence of rule learning. The addition of a third, middle option to our test would allow us to determine whether proactive interference errors (i.e., peck the previously rewarded buttons) or exploration errors (i.e., peck the button not previously rewarded to collect more data) occurred (reviewed in: Izquierdo *et al*., 2017).

Overall, our results suggest that lower elevation great tits perform more accurately on a serial spatial reversal learning task than higher elevation birds. This difference was detected only over multiple reversals, possibly because of increased statistical power compared with an analysis on a single reversal, pointing to the usefulness of serial reversal learning as a paradigm to examine differences in reversal learning performance. More generally, our results suggest that the relationship between food variability, food scarcity, and cognition may not be consistent across all species or all harshness gradients. Furthermore, trade-offs between useful cognitive abilities such as spatial memory (Croston *et al*., 2017) and alternative factors, such as social complexity (Bond *et al*., 2007), may also drive variation in reversal learning ability. On their own, these results do not allow rejection of the "harsh environment" hypothesis, but suggests that we need a clearer definition of what constitutes a "harsh" environment. Future studies would benefit from measuring the spatial and temporal variation in ecological variables to properly define the "harshness" of a system, so as to build a greater understanding of the links between cognition and the environment (Kozlovsky *et al*., 2017).

Supplementary

S1.1. Differences in snow cover between high and low elevation capture sites.

We assessed differences in snow cover using version 6 of the MODIS/Terra Snow Cover 8-Day L3 Global 500m Grid product (Hall and Riggs 2016; Sayol *et al*. 2016). This product uses a visible/near infrared sensor to detect snow, and pools data over 8-days as single day measures can be disrupted by cloud cover. We extracted snow cover data from the MODIS product using a 500 m buffer around the capture sites as this contains most of the winter ranges of great tits reported in the literature (East & Hofer, 1985; Gosler, 1993; Krebs, 1971; Krištín & Kaňuch, 2017); and a 1500 m buffer as this was the maximum great tit winter range reported in the literature (Krištín & Kaňuch, 2017). We defined "winter" as September 5-6 to March 4-5 of the next year, as September is when juvenile great tit flocks stabilize, and adult great tits form territories (Gosler, 1993) and the beginning of March as it is the approximate start of the breeding season in great tits (Salle, 2016). We extracted snow cover data from September 5-6, 2000 until March 4-5, 2016 for a total of 16 winters. The presence of snow within the buffer during an 8-day period was coded as 1, and the presence of no snow in the buffer during an 8-day period was coded as 0. We examined the differences in snow cover between elevations with 500 m and 1500 m buffers using a generalized linear mixed model (GLMM). We included number of days from September 5-6 (September 5-6 is Day 1) as a quadratic fixed effect to control for the effect of date on snow cover during the winter. Number of days was standardized by grand mean centering and dividing by one standard deviation. Elevation (high, low) was included as a fixed effect to test for differences between high and low sites. Site and winter year (e.g. Winter 2000-2001: 1, 2001-2002: 2,…2015-2016:16) were included as random intercepts. GLMMs were fit using the statistical package *lme4* (Bates *et al*., 2015) in *R* (version 3.4.4). We

tested for significance of elevation using a Wald χ^2 test (Bolker *et al*., 2008) from the statistical package *car* (Fox *et al*., 2017).

The effect of elevation (GLMM: Low elevation ± SE = -0.99 ± 0.17, χ^2= 32.36, P<0.001) and day (GLMM: day^2 ± SE = -3.67 ± 0.83, χ^2= 19.44,P<0.001) were significant using the 500 m buffers, and using the 1500 m buffers (GLMM: Low elevation ± SE = -0.79 ± 0.14 , χ^2= 31.00,P<0.001) ; GLMM: day^2 ± SE = -4.88 ± 0.48, χ^2= 104.32, P<0.001). Therefore, there is greater snow cover at high elevation sites than low elevation sites during the winter.

S1.2. The number of birds captured from high and low sites during the winter (February 15, 2016 to March 1, 2016) and fall (October 7, 2016 to November 22, 2016) captures. Number of birds is split by sex (male, female), age (juvenile: < 1 year old, adult: >1 year old), and the number of individuals housed in the aviary in groups of two, three, four, and five birds. One low elevation bird from winter capture died during habituation. Two low elevation birds from the winter capture, two low elevation birds from the fall capture and one high elevation bird from the fall capture had faulty tags and were removed from the analysis.

	Sex		Age		Group Size			
	Male	Female	Juvenile	Adult	2	3	4	5
Winter								
High	12	10	19	3	0	3	9	10
Low	13	1	8	6	1	0	3	10
Fall								
High	3	3	5	1	0	0	6	0
Low	8	2	6	4	0	0	10	0

S3. Models within <Δ2 AICc of the top model for initial reversal learning accuracy (1: correct, 0: incorrect) of high elevation sites, degrees of freedom, AICc, Δi, and Akaike weights.

Candidate Models	df	AICc	Δi	ωi
Season+Trial+(1|ID)	4	1199.46	0.00	0.36
Season+Sex+Trial+(1|ID)	5	1200.48	1.02	0.21
Season+Site+Trial+(1|ID)	5	1201.08	1.61	0.16
Site+Trial+(1|ID)	4	1201.42	1.95	0.13
Age+Season+Trial+(1|ID)	5	1201.43	1.97	0.13

S4. Summary of results of model averaged coefficients for initial reversal learning accuracy (1: correct, 0: incorrect) at high elevation sites, after model averaging. Estimates are from the full average of five models <Δ2 AIC of the top model. Full model: Correct/Incorrect = Sex + Age + Season + Site+ Trial + (1|ID).

Parameter	Estimate	Unconditional SE	Confidence interval	Relative importance	*P*
Intercept	0.43	0.30	(-0.16, 1.01)		0.16
Season (Fall)	-0.67	0.42	(-1.46, -0.09)	0.87	0.11
Trial	2.64	0.36	(1.94, 3.34)	1.00	<0.001
Sex (Female)	0.07	0.19	(-0.29, 0.90)	0.21	0.73
Site (Cap Sour)	0.13	0.31	(-0.39, 1.28)	0.29	0.67
Age	-0.02	0.26	(-1.54, 1.21)	0.13	0.93

S5. Models within <Δ2 AICc of the top model for initial reversal learning accuracy (1: correct 0: incorrect) of low elevation sites, degrees of freedom, AICc, Δi, and Akaike weights.

Candidate Models	df	AICc	Δi	ωi
Sex+Trial+(1|ID)	4	716.71	0.00	0.49
Season+Sex+Trial+(1|ID)	5	718.01	1.30	0.26
Age+Sex+Trial+(1|ID)	5	718.07	1.36	0.25

S6. Summary of results of model averaged coefficients for initial reversal learning accuracy (1: correct, 0: incorrect) at low elevation sites, after model averaging. The estimates are from the full average of three models <Δ2 AIC of the top model. Site was not included in the top models. Full model: Correct/Incorrect = Sex + Age + Season + Site+ Trial + (1|ID).

Parameter	Estimate	Unconditional SE	Confidence interval	Relative importance	*P*
Intercept	0.27	0.41	(-0.53, 1.06)		0.51
Sex (Female)	2.17	1.02	(0.16, 4.18)	1.00	0.034
Trial	3.77	0.54	(2.70 4.84)	1.00	<0.001
Season (Fall)	0.14	0.41	(-0.72, 1.82)	0.26	0.73
Age (Adult)	0.14	0.43	(-0.78 1.94)	0.25	0.74

S7. Models within <Δ2 AICc of the top model for serial reversal learning accuracy (1: correct 0: incorrect) including the Elevation*Reversal Number interaction, degrees of freedom, AICc, Δi, and Akaike weights.

Candidate Models	df	AICc	Δi	ωi
Elevation+Season+Reversal Number+Trial+(1+Reversal Number|ID)	8	24091.56	0.00	0.18
Elevation+Reversal Number+Trial+(1+Reversal Number|ID)	7	24091.75	0.18	0.16
Elevation+Season+Group Size+Reversal Number+Trial+(1+Reversal Number|ID)	9	24091.99	0.42	0.15
Elevation+Season+Reversal Number+Trial+Elevation*Reversal Number+(1+Reversal Number|ID)	9	24092.54	0.98	0.11
Elevation+Reversal Number+Trial+Elevation*Reversal Number+(1+Reversal Number|ID)	8	24092.73	1.16	0.10
Elevation+Season+Group Size+Reversal Number+Trial+Elevation*Reversal Number+(1+Reversal Number|ID)	10	24093.01	1.45	0.09
Elevation+Group Size+Reversal Number+Trial+(1+Reversal Numberr|ID)	8	24093.34	1.77	0.07
Elevation+Season+Sex+Reversal Number+Trial+(1+Reversal Number|ID)	9	24093.54	1.98	0.07

Age+Elevation+Season+Reversal Number+Trial+(1+Reversal Number|ID)	9	24093.55	1.98	0.07

S8. Summary of results of model averaged coefficients for serial reversal learning accuracy (1: correct 0: incorrect) including the Elevation*Reversal Number after model averaging: effects of each parameter on the accuracy (1: correct, 0: incorrect) of choice during reversals 1-18. The estimates are from the full average of nine models <Δ2 AIC of the top model. Full model: Correct/Incorrect = Group Size + Age + Sex + Season + Trial + Elevation * Reversal Number + (1+Reversal Number|ID).

Parameter	Estimate	Unconditional SE	Confidence interval	Relative importance	*P*
Intercept	0.11	0.07	(-0.02, 0.24)		0.09
Elevation (Low)	0.24	0.08	(0.08, 0.41)	1.00	0.004
Season (Fall)	-0.09	0.09	(-0.30, 0.03)	0.66	0.33
Reversal Number	0.56	0.17	(0.22, 0.90)	1.00	0.001
Trial	3.40	0.11	(3.19, 3.61)	1.00	<0.001
Group Size	-0.06	0.14	(-0.56, 0.16)	0.31	0.66
Elevation * Reversal Number	-0.09	0.21	(-0.86, 0.27)	0.30	0.67
Sex (Female)	-0.001	0.02	(-0.18, 0.16)	0.07	0.97
Age (Adult)	0.001	0.03	(-0.19, 0.22)	0.07	0.97

References

Amy, M., van Oers, K., & Naguib, M. (2012). Worms under cover: relationships between performance in learning tasks and personality in great tits (*Parus major*). *Animal Cognition*, *15*, 763–770.

Ashton, B. J., Ridley, A. R., Edwards, E. K., & Thornton, A. (2018). Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature*, *554*, 364–367.

Bartoń, K. (2018). Multi-model inference. R package version 1.40.4.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J Stat Softw.*, *67*, 1–48.

Bebus, S. E., Small, T. W., Jones, B. C., Elderbrock, E. K., & Schoech, S. J. (2016). Associative learning is inversely related to reversal learning and varies with nestling corticosterone exposure. *Animal Behaviour*, *111*, 251–260.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2008). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution*, *24*(3), 127–135.

Bond, A. B., Kamil, A. C., & Balda, R. P. (2007). Serial reversal learning and the evolution of behavioral flexibility in three species of North American corvids (*Gymnorhinus cyanocephalus, Nucifraga columbiana, Aphelocoma californica*). *Journal of Comparative Psychology*, *121*(4), 372–379.

Boyle, A. W., Sandercock, B. K., & Martin, K. (2016). Patterns and drivers of intraspecific variation in avian life history along elevational gradients: a meta-analysis. *Biological Reviews*, *91*, 469–482.

Bruendl, A. C. (2017). *Parental Investment across an Altitudinal Gradient in Blue Tits (Cyanistes caeruleus)*. University of Exeter.

Calandreau, L., Bertin, A., Boissy, A., Arnould, C., Constantin, P., Desmedt, A., Guémené, D., Nowak, R., & Leterrier, C. (2011). Effect of one week of stress on emotional reactivity and learning and memory performances in Japanese quail. *Behavioural Brain Research*, *217*, 104–110.

Cauchoix, M., Hermer, E., Chaine, A. S., & Morand-Ferron, J. (2017). Cognition in the field: comparison of reversal learning performance in captive and wild passerines. *Scientific Reports*, *7*(12945), 1–10.

Chow, P. K. Y., Leaver, L. A., Wang, M., & Lea, S. E. G. (2017). Touch screen assays of behavioural flexibility and error characteristics in Eastern grey squirrels (*Sciurus carolinensis*). *Animal Cognition*.

Croston, R., Branch, C. L., Kozlovsky, D. Y., Roth II, T. C., LaDage, L. D., Freas, C. A., & Pravosudov, V. V. (2015). Potential mechanisms driving population variation in spatial memory and the hippocampus in food-caching chickadees. *Integrative and Comparative Biology*, *55*(3), 354–371.

Croston, R., Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Bridge, E. S., Parchman, T. L., & Pravosudov, V. V. (2017). Predictably harsh environment is associated with reduced cognitive flexibility in wild food-caching mountain chickadees. *Animal Behaviour*, *123*, 139–149.

Dukas, R. (2009). Learning mechanisms, ecology, and evolution. In R. Dukas & J. M. Ratcliff (Eds.), *Cognitive Ecology II* (pp. 7–26). Chicago: University of Chicago Press.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, *9*(10), 178–190.

Dunlap, A. S., & Stephens, D. W. (2016). Reliability, uncertainty, and costs in the evolution of animal learning. *Current Opinion in Behavioral Sciences*, *12*, 73–79.

East, M. L., & Hofer, H. (1985). The use of radio-tracking for monitoring Great Tit *Parus major* behaviour: a pilot study. *Ibis*, (128), 103–114.

Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-Bovy, G., Ellison, S., Firth, D., Friendly, M., Gregor, G., Graves, S., Heiberger, R., Laboissiere, R., Monette, G., Murdoch, Duncan, Nilsson, H., Ogle, D., Ripley, B., Venables, W., Zeileis, A., & R-Core. (2017). Companion to Applied Regression.

Freas, C. A., Bingman, K., LaDage, L. D., & Pravosudov, V. V. (2013). Untangling elevation-related differences in the hippocampus in food-caching mountain chickadees: The effect of a uniform captive environment. *Brain, Behavior and Evolution*, *82*, 199–209.

Freas, C. A., Roth II, T. C., LaDage, L. D., & Pravosudov, V. V. (2013). Hippocampal neuron soma size is associated with population differences in winter climate severity in food-caching chickadees. *Functional Ecology*, *27*, 1341–1349.

Gaalema, D. E. (2011). Visual discrimination and reversal learning in rough-necked monitor lizards (*Varanus rudicollis*). *Journal of Comparative Psychology*, *125*(2), 246–249.

Gonzalez, R. C., Behrend, E. R., & Bitterman, M. E. (1967). Reversal learning and forgetting in bird and fish. *Science.*, *158*, 519–521.

Gosler, A. G. (1993). *The great tit*. London: Hamlyn.

Grueber, C. E., Nakagawa, S., Laws, R. J., & Jamieson, I. G. (2011). Multimodel inference in ecology and evolution: challenges and solutions. *Journal of Evolutionary Biology*, *24*, 699–711.

Hall, D. K., & Riggs, G. A. (2016). MODIS/Terra Snow Cover 8-Day L3 Global 500m Grid, Version 6 [September 2000-December 2016].

Hassett, T. C., & Hampton, R. R. (2017). Change in the relative contributions of habit and working memory facilitates serial reversal learning expertise in rhesus monkeys. *Animal Cognition*, *20*(3), 485–497.

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: an updated perspective. *Neuroscience*, *345*, 12–26.

Janmaat, K. R. L., Boesch, C., Byrne, R., Chapman, C. A., Goné Bi, Z. B., Head, J. S., Robbins, M. M., Wrangham, R. W., & Polansky, L. (2016). Spatio-temporal complexity of chimpanzee food: how cognitive adaptations can counteract the ephemeral nature of ripe fruit. *American Journal of Primatology*, *78*(6), 626–645.

Körner, C. (2007). The use of "altitude" in ecological research. *TRENDS in Ecology and Evolution*, *22*(11), 569–574.

Kotrschal, A., & Taborsky, B. (2010). Environmental change enhances cognitive abilities in fish. *PLoS Biology*, *8*(4).

Kozlovsky, D. Y., Weissgerber, E. A., & Pravosudov, V. V. (2017). What makes specialized food-caching mountain chickadees successful city slickers? *Proceedings of the Royal Society B*, *284*, 20162613.

Krebs, J. R. (1971). Territory and breeding density in the great tit, *Parus major* L. *Ecology*, *52*(1), 2–22.

Krebs, J. R. (1990). Food-storing birds: adaptive specialization in brain and behaviour? *Phil. Trans. R. Soc. Lond. B*, *329*, 153–160.

Krištín, A., & Kaňuch, P. (2017). Stay or go? Strong winter feeding site fidelity in small woodland passerines revealed by a homing experiment. *Journal of Ornithology*, *158*(1), 53–61.

Liu, Y., Day, L. B., Summers, K., & Burmeister, S. S. (2016). Learning to learn: advanced behavioural flexibility in a poison frog. *Animal Behaviour*, *111*, 167–172.

López-Moreno, J. I., & Nogués-Bravo, D. (2005). A generalized additive model for the spatial distribution of snowpack in the Spanish Pyrenees. *Hydrological Processes*, *19*, 3167–3176.

Mery, F., & Burns, J. G. (2010). Behavioural plasticity: an interaction between evolution and experience. *Evolutionary Ecology*, *24*, 571–583.

Milton, K. (1981). Distribution patterns of tropical plant foods as an evolutionary stimulus to primate mental development. *American Anthropologist*, *83*(3), 534–548.

Morand-Ferron, J., Hamblin, S., Cole, E. F., Aplin, L. M., & Quinn, J. L. (2015). Taking the operant paradigm into the field: associative learning in wild great tits. *PloS ONE*, *10*(8), e0133821.

Ninot, J. M., Carrillo, E., & Ferré, A. (2017). Chapter 8 The Pyrenees. In J. Loidi (Ed.), *The Vegetation of the Iberian Peninsula* (Vol. 12, pp. 323–366). Cham: Springer International Publishing.

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

Pausas, J. G., & Austin, M. P. (2001). Patterns of plant species richness in relation to different environments : an appraisal. *Journal of Vegetation Science*, *12*, 153–166.

Pausas, J. G., Carreras, J., Ferré, A., & Font, X. (2003). Coarse-scale plant species richness in relation to environmental heterogeneity. *Journal of Vegetation Science*, *14*, 661–668.

Pravosudov, V. V., & Roth II, T. C. (2013). Cognitive ecology of food hoarding: the evolution of spatial memory and the hippocampus. *Annual Review of Ecology, Evolution, and Systematics*, *44*, 173–193.

Pravosudov, V. V., Roth, T. C., LaDage, L. D., & Freas, C. A. (2015). Environmental influences on spatial memory and the hippocampus in food-caching chickadees. *Comparative Cognition & Behavior Reviews*, *10*, 25–43.

R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rayburn-Reeves, R. M., Laude, J. R., & Zentall, T. R. (2013). Pigeons show near-optimal win-stay/lose-shift performance on a simultaneous-discrimination, midsession reversal task with short intertrial intervals. *Behavioural Processes*, *92*, 65–70.

Roth II, T. C., Chevalier, D. M., LaDage, L. D., & Pravosudov, V. V. (2013). Variation in hippocampal glial cell numbers in food-caching birds from different climates. *Developmental Neurobiology*, *73*, 480–485.

Roth II, T. C., & Pravosudov, V. V. (2009). Hippocampal volumes and neuron numbers increase along a gradient of environmental harshness: a large-scale comparison. *Proceedings of the Royal Society B*, *276*, 401–405.

Salle, L. (2016). *Altitudinal shift in life history strategies of a hole-nesting passerine*. University of Exeter.

Sayol, F., Maspons, J., Lapiedra, O., Iwaniuk, A. N., Székely, T., & Sol, D. (2016). Environmental variation and the evolution of large brains in birds. *Nature Communications*, *7*.

Schielzeth, H., & Forstmeier, W. (2008). Conclusions beyond support: overconfident estimates in mixed models. *Behavioral Ecology*, *20*, 416–420.

Sherry, D. F. (1998). The ecology and neurobiology of spatial memory. In R. Dukas (Ed.), *Cognitive Ecology* (1st ed., pp. 261–296). Chicago: University of Chicago Press.

Shettleworth, S. J. (2010). Discrimination, classification and concepts. In *Cognition, evolution and behaviour* (2nd ed., pp. 186–188). New York: Oxford University Press.

Sol, D., Lefebvre, L., & Rodríguez-Teijeiro, J. D. (2005). Brain size, innovative propensity and migratory behaviour in temperate Palaearctic birds. *Proceedings of the Royal Society B*, *272*, 1433–1441.

Strang, C. G., & Sherry, D. F. (2014). Serial reversal learning in bumblebees (*Bombus impatiens*). *Animal Cognition*, *17*, 723–734.

Svensson, L. (1992). *Identification guide to european passerines* (4th ed.). British Trust for Ornithology.

Tebbich, S., Taborsky, M., Fessl, B., & Dvorak, M. (2002). The ecology of tool-use in the woodpecker finch (*Cactospiza pallida*). *Ecology Letters*, *5*, 656–664.

Tebbich, S., & Teschke, I. (2014). Coping with uncertainty: woodpecker finches (*Cactospiza pallida*) from an unpredictable habitat are more flexible than birds from a stable habitat. *PLoS ONE*, *9*(3).

Teschke, I., Cartmill, E. A., Stankewitz, S., & Tebbich, S. (2011). Sometimes tool use is not the key: no evidence for cognitive adaptive specializations in tool-using woodpecker finches. *Animal Behaviour*, *82*, 945–956.

Thai, C. A., Zhang, Y., & Howland, J. G. (2013). Effects of acute restraint stress on set-shifting and reversal learning in male rats. *Cognitive, Affective, & Behavioral Neuroscience*, *13*, 164–173.

Titulaer, M., van Oers, K., & Naguib, M. (2012). Personality affects learning performance in difficult tasks in a sex-dependent way. *Animal Behaviour*, *83*, 723–730.

Chapter 2: *Great tits who remember more accurately have difficulty forgetting, but variation is not driven by environmental harshness*

Adapted from:

Hermer, E., Murphy, B., Chaine, A. S. & Ferron, J. M. (2021). Great tits who remember more accurately have difficulty forgetting, but variation is not driven by environmental harshness. *Scientific Reports* 1–11.

Introduction

There is growing evidence that wild animals differ in their ability to learn and retain information, these differences are partly heritable (Boogert, Madden, Morand-Ferron, & Thornton, 2018; Croston, Branch, Kozlovsky, Dukas, & Pravosudov, 2015; Langley *et al.*, 2020), and they can impact fitness (Benedict *et al.*, 2020; Cauchoix & Chaine, 2016; Shaw, MacKinlay, Clayton, & Burns, 2019; Sonnenberg, Branch, Pitera, Bridge, & Pravosudov, 2019). However, which cognitive abilities are beneficial and in which contexts they are beneficial has been examined in a limited number of species. Spatial memory, or the ability to memorize where objects are in space, is utilized by animals to remember where food sources are located (e.g., insects: Collett, Chittka, & Collett, 2013; birds: Hampton & Shettleworth, 1996; primates: Janmaat *et al.*, 2016; lizards: LaDage, Roth, Cerjanic, Sinervo, & Pravosudov, 2012). It has been hypothesized that as environments change, the ability to accurately remember where previously available food sources were in space could allow animals to more quickly return to these sources instead of enduring the cost of searching for a new patch as food availability decreases (Janmaat *et al.*, 2016; Milton, 1981; *Shaw et al.*, 2019; Thornton & Boogert, 2019). Therefore, environments where food availability fluctuates, such as seasonal habitats, may drive selection for more accurate spatial memory (Pravosudov & Clayton, 2002; Sonnenberg *et al*., 2019). This hypothesized relationship between seasonally fluctuating food availability and cognition is termed the harsh environment hypothesis (Pravosudov & Clayton, 2002; Sonnenberg *et al.*, 2019).

Both latitude and elevation have been used to investigate the harsh environment hypothesis, as increases in latitude and elevation are related to more snow cover, lower

temperatures, and greater seasonality, which leads to greater food scarcity and variability, as well as higher metabolic costs during winter (Boyle, Sandercock, & Martin, 2016; Hermer, Cauchoix, Chaine, & Morand-Ferron, 2018; Morand-Ferron, Hermer, Jones, & Thompson, 2019; Roth II & Pravosudov, 2009; but see: Körner, 2007). Along these two gradients, the harsh environment hypothesis has received support in winter resident, scatter hoarding birds that store food in multiple areas and use spatial memory to return to these caches when food is scarce during the winter (Roth II, LaDage, & Pravosudov, 2010; Tello-Ramos, Branch, Kozlovsky, Pitera, & Pravosudov, 2018). Indeed, spatial memory accuracy, and the size and neuron density of the primary brain structure responsible for spatial memory (i.e., hippocampus) increase along gradients of elevation and latitude (Roth II *et al.*, 2010; Tello-Ramos, Branch, Kozlovsky, *et al*., 2018). Moreover, juveniles with more accurate spatial memory (i.e., make fewer errors on a spatial task) had greater overwinter survival, providing evidence that decreased overwinter food availability may select for more accurate spatial memory in scatter-hoarders (Sonnenberg *et al*., 2019). Spatial memory could also aid non-scatter hoarders foraging in a harsh environment (Janmaat *et al.*, 2016; Milton, 1981; Shaw *et al.*, 2019; Thornton & Boogert, 2019) but no study, to our knowledge, has investigated whether increased environmental harshness is linked to increased spatial memory ability in non-hoarders.

Although spatial memory can contribute to foraging success, increased investment in memory also has potential costs. One hypothesized cost is proactive interference, with more accurate memories being more difficult to forget and interfering with the formation of new memories (Gonzalez, Behrend, & Bitterman, 1967; Tello-Ramos *et al.*, 2018; but see: Strang & Sherry, 2014). Proactive interference can lead to repeating incorrect responses instead of flexibly changing them as the situation changes or new information is presented (e.g., humans: Herszage

& Censor, 2018, other vertebrates: Gonzalez *et al.*, 1967; Missaire *et al.*, 2017; Miyashita, Nakajima, & Imada, 2000; Squier, 1969, and invertebrates: Bublitz, Weinhold, Strobel, Dehnhardt, & Hanke, 2017; Chittka, 1998; but see: Strang & Sherry, 2014). For example, an individual may repeatedly return to an empty food patch instead of learning the location of a new patch as their previous memory is interfering with their ability to learn new information. However, evidence for a correlation between increased spatial memory accuracy and increased proactive interference is not clear. Lab studies have compared rodents and birds whose neurobiology has been modified to those who were not modified (e.g., drug intake: Chrobak, Hinman, & Sabolek, 2008; gene expression: Joseph *et al.*, 2015; Malleret *et al.*, 2010; cannabinoid receptor blockage: Shiflett, Rankin, Tomaszycki, & DeVoogd, 2004; nutrient uptake: Meck & Williams, 1999) and found concurrent, positive changes in spatial memory and proactive interference (Chrobak *et al.*, 2008; Joseph *et al.*, 2015; Malleret *et al.*, 2010; but see: Meck & Williams, 1999). Comparative studies have found that spatial memory accuracy is generally higher in scatter hoarders as compared to non-scatter hoarders (Clayton & Krebs, 1994; Hampton & Shettleworth, 1996; McGregor & Healy, 1999; but see: Healy, 1995; Healy & Krebs, 1992) but scatter hoarders express less, not more, proactive interference on spatial tasks than non-scatter hoarders (Clayton & Krebs, 1994; Hampton & Shettleworth, 1996; but see: Hampton, Shettleworth, & Westwood, 1998). However, two studies of mountain chickadees (*Poecile gambeli*) in the wild found that high elevation chickadees had more accurate spatial memory, but committed a greater number of errors on a previously rewarded feeder compared to low elevation birds (Croston *et al.*, 2017; Tello-Ramos, Branch, Pitera, *et al.*, 2018), In one of these studies, it was also found that the individual's mean number of errors was positively associated with the tendency to return to the previously rewarded feeder (Croston *et al.*, 2017).

Therefore, the evidence showing a positive correlation between spatial memory and proactive interference is not found in between-species comparisons, and only found at the phenotypic level in within-species comparisons. These phenotypic correlations may only be evidence that proactive interference increases if an individual's spatial memory increases, and not that individuals with more accurate spatial memory, on average, have greater proactive interference compared to other individuals, otherwise known as the among-individual correlation (Careau & Wilson, 2017; Niemelä & Dingemanse, 2018). There is still no direct empirical evidence from behavioural studies in wild animal populations for an among-individual correlation between spatial memory and proactive interference that would be indicative of a trade-off.

In this study, we compared the performance of non-scatter hoarding, great tits (*Parus major*), that feed on patches of seeds during the winter (Gosler, 1993), from several high and low elevation sites that differ in harshness (Bründl *et al*., 2019; Hermer *et al*., 2018; Lejeune *et al*., 2019) on a spatial memory task and a single spatial reversal task designed to measure proactive interference (Tello-Ramos, Branch, Kozlovsky, *et al.*, 2018; Figure 2.1). The spatial memory portion of this task consists of an information stage where birds are shown the location of a food reward in one tree out of three (Thompson & Morand-Ferron, 2019). Their memory for the location of this reward is then tested over seven trials with memory accuracy being counted as the number of errors before finding the food reward (Roth II, LaDage, Freas, & Pravosudov, 2012; Thompson & Morand-Ferron, 2019). The reversal task to measure proactive interference consists of a single information stage wherein the food reward was moved to a previously unrewarded tree, and five trials to measure the ratio of errors made on the previously rewarded tree over the other two trees (Croston *et al*., 2017). We predict that high elevation birds will commit fewer errors than low elevation birds on a spatial memory task if spatial memory aids in

foraging in a harsh environment. We predict that if there is a relationship between accurate memory and increased proactive interference, during the reversal, high elevation birds will also commit a greater ratio of errors on the previously rewarded tree than low elevation birds (i.e., greater proactive interference). Finally, if accurate memory correlates with greater proactive interference, we predict that performance on both tasks will negatively co-vary at the among-individual level (Careau & Wilson, 2017). In order to assess this correlation, we also quantify repeatability (i.e., consistency of individual differences) of accuracy during the spatial memory and proactive interference trials (Griffin, Guillette, & Healy, 2015).

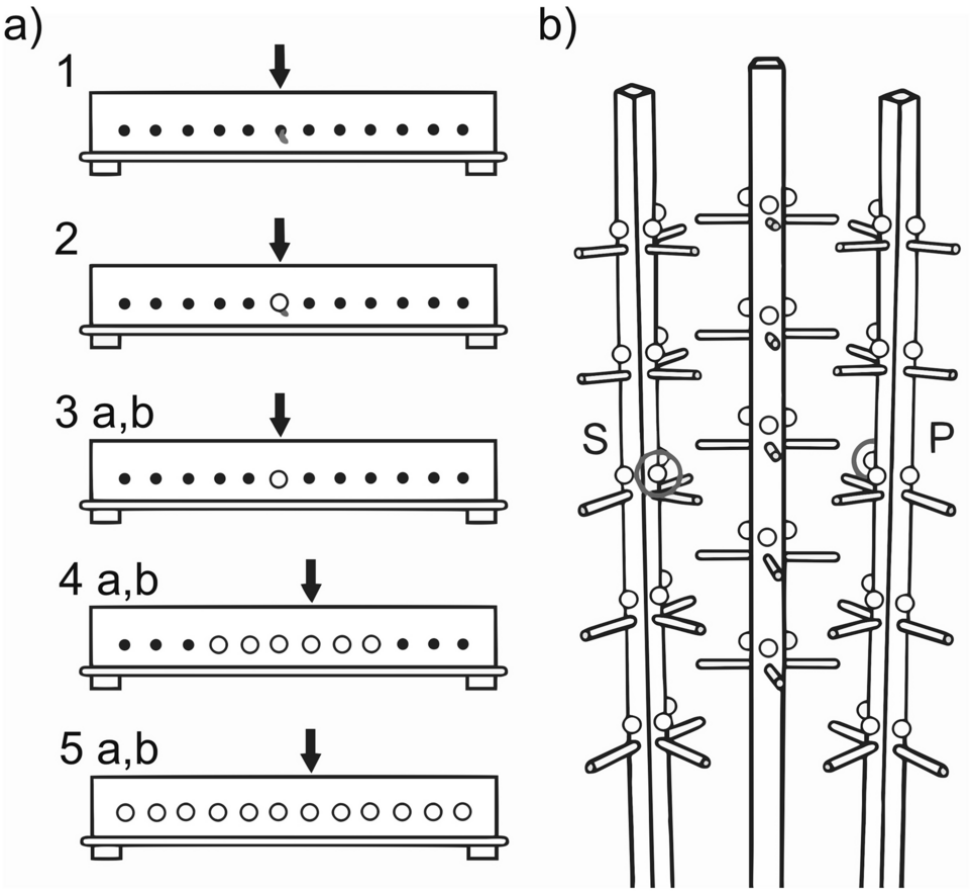


Figure 2.1. **a.** The Five Motor Training Stages (eight stages total) as they appeared on the training board. The black dots represent empty holes. The white dots represent holes covered by white, acrylic 0.5 inch pompoms. The arrows denote the location of the mealworm (*Tenebrio molitor*) reward. The mealworm reward is pictured uncovered in stage 1, half covered in stage 2, and completely covered by a pompom after stage 3. In each of the stages, the bird had to consume the mealworm reward from the training panel to progress to the next stage. If it failed the stage (i.e., did not consume the meal worm within 15 min), it regressed to the previous stage. **b**. Location of the mealworm reward during the spatial memory trials (S) and proactive interference trials (P). The location of the worm alternated sides across cages, and only one side is pictured here. It proceeded with the worm having no pompom over it (information stage 1). If the bird successfully retrieved the worm, another worm was placed in the hole (information stage

2–1) and two mealworms were placed in the hole if the initial worm was not retrieved (information stage 2–2). The bird continued onto information stage 3 if the mealworms were consumed or stayed at information stage 2–2 if they were not. For information stage 3, the bird retrieved a mealworm half covered by a pompom, then retrieved a mealworm completely covered by a single pompom for information stages 4–5. Every hole was covered with a pompom during the spatial memory and proactive interference trials.

Methods

Capture and housing

Wild great tits (*Parus major*) were captured from three high (800–900 m; n = 34; n = 17 males, n = 17 females, n = 22 juveniles, n = 12 adults) and three low elevation sites (400–500 m; n = 36; n = 22 males, n = 14 females, n = 24 juveniles, n = 12 adults) near the Experimental and Theoretical Ecology Research Station in Moulis, France between October 24, 2017, and February 25, 2018 (Supplementary S2.1). High elevation sites are characterized by longer weekly snow cover, and lower temperatures relative to low elevation sites (Bründl *et al.*, 2020, 2019; Hermer *et al*., 2018; Lejeune *et al*., 2019). Great tits were captured in batches of 4–6 individuals, using mist nests and marked with a CRBPO (Centre de Recherches sur la Biologie des Populations d'Oiseaux) metal band. We used plumage to sex (male/female) and age (juvenile/adult) great tits (Svensson, 1992). Birds that experienced a previous cognitive test were released and not used in testing. Birds were transported to outdoor aviaries in cloth bags and housed individually (1 × 4 × 3 m) in every second cage to visually isolate them. High and low elevation birds were housed in the same cages, on the same side of the aviary, but the experimenters were not blind to the bird's identity during placement. Each aviary contained foliage for cover in the non-testing area, two roosting boxes, and two horizontal perches between the foliage and testing area.

Acclimatization

For six days after capture, birds were acclimated to the aviary and testing environment (Supplementary S2.2). The birds had access to ad libitum black oil sunflower seeds, fat balls, meal worms, and water. A small heated (25 °C) room inside of the aviary building was left open

and contained a second source of ad libitum food and water and constant light for the first two days to encourage feeding. Three un-baited testing trees and one un-baited motor training panel were in the aviary to acclimatize the birds to the testing devices. The testing trees, motor training board, and food were removed during the final night of acclimatization (day 6). To reduce stress to the birds caused by recapture, weight was not measured during acclimatization or during testing. All training and testing occurred concurrently for each batch of birds.

Motor training

On day 7, the birds were trained to approach and remove a pompom (0.5 inch diameter white acrylic ball) from a 0.5 cm hole and retrieve a mealworm reward hidden underneath the pompom across eight stages on a training board (Fig. 2.1; Supplementary S2.2). 15-min motor training sessions occurred one after another from 08 h to 11 h 30. Ad libitum food was given from 11 h 30 to 12 h 30. Training resumed until 16 h 00, or until the birds passed motor training. Ad libitum food was returned to the cage and the motor training panel was removed afterwards. Fall and Winter birds underwent a slightly different motor training protocol due to slight differences in methodology. When fall birds failed to pass stage 2, they did not revert to stage 1 but rather stayed at stage 2 (low: $n = 3$, high: $n = 1$). Also, one bird experienced 5a twice after 3b, was returned to 4a and then proceeded to pass 4a–5b. We included these extra trials in the motor training analysis. Removing this individual (bird id: 123) did not qualitatively change the results. All birds that successfully passed motor training were kept in the analysis and began information trials on day 8 (Supplementary S2.3, S2.4).

Information trial, spatial memory, and proactive interference

The night of day 7, the three testing trees were added back to the cage, and the food was removed (Fig. 2.1; Supplementary S2.2). On day 8 (07h00), birds had to complete five information trials before proceeding to the spatial memory task. During the information trials, the birds learned the rewarded location by repeatedly retrieving a worm from the same location on one of the testing trees: twice the worm was uncovered, once half covered, and twice completely covered by a pompom. All birds that consumed all the worms during the information trials were kept in the analysis and proceeded to spatial memory tests (Supplementary S2.3, S2.4).

The spatial memory task (Roth II *et al.*, 2012; Thompson & Morand-Ferron, 2019) started on day 8 at 10 h (Fig. 2.1; Supplementary S2.2). Before each spatial memory trial, birds were food deprived for 30-min. To decrease the usefulness of potential social cues from the experimenters, we mimed placing the worm into each hole by covering the hole with a hand and motioning as if placing the worm in the hole underneath before placing the pompom. Pompoms were placed into all of the trees' holes (n = 45) in the same order throughout trials, and a mealworm was placed into the same rewarded hole as in the information stage (Fig. 2.1). Birds had 1 h to find the worm and all pompoms pulled before finding the worm were considered errors (Pravosudov & Clayton, 2002; Thompson & Morand-Ferron, 2019). Ad libitum food was returned after the first spatial memory trial for 30-min. Two more spatial memory trials followed, with a 30-min deprivation period occurring in between. At the end of these two trials, ad libitum food was returned to the cages, and removed at night.

Four spatial memory trials occurred on day 9 (Supplementary S2.2). The fourth trial started at 08h00, and trial 5 followed after a 30-min deprivation. The rest of the trials followed the same schedule as the previous day. If the worm was not found during a trial, it was left in the rewarded hole after the pompoms were removed. Birds that did not consume the open mealworm were excluded from further trials (Supplementary S2.3, S2.4). If the worm was found, the intertrial interval was calculated from when the worm was found, and only the trial where the bird did not find the worm was excluded. If a bird failed to consume the worm twice during the spatial memory tests, we only kept trials up to the second missed trial as we assumed the bird was not motivated to complete the task.

Proactive interference trials followed the same protocol as the second day of the spatial memory task (day 10; Supplementary S2.2). However, the reward was now located on the tree opposite to the previously rewarded tree, in an inward facing hole (Fig. 2.1). The first trial was a single information stage followed by three proactive interference trials. Two more proactive interference trials followed on the next day (day 11). Errors made on the previously rewarded tree indicate that birds did not extinguish the positive association with the previously rewarded location, while errors on the other two trees are assumed to be due to the newly learned association with the currently rewarded tree, or exploration errors made to the never rewarded tree. Therefore, proactive interference was quantified as the ratio of errors made on the previously rewarded tree, over the errors made on the other two trees (Croston *et al*., 2017). Some birds became acclimated to the tester and would start removing pompoms before all pompoms were placed in the trees (n = 4 trials; Supplementary S2.4). We did not count these errors in our analysis, but results were qualitatively the same with or without these trials.

Video analysis

Data from the information trial, spatial memory and proactive interference trials were extracted using BORIS video analysis software by three observers (EH, AR, JH; Friard & Gamba, 2016). A blind procedure was used with observers watching muted videos labelled by dates or batch with no identifying information viewable on the screen. Intertrial interval was quantified as the time difference (minutes) between the moment when a bird found a mealworm in the previous trial, and when the experimenter left the cage after preparing it for the start of the next trial. Videos for some birds were lost (n = 2 birds); only their trials up to the missing videos were kept in the analysis. (Supplementary S2.3, S2.4). The number of errors for some trials exceeded the possible number of errors (>15 errors on a tree: n = 9/284 PI trials, >45 total errors: n = 1/284 PI trials). These extra errors were kept in the analysis as we assume the observer randomly overcounted errors across all high and low elevation birds, and only removing detectable overcounts would be artificially lowering only high error count videos.

Statistical analysis

We compared motor training speed between elevations by comparing the number of trials it took to consume a worm half covered by a pompom (stage 2; Fig. 2.1), as well as the number of trials to pass motor training using a non-paired Wilcoxon Rank Sum Test as the data distribution did not fit the assumption of normality. We also compared the proportion of high and low elevation birds that passed motor training using a Fisher's exact test.

We analyzed whether the birds had learned the location of the reward during the spatial memory trials by comparing the mean number of spatial errors to the mean number of errors

predicted by random searching using a Wilcoxon signed rank test (Tillé, Newman, & Healy, 1996). We then analyzed the relationship between elevation and spatial memory using a linear mixed model (LMM; lme4 1.1-25: Bates *et al*., 2015; lmerTest3.1-3: Kuznetsova, P B, & Christensen, 2017; R Version 4.0.3: R Core Team, 2020) with log transformed number of errors as the response to meet the assumption of normality of the residuals (Warton, Lyons, Stoklosa, & Ives, 2016). Elevation (high/low), age (juvenile/adult), sex (male/female), capture order (1–12), which tree the reward was on (left/right), trial number (1–7), video observer (EH, AR, JH) and inter-trial interval (minutes) were included as fixed effects. Bird ID was included as a random intercept. To control for the effect of capture site we included site as a random effect, but the model would not converge. Therefore, we ran a separate high and low elevation models and included site as a fixed effect. Site was not significant in either model and it was excluded from further analysis (Supplementary S2.5, S2.6).

We analyzed whether the birds experienced proactive interference by comparing the mean ratio of errors to the ratio of errors that we would expect given random sampling of the three trees [(1/3)/(2/3)] using a Wilcoxon signed rank test. We also compared trial 1 errors between the spatial memory task and proactive interference task using a one-tailed, paired sample t-test. We analyzed the relationship between elevation and proactive interference using an LMM with the log transformed ratio of errors to meet the assumption of normality of the residuals. We utilized the same fixed and random effects as the above model except trial number went from 1 to 5. The fit was singular with site as a random effect. Therefore, we ran the separate high and low elevation models and included site as a fixed effect. The high elevation model would not run with observer included and so it was dropped from the model. Site was not significant in either model and was excluded from further analysis (Supplementary S2.7, S2.8).

All continuous predictor variables were standardized by grand mean centering and dividing by 1 standard deviation. Assumptions of normality and homogeneity were visually assessed using histograms, Q-Q plots, and residual versus fitted plots, respectively. The analysis was not performed blind and sample sizes were not calculated a priori.

Adjusted repeatabilities (Wilson, 2018) were calculated using the same models as the LMMs without sex and age using rptR (rptR 0.9.22: Stoffel, Nakagawa, & Schielzeth, 2017). A multivariate mixed model was utilized to calculate the among-individual covariance between performance on the spatial memory and proactive interference measures (MCMCglmm 2.29: Hadfield, 2010). The log transformed number of errors from the spatial memory trials and the log transformed ratio of errors from the proactive interference trials were included as traits with Gaussian error structures. We included the same fixed effects as above. Bird ID was included as a random effect for both traits. Family was defined as 'Gaussian' and residual variance at the limit was set to 1. The random effect variance structure (*G*) used in the prior included a variance set to 1 and a degree of belief (nu) set to 0.002. Burn in was set to 20,000, the number of iterations was 420,000, thin was set to 100 (Houslay & Wilson, 2017). Convergence of the model was assessed by visual inspection of traces.

Ethics

Trapping and marking of wild great tits was performed under permits from the French ringing office (CRBPO, project 576; permit 13619). Capture and holding birds from the wild was approved by the Région Midi-Pyrenées (DIREN, n°2012-07) in the Moulis experimental aviaries (Préfecture de l'Ariège, institutional permit n°SA-12-MC-054; Préfecture de l'Ariège, Certificat de Capacite, n°09-321). This study was approved by the Animal Care Committee at the University of Ottawa (protocol: 1758). Testing complied with the ARRIVE Essential 10 guidelines (Kilkenny, Browne, Cuthill, Emerson, & Altman, 2010). All methods were performed in accordance with the relevant guidelines and regulations.

Results

Motor training

All high and low elevation birds approached the task and consumed a half-covered mealworm (stage 2; Fig. 2.1). There was no significant difference between elevations in the number of attempts to consume a half-covered mealworm (Mean ± s.d.: high: 1.32 ± 1.72; low: 1.33 ± 0.79; Wilcoxon rank sum test, high vs. low: $W = 531$, $P = 0.10$, $n = 70$). There was no significant difference between elevations in the number of stages required to successfully pass motor training (Mean ± s.d.: high: 11.3 ± 3.68; low: 11.2 ± 3.21; Wilcoxon rank sum test, high vs. low: $W = 537$, $P = 0.76$, $n = 67$). There was no significant difference between elevations in the proportion of birds that passed motor training (High: 33/34; low: 34/36; Fisher's exact test, contingency table: CI = [0.01, 10.43], odds ratio = 0.52, $P = 1$). Therefore, high and low elevation birds were both successfully trained to remove pompoms covering food rewards and exhibited no apparent difference in motivation to consume the mealworms.

Spatial memory

The birds made significantly fewer errors than expected by random searching (chance = 23 following negative hypergeometric distribution; One-tailed Wilcoxon signed-rank test, $N = 62$, mean ± s.d. = 12.48 ± 5.40, $P < 0.001$). This indicates the birds had learned the location of the reward. Trial number was significant and had a negative coefficient, indicating that birds improved in accuracy over trials. Elevation was non-significant (Table 2.1; Figure 2.2). This indicates that high and low elevation birds did not differ in their spatial memory accuracy.

Proactive interference

The mean ratio of errors to the previously rewarded tree (i.e., errors on the previously rewarded tree/rewarded and unrewarded tree) was significantly higher than expected by random sampling (chance = 0.5; One-tailed Wilcoxon signed-rank test, n = 59, mean ± s.d. = 0.86 ± 0.45, $P < 0.001$). The number of errors was significantly higher in the first proactive interference trial than in the first spatial memory trial (One-tailed paired t-test, $t = -4.80$, CI = − 6.31, mean of the difference = − 9.68, $P < 0.001$; n = 57), suggesting that the birds did show evidence for proactive interference. All fixed effects, including elevation, were non-significant (Table 2.2; Figure 2.3). This indicates that high and low elevation birds did not differ in their intensity of proactive interference.

Among-individual trade-off

The number of errors was significantly and moderately repeatable across spatial memory trials (R = 0.23 ± 0.05, CI = [0.14, 0.34], $P < 0.001$), as well as across proactive interference trials (R = 0.30 ± 0.07, CI = [0.17, 0.43], $P < 0.001$). There was strong evidence that the number of errors in the spatial memory task was negatively correlated to the ratio of errors in the proactive interference task at the among-individual level ($r_{ind} = -0.68 \pm 0.15$; 95% CI = [− 0.93, − 0.39]; Figure 2.4). This indicates that individuals who performed more accurately on the spatial memory task showed greater proactive interference.

Table 2.1. Predictors of the log transformed number of errors made by birds (n = 62; n = 423 trials) across seven spatial memory trials fitted with a linear mixed effect model with trial, capture order, intertrial interval (minutes), elevation (high/low), sex (male/female), age (juvenile/adult), rewarded side of the tree (left/right), and observer (EH, JH, AR) included as fixed effects.

Predictors	Estimate ± SE	F-statistic	*P*
Intercept	2.16 ± 0.26		
Trial	− 0.21 ± 0.03	38.95	< .0001
Elevation (Low)	− 0.14 ± 0.13	1.18	0.28
Sex (Male)	0.02 ± 0.12	0.03	0.856
Age (Juvenile)	− 0.05 ± 0.14	0.12	0.73
Capture order	0.24 ± 0.11	4.33	0.04
Intertrial interval	0.02 ± 0.033	0.37	0.545
Correct side (Right)	− 0.11 ± 0.12	0.87	0.36
Observer (EH)	0.45 ± 0.27	1.37	0.26
Observer (JH)	0.24 ± 0.26		

Bird ID was included as a random intercept.

Table 2.2. Predictors of the log transformed ratio of errors made by birds (n = 59; n = 284 trials) across five proactive interference trials fitted with a linear mixed effect model with trial, capture order, intertrial interval (minutes), and elevation (high/low), sex (male/female), age (juvenile/adult) rewarded side of the tree (left/right), and observer (EH, JH, AR) included as fixed effects.

Predictors	Estimate ± SE	F-statistic	*P*
Intercept	0.61 ± 0.08		
Trial	0.01 ± 0.02	0.58	0.45
Elevation (Low)	0.10 ± 0.06	2.79	0.10
Sex (Male)	− 0.01 ± 0.06	0.05	0.83
Age (Juvenile)	− 0.07 ± 0.07	1.12	0.30
Capture order	0.02 ± 0.03	0.68	0.41
Intertrial interval	0.01 ± 0.02	0.18	0.67
Correct side (Right)	− 0.08 ± 0.06	2.06	0.16
Observer (EH)	− 0.10 ± 0.15	0.31	0.74
Observer (JH)	0.03 ± 0.08		

Bird ID was included as a random intercept.

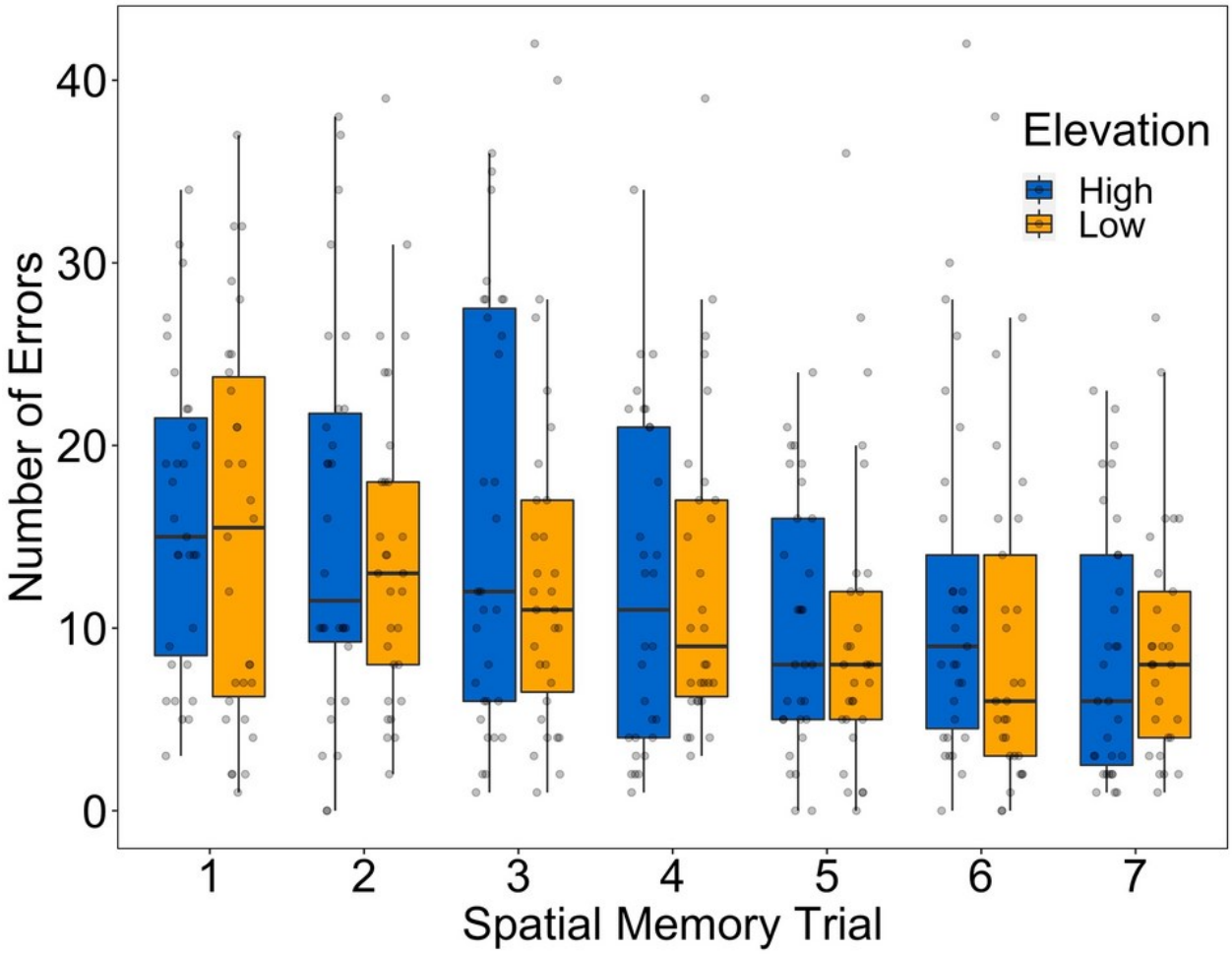


Figure 2.2. Boxplot of the untransformed number of spatial memory errors across spatial memory trials. High elevation birds are in blue, low elevation birds are in orange. Raw number of errors are plotted in grey.

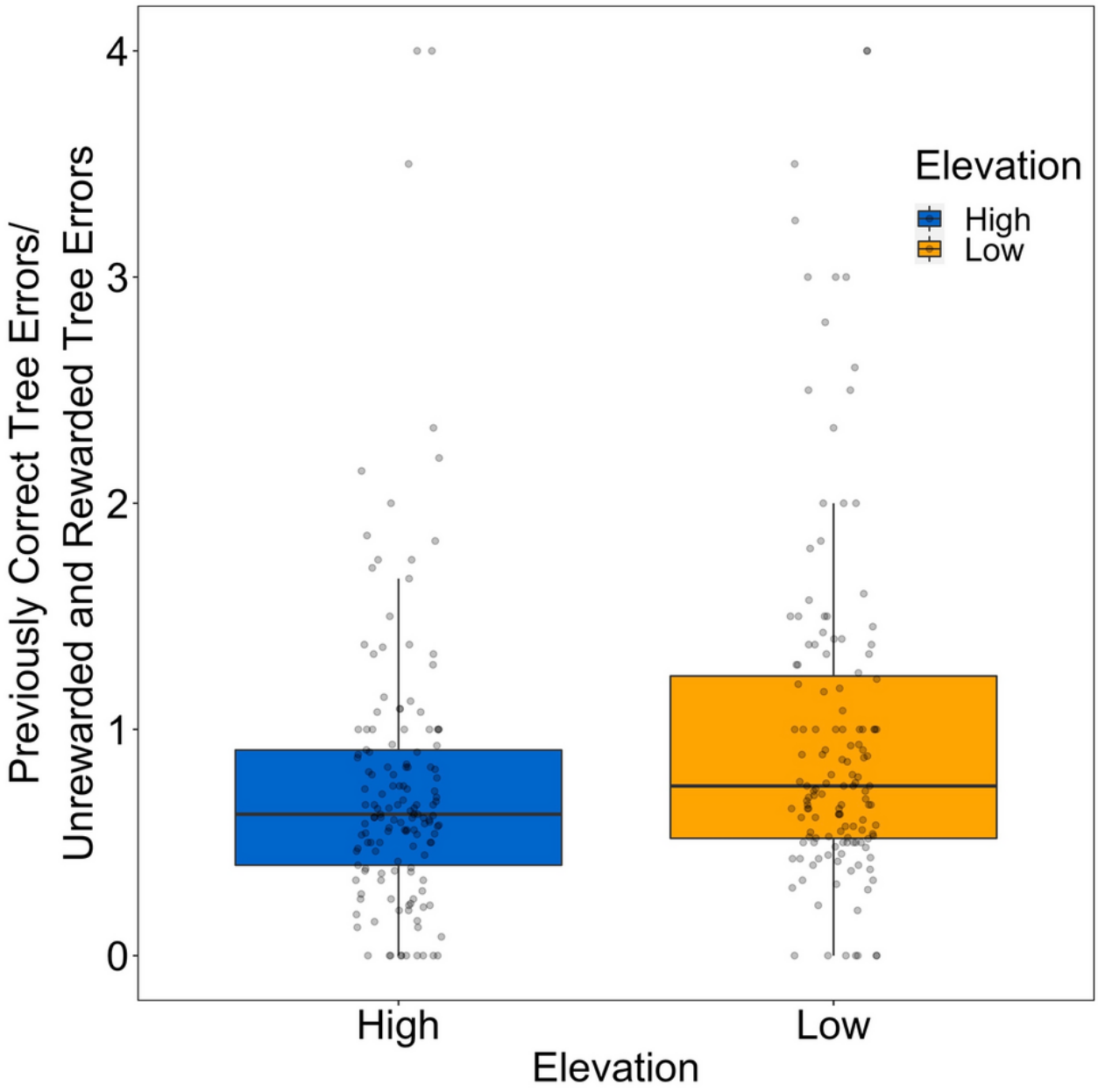


Figure 2.3. Boxplot of the untransformed previously correct tree errors/unrewarded and rewarded tree errors for high (blue) and low (orange) elevation birds. Raw number of errors are plotted in grey.

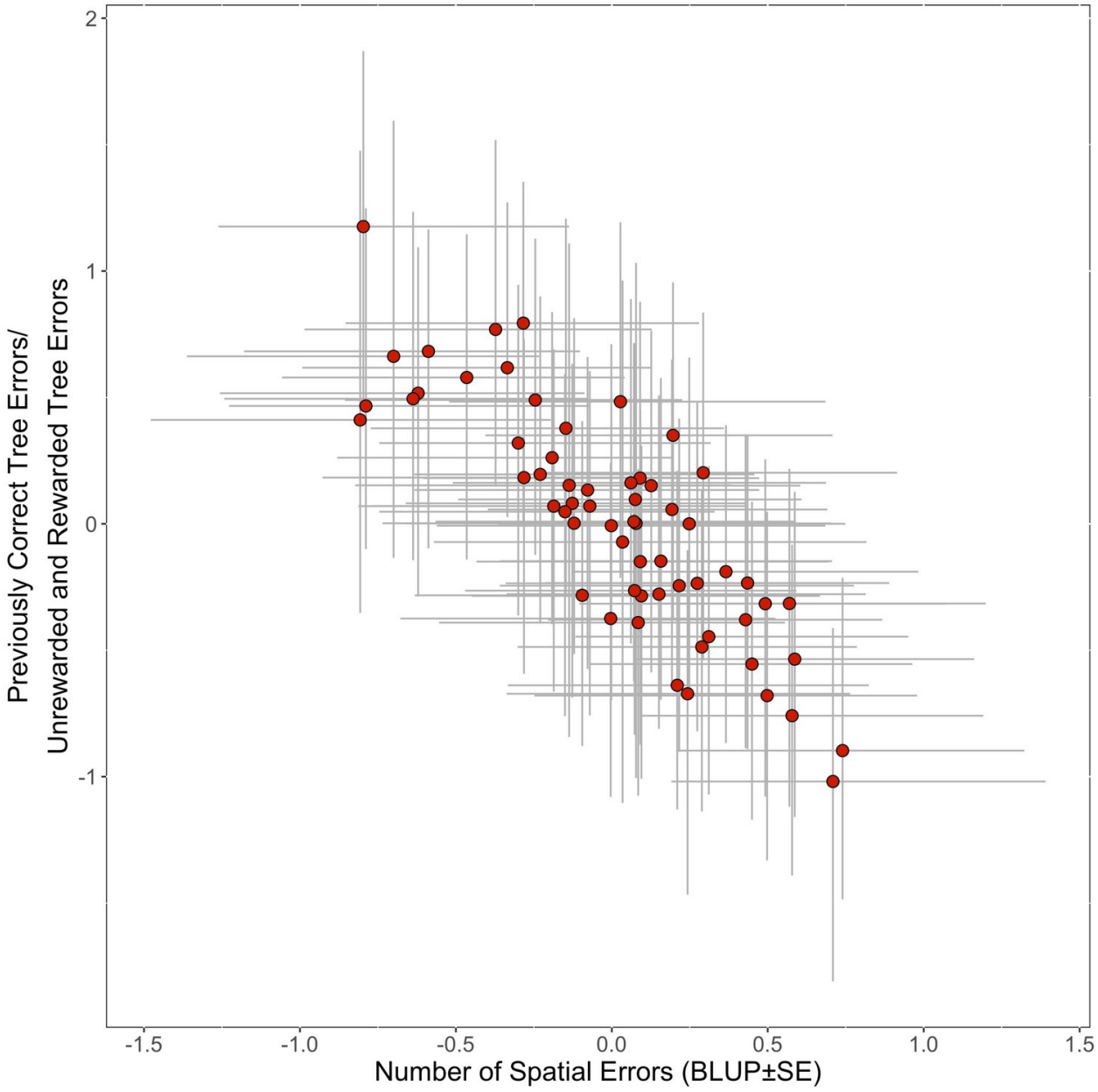


Figure 2.4. Among-individual correlation (red) between the standardized number of spatial errors and the standardized number of errors made to the previously rewarded tree/unrewarded and rewarded trees in great tits (*Parus major*) from high and low elevations. Individual deviations from the population mean are illustrated using best linear unbiased predictors (BLUP ± SE) associated with the random effect of bird identity.

Discussion

Although we know that there are individual differences in wild animal cognition, and these differences may affect fitness (Benedict *et al*., 2020; Cauchoix & Chaine, 2016; Shaw *et al.*, 2019; Sonnenberg *et al.*, 2019), we still do not fully understand how the environment may impact cognition in natural populations (Ashton, Thornton, & Ridley, 2018; Croston, Branch, Kozlovsky, Dukas, & Pravosudov, 2015; Thornton & Boogert, 2019). Harsh environments may be one driver of population differences in cognition (Morand-Ferron *et al.*, 2019; Pravosudov & Clayton, 2002). Accurate spatial memory could aid foraging in harsh environments, but accurate memories are predicted to inhibit the formation of similar, new memories, creating a trade-off between spatial memory performance and proactive interference (Tello-Ramos, Branch, Kozlovsky, *et al.*, 2018). We explored these relationships for the first time in a wild population of a non-hoarding species. We found no differences in spatial memory accuracy or proactive interference performance among great tits from low and high elevations which differ in harshness, but we did find that birds that were more accurate on a spatial task also had greater proactive interference. This among-individual correlation between performance in both tasks reveals a trade-off at the level at which selection can act.

Studies of the harsh environment hypothesis have consistently found that spatial memory is more accurate in scatter hoarders from high than low elevations (Tello-Ramos, Branch, Kozlovsky, *et al.*, 2018), and we predicted the same relationship in non-scatter hoarders. However, we did not find a significant difference in spatial memory accuracy between high and low elevation great tits. A first possible explanation is that the two elevations we sampled do not differ enough in harshness. However, this is unlikely as previous studies that sampled birds from

the same population and elevations found a significant cognitive difference (i.e., laboratory test of serial reversal learning: Hermer *et al.*, 2018), as well as differences in breeding phenology (Bründl *et al.*, 2020) and parental care (Lejeune *et al.*, 2019) between high and low elevation birds. Therefore, some environmental difference between high and low elevations seems to be driving behavioural differences. Second, the differences between our results and those from previous studies could be due to a difference in selection pressure between scatter hoarders and non-scatter hoarders. Scatter hoarders, specifically mountain and black-capped chickadees (*Poecile atricapillus*), from harsher environments have a higher propensity to cache food than scatter hoarders from less harsh habitats (Freas, LaDage, Roth II, & Pravosudov, 2012; Pravosudov & Clayton, 2002). Caching utilizes spatial memory and the greater need to remember caches creates strong selection for spatial memory as harshness increases (Benedict *et al.*, 2020; Croston, Branch, Kozlovsky, Roth II, *et al.*, 2015; Pravosudov & Clayton, 2002; Pravosudov & Roth II, 2013; Sonnenberg *et al.*, 2019). Non-scatter hoarders may still utilize accurate spatial memory when remembering and returning to food patches (Collett *et al.*, 2013; Hampton & Shettleworth, 1996; Janmaat *et al.*, 2016; LaDage *et al.*, 2012) but this need may be similar to that of scatter hoarders that need to find and return to the food when initially foraging for food to cache. Therefore, the adaptive value of accurate spatial memory may not be as high for non-scatter hoarders and any differences among elevations may be small. A third potential explanation is that greater spatial memory accuracy may also be just as helpful in low than high harshness environments. For example, our lower elevation field sites are characterized as having more diversity in food sources than high elevation (Lejeune *et al*., 2019). In areas of high diversity, it may be beneficial to remember where the high-quality food items are and return to the best food source (Janmaat *et al.*, 2016). Finally, it is possible that our spatial memory task

was not difficult enough to detect a difference between elevations. For instance, we used retention intervals of 1 h as well as overnight. Increasing the retention interval to weeks instead of hours or days may make the task difficult enough for even small differences in spatial memory across elevations to show (e.g., 17 days: Kozlovsky, Weissgerber, & Pravosudov, 2017). Testing for a correlation between spatial memory performance and over-winter survival may allow for a better understanding of spatial memory's usefulness to non-scatter hoarders (Benedict *et al.*, 2020).

We predicted that high elevation great tits should suffer greater proactive interference and make a larger proportion of errors on the previously correct tree compared with low elevation birds. Instead, and in accordance with a lack of spatial memory accuracy differences, we found no difference between high and low elevation birds in their proactive interference. Interestingly, in a previous study on great tits collected from the same population, we found that birds from high elevation performed less accurately on a serial spatial reversal learning task than low elevation great tits (Hermer *et al.*, 2018). We surmised that one potential explanation for this result was that spatial memory may have been more accurate in high elevation great tits, which would lead to greater spatial proactive interference and worse reversal performance, as found in the scatter hoarding system (Croston *et al.*, 2017; Tello-Ramos, Branch, Kozlovsky, *et al.*, 2018). Given that there is no difference in proactive interference between high and low elevation birds in the current study, we have preliminary evidence to rule out this explanation. Instead, our previous results may have been due to differences in win-stay/lose-shift rule learning between elevations, where the animal is not memorizing associations per se but is changing its choice when it is incorrect and maintaining that choice when it is correct (Izquierdo, Brigman, Radke, Rudebeck, & Holmes, 2017). As the great tits did not reach a single error switch in our serial

reversal learning task, we were unable to measure this (Hermer *et al.*, 2018). This explanation remains to be tested empirically.

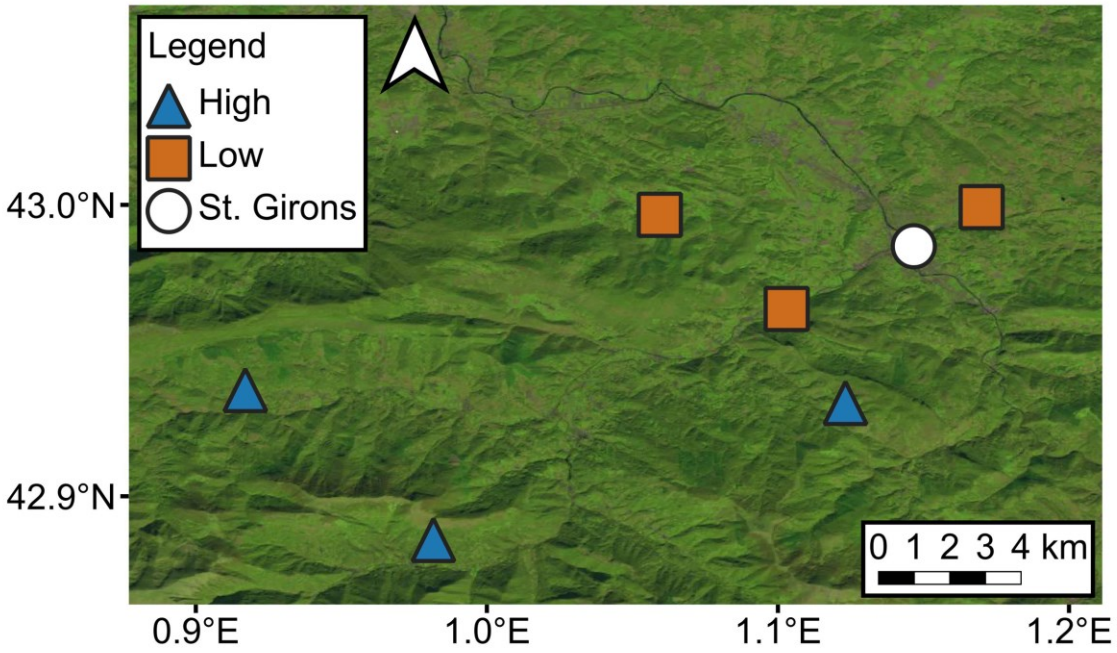
We found a positive among-individual correlation between spatial memory accuracy and proactive interference. This positive among-individual correlation indicates that individuals that have more accurate spatial memory on average, also have higher proactive interference on average (Niemelä & Dingemanse, 2018). This correlation would traditionally be assessed at the unpartitioned, phenotypic level, by collecting one measure of spatial memory, and one measure of proactive interference for multiple individuals. However, phenotypic correlations are influenced by within-individual variance, which reflects how two traits change with each other within the individual over repeated measurements. For example, if an individual great tit's spatial memory accuracy increases with state or age, a positive within-individual correlation would indicate that its proactive interference should also increase with state or age (Careau & Wilson, 2017; Niemelä & Dingemanse, 2018). To avoid this 'individual gambit', multiple measures of each test are used to partition variance to among-individual and within-individual levels, and directly assess among-individual correlation (Niemelä & Dingemanse, 2018). We found a positive among-individual correlation between spatial memory accuracy and proactive interference performance with great tits who, on average, made a lower number of errors during their spatial memory trials, also on average made a greater proportion of their errors on the tree that was previously rewarded during the spatial task. In other words, birds who remember well also have a difficult time forgetting and learning a new reward location. To our knowledge, this is the first examination of an among-individual correlation between spatial memory accuracy and proactive interference in wild animals.

Our among-individual correlation generally agrees with other lab experiments (Chrobak *et al.*, 2008; Joseph *et al.*, 2015; Malleret *et al.*, 2010) and field studies (Croston *et al.*, 2017; Tello-Ramos, Branch, Pitera *et al.*, 2018), that show evidence for a trade-off between spatial memory accuracy and proactive interference performance. However, our population comparisons did not match the results found in the scatter hoarding systems (Croston *et al.*, 2017; Tello-Ramos, Branch, Pitera, *et al.*, 2018). We believe this indicates that although there are no differences between high and low elevation great tits in either of these behaviours, there is preliminary evidence that this trade-off is present in the overall population. If a change in selection pressure occurs that leads to an increase in spatial memory accuracy in either high or low elevation great tits, their proactive interference could be expected to change in kind. However, we do caution that this result would be more robust with additional testing to see if this relationship holds. Our multiple measures for spatial memory and proactive interference came from the same cognitive task and may thus suffer from a lack of independence. In the future, a more robust test should alternate measuring spatial memory and proactive interference, ideally using a different experimental set-up the second time (i.e., contextual repeatability: Cauchoix *et al.*, 2018; e.g., spatial task and reversal using a set of automated feeders: Croston *et al.*, 2016). Finally, increasing the number of cues tested (e.g., spatial and colour) could increase our understanding of the relationship between learning and proactive interference in general.

Overall, we did not find any population differences in either spatial memory accuracy or proactive interference performance measures. We found that individual great tits' proactive interference and spatial memory accuracy are both significantly and moderately repeatable and are traded-off at the among-individual level. Therefore, our results show that spatial memory may not be under increased selection at high elevations as it seems to be in some scatter-

hoarding birds, but the material is there for selection to concurrently act upon spatial memory and proactive interference. Selection may act differently depending on a species' or population's functional behaviour and ecology, and cognitive ecology research should continue to open up the breadth of study systems examined.

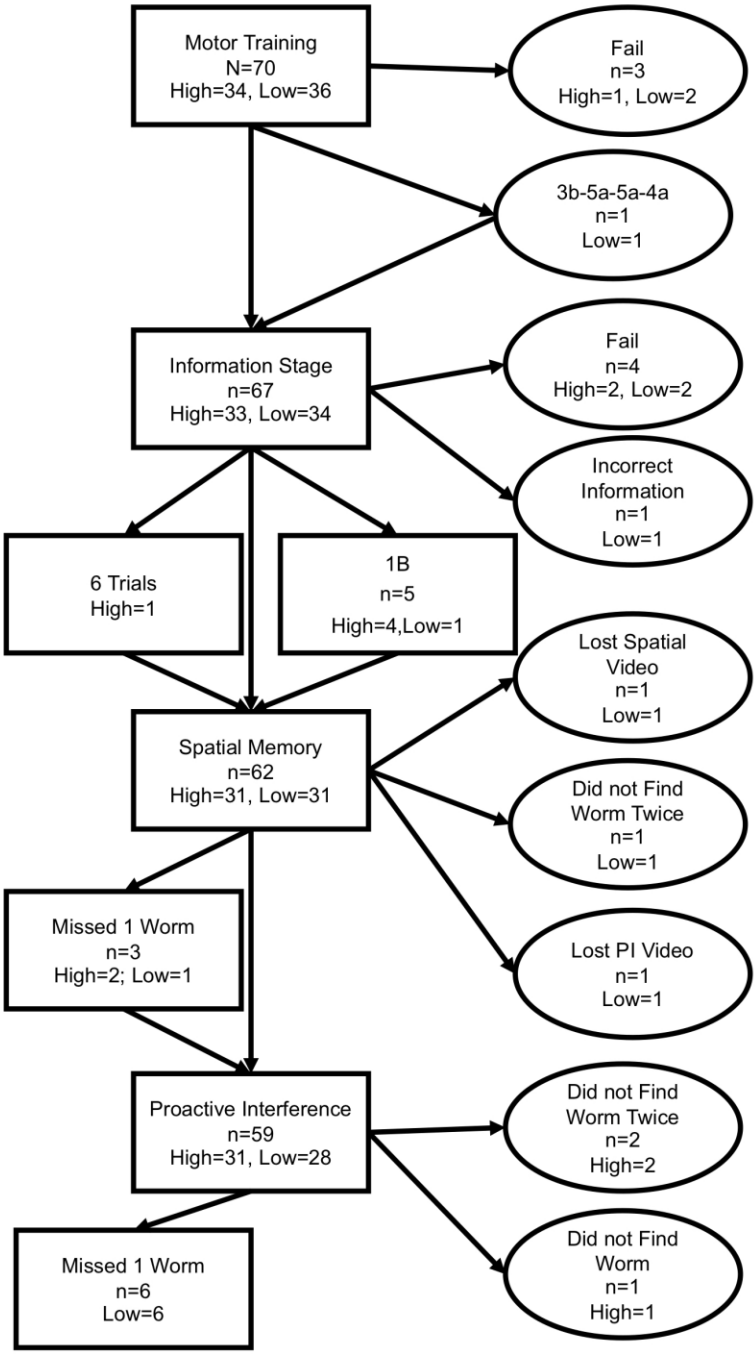
Supplementary



S2.1. Map of the high-elevation sites (blue triangles; 800-900m) and low-elevation sites (orange squares; 400-500m) near the Experimental and Theoretical Ecology Research Station near St. Girons, France (white circle). Landsat 8 OLI-TIRS Collection 1 scene LC81990302018278LGN00 provided courtesy of U.S. Geological Survey.

S2.2. Testing schedule for a single batch of birds.

Days 1-6	Day 7	Day 8	Day 9	Day 10	Day 11
Acclimatization	Motor Training	Information Stages and Three Spatial Memory Trials	Four Spatial Memory Trials	Proactive Interference Information Stage and Three Proactive Interference Trials	Two Proactive Interference Trials



S2.3. Decision tree for which birds were kept in testing. Squares indicate birds kept in testing; circles denote birds that were removed from testing at that stage.

S2.4. Individual bird ID's and motor training, spatial memory and proactive interference comments.

Task	Bird ID	Elevation	Comments
Motor Training	1009	High	Did not pass motor training.
Motor Training	117	Low	Did not pass motor training.
Motor Training	923	Low	Did not pass motor training.
Motor Training	123	Low	Went from 3b-5a-5a-4a.
Information Stage	801	High	Did not find worm.
Information Stage	805	High	Did not find worm.
Information Stage	717	Low	Did not find worm.
Information Stage	721	Low	Did not find worm.
Information Stage	713	Low	Given incorrect information during information trial.
Information Stage	811	High	Did not remove worm until second 1b trial.
Information Stage	1001	High	Had 1B for information trial.

Information Stage	809	High	Had 1B for information trial.
Information Stage	715	High	Had 1B for information trial.
Information Stage	917	Low	Had 1B for information trial.
Information Stage	1123	High	Had 1B for information trial.
Spatial Memory	601	Low	Lost Second Day Videos, only has trials 1-3.
Spatial Memory	1005	High	Did not find worm at trial 4 but found it during deprivation.
Spatial Memory	513	High	Did not find worm in trial 2 but found it during deprivation.
Spatial Memory	921	Low	Did not find worm during trial 1 but found it during deprivation.

Spatial Memory	411	Low	Did not find worm at trial 2 and trial 5.
Proactive Interference	119	Low	Lost first day of proactive interference videos.
Proactive Interference	319	High	Did not find worm in the information stage and trial 2.
Proactive Interference	811	High	Did not find worm in the information stage and trial 2.
Proactive Interference	201	High	Does not find worm during trial 5.
Proactive Interference	123	Low	Does not find worm during the information stage, finds during deprivation.
Proactive Interference	405	Low	Does not find worm during the information stage,

			finds during deprivation.
Proactive Interference	915	Low	Does not find worm during the information stage, finds during deprivation.
Proactive Interference	919	Low	Does not find worm during the information stage, finds during deprivation.
Proactive Interference	719	Low	Does not find worm during the information stage, finds during deprivation.
Proactive Interference	1211	Low	Does not find worm during trial 2, finds during deprivation.

Proactive Interference	609	Low	Pulled pompoms early in trial 3 and trial 5.
Proactive Interference	607	Low	Pulled pompoms early in trial 3.
Proactive Interference	121	Low	Pulled pompoms early in trial 3.

S2.5. Predictors of the log transformed number of errors made by high elevation birds (n=31; n=215 trials) across seven spatial memory trials fitted with a linear mixed effect model with trial, capture order, intertrial interval (min), and site (Cap Sour, Galey, Balacet), sex (male/female), age (juvenile/adult) rewarded side of the tree (left/right), and observer (EH, JH, AR) included as fixed effects. Bird ID was included as a random intercept.

Predictors	Estimate ± SE	F-statistic	*P*
Intercept	2.30 ± 0.77		
Trial	-0.22 ± 0.05	20.91	<.0001
Sex (M)	0.37 ± 0.19	4.02	0.06
Age (Juvenile)	0.11 ± 0.21	0.25	0.62
Capture Order	0.07 ± 0.75	0.01	0.93
Intertrial interval (min)	-0.01 ± 0.05	0.06	0.81
Correct Side (Right)	-0.31 ± 0.18	-3.03	0.10
Observer (EH)	-0.01 ± 1.56	0.04	0.96
Observer (JH)	0.10 ± 0.57		
Site (Cap Sour)	-0.01 ± 0.52	0.37	0.70
Site (Galey)	-0.21 ± 0.29		

S2.6. Predictors of the log transformed number of errors made by low elevation birds (n=31; n=208 trials) across seven spatial memory trials fitted with a linear mixed effect model with trial, capture order, intertrial interval (min), and site (Grotte D'Aliou, Montjoie, Aubert), sex (male/female), age (juvenile/adult) rewarded side of the tree (left/right), and observer (EH, JH, AR) included as fixed effects. Bird ID was included as a random intercept.

Predictors	Estimate ± SE	F-statistic	*P*
Intercept	1.82 ± 0.47		
Trial	-0.20 ± 0.05	17.92	<.0001
Sex (Male)	-0.29 ± 0.17	2.90	0.10
Age (Juvenile)	-0.09 ± 0.22	0.16	0.70
Capture Order	0.20 ± 0.13	2.51	0.13
Intertrial Interval (min)	0.05 ± 0.05	1.30	0.26
Correct Side (Right)	-0.02 ± 0.16	0.010	0.93
Observer (EH)	0.80 ± 0.35	2.77	0.08
Observer (JH)	0.57 ± 0.44		
Site (Grotte D'Aliou)	0.12 ± 0.26	0.31	0.74
Site (Montjoie)	0.16 ± 0.22		

S2.7. Predictors of the log transformed ratio of errors made by high elevation birds (n=31; n=145 trials) across five reversal trials fitted with a linear mixed effect model with trial, capture order, intertrial interval (min), and site (Cap Sour, Galey, Balacet), sex (male/female), age (juvenile/adult), and rewarded side of the tree (left/right) included as fixed effects. Bird ID was included as a random intercept.

Predictors	Estimate ± SE	F-statistic	*P*
Intercept	0.68 ± 0.12		
Trial	0.02 ± 0.02	0.57	0.45
Sex (Male)	-0.05 ± 0.09	0.27	0.61
Age (Juvenile)	-0.12 ± 0.09	1.54	0.23
Capture Order	-0.02 ± 0.05	0.11	0.74
Intertrial Interval	-0.04 ± 0.02	3.19	0.08
Side (Right)	-0.15 ± 0.08	3.48	0.08
Site (Cap Sour)	0.05 ± 0.10	0.14	0.87
Site (Galey)	0.01 ± 0.10		

S2.8. Predictors of the ratio of errors made by low elevation birds (n=28; n=139 trials) across five reversal trials fitted with a linear mixed effect model with trial, capture order, intertrial interval (min), and site (Grotte D'Aliou, Montjoie, Aubert), sex (male/female), age (juvenile/adult) rewarded side of the tree (left/right), and observer (JH, AR) included as fixed effects. Bird ID was included as a random intercept.

Predictors	Estimate ± SE	F-statistic	*P*
Intercept	0.56 ± 0.13		
Trial	0.01 ± 0.03	0.14	0.71
Sex (Male)	-0.001 ± 0.10	0.0002	0.99
Age (Juvenile)	0.03 ± 0.13	0.05	0.82
Capture Order	0.06 ± 0.05	1.27	0.27
Intertrial Interval	0.06 ± 0.03	4.72	0.03
Side (Right)	-0.02 ± 0.09	0.04	0.85
Observer (JH)	0.11 ± 0.18	0.39	0.54
Site (Grotte D'Aliou)	0.10 ± 0.16	0.21	0.82
Site (Montjoie)	0.03 ± 0.12		

References

Ashton, B. J., Thornton, A., & Ridley, A. R. (2018). An intraspecific appraisal of the social intelligence hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756), 20170288.

Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software*, *67*(1), 1–48.

Benedict, L. M., Pitera, A. M., Branch, C. L., Kozlovsky, D. Y., Sonnenberg, B. R., Bridge, E. S., & Pravosudov, V. V. (2020). Elevation-related differences in annual survival of adult food-caching mountain chickadees are consistent with natural selection on spatial cognition. *Behavioral Ecology and Sociobiology*, *74*(4).

Boogert, N. J., Madden, J. R., Morand-Ferron, J., & Thornton, A. (2018). Measuring and understanding individual differences in cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756).

Boyle, A. W., Sandercock, B. K., & Martin, K. (2016). Patterns and drivers of intraspecific variation in avian life history along elevational gradients: a meta-analysis. *Biological Reviews*, *91*, 469–482.

Bründl, A. C., Sallé, L., Lejeune, L. A., Sorato, E., Thiney, A. C., Chaine, A. S., & Russell, A. F. (2020). Elevational gradients as a model for understanding associations among temperature, breeding phenology and success. *Frontiers in Ecology and Evolution*, *8*.

Bründl, A. C., Sorato, E., Sallé, L., Thiney, A. C., Kaulbarsch, S., Chaine, A. S., & Russell, A. F. (2019). Experimentally induced increases in fecundity lead to greater nestling care in blue tits. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1905).

Bublitz, A., Weinhold, S. R., Strobel, S., Dehnhardt, G., & Hanke, F. D. (2017). Reconsideration of serial visual reversal learning in octopus (*Octopus vulgaris*) from a methodological perspective. *Frontiers in Physiology*, *8*, 1–11.

Careau, V., & Wilson, R. S. (2017). Of uberfleas and krakens: Detecting trade-offs using mixed models. *Integrative and Comparative Biology*, *57*(2), 362–371.

Cauchoix, M., & Chaine, A. S. (2016). How can we study the evolution of animal minds? *Frontiers in Psychology*, *7*, 1–18.

Cauchoix, M., Chow, P. K. Y., van Horik, J. O., Atance, C. M., Barbeau, E. J., Barragan-Jason, G., Bize, P., Boussard, A., Buechel, S. D., Cabirol, A., Cauchard, L., Claidière, N., Dalesman, S., Devaud, J.M., Didic, M., Doligez, B., Fagot, J., Fichtel, C., Henke-von der Malsburg, J., Hermer, E., Huber, L., Huebner, F., Kappeler, P. M., Klein, S., Langbein, J., Langley, E. J. G., Lea, S. E. G., Lihoreau, M., Lovlie, H., Matzel, L. D., Nakagawa, S., Nawroth, C., Oesterwind, S., Sauce, B., Smith, E. A., Sorato, E., Tebbich, S., Wallis, L. J., Whiteside, M. A., Wilkinson, A., Chaine, A. S., Morand-Ferron, J. (2018). The repeatability of cognitive performance: a meta-analysis, *Philosophical Transactions of the Royal Society B: Biological Sciences. 373*, 20170281.

Chittka, L. (1998). Sensorimotor learning in bumblebees: long-term retention and reversal training. *Journal of Experimental Biology*, *201*(4), 515–524.

Chrobak, J. J., Hinman, J. R., & Sabolek, H. R. (2008). Revealing past memories: proactive interference and ketamine-induced memory deficits. *Journal of Neuroscience*, *28*(17), 4512–4520.

Clayton, N. S., & Krebs, J. R. (1994). One-trial associative memory: Comparison of food-storing and nonstoring species of birds. *Animal Learning & Behavior*, *22*(4), 366–372.

Collett, M., Chittka, L., & Collett, T. S. (2013). Spatial memory in insect navigation. *Current Biology*, *23*(17), R789–R800.

Croston, R., Branch, C. L., Kozlovsky, D. Y., Dukas, R., & Pravosudov, V. V. (2015). Heritability and the evolution of cognitive traits. *Behavioral Ecology*, *26*(6), 1447–1459.

Croston, R., Branch, C. L., Kozlovsky, D. Y., Roth II, T. C., LaDage, L. D., Freas, C. A., & Pravosudov, V. V. (2015). Potential mechanisms driving population variation in spatial memory and the hippocampus in food-caching chickadees. *Integrative and Comparative Biology*, *55*(3), 354–371.

Croston, R., Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Bridge, E. S., Parchman, T. L., & Pravosudov, V. V. (2017). Predictably harsh environment is associated with reduced cognitive flexibility in wild food-caching mountain chickadees. *Animal Behaviour*, *123*, 139–149.

Croston, R., Kozlovsky, D. Y., Branch, C. L., Parchman, T. L., Bridge, E. S., & Pravosudov, V. V. (2016). Individual variation in spatial memory performance in wild mountain chickadees from different elevations. *Animal Behaviour*, *111*, 225–234.

Freas, C. A., LaDage, L. D., Roth II, T. C., & Pravosudov, V. V. (2012). Elevation-related differences in memory and the hippocampus in mountain chickadees, Poecile gambeli. *Animal Behaviour*, *84*, 121–127.

Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*(11), 1325–1330.

Gonzalez, R. C., Behrend, E. R., & Bitterman, M. E. (1967). Reversal learning and forgetting in bird and fish. *Science.*, *158*, 519–521.

Gosler, A. G. (1993). *The great tit*. London: Hamlyn.

Griffin, A. S., Guillette, L. M., & Healy, S. D. (2015). Cognition and personality: An analysis of an emerging field. *Trends in Ecology and Evolution*, *30*(4), 207–214.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, *33*(2), 1–22.

Hampton, R. R., & Shettleworth, S. J. (1996). Hippocampus and memory in a food-storing and in a nonstoring bird species. *Behavioral Neuroscience*, *110*(5), 946–964.

Hampton, R. R., Shettleworth, S. J., & Westwood, R. P. (1998). Proactive interference, recency, and associative strength: comparisons of black-capped chickadees and dark-eyed juncos. *Animal Learning & Behavior*, *26*(4), 475–485.

Healy, S. D. (1995). Memory for objects and positions: delayed non-matching-to-sample in storing and non-storing tits. *The Quarterly Journal of Experimental Psychology Section B*, *48*(2), 179–191.

Healy, S. D., & Krebs, J. R. (1992). Delayed-matching-to-sample by marsh tits and great tits. *The Quarterly Journal of Experimental Psychology Section B*, *45*(1), 33–47.

Hermer, E., Cauchoix, M., Chaine, A. S., & Morand-Ferron, J. (2018). Elevation-related difference in serial reversal learning ability in a nonscatter hoarding passerine. *Behavioral Ecology*, *29*(4), 840–847.

Herszage, J., & Censor, N. (2018). Modulation of learning and memory: a shared framework for interference and generalization. *Neuroscience*, *392*, 270–280.

Houslay, T. M., & Wilson, A. J. (2017). Avoiding the misuse of BLUP in behavioural ecology. *Behavioral Ecology*, *28*(4), 948–952.

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: an updated perspective. *Neuroscience*, *345*, 12–26.

Janmaat, K. R. L., Boesch, C., Byrne, R., Chapman, C. A., Goné Bi, Z. B., Head, J. S., Robbins, M. M., Wrangham, R. W., & Polansky, L. (2016). Spatio-temporal complexity of chimpanzee food: how cognitive adaptations can counteract the ephemeral nature of ripe fruit. *American Journal of Primatology*, *78*(6), 626–645.

Joseph, M. A., Fraize, N., Ansoud-Lerouge, J., Sapin, E., Peyron, C., Arthaud, S., Libourel, P. A., Parmentier, R., Salin, P. A., & Malleret, G. (2015). Differential involvement of the dentate gyrus in adaptive forgetting in the rat. *PLoS ONE*, *10*(11), 1–17.

Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M., & Altman, D. G. (2010). Improving bioscience research reporting: The arrive guidelines for reporting animal research. *PLoS Biology*, *8*(6), 6–10.

Körner, C. (2007). The use of "altitude" in ecological research. *TRENDS in Ecology and Evolution*, *22*(11), 569–574.

Kozlovsky, D. Y., Weissgerber, E. A., & Pravosudov, V. V. (2017). What makes specialized food-caching mountain chickadees successful city slickers? *Proceedings of the Royal Society B*, *284*, 20162613.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26.

LaDage, L. D., Roth, T. C., Cerjanic, A. M., Sinervo, B., & Pravosudov, V. V. (2012). Spatial memory: are lizards really deficient? *Biology Letters*, *8*(6), 939–941.

Langley, E. J. G., Adams, G., Beardsworth, C. E., Dawson, D. A., Laker, P. R., van Horik, J. O., Whiteside, M. A., Wilson, A. J., & Madden, J. R. (2020). Heritability and correlations among learning and inhibitory control traits. *Behavioral Ecology*, 1–9.

Lejeune, L., Savage, J. L., Bründl, A. C., Thiney, A., Russell, A. F., & Chaine, A. S. (2019). Environmental effects on parental care visitation patterns in blue tits *Cyanistes caeruleus*. *Frontiers in Ecology and Evolution*, *7*, 1–15.

Malleret, G., Alarcon, J. M., Martel, G., Takizawa, S., Vronskaya, S., Yin, D., Chen, I, Z., Kandel, E. R., & Shumyatsky, G. P. (2010). Bidirectional regulation of hippocampal long-term synaptic plasticity and its influence on opposing forms of memory. *Journal of Neuroscience*, *30*(10), 3813–3825.

McGregor, A., & Healy, S. D. (1999). Spatial accuracy in food-storing and nonstoring birds. *Animal Behaviour*, *58*, 727–734.

Meck, W. H., & Williams, C. L. (1999). Choline supplementation during prenatal development reduces proactive interference in spatial memory. *Developmental Brain Research*, *118*(1–2), 51–59.

Milton, K. (1981). Distribution patterns of tropical plant foods as an evolutionary stimulus to primate mental development. *American Anthropologist*, *83*(3), 534–548.

Missaire, M., Fraize, N., Joseph, M. A., Hamieh, A. M., Parmentier, R., Marighetto, A., Salin, P. A., & Malleret, G. (2017). Long-term effects of interference on short-term memory performance in the rat. *PLoS ONE*, *12*(3), 1–18.

Miyashita, Y., Nakajima, S., & Imada, H. (2000). Differential outcome effect in the horse. *Journal of the Experimental Analysis of Behavior*, *74*(2), 245–253.

Morand-Ferron, J., Hermer, E., Jones, T. B., & Thompson, M. J. (2019). Environmental variability, the value of information, and learning in winter residents. *Animal Behaviour*, *147*, 137–145.

Niemelä, P. T., & Dingemanse, N. J. (2018). On the usage of single measurements in behavioural ecology research on individual differences. *Animal Behaviour*, *145*, 99–105.

Pravosudov, V. V., & Clayton, N. S. (2002). A test of the adaptive specialization hypothesis: population differences in caching, memory, and the hippocampus in black-capped chickadees (*Poecile atricapilla*). *Behavioral Neuroscience*, *116*(4), 515–522.

Pravosudov, V. V., & Roth II, T. C. (2013). Cognitive ecology of food hoarding: the evolution of spatial memory and the hippocampus. *Annual Review of Ecology, Evolution, and Systematics*, *44*, 173–193.

R Core Team. (2020). *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Roth II, T. C., LaDage, L. D., Freas, C. A., & Pravosudov, V. V. (2012). Variation in memory and the hippocampus across populations from different climates: a common garden approach. *Proc. R. Soc. B*, *279*(1727).

Roth II, T. C., LaDage, L. D., & Pravosudov, V. V. (2010). Learning capabilities enhanced in harsh environments: a common garden approach. *Proceedings of the Royal Society B*, *277*, 3187–3193.

Roth II, T. C., & Pravosudov, V. V. (2009). Hippocampal volumes and neuron numbers increase along a gradient of environmental harshness: a large-scale comparison. *Proceedings of the Royal Society B*, *276*, 401–405.

Shaw, R. C., MacKinlay, R. D., Clayton, N. S., & Burns, K. C. (2019). Memory performance influences male reproductive success in a wild bird. *Current Biology*, *29*(9), 1498–1502.

Shiflett, M. W., Rankin, A. Z., Tomaszycki, M. L., & DeVoogd, T. J. (2004). Cannabinoid inhibition improves memory in food-storing birds, but with a cost. *Proceedings of the Royal Society B: Biological Sciences*, *271*(1552), 2043–2048.

Sonnenberg, B. R., Branch, C. L., Pitera, A. M., Bridge, E., & Pravosudov, V. V. (2019). Natural selection and spatial cognition in wild food-caching mountain chickadees. *Current Biology*, *29*, 1–7.

Squier, L. H. (1969). Reversal learning improvement in the fish *Astronotus ocellatus* (Oscar). *Psychonomic Science*, *14*(3), 143–144.

Stoffel, M. A., Nakagawa, S., & Schielzeth, H. (2017). rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *8*(11), 1639–1644.

Strang, C. G., & Sherry, D. F. (2014). Serial reversal learning in bumblebees (*Bombus impatiens*). *Animal Cognition*, *17*, 723–734.

Svensson, L. (1992). *Identification guide to european passerines* (4th ed.). British Trust for Ornithology.

Tello-Ramos, M. C., Branch, C. L., Kozlovsky, D. Y., Pitera, A. M., & Pravosudov, V. V. (2018). Spatial memory and cognitive flexibility trade-offs: to be or not to be flexible, that is the question. *Animal Behaviour*, 1–8.

Tello-Ramos, M. C., Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Bridge, E. S., & Pravosudov, V. V. (2018). Memory in wild mountain chickadees from different elevations: comparing first-year birds with older survivors. *Animal Behaviour*, *137*, 149–160.

Thompson, M. J., & Morand-Ferron, J. (2019). Food caching in city birds: urbanization and exploration do not predict spatial memory in scatter hoarders. *Animal Cognition*, *22*(5), 743–756.

Thornton, A., & Boogert, N. J. (2019). Animal cognition: the benefits of remembering. *Current Biology*, *29*(9), R324–R327.

Tillé, Y., Newman, J. A., & Healy, S. D. (1996). New tests for departures from random behavior in spatial memory experiments. *Animal Learning and Behavior*, *24*(3), 327–340.

Warton, D. I., Lyons, M., Stoklosa, J., & Ives, A. R. (2016). Three points to consider when choosing a LM or GLM test for count data. *Methods in Ecology and Evolution*, *7*(8), 882–890.

Wilson, A. J. (2018). How should we interpret estimates of individual repeatability? *Evolution Letters*, *2*(1), 4–8.

Chapter 3: *Serial reversal learning, spatial memory, and the greater male variability phenomenon*

Introduction

Differences in variability between sexes is a phenomenon reported across a variety of morphological (e.g., feather ornament size, body size, metabolism, organ weight, brain size) and behavioural measures (e.g., song length) in lab (Becker, Prendergast, & Liang, 2016; DeCasien, Sherwood, Schapiro, & Higham, 2020; Smarr, Grant, Zucker, Prendergast, & Kriegsfeld, 2017; Zajitschek *et al*., 2020) and wild animal populations (e.g., aves, insects, mammals: Cuervo & Møller, 1999; Pomiankowski & Møller, 1995; Reinhold & Engqvist, 2013; Wyman & Rowe, 2014; but see Harrison, Noble & Jennions, 2021). These differences in variability were first reported as part of the lek paradox, wherein sexual selection was expected to decrease the variance of a sexually selected trait relative to its non-sexually selected counterpart, but paradoxically the opposite was found (Cuervo & Møller, 1999; Pomiankowski & Møller, 1995).

The first explanation for sex differences in variability (SDV) came out of this lek paradox research, and the genic capture hypothesis (Rowe & Houle, 1996; Tomkins, Radwan, Kotiaho, & Tregenza, 2004). The genic capture hypothesis is based on two principles. First, most sexually selected traits are condition dependent (Cuervo & Møller, 2001; Jennions, Møller, & Petrie, 2001). Second, traits that express condition are dependent on multiple other loci and traits, and they "capture," or express all of that variance (Kotiaho, Simmons, & Tomkins, 2001; Rowe & Houle, 1996; reviewed in: Tomkins *et al*., 2004). Therefore, if a trait is under sexual selection and is condition dependent in males, but not females, it may show greater male variance (GMV). Additionally, any system with prominent sex-biased body size sexual dimorphism may increase this variance by amplifying the condition dependent expression (reviewed in: DeCasien *et al*., 2020). In support of this sexual selection hypothesis, male sexually selected traits have been found to be more variable than a similar non-sexually selected trait in the same individuals

(Cuervo & Møller, 1999, 2001) or when compared with equivalent non-sexually selected female traits (Pomiankowski & Møller, 1995). However, the opposite has been found, with female and male trait variance not differing when one is under sexual selection (Cotton, Fowler, & Pomiankowski, 2004; no difference after Bonferroni correction: Cuervo & Møller, 1999) and sexual dimorphism, utilized as a measure of the difference in sexual selection, did not explain the magnitude of male and female variance differences in animal personality (Harrison *et al.* 2021).

The other main explanation for SDVs is that the trait is influenced by the sex chromosome pair (Reinhold & Engqvist, 2013). In animals with an XY or ZW sexual determination system, the homogametic sex has two chromosomes that code for traits, while the heterogametic sex only has one. Variation is decreased by the presence of these two chromosomes as the two chromosomes can each be expressed throughout the body, creating an average of the two throughout. Also, the doubled number of available chromosomes decreases the probability of rare, extreme phenotypes being expressed. Therefore, the heterogametic sex has a greater chance to express more extreme phenotypes, and to show larger variability in phenotypes, while the homogametic sex has a lower chance to express more extreme phenotypes and should instead express more average traits. Evidence for this sex chromosome hypothesis has been found with the heterogametic males (mammals, insects) and females (birds) showing greater body size variability than the homogametic sex (Reinhold & Engqvist, 2013), and in humans where males are shown to generally have more variable intelligence scores than females and some measures of intelligence may be related to genes on the X-chromosome (reviewed in: Johnson, Deary, & Carothers, 2009).

This phenomenon is termed the greater male variability (GMV) hypothesis (or phenomenon: Branch *et al*., 2020) in cognition research, as human males are generally found to

have greater variability in their intelligence scores than in females (Hedges & Nowell, 1995; Johnson, Deary, & Carothers, 2008; Lehre, Lehre, Laake, & Danbolt, 2009; Zechner *et al.*, 2001 but see: Harrison *et al.* 2021). There are very few SDV studies in animal cognition, with only one captive chimpanzee (*Pan troglodytes*) study finding that male brain structures were more variable in size than female brain structures (DeCasien *et al.*, 2020) and a study in lab rats finding no difference between male and female variance in neurochemistry, histology, electrophysiology, or behaviour, while cardiovascular measures were more variable in females (Becker *et al*., 2016). A study comparing male and female variance in spatial memory accuracy and reversal learning performance in wild mountain chickadees (*Poecile gambeli*) found that spatial memory accuracy, a trait shown to be under sexual selection in this population, did not show SDV. However, reversal learning accuracy, which does not have evidence for sexual selection acting on it, showed GMV (Branch *et al.*, 2020). Interestingly, this study provides evidence for the GMV as a phenomenon, but does not support either the sexual selection hypothesis, which would predict only the cognitive abilities under sexual selection to show SDV, nor the sex chromosome hypothesis, which predicts greater variation in females (the heterogametic sex in birds). I broaden the focus of these studies to other non-mammalian wild species to better understand where and which SDVs may be found.

I have quantified serial spatial reversal learning performance (*Ch. 1*: Hermer, Cauchoix, Chaine, & Morand-Ferron, 2018), and spatial memory accuracy (*Ch. 2*: Hermer, Murphy, Chaine, & Ferron, 2021) in a wild population of great tits, and found an elevation-related difference in serial reversal learning accuracy but no elevation-related difference in spatial memory accuracy. I propose to use these two sets of spatial cognition data to examine the sexual selection and sex chromosome hypothesis in these great tits.

Spatial cognition is thought to be a target for sexual selection in a few species. Greater spatial cognition in males has been linked to increased clutch size, and fledgling weight in two avian species (Branch, Pitera, Kozlovsky, Bridge, & Pravosudov, 2019; Medina-García & Wright, 2021), and increased male attractiveness in meadow voles (*Microtus pennsylvanicus*: Spritzer, Meikle, & Solomon, 2005). These studies suggest potential indirect genetic benefits of reproducing with mates possessing greater spatial abilities, thereby explaining the increased attractiveness of these males. Although I have not tested whether female great tits find males with greater spatial cognition to be more attractive, it has been hypothesized that spatial memory (reviewed in: Thornton & Boogert, 2019) and reversal learning could generally improve foraging ability (Janmaat *et al*., 2016). Furthermore, it has been hypothesized that traits that increase foraging ability could be signalled indirectly via conspicuous ornaments that reflect foraging ability, and influence female mate choice (reviewed in: Boogert, Fawcett, & Lefebvre, 2011). In great tits, it is believed that males are under sexual selection and their quality may be judged through their conspicuous plumage. The size of the black breast stripe and the yellow side plumage are shown during mating displays (Gosler, 1993). Females lay larger clutches and commence breeding earlier in the season when breeding with males with a larger black breast stripe (Norris, 1990a). Additionally, males with larger breast stripes produce heavier fledglings with a higher probability of survival (Norris, 1990b). This indicates that the black breast stripe may be utilized by females when making a mate choice, and it may signal increased foraging ability that leads to greater reproductive success. Great tits also mate assortatively based on the lightness of the yellow flank (Ferns & Hinsley, 2008). Males with higher values of hue in the yellow side plumage fed their offspring a higher portion of spiders, which are correlated to improved body condition (Pagani-Núñez & Senar, 2014). Carotenoid based signals, like the

yellow side plumage, have been implicated across species as an indicator of foraging ability, immune status and parasite load, and detoxification potential (Møller *et al*., 2000). However, the yellow side plumage is not correlated to motor learning ability, but its correlation to spatial cognition has not been tested (Cauchard, Doucet, Boogert, Angers, & Doligez, 2017). Finally, cognitive ability is generally shown to be condition dependent in many avian species (reviewed in: Buchanan, Grindstaff, & Pravosudov, 2013) and great tits compete for territories (Gosler, 1993) which may amplify the condition dependence of their traits. If spatial cognition is useful in foraging and signalled by the black breast stripe or yellow side plumage, or if spatial cognitive performance can be directly assessed by females, spatial cognitive ability could be under sexual selection in the great tit and variance in males could be amplified due to its condition dependence.

If cognitive ability is condition dependent, and if it is indirectly sexually selected for through plumage or another aspect of visible foraging ability such as observed foraging efficiency during the winter, I predict this to manifest as GMV. Conversely, if differences in variance are dependent on the sex chromosome pair, I predict that females, the heterogametic sex in great tits (ZW) will have greater variability than males. Here I test both the sexual selection and sex chromosome hypotheses and their associated predictions (Table 3.1) using data I collected for *Chapter 1* and *2* on spatial reversal learning accuracy and spatial memory accuracy.

Table 3.1. Three potential conclusions and their interpretation

Greater Male Variability	Greater Female Variability	No Difference
Could be indicative of the cognitive ability being condition dependent, and under sexual selection (sexual selection hypothesis).	The cognitive ability could be linked to the sex chromosomes in great tits (sex chromosome hypothesis).	No indication of sexual selection or chromosome link in this cognitive ability.

Methods

I am utilizing spatial serial reversal learning (Hermer *et al*., 2018) and spatial memory data (Hermer *et al*., 2021) from two previously published papers. Both papers sampled great tits (Serial reversal learning: n=31, females: n=12, males: n=19; spatial memory: n=62 females: n=27, males: n=35) in different years in the Pyrénées Mountain Range, near St. Girons, France. Both tasks were tested in a semi-natural outdoor captive environment after a period of habituation (7-12 days for serial reversal learning, 6 days for spatial memory). Birds were deprived of their regular food before and during the tasks. The serial reversal task measured the number of trials to complete the task across 18 reversals wherein they had to peck the correct pecking key (left or right) to receive a meal worm reward which was switched after reaching the correct criterion (achieving 9 correct pecks across 10 trials). The spatial memory task measured spatial accuracy by providing information on the location of a mealworm reward that was in one of 45 holes spread across three artificial trees (15 holes per tree) and then measuring the number of errors birds made before retrieving the worm in each of seven trials. Serial reversal learning trials to criterion has significant, low temporal unadjusted repeatability (Cauchoix *et al.* 2018) and number of spatial memory errors has significant, moderate temporal adjusted repeatability (Hermer *et al*. 2021).

I utilized trials to criterion for each of the reversals (n = 18) in the serial reversal learning test, and log-transformed number of errors for each trial (n = 7) of the spatial memory data to better meet the assumptions of a Gaussian distribution. To test for SDVs I utilized the maximal models fixed effects from each paper (serial reversal learning: Table 2; spatial memory accuracy: Table 3) and added a sex x individual bird ID random interaction (i.e., a heterogeneous variance model). This interaction allows for the residual or within-sex variance, and the among-sex

variance to be calculated. This is similar to adding in a random effect of ID which allows for the calculation of among and within-individual variance. Both models were fitted using the Bayesian MCMCglmm package (Version 2.29: Hadfield, 2010) in R (Version 4.0.3: R Core Team, 2020). The serial reversal learning model was run with 1300000 iterations, 1000 thinning interval, and a burn in of 30000 to meet visual model diagnostics, auto-correlation diagnostics, and effective sample sizes. The spatial memory model was run with 1300000 iterations, a thinning interval of 100, and a burn in of 300000 to meet visual model diagnostics and effective sample sizes.

By utilizing a Bayesian approach, I can directly compare the difference in female and male within-sex variance by subtracting the posterior estimate of the residuals and then calculating the HPDI (high probability density interval; analogous to the 95% confidence interval) for the female and male residual difference. By comparing the male and residual variance directly in the model, it avoids the potential loss of information relating to uncertainty, and overestimation of effects created when extracting the residuals from a separate model and comparing them in another separate model (see for a discussion on overestimation caused by two step analysis: Houslay & Wilson, 2017).

Results

There is no evidence that the residual variance for males and females differs for trials to criterion during the reversal learning test (male-female, estimate of difference in variance: -23.67, HPDI: [-135.04, 74.17]; Table 3.1; Figure 3.1). One individual had been motor trained on the serial reversal task in a previous year and thus had additional experience with the test content compared with other subjects. I thus re-ran the above model without this individual and an increased burn in (300000) and found qualitatively the same result (male-female difference, posterior mode: -8.56, HPDI: [-152.41, 71.01]).

There is evidence that females have greater residual variance for log transformed number of errors in the spatial task (male-female, estimate of difference in variance: -0.172, HPDI: [-0.34, -0.03]; Table 3.2; Figure 3.2). One individual had been trained differently on the spatial memory task; I re-ran the above model without this individual and the same above settings and found qualitatively the same result (male-female difference, posterior mode: -0.18, HPDI: [-0.34, -0.05]).

Table 3.2. Summary of the serial spatial reversal learning model's fixed effects' posterior means and HPDI's. Stars note HPDI's that do not cross zero. Fixed effects: Age (juvenile/adult) + Sex male/female) + Elevation (high/low) + Stage + Group Size + Capture Season (winter/fall). Random Effects: Sex*ID.

	Posterior Mean	Lower 95% HDPI	Upper 95% HDPI
Intercept*	45.77	37.34	53.16
Age (Juvenile)	-0.23	-5.64	5.59
Sex (Male)	-1.73	-8.17	4.29
Elevation (Low)*	-10.61	-15.81	-5.69
Stage (1-18)*	-3.75	-5.55	-1.90
Group Size (3-5)*	4.42	1.38	7.41
Capture (Winter)*	-13.55	-19.14	-8.81

Table 3.3. Summary of the spatial memory model's fixed effects' posterior means and HPDI's. Stars note HPDI's that do not cross zero. Fixed effects: Trial + Elevation (high/low) + Sex (male/female) + Age (juvenile/adult) + Batch + Intertrial Interval+ Side (left/right) + Observer (AR, EH, JH). Random Effects: Sex*ID.

	Posterior Mean	Lower 95% HDPI	Upper 95% HDPI
Intercept	2.16	1.64	2.68
Trials (1-7)*	-0.22	-0.28	-0.15
Elevation (Low)	-0.13	-0.397	0.127
Sex (Male)	0.02	-0.22	0.27
Age (Juvenile)	-0.05	-0.31	0.24
Batch (1-12)*	0.23	0.01	0.45
Intertrial Interval (s)	0.02	-0.04	0.08
Side (Right)	-0.11	-0.35	0.12
Observer (EH)	0.44	-0.08	0.99
Observer (JH)	0.23	-0.26	0.75

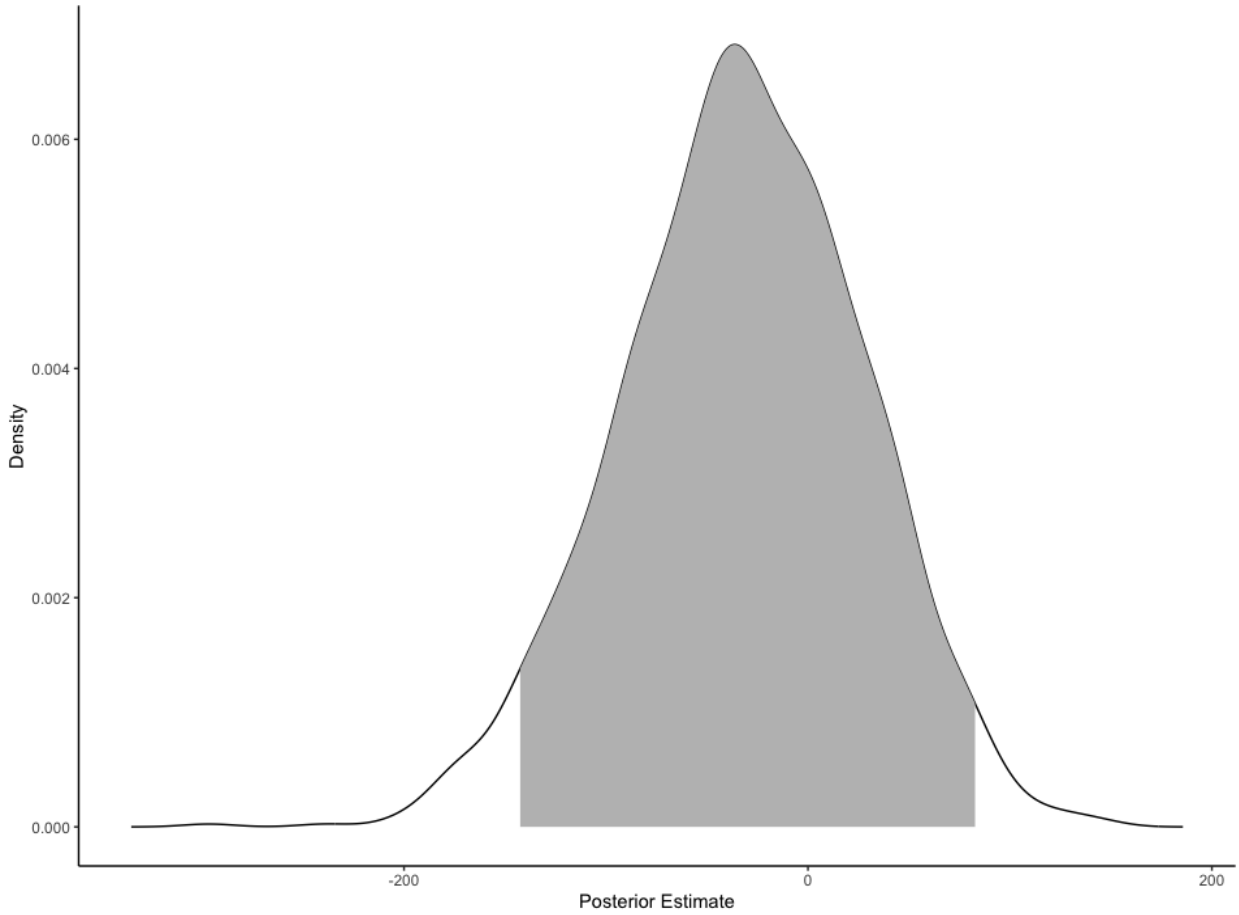


Figure 3.1. Density plot of the posterior estimate for the difference in male and female serial reversal learning variance. The HPDI is filled in gray.

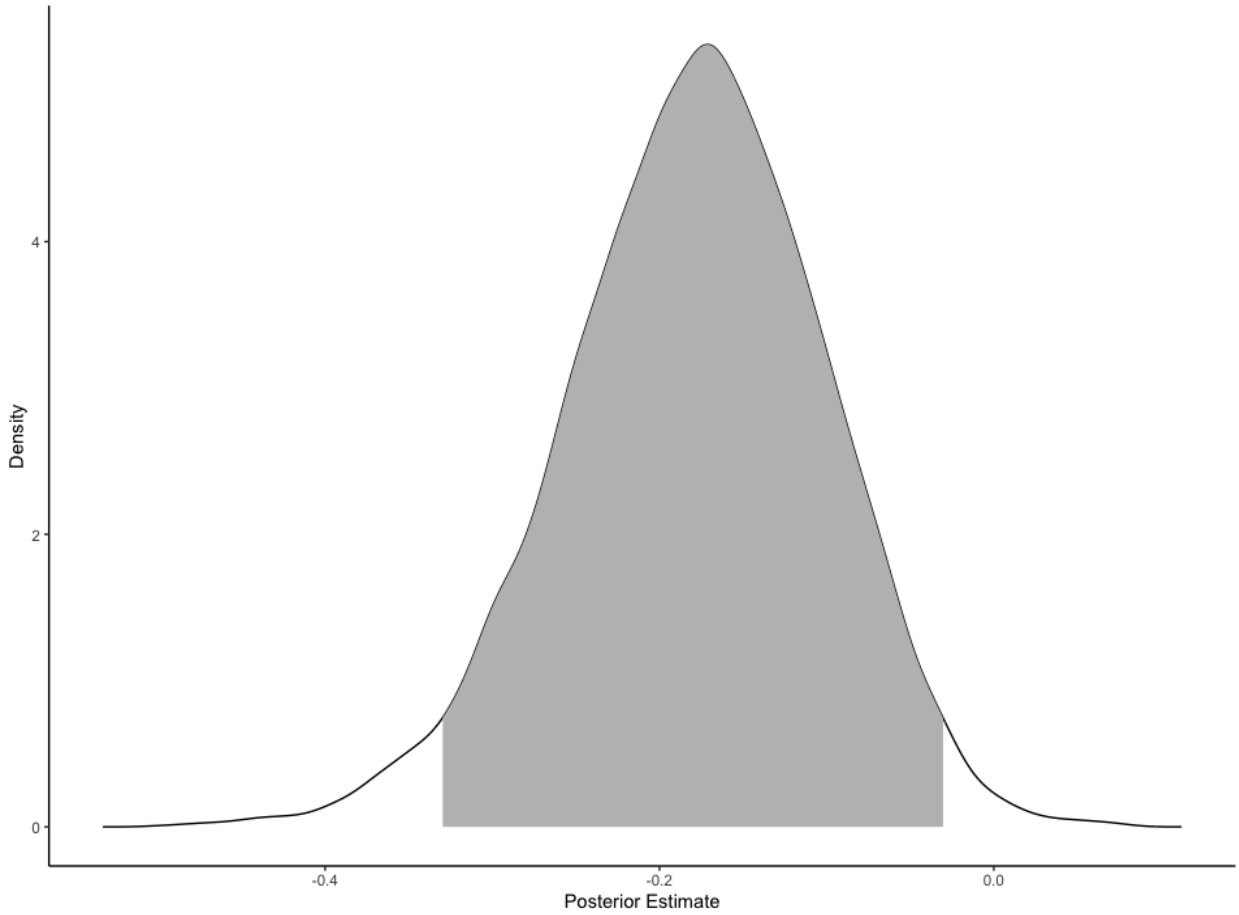


Figure 3.2. Density plot of the posterior estimate for the difference in male and female spatial memory variance. The HPDI is filled in gray.

Discussion

Differences in variance for cognition have been observed in humans (e.g., Hedges & Nowell, 1995; Johnson *et al.*, 2008; Lehre *et al.*, 2009; Zechner *et al.*, 2001) but only one study has compared male and female variability in cognition in wild animals (Branch *et al.*, 2020). To increase the breadth of non-human species studied, I compared the variability of males and females cognitive task performance on a spatial serial reversal learning task and a spatial memory task. I found that males did not have greater variability in either trials to criterion for serial reversal learning or log-transformed number of errors for spatial memory. In fact, females showed greater variability in spatial memory accuracy than males. Therefore, I found no evidence for a widespread GMV phenomena, and no evidence for sexual selection to potentially be acting on either cognitive ability. Instead, I found preliminary evidence that spatial memory may be linked to the sex chromosomes in great tits.

The two hypotheses set out above give contrasting predictions when tested in great tits. The sexual selection hypothesis predicts that sexually selected traits will be more variable than non-sexually selected traits, and so we should expect greater male variability if this is the case in great tits, because sexual selection is generally stronger in males in this species (Tomkins *et al*., 2004). The sex chromosome hypothesis predicts that the heterogametic sex (females in birds) will be more variable if the behaviour is linked to the sex chromosome (Reinhold & Engqvist, 2013). In avian systems such as the great tits here, and the mountain chickadee (Branch *et al.*, 2020) both contrasting hypotheses are possible. However, Branch *et al.*'s (2020) results contradict both hypotheses. Branch *et al.* (2020) tested for SDV's in reversal learning, a task not thought to be under sexual selection in mountain chickadees, and spatial memory, which may be under sexual selection (Branch *et al.*, 2019). Interestingly, Branch *et al.* (2020) found GMV in

reversal learning, and no SDV in spatial memory. Branch *et al.* (2020) suggest that these contradictory results may be due to natural selection. Reversal learning is not believed to be important for mountain chickadee survival, and differences in reversal learning between conspecifics are believed to be consequences of variation in memory strength, not of any direct selective pressure, while spatial memory is under strong natural selection through its link with food caching (Sonnenberg, Branch, Pitera, Bridge, & Pravosudov, 2019; Tello-Ramos, Branch, Kozlovsky, Pitera, & Pravosudov, 2018). Given the above, Branch *et al.* (2020) raise the possibility that individuals would invest as many resources as possible in spatial memory, narrowing its variance and producing no sex difference in variability. However, reversal learning's lack of importance for survival in mountain chickadees may lead to individuals only investing in it when they have an excess of resources, thereby increasing variance among-individuals. In Branch *et al.* (2020), females show evidence of the spatial memory-reversal learning trade-off while males do not, which they tentatively provide as evidence for natural selection acting more strongly on spatial memory in females as opposed to males. Therefore, males of high quality may invest in reversal learning and amplify their phenotype, while males of lower quality cannot, reducing their expressed phenotype, and increasing variance. However, females must invest so much in spatial memory, that there are fewer resources available within females to invest in reversal learning, reducing the variance below that of males. Therefore, these authors suggest natural selection seems to primarily drive differences in variance in the mountain chickadee system (Branch *et al.*, 2020).

In contrast, I show an opposite trend, with serial reversal learning showing no difference in variance, and spatial memory showing greater female variance. Genes related to learning and memory in great tits have been targets for positive selection in many European great tit

populations, and so broadly speaking natural selection may act on cognition in great tits (Laine *et al.*, 2016). In my great tit study system, high and low elevation great tits did differ in serial reversal learning ability, with low elevation being more accurate overall than high elevation birds (Hermer *et al.*, 2018). This could suggest that natural selection is acting differentially on serial reversal learning ability in low versus high elevation great tits. If this is the case, I may have found no difference in serial reversal learning variance between male and female great tits because natural selection is reducing the variance in both sexes, assuming that reversal learning is equally important for both sexes in this population.

However, I do not have evidence for natural selection acting on spatial memory as I have not found an elevation difference in spatial memory in these great tits (Hermer *et al.*, 2021). If my assumption that natural selection is not acting on spatial memory is true, natural selection may not be reducing variance in spatial memory, and other potential mechanisms, such as those outlined in the sex chromosome hypothesis may rather be determining variance differences in these great tits. Therefore, although the results related to SDV's in cognitive abilities may differ between Branch *et al.* (2020) and this study, similar mechanisms could explain both Branch *et al.* (2020) and my results.

Another possible explanation for the lack of difference in the serial reversal learning SDV is the compound nature of serial reversal learning performance. Serial spatial reversal learning performance is driven by a series of cognitive abilities, such as inhibition, rule-learning, and proactive interference (Izquierdo, Brigman, Radke, Rudebeck, & Holmes, 2017; Nilsson, Alsiö, Somerville, & Clifton, 2015). If sexual selection is acting on each of these components in contrasting ways, it could dampen any effects sexual selection may have on serial reversal performance, especially considering the polygenic nature of cognitive abilities (reviewed in:

Croston, Branch, Kozlovsky, Dukas, & Pravosudov, 2015). Overall, my null serial reversal learning result provides no evidence for either the sexual selection or sex chromosome hypothesis. Future studies should investigate the variance of serial reversal learning's separate cognitive components to study whether the differences in variance between males and females are similar across its components which could provide more precise evidence for SDV's occurring in serial reversal learning measures.

I found evidence for greater female variance in spatial memory accuracy. Therefore, I have no evidence for the sexual selection hypothesis. However, these results support the sex chromosome hypothesis, as female great tits are the heterogametic sex (ZW). Therefore, my results provide preliminary evidence that some aspect of spatial memory accuracy is coded on the sex chromosomes of great tits. Recent studies in great tits have sequenced its genome (Bosse *et al.*, 2017; Laine *et al.*, 2016), and future studies should utilize this assembly and measures of cognitive ability to pursue the question of whether spatial cognition is coded for on the sex chromosomes and pinpoint the potential genetic basis for this phenotypic difference.

Overall, my results provide evidence for the sex chromosome hypothesis regarding spatial memory accuracy, but no evidence for the sexual selection hypothesis in either cognitive performance measure. When taken with other contradictory results found in humans (e.g., Lehre *et al.*, 2009), chimpanzees (DeCasien *et al.*, 2020), rats (Becker *et al.*, 2016) and even other avian species (Branch *et al.*, 2020), there does not seem to be a global trend of greater variability in either sex for cognitive measures such as ability or brain size. Therefore, there may be some general hypotheses that can explain differences in cognitive variation, but they may be species and situation specific, instead of sex specific. To better understand the phenomena cognitive SDV's, future studies of cognitive ecology should include a sex*individual random interaction to

test for these differences. Additionally, the meta-analytic approach is common in non-cognitive studies of SDVs (e.g., Reinhold & Engqvist, 2013; Zajitschek *et al.*, 2020; Harrison *et al.* 2021) and could be used to investigate cognitive SDV's in wild animals.

References

Becker, J. B., Prendergast, B. J., & Liang, J. W. (2016). Female rats are not more variable than male rats: A meta-analysis of neuroscience studies. *Biology of Sex Differences*, *7*(1), 1–7.

Boogert, N. J., Fawcett, T. W., & Lefebvre, L. (2011). Mate choice for cognitive traits: A review of the evidence in nonhuman vertebrates. *Behavioral Ecology*, *22*(3), 447–459.

Bosse, M., Spurgin, L. G., Laine, V. N., Cole, E. F., Firth, J. A., Gienapp, P., Gosler, A. G., McMahon, K., Poissant, J., Verhagen, I., Groenen, M. A. M., van Oers, K., Sheldon, B. C., Visser, M. E., & Slate, J. (2017). Recent natural selection causes adaptive evolution of an avian polygenic trait. *Science*, *358*(6361), 365-368.

Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Bridge, E. S., & Pravosudov, V. V. (2019). Smart is the new sexy: female mountain chickadees increase reproductive investment when mated to males with better spatial cognition. *Ecology Letters*, *22*(6), 897–903.

Branch, C. L., Sonnenberg, B. R., Pitera, A. M., Benedict, L. M., Kozlovsky, D. Y., Bridge, E. S., & Pravosudov, V. V. (2020). Testing the greater male variability phenomenon: Male mountain chickadees exhibit larger variation in reversal learning performance compared with females. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1931), 13–16.

Buchanan, K. L., Grindstaff, J. L., & Pravosudov, V. V. (2013). Condition dependence, developmental plasticity, and cognition: Implications for ecology and evolution. *Trends in Ecology and Evolution*, *28*(5), 290–296.

Cauchard, L., Doucet, S. M., Boogert, N. J., Angers, B., & Doligez, B. (2017). The relationship between plumage colouration, problem-solving and learning performance in great tits Parus major. *Journal of Avian Biology*, *48*(9), 1246–1253.

Cauchoix, M., Chow, P. K. Y., van Horik, J. O., Atance, C. M., Barbeau, E. J., Barragan-Jason, G., Bize, P., Boussard, A., Buechel, S. D., Cabirol, A., Cauchard, L., Claidière, N., Dalesman, S., Devaud, J.M., Didic, M., Doligez, B., Fagot, J., Fichtel, C., Henke-von der Malsburg, J., Hermer, E., Huber, L., Huebner, F., Kappeler, P. M., Klein, S., Langbein, J., Langley, E. J. G., Lea, S. E. G., Lihoreau, M., Lovlie, H., Matzel, L. D., Nakagawa, S., Nawroth, C., Oesterwind, S., Sauce, B., Smith, E. A., Sorato, E., Tebbich, S., Wallis, L. J., Whiteside, M. A., Wilkinson, A., Chaine, A. S., Morand-Ferron, J. (2018). The repeatability of cognitive performance: a meta-analysis, *Philosophical Transactions of the Royal Society B: Biological Sciences. 373*, 20170281.

Cotton, S., Fowler, K., & Pomiankowski, A. (2004). Condition dependence of sexual ornament size and variation in the stalk-eyed fly *Cyrtodiopsis dalmanni* (Diptera: Diopsidae). *Evolution*, *58*(5), 1038–1046.

Croston, R., Branch, C. L., Kozlovsky, D. Y., Dukas, R., & Pravosudov, V. V. (2015). Heritability and the evolution of cognitive traits. *Behavioral Ecology*, *26*(6), 1447–1459.

Cuervo, J. J., & Møller, A. P. (1999). Phenotypic variation and fluctuating asymmetry in sexually dimorphic feather ornaments in relation to sex and mating system. *Biological Journal of the Linnean Society*, *68*(4), 505–529.

Cuervo, J. J., & Møller, A. P. (2001). Components of phenotypic variation in avian ornamental and non-ornamental feathers. *Evolutionary Ecology*, *15*(1), 53–72.

DeCasien, A. R., Sherwood, C. C., Schapiro, S. J., & Higham, J. P. (2020). Greater variability in chimpanzee (*Pan troglodytes*) brain structure among males. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1925), 1–10.

Ferns, P. N., & Hinsley, S. A. (2008). Carotenoid plumage hue and chroma signal different aspects of individual and habitat quality in tits. *Ibis*, *150*(1), 152–159.

Gosler, A. G. (1993). *The great tit*. London: Hamlyn.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, *33*(2), 1–22.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, *269*(5220), 41–45.

Hermer, E., Cauchoix, M., Chaine, A. S., & Morand-Ferron, J. (2018). Elevation-related difference in serial reversal learning ability in a nonscatter hoarding passerine. *Behavioral Ecology*, *29*(4), 840–847.

Hermer, E., Murphy, B., Chaine, A. S., & Ferron, J. M. (2021). Great tits who remember more accurately have difficulty forgetting , but variation is not driven by environmental harshness. *Scientific Reports*, 1–11.

Houslay, T. M. & Wilson, A. J. (2017). Avoiding the misuse of BLUP in behavioural ecology: I. Multivariate modelling for individual variation. *Behavioural. Ecology* 1–16 (2017).

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: an updated perspective. *Neuroscience*, *345*, 12–26.

Janmaat, K. R. L., Boesch, C., Byrne, R., Chapman, C. A., Goné Bi, Z. B., Head, J. S., Robbins M. M., Wrangham, R. W., & Polansky, L. (2016). Spatio-temporal complexity of chimpanzee food: how cognitive adaptations can counteract the ephemeral nature of ripe fruit. *American Journal of Primatology*, *78*(6), 626–645.

Jennions, M. D., Møller, A. P., & Petrie, M. (2001). Sexually selected traits and adult survival: a meta-analysis. *The Quarterly Review of Biology*, *76*(1), 3–36.

Johnson, W., Deary, I. J., & Carothers, A. (2008). Sex differences in variability in general intelligence: a new look at the old question. *Perspectives on Psychological Science*, *3*(6), 518–531.

Johnson, W., Deary, I. J., & Carothers, A. (2009). A Role for the X Chromosome in Sex Differences in Variability in General Intelligence? *Perspectives on Psychological Science*, *4*(6), 598–611.

Kotiaho, J. S., Simmons, L. W., & Tomkins, J. L. (2001). Towards a resolution of the lek paradox. *Nature*, *410*(6829), 684–686.

Laine, V. N., Gossmann, T. I., Schachtschneider, K. M., Garroway, C. J., Madsen, O., Verhoeven, K. J., de Jager, V., Megens, H. J., Warren, W. C., Minx, P., Crooijmans, R. P. M. A., Corcoran, P., The Great Tit HapMap Consortium, Sheldon, B. C., Slate, J., Zeng, K., van Oers, K., Visser, M. E., & Groenen, M. A. (2016). Evolutionary signals of selection on cognition from the great tit genome and methylome. *Nature Communications*, *10474*, 1–9.

Lehre, A. C., Lehre, K. P., Laake, P., & Danbolt, N. C. (2009). Greater intrasex phenotype variability in males than in females is a fundamental aspect of the gender differences in humans. *Developmental Psychobiology*, *51*(2), 198–206.

Medina-García, A., & Wright, T. F. (2021). An integrative measure of cognitive performance, but not individual task performance, is linked to male reproductive output in budgerigars. *Scientific Reports*, *11*(1), 1–10.

Møller, A. P., Biard, C., Blount, J. D., Houston, D. C., Ninni, P., Saino, N., & Surai, P. F. (2000). Carotenoid-dependent signals: Indicators of foraging efficiency, immunocompetence or detoxification ability? *Avian and Poultry Biology Reviews*, *11*(3), 137–159.

Nilsson, S. R. O., Alsiö, J., Somerville, E. M., & Clifton, P. G. (2015). The rat's not for turning: Dissociating the psychological components of cognitive inflexibility. *Neuroscience and Biobehavioral Reviews*, *56*, 1–14.

Norris, K. J. (1990a). Female choice and the evolution of the conspicuous plumage coloration of monogamous male great tits. *Behavioral Ecology and Sociobiology*, *26*(2), 129–138.

Norris, K. J. (1990b). Female choice and the quality of parental care in the great tit *Parus major*, *27*, 275–281.

Pagani-Núñez, E., & Senar, J. C. (2014). Are colorful males of great tits *Parus major* better parents? Parental investment is a matter of quality. *Acta Oecologica*, *55*, 23–28.

Pomiankowski, A., & Møller, A. P. (1995). A resolution of the lek paradox. *Proceedings: Biological Sciences*, *260*(1357), 21–29.

R Core Team. (2020). *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reinhold, K., & Engqvist, L. (2013). The variability is in the sex chromosomes. *Evolution*, *67*(12), 3662–3668.

Rowe, L., & Houle, D. (1996). The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society B: Biological Sciences*, *263*(1375), 1415–1421.

Smarr, B. L., Grant, A. D., Zucker, I., Prendergast, B. J., & Kriegsfeld, L. J. (2017). Sex differences in variability across timescales in BALB/c mice. *Biology of Sex Differences*, *8*(1), 1–7.

Sonnenberg, B. R., Branch, C. L., Pitera, A. M., Bridge, E., & Pravosudov, V. V. (2019). Natural selection and spatial cognition in wild food-caching mountain chickadees. *Current Biology*, *29*, 1–7.

Spritzer, M. D., Meikle, D. B., & Solomon, N. G. (2005). Female choice based on male spatial ability and aggressiveness among meadow voles. *Animal Behaviour*, *69*(5), 1121–1130.

Tello-Ramos, M. C., Branch, C. L., Kozlovsky, D. Y., Pitera, A. M., & Pravosudov, V. V. (2018). Spatial memory and cognitive flexibility trade-offs: to be or not to be flexible, that is the question. *Animal Behaviour*, 1–8.

Thornton, A., & Boogert, N. J. (2019). Animal cognition: the benefits of remembering. *Current Biology*, *29*(9), R324–R327.

Tomkins, J. L., Radwan, J., Kotiaho, J. S., & Tregenza, T. (2004). Genic capture and resolving the lek paradox. *Trends in Ecology and Evolution*, *19*(6), 323–328.

Wyman, M. J. & Rowe, L. (2014). Male bias in distributions of additive genetic, residual, and phenotypic variances of shared traits. *Am. Nat. 184*, 326–337.

Zajitschek, S. R. K., Zajitschek, F., Bonduriansky, R., Brooks, R. C., Cornwell, W., Falster, D. S., Lagisz, M., Mason, J., Senior, A. M., Noble, D. W. A., & Nakagawa, S. (2020). Sexual dimorphism in trait variability and its eco-evolutionary and statistical implications. *ELife*, *9*, 1–17.

Zechner, U., Wilda, M., Kehrer-Sawatzki, H., Vogel, W., Fundele, R., & Hameister, H. (2001). A high density of x-linked genes for general cognitive ability: a run-away process shaping human evolution? *TRENDS in Genetics*, *17*(12), 697–701.

Chapter 4: *The relationship between foraging, cognition, and fitness in a vulnerable seabird*

Introduction

The causes and consequences of intra and interspecific variation in cognition, "or the mechanisms by which animals acquire, process, store and act on information from the environment," remain uncertain (Boogert, Madden, Morand-Ferron, & Thornton, 2018; Shettleworth, 2010). Many contemporary studies utilize the intraspecific approach which allows for insight into how cognition can directly affect an individual's behaviour and fitness, how variation is maintained, and whether selection acts on cognition in contemporary populations (reviewed in: Boogert *et al.*, 2018). Evidence for the environment influencing cognition (Ashton, Ridley, Edwards, & Thornton, 2018; e.g., Croston *et al.*, 2017; Hermer, Cauchoix, Chaine, & Morand-Ferron, 2018; Tebbich & Teschke, 2014) and cognition influencing fitness have been found (e.g., Ashton *et al.*, 2018; Madden, Langley, Whiteside, Beardsworth, & van Horik, 2018; Raine & Chittka, 2008; Shaw, MacKinlay, Clayton, & Burns, 2019), but the links between cognition, foraging ability and reproductive success are rarely studied (e.g., Raine & Chittka, 2008; Sonnenberg, Branch, Pitera, Bridge, & Pravosudov, 2019; reviewed in: Madden *et al.*, 2018). Therefore, the complete within-species link between behaviour-cognition-fitness is still not well studied.

The black-legged kittiwake (*Rissa tridactyla*) is an IUCN vulnerable, surface feeding, colonial, long-lived, central place foraging seabird whose range extends throughout the northern hemisphere (Birdlife International, 2018). The kittiwake colony that nests on Middleton Island, Alaska has been part of an almost 50-year long-term study that monitors 300-1200 banded pairs of nesting kittiwakes a year (Hatch, 2013). Approximately 20 years of detailed fledgling success, sex, age, chick body mass, adult body mass, food availability, and pedigree data have been collected (Hatch, 2013). There are low levels of nest predation and fledgling success has been

directly linked to food availability in this colony, indicating a strong selective force on foraging ability for Middleton kittiwakes (Gill & Hatch, 2002; Hatch, 2013; Kitaysky, *et al.*, 2010). Finally, GPS and accelerometers have allowed for the tracking of kittiwake foraging behaviour and have shown variation in both foraging time, and distance travelled when searching for reliable food patches of fish or cephalopods (Chivers, Hatch, & Elliott, 2016; Kotzerka, Garthe, & Hatch, 2010; Osborne *et al.*, 2020). Visual associative learning, or the ability to associate visual cues with food, is hypothesized to aid foraging kittiwakes in learning which visual cues are indicative of food sources on the surface of the water which would increase foraging success (Kitaysky, Kitaiskaia, Piatt, & Wingfield, 2006). Spatial memory is hypothesized to aid foraging animals in remembering the location of previous food patches that contain an abundance of food and return to them more efficiently (Thornton & Boogert, 2019). An accurate spatial memory could aid black legged kittiwakes in quickly returning to areas where food is abundant, such as along coastal shelves or in fish spawning areas (Hatch, 2013; Hatch, Byrd, Irons, & Hunt, 1993), and decrease travel time. However, whether these cognitive abilities are correlated to foraging behaviour in seabirds is unknown, and the connection between cognition and fitness is poorly studied. If these cognitive abilities could be measured in these kittiwakes this presents a prime opportunity to study whether learning and memory correlate to foraging efficiency in seabirds, whether these differences affect offspring and parental survival, as well as questions of assortative mating relative to cognitive ability in a long-term monitored population.

To take the first step towards answering these questions, I designed an experiment that could be used to measure learning and memory on the kittiwake's nest during the breeding season at Middleton Island. I first acclimatize the birds to receiving fish rewards through a feeding hole in the panel. I then conduct a three stage novelty response training task to measure

how long it takes the kittiwakes to acclimatize to the testing apparatus and reduce the effects of neophobia on their motor training performance. Finally, I test whether kittiwakes can complete the motor training portion of a button pressing task (e.g., Morand-Ferron, Hamblin, Cole, Aplin, & Quinn, 2015). The original goal was to continue developing this protocol by adding tests of associative learning and spatial memory. However, further testing was halted due to the ongoing COVID-19 pandemic such that a second field season could not be completed before the end of my thesis.

Methods

Species and study sites

Black-legged kittiwakes were tested on their nests (n=39 nest sites, n = 94 individuals, including non-nest residents) in Middleton Island, Alaska, United States (59.4° N, 146.3° W). The kittiwake nests are located on a derelict U.S. Air Force radar tower that has bordered wooden ledges that allow a single pair to nest, be observed through one-way mirrored glass, and be fed through a tube on the right of the nest (Figure 4.1; Collins, Hatch, Elliott, & Jacobs, 2019; Gill & Hatch, 2002). During the experiment, adult birds were captured periodically to take body weight and other physiological measurements and nests were observed three times a day to collect laying date and nest occupancy data as part of a long-term monitoring project.

Acclimatization

Kittiwakes were fed while they built their nests, between April 20, 2019- May 18, 2019, and May 23, 2019 - June 6, 2019 (feeding times: 9 am - 11 am; 2 pm – 4 pm; 6 pm – 8 pm) through the feeding hole to acclimate them to feeding (Figure 4.2). Kittiwakes were fed until they stopped taking whole capelin (*Mallotus villosus*), the fish species used to feed the kittiwakes in the tower. 15 nests from the first feeding period were fed for one extra day during the second feeding period (n= 27 individuals). During the second feeding, one window did not have consistent resident kittiwakes and was removed from testing (n = 6 individuals), and two windows had two nests and were removed from testing (n = 6 individuals). Nine birds were fed once but were not resident birds and were excluded from testing. The remaining kittiwakes were fed between 1-13 days (mean ± standard deviation = 9.97 ± 1.92), with 5-37 feedings (mean ± standard deviation = 24.24 ± 6.28) occurring over these days. These kittiwakes consumed between 0-11 whole capelin per feeding (mean ± standard deviation*;* 2.33 ± 1.54). Seven fed

windows (n = 14 individuals) were not included in the next part of the experiment due to time constraints. One window (n = 1 individual) was removed as its window became blocked before testing could start. One window switched its residents and started the test without being fed (n = 2).

Novelty Response Training

Novelty response training started after the first egg was laid, between 4-36 days after the last feeding and started at the same time as the previous acclimatization schedule (May 31 – July 11; session times: 9 am - 12 am; 2 pm – 5 pm; 6 pm – 9 pm). Kittiwakes were tested individually, while the mate was off the nest. The novelty response test had three separate stages wherein the bird had to feed in front of the panel (stage 1), feed from the open panel (stage 2), and feed from the open panel after a mirrored panel was removed from in front of the hole (stage 3; Figure 4.2). Each of these stages had to be passed in three uninterrupted testing sessions. A test session started with the kittiwake being presented one 4-6 g piece of capelin through the feeding hole for a minute (e.g., Shaw, Boogert, Clayton, & Burns, 2015). If it did not consume the capelin, the capelin was returned through the feeding hole, and if it was not taken after the second presentation, this session was considered a failure as the bird was unmotivated. After taking the food, the mirrored glass panel (Fig 4.1.) was replaced with the testing panel (Fig. 4.2), and the observer waited five minutes for the bird to return to the nest, if it flushed off the nest. If the bird did not return after five minutes, the bird failed the session. Upon the bird's return, the observer waited 1-2 minutes. After waiting, a piece of capelin was either presented through the feeding hole if it was novelty test one, or through the centre hole in the panel for novelty tests 2 and 3 for one minute. If the bird did not consume the capelin, the fish was presented again for one minute, and the session was considered a failure if it did not consume the capelin. Once the

bird had consumed the capelin in this way, three times, a piece of capelin was presented through the feeding hole and the trial was ended (e.g., Shaw *et al*., 2015). Additionally, trials were ended and considered a failure if the bird was interrupted by another bird landing on the nest or pecking the bird for more than five seconds, if the bird dropped the food off the nest, or if the trial was interrupted by technical difficulties. Two birds passed the second novelty response test twice; these extra trials were removed (n = 3 and n = 5 trials). Six pairs of birds (n = 12) were part of pre-trials, and only their performance on the first novelty test stage was included as they were tested on a variety of different methods before the current novelty test stage 2. Six birds briefly experienced an alternative novelty response test stage 2 that involved moving doors instead of the opening. Their novelty stage 2 data was removed. One bird would not return to the normal testing panel and had an alternative novelty test 1. This bird was removed. I utilized 4-6 g pieces of capelin as these were about 0.33 of a full capelin, and a kittiwake would receive only 1.67 capelin per testing session. This was just over half of the mean number of capelin an individual kittiwake would consume during one feeding period when fed to satiation. It was also much less fish (60-90g of capelin) than what they have been recorded to consume when fed to satiation three times a day (~207 g/day of herring (*Clupea pallasi*); Gill *et al*. 2002). Additionally, when utilizing this amount, kittiwakes still left the nest to forage for food, indicating that this did not represent their total nutritional needs and allowing for individual testing.

Motor Training

After passing novelty response training, the birds underwent motor training trials to train them to peck an opening in the panel to receive a piece of capelin. Motor training proceeded on the same schedule as the previous novelty response task with a free piece of capelin being presented before and after the trials. For the motor training phase, trials were passed sequentially,

and did not have to be passed consecutively within one session. During these tests, interrupted trials, not consuming the final capelin, and other issues specified above ended a session but did not count as a failure. Only not consuming the capelin during a trial was considered a failure. Each motor training stage had to be passed three times, and if birds did not pass a stage, they reverted to the previous stage, except for the first motor training stage wherein failing the first trial of it would only return them to trial 1 of the first motor training stage. For motor training stage 1, the kittiwake had to consume the piece of capelin that rested against the middle opening, after a mirrored sheet was removed. For motor training stage 2 the kittiwake had to consume a piece of capelin that was sticking half inside of the opening after a mirrored panel was removed from the centre hole. Finally, the bird had to peck the middle opening when the mirrored panel was removed to receive a piece of capelin. During the third motor training stage, one bird was erroneously given the novelty response stage 2 during the third motor training stage and these three trials were removed. Four birds were given one extra trial, one was given two extra trials, and one was given three extra trials during motor training stage 3. These trials were removed. Completing testing of all birds to motor training stage 3, or other cognitive tasks was interrupted by chick's hatching (n = 7 birds) and time constraints.

Figure 4.1. **a**. A black-legged kittiwake standing on its nest. In front of the kittiwake is the removable mirrored glass panel that was removed and replaced by the cognitive test board during testing. On the left side of the kittiwake is a feeding hole where capelin would be presented through. Feeding holes were positioned randomly on the right or left of the nest. **b**. The tower that housed the kittiwakes tested in this experiment. Photo credit: Ethan Hermer

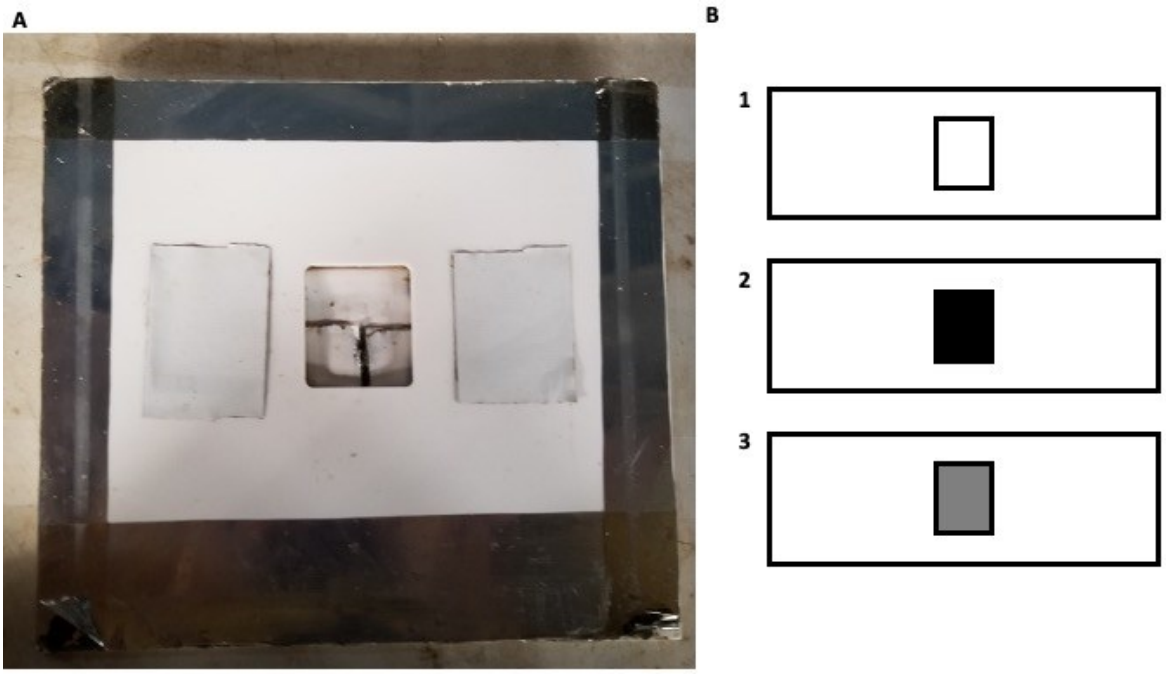


Figure 4.2. **a**. Training panel with an open centre hole. The panel is made of acrylic and is covered on the edges by reflective tape. The white testing area is 20 x 15 cm. In the centre is the duct tape hole or "button," that the bird had to peck to receive capelin. **b**. The three novelty response tests. In test 1, the opening is covered on the outside by a piece of white duct tape (white square). In test 2, the hole is open (black square), as seen in the A panel. In test 3, the opening is covered by a panel made of reflective tape (grey square).

Ethics

The experiment occurred in accordance with the protocols of the McGill University Animal Care Program (2015-7599), the Alaska Department of Fish and Game (19-142), the United States Fish and Wildlife Service (85004-C), and the University of Ottawa Animal Care Committee (protocol: 1758).

Results

Novelty Response

43/55 birds passed the first novelty response stage. These birds required between 3-30 trials (mean ± standard deviation = 10.07 ± 6.93) to pass the first novelty response stage. 22/30 birds passed the second novelty response stage with between 3-17 trials (mean ± standard deviation = 5.82 ± 4.11). 26/28 birds passed the third novelty response stage with between 3-4 trials (mean ± standard deviation 3.13 ± 0.34).

Motor Training

24/26 birds passed the first motor training stage in 3-10 trials (mean ± standard deviation: 3.88 ± 1.78). 21/24 birds passed the second motor training stage in 6-18 trials (mean ± standard deviation: 8.19 ± 3.20). 16/21 birds passed the third motor training stage in 11-41 trials (mean ± standard deviation: 21.69 ± 8.14).

Discussion

There are only a few studies investigating the links between cognition, behaviour, and fitness within one species (e.g., Ashton *et al*., 2018; e.g., Raine & Chittka, 2008; Sonnenberg *et al.,* 2019: reviewed in: Madden *et al.*, 2018). To increase the breadth of species these questions are studied in, I undertook a preliminary study to test the ability for a seabird, the black-legged kittiwake, to complete a motor training task on their nest. I found that kittiwakes were able to complete the task, with a participation level of 78% (43/55 birds completed the first novelty response stage) attained with one observer in the first field season. However, time constraints in the first field season, and then the COVID pandemic prevented me from taking enough birds to the next steps which were aimed at measuring associative learning and spatial memory.

There are a few modifications that could reduce the time constraints and allow for other learning tasks to be implement. During all stages of the task, the glass panel in front of the kittiwake's nest was removed, and an experimental panel replaced it. The only other time the glass panel would have been removed was when the kittiwakes or their offspring were captured, and it also opened the front of the nest to the inside of the tower. This caused many of the kittiwakes to leave their nest when it was removed, and the stressfulness of this is reflected in how it took around 10 and at maximum 30 trials to acclimate the birds to this action. To reduce this disruption, reduce the number of novelty response trials, and focus the novelty response test to just the panel, a semi-permanent one-way mirrored experimental panel that contains an internal small testing surface could be used. By not removing the whole panel, and instead opening a small testing surface, the kittiwakes would not be exposed to the inside of the tower, which should decrease flushing from their nest, as well as neophobia. This in turn should reduce the number of trials needed to have the kittiwakes start working on the actual cognitive task and

measure neophobia more accurately. Furthermore, the first acclimatization phase where the birds were feeding from the adjacent hole could also be removed by starting the method with feeding them from the testing surface.

Given the short time constraints for testing the task, it is difficult to measure how many birds would have truly completed the task and continued to complete an associative learning task. However, 78% of birds passed the first novelty response stage, and 62% of birds that made it to the motor training portion of the task completed the task. This level of participation is higher than what is found in many other cognitive tasks in wild animals (reviewed in: van Horik, Langley, Whiteside, & Madden, 2017). Additionally, I had planned to return with a second observer two months earlier than I had in this attempt, doubling the time spent testing and number of birds that could be tested. Finally, black-legged kittiwake chicks can complete an associative colour learning and memory task (Kitaysky, A. S., Kitaiskaia, Piatt, & Wingfield, 2003). Based on these results, I had planned to utilize a two-choice shape, spatial, or colour associative learning task modified from methods used in studies of general intelligence (e.g., Shaw *et al.*, 2015) and avian cognition (e.g., Morand-Ferron *et al.*, 2015). The kittiwakes were to be exposed to two side openings that could also be different colours or have different shapes overlaying the openings. Given the above modifications, the high participation rates, and the fact that kittiwakes have been shown to complete associative learning tasks, I believe the plan to measure associative learning and memory in black-legged kittiwakes was feasible. However, these modifications and plans were cancelled due to the COVID-19 pandemic. If successful, these tasks would have been used to measure the link between cognitive ability, detailed foraging behaviour as measured through GPS loggers (Lescroël *et al.*, 2019), chick survival, offspring survival and the across year repeatability (Boogert *et al.*, 2018) of these measures in 2020 and

2021. This task would have been the first to do so in a seabird, but also the first to have such fine scale detail when correlating foraging behaviour, cognition, and fitness.

References

Ashton, B. J., Ridley, A. R., Edwards, E. K., & Thornton, A. (2018). Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature*, *554*, 364–367.

Birdlife International. (2018). *Rissa tridactyla,* Black-legged Kittiwake. *The ICUN Red List of Threatened Species* (Vol. 8235).

Boogert, N. J., Madden, J. R., Morand-Ferron, J., & Thornton, A. (2018). Measuring and understanding individual differences in cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756).

Chivers, L. S., Hatch, S. A., & Elliott, K. H. (2016). Accelerometry reveals an impact of short-term tagging on seabird activity budgets. *The Condor*, *118*(1), 159–168.

Collins, S. M., Hatch, S. A., Elliott, K. H., & Jacobs, S. R. (2019). Boldness, mate choice and reproductive success in *Rissa tridactyla*. *Animal Behaviour*, *154*, 67–74.

Croston, R., Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Bridge, E. S., Parchman, T. L., & Pravosudov, V. V. (2017). Predictably harsh environment is associated with reduced cognitive flexibility in wild food-caching mountain chickadees. *Animal Behaviour*, *123*, 139–149.

Gill, V. A., & Hatch, S. A. (2002). Components of productivity in black-legged kittiwakes Rissa tridactyla: Response to supplemental feeding. *Journal of Avian Biology*, *33*(2), 113–126.

Hatch, S. A. (2013). Kittiwake diets and chick production signal a 2008 regime shift in the Northeast Pacific. *Mar. Ecol. Prog. Ser*, *477*, 271–284.

Hatch, S. A., Byrd, G. V., Irons, D. B., & Hunt, G. L. (1993). Status and ecology of kittiwakes (*Rissa tridactyla* and *R. brevirostris*) in the North Pacific. In *The Status, Ecology, and Conservation of Marine Birds of the North Pacific* (pp. 140–153).

Hermer, E., Cauchoix, M., Chaine, A. S., & Morand-Ferron, J. (2018). Elevation-related difference in serial reversal learning ability in a nonscatter hoarding passerine. *Behavioral Ecology*, *29*(4), 840–847.

Kitaysky, A. S., Kitaiskaia, E. V., Piatt, J. F., & Wingfield, J. C. (2003). Benefits and costs of increased levels of corticosterone in seabird chicks. *Hormones and Behavior*, *43*(1), 140–149.

Kitaysky, Alexander S., Kitaiskaia, E. V., Piatt, J. F., & Wingfield, J. C. (2006). A mechanistic link between chick diet and decline in seabirds? *Proceedings of the Royal Society B: Biological Sciences*, *273*(1585), 445–450.

Kitaysky, Alexander S., Piatt, J. F., Hatch, S. A., Kitaiskaia, E. V., Benowitz-Fredericks, Z. M., Shultz, M. T., & Wingfield, J. C. (2010). Food availability and population processes: severity of nutritional stress during reproduction predicts survival of long-lived seabirds. *Functional Ecology*, *24*(3), 625–637.

Kotzerka, J., Garthe, S., & Hatch, S. A. (2010). GPS tracking devices reveal foraging strategies of Black-legged kittiwakes. *Journal of Ornithology*, *151*(2), 459–467.

Lescroël, A., Ballard, G., Massaro, M., Dugger, K., Jennings, S., Pollard, A., Porzig, E., Shmidt, A., Varasani, A., Grémillet & Ainley, D. (2019). Evidence of age-related improvement in the foraging efficiency of Adélie penguins. *Scientific Reports*, *9*(1), 1–13.

Madden, J. R., Langley, E. J. G., Whiteside, M. A., Beardsworth, C. E., & van Horik, J. O. (2018). The quick are the dead: pheasants that are slow to reverse a learned association survive for longer in the wild. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756), 20170297.

Morand-Ferron, J., Hamblin, S., Cole, E. F., Aplin, L. M., & Quinn, J. L. (2015). Taking the operant paradigm into the field: associative learning in wild great tits. *PloS ONE*, *10*(8), e0133821.

Osborne, O. E., Hara, P. D. O., Whelan, S., Zandbergen, P., Hatch, S. A., & Elliott, K. H. (2020). Breeding seabirds increase foraging range in response to an extreme marine heatwave. *Marine Ecology Progress Series*, *646*, 161–173.

Raine, N. E., & Chittka, L. (2008). The correlation of learning speed and natural foraging success in bumble-bees. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1636), 803–808.

Shaw, R. C., Boogert, N. J., Clayton, N. S., & Burns, K. C. (2015). Wild psychometrics: Evidence for "general" cognitive performance in wild New Zealand robins, *Petroica longipes*. *Animal Behaviour*, *109*, 101–111.

Shaw, R. C., MacKinlay, R. D., Clayton, N. S., & Burns, K. C. (2019). Memory performance influences male reproductive success in a wild bird. *Current Biology*, *29*(9), 1498–1502.

Shettleworth, S. J. (2010). Discrimination, classification and concepts. In *Cognition, evolution and behaviour* (2nd ed., pp. 186–188). New York: Oxford University Press.

Sonnenberg, B. R., Branch, C. L., Pitera, A. M., Bridge, E., & Pravosudov, V. V. (2019). Natural selection and spatial cognition in wild food-caching mountain chickadees. *Current Biology*, *29*, 1–7.

Tebbich, S., & Teschke, I. (2014). Coping with uncertainty: woodpecker finches (*Cactospiza pallida*) from an unpredictable habitat are more flexible than birds from a stable habitat. *PLoS ONE*, *9*(3).

Thornton, A., & Boogert, N. J. (2019). Animal cognition: the benefits of remembering. *Current Biology*, *29*(9), R324–R327.

van Horik, J. O., Langley, E. J. G., Whiteside, M. A., & Madden, J. R. (2017). Differential participation in cognitive tests is driven by personality, sex, body condition and experience. *Behavioural Processes, 134,* 22-30.

Conclusion

My research presented here expands upon previous research on the 'harsh environment' hypothesis by comparing the serial reversal learning accuracy, spatial memory accuracy, and proactive interference performance of high and low elevation non-scatter hoarders. Additionally, by using a variance partitioning approach, I tested for the first among-individual correlation between spatial memory and proactive interference, and for variance differences in serial reversal learning performance and spatial memory accuracy in males and females. Finally, I tested a proof-of-concept methodology to measure learning and memory in a seabird. In *Chapter 1*, I show that low elevation great tits (*Parus major*) performed more accurately on a serial reversal learning task than high elevation great tits. In *Chapter 2*, I show that the relationship between elevation and spatial memory differs between hoarders and non-hoarders as high and low elevation great tits do not differ in their spatial memory accuracy or proactive interference performance. However, I do show that individual great tits who have more accurate spatial memory also experience greater proactive interference. In *Chapter 3*, I show that female great tits have greater variance than male great tits in their spatial memory accuracy, but not in their spatial serial reversal learning performance. In *Chapter 4*, I show that black-legged kittiwakes (*Rissa tridactyla*) can complete the motor training portion of a cognitive task. Overall, this thesis shows that the relationships between cognition, the environment, and sex are more species specific than previously thought. This highlights the need for expanding the breadth of species utilized by cognitive ecology to better understand not just how and why cognition varies, but where we expect the how's and why's to differ.

The 'harsh environment' hypothesis states that certain cognitive abilities may aid in coping with environments with low or variable food availability and are predicted to increase

with harshness (Pravosudov & Clayton, 2002). Previous research into the harsh environment hypothesis primarily focused on scatter hoarders (reviewed in: Pravosudov & Roth II, 2013; Sonnenberg, Branch, Pitera, Bridge, & Pravosudov, 2019) and only one other study to date has investigated this outside of the scatter hoarding paradigm (Tebbich & Teschke, 2014). To broaden the number of species and cognitive abilities this hypothesis was tested in, I performed the first comparison of serial reversal learning accuracy (*Ch. 1*: Hermer, Cauchoix, Chaine, & Morand-Ferron, 2018), spatial memory accuracy, and proactive interference performance (Hermer, Murphy, Chaine, & Ferron, 2021) in a non-scatter hoarder. I found that contrary to the 'harsh environment' hypothesis, lower elevation great tits had more accurate serial reversal learning ability, and there was no difference between high and low elevation great tits spatial memory accuracy nor their proactive interference performance. When taken together, these results do not provide evidence for the harsh environment hypothesis acting on great tit cognition in the Pyrénées mountains. However, there is still a difference among elevations, indicating that unlike in scatter hoarders (e.g., Benedict *et al.*, 2020; Roth II, LaDage, & Pravosudov, 2010), and another non-scatter hoarding species (Tebbich & Teschke, 2014), an environmental pressure other than harshness may be driving differences in spatial serial reversal learning accuracy. In *Chapter 1*, I discuss the possibility that there may be an increased density of individuals at lower elevations which could drive the elevation difference in spatial reversal learning accuracy (Hermer *et al.*, 2018). The social intelligence hypothesis states that the social pressures of living in large groups might increase cognitive abilities that are useful in maneuvering through these developmental and/or evolutionary pressures (Ashton, Thornton, & Ridley, 2018; Dunbar, 1998; but see: Holekamp, 2007). Serial reversal learning may be useful in a group as it could aid an individual in flexibly changing their behaviour as the group dynamics change, or an individual

must modify their behaviour according to which conspecifics they are interacting with (Dunbar, 1998). If this serial reversal learning ability is generally useful for navigating the difficulties of social groups, it could explain *Chapter 1's* results.

However, when combined with *Chapter 2's* results, the social intelligence hypothesis is not a clear explanation for the elevation differences in cognitive ability. The original social intelligence hypothesis proposes a general increase in all cognitive abilities with group size (Dunbar, 1998), and recent research has provided evidence that social pressures have a general effect on cognition, termed *g,* or general intelligence (e.g., Ashton, Ridley, Edwards, & Thornton, 2018). *g* is the hypothesized underpinning of all cognitive abilities, that is shown through positive correlations between all measures of cognitive ability (reviewed in: Poirier, Kozlovsky, Morand-Ferron, & Careau, 2020). However, the combined results of *Chapters 1* and *2* points towards elevation acting specifically on each cognitive ability, instead of generally as predicted by the social intelligence hypothesis, as there was a significant difference in serial reversal learning task performance but not spatial memory accuracy in high and low elevation great tits. Therefore, if the social intelligence hypothesis is the mechanism for these elevation differences, then it is not acting on general cognition, but it is acting on specific cognitive abilities. Future studies in this system should measure flock size and dynamics and compare it between high and low elevation great tits while measuring multiple cognitive abilities within the same individuals to discern whether social pressures are affecting cognitive ability in this population, and if it acts on *g* or specific cognitive abilities.

My results in *Chapters 1* and *2* may have also been due to differences in condition between high and low elevation birds (e.g., van Horik, Langley, Whiteside, and Madden, 2017). However, if condition was driving cognitive differences between the two elevations', we would

have expected to see the same elevation difference across both tests. As this is not the case, the lack of agreement between *Chapters 1* and *2* provide evidence the differences between elevations are not due to differences in condition.

In *Chapters 2* and *3* I utilized a variance partitioning approach to measure an among-individual correlation between spatial memory accuracy and proactive interference performance, and the difference in variance between male and female great tits cognitive performance. In *Chapter 2*, I found that more accurate spatial memory has a cost of increased proactive interference, and both are moderately repeatable. The spatial memory-proactive interference trade-off matches phenotypic correlations found in scatter hoarding birds (e.g., Tello-Ramos, Branch, Kozlovsky, Pitera, & Pravosudov, 2018) and lab experiments in other species (e.g., Chrobak, Hinman, & Sabolek, 2008; Malleret *et al.*, 2010), providing ample evidence for proactive interference being a consistent cost to more accurate spatial memory. The implication of this trade-off is that cognitive abilities do not just have energetic or fitness costs, but cognitive costs as well, as investing in one may hinder the other. This provides further evidence against the proposed presence of *g* which predicts an overall increase in cognitive ability (e.g., Shaw, Boogert, Clayton, & Burns, 2015; reviewed in: Poirier *et al.*, 2020). Given the costliness of cognition, it may be expected that differences in cognition between populations reflects differential investment into specific cognitive abilities (e.g., adaptive specialization hypothesis; Krebs, 1990) as opposed to general investment in cognitive ability (*g*).

Finding trade-offs also begs the question of what is being measured in more multifaceted tests like serial reversal learning tasks (Izquierdo, Brigman, Radke, Rudebeck, & Holmes, 2017; Nilsson, Alsiö, Somerville, & Clifton, 2015). I did not find the same link between serial reversal learning and elevation, as I did for spatial memory and proactive interference. Differences in

proactive interference are a proposed mechanism for differences in performance in reversal tasks (Gonzalez, Behrend, & Bitterman, 1967; Tello-Ramos *et al.*, 2018), but together my results imply that the elevation-related differences found in my serial reversal learning task were probably not due to differences in proactive interference or memory. This, taken along with a recently published study that shows no correlation between two spatial reversals that differ in scale (Troisi et al., 2021), provides further evidence that reversal learning tasks may be measuring a variety of aspects of cognitive ability instead of one particular cognitive ability (Nilson *et al*. 2015; Izquierdo *et al.* 2017). Studies like Troisi *et al*. (2021) should help better tease apart the underpinnings of reversal learning.

The moderate repeatability of spatial memory performance in my task provides preliminary evidence that some cognitive abilities may be relatively stable traits. This concurs with recent evidence of a genetic basis for spatial memory in mountain chickadees (Branch, Semenov, Wagner, Sonnenberg, Pitera, Bridge, Taylor, and Pravosudov, 2021). However, recent research in great tits shows that the number of neurons in the hippocampus, and other areas of the brain changes across seasons (Lange, Walker, Orell, & Smulders, 2021). Given this change in hippocampal neuron density, it would be interesting to study whether spatial memory is repeatable across seasons, or if individuals vary in the plasticity of their spatial memory across breeding and non-breeding seasons. It would be interesting to see if this potential change in spatial memory across season changes the trade-off between spatial memory and proactive interference. Therefore, spatial memory, proactive interference and their trade-off should also be tested across seasons to further investigate the repeatability of the trade-off. In addition, testing the repeatability of these traits and their relationships does move away from the "phenotypic gambit" (Grafen 1984) that has been utilized in these chapters. However, repeatability is the

"upper limit" of heritability, and finding that spatial memory is repeatable does not provide conclusive evidence that spatial memory has a heritable component (reviewed in Wilson, 2018). Therefore, the genetic basis of these cognitive abilities and their trade-offs should also be investigated to fully understand whether these abilities and trade-offs are truly heritable (Grafen 1984; e.g., Branch *et al.* 2021)

Another question answered through variance partitioning is how sex influences differences in cognitive ability. I utilized mixed models to partition the within-sex variance and compare it between male and female birds, directly measuring for differences in variance between the sexes. I showed that there was no difference in serial reversal learning trials to criterion variability, but females showed greater variability in spatial memory accuracy than males. Broadly, this implies that greater male variability is not the global trend in cognition, or even in birds (Branch *et al.*, 2020), but these results still provide some evidence of sex differences in cognitive ability. Additionally, we found greater female variability in the heterogametic sex, which provides preliminary evidence for the sex chromosome hypothesis, and a genetic basis of sex differences in variability that need to be tested. Given the number of cognitive tests that have occurred in wild animals with sex as a measured variable, a meta-analytic approach could be used to test for greater variability in cognitive abilities across a variety of taxa to test whether differences in variance are common across cognitive abilities in wild animals and tease apart which of these hypothesises (i.e., sexual selection or sex chromosome) is generally supported, and where they may specifically have support. This is especially important as the sparse but cumulative evidence in wild, non-human animals shows a variety of outcomes that point towards a species-specific set of outcomes as opposed to sex specific outcomes. Finally, individuals could be tested across life stages (e.g., Ashton *et al*. 2018)

and the variance partitioning approach can be used to measure the change in among-sex variance to test whether differences in female and male development may influence sex differences in variability. Overall, these three chapters show that the relationship between cognition, environmental factors and sex are probably species specific. Studies should continue to utilize variance partitioning approaches to look at how cognition correlates to environmental factors, but also to sex, and among-individuals. My results concerning these correlations do not match those found in other species, so these hypotheses should continue to be tested in a variety of species to further understand when and where they may affect cognitive ability.

The first three chapters of my thesis highlight the necessity to continue to test the causes and consequences of variation in cognitive ability across a variety of species and contexts. In *Chapter 4*, I tested the motor training portion of a task that could be used to test for learning and memory in a long-lived seabird, black-legged kittiwakes. Until now, most cognitive studies have utilized either primates, insects, or songbirds. Seabirds are under-represented in cognitive studies (e.g., Kitaysky, Kitaiskaia, Piatt, & Wingfield, 2003; Kitaysky, Kitaiskaia, Piatt, & Wingfield, 2006), which is surprising given many are long lived, take four to five years to forage effectively, and travel tens to hundreds of kilometers from their breeding sites to find difficult to perceive prey (reviewed in: Irons, 1998). Taken together, this provides a species wherein associative learning seems to be very important for their survival, as do spatial memory abilities to remember where distant patches are. Black-legged kittiwakes are especially interesting seabirds as they broadly exhibit site fidelity within years (Irons, 1998) and across years if food availability is stable (Osborne *et al*., 2020), indicating a potential component of long-term spatial memory in their foraging behaviour. They also exhibit interindividual variation in site fidelity, indicating a variety of foraging strategies that could correlate to individual variations in cognitive

ability (Irons, 1998). Kittiwakes tend to forage alone more often than in flocks, indicating a potential reliance on personal information gained through learning instead relying on social information (Irons, 1998). Finally, kittiwakes change their foraging locations throughout the day to match the tides (Irons, 1998), and within years to match local food availability indicating a temporal aspect to their memories (Osborne et al., 2020). Specifically on Middleton Island, Alaska, foraging and finding food is the determining factor in chick survival (Hatch, 2013; Irons, 1998). Therefore, it should be expected in this population that kittiwakes that have accurate and long-lasting visuo-spatial learning and memory abilities that may aid them in foraging could have greater chick survival. The importance of the specific Middleton Island study system compared to others is that very detailed foraging behaviour for each individual can be measured through GPS and accelerometry tracking devices and correlated to their cognitive performance and offspring fitness. As opposed to other species, which have inferred connections between cognition, foraging behaviour, and their offspring's fitness (reviewed in: Madden, Langley, Whiteside, Beardsworth, & van Horik, 2018). Additionally, by measuring an individual's cognitive ability and foraging strategies, we may see that it's not necessarily one set of cognitive abilities that are useful, but various combinations of cognition and behaviour that are useful for kittiwakes. Heat waves and other climactic events can lead to lasting changes in foraging behaviour (Osborne *et al*., 2020) in foraging kittiwakes and it may be possible to measure the effects changing foraging conditions have on foraging behaviour and cognition. No study system has yet to check all these boxes at once, indicating the importance of focusing on Middleton Island black-legged kittiwakes for future cognitive studies on the causes and consequences of interindividual variation in cognitive ability.

Overall, this thesis shows that the relationship between cognition, the environment and sex is complicated, and much more specific than previously thought. Additionally, it utilizes increasingly popular variance partitioning techniques to do so. Finally, it tests a battery of cognitive tasks in a novel study system that is ripe with opportunity for answering the questions concerning the causes and consequences of cognitive variation. The evidence as it stands, points towards a mosaic of possibilities for how and why these differences exist. To further disentangle these relationships, new species and powerful variance portioning approaches should be utilized. With more study, we may find that the causes and consequences of individual variation in cognitive ability could be as unique as the species themselves.

References

Ashton, B. J., Ridley, A. R., Edwards, E. K., & Thornton, A. (2018). Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature*, *554*, 364–367.

Ashton, B. J., Thornton, A., & Ridley, A. R. (2018). An intraspecific appraisal of the social intelligence hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756), 20170288.

Benedict, L. M., Pitera, A. M., Branch, C. L., Kozlovsky, D. Y., Sonnenberg, B. R., Bridge, E. S., & Pravosudov, V. V. (2020). Elevation-related differences in annual survival of adult food-caching mountain chickadees are consistent with natural selection on spatial cognition. *Behavioral Ecology and Sociobiology*, *74*(4).

Branch, C. L., Sonnenberg, B. R., Pitera, A. M., Benedict, L. M., Kozlovsky, D. Y., Bridge, E. S., & Pravosudov, V. V. (2020). Testing the greater male variability phenomenon: Male mountain chickadees exhibit larger variation in reversal learning performance compared with females. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1931), 13–16.

Branch, C. L, Semenov, G. A., Wagner, D. N., Sonnenberg, B. R., Pitera, A. M., Bridge, E. S., Taylor, S. A.& Pravosudov V. V. (2021). The genetic basis of spatial cognitive variation in a food-caching bird. *Curr. Biol.* 1–10.

Chrobak, J. J., Hinman, J. R., & Sabolek, H. R. (2008). Revealing past memories: proactive interference and ketamine-induced memory deficits. *Journal of Neuroscience*, *28*(17), 4512–4520.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, *9*(10), 178–190.

Gonzalez, R. C., Behrend, E. R., & Bitterman, M. E. (1967). Reversal learning and forgetting in bird and fish. *Science.*, *158*, 519–521.

Grafen, A. (1984). Natural selection, kin selection and group selection. *Behavioural Ecology: an Evolutionary Approach,* 62–84.

Hatch, S. A. (2013). Kittiwake diets and chick production signal a 2008 regime shift in the Northeast Pacific. *Mar. Ecol. Prog. Ser*, *477*, 271–284.

Hermer, E., Cauchoix, M., Chaine, A. S., & Morand-Ferron, J. (2018). Elevation-related difference in serial reversal learning ability in a nonscatter hoarding passerine. *Behavioral Ecology*, *29*(4), 840–847.

Hermer, E., Murphy, B., Chaine, A. S., & Ferron, J. M. (2021). Great tits who remember more accurately have difficulty forgetting , but variation is not driven by environmental harshness. *Scientific Reports*, 1–11.

Holekamp, K. E. (2007). Questioning the social intelligence hypothesis. *Trends in Cognitive Sciences*, *11*(2), 65–69.

Irons, D. B. (1998). Foraging area fidelity of individual seabirds in relation to tidal cycles and flock feeding. *Ecology*, *79*(2), 647–655.

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: an updated perspective. *Neuroscience*, *345*, 12–26.

Kitaysky, A. S., Kitaiskaia, E. V., Piatt, J. F., & Wingfield, J. C. (2003). Benefits and costs of increased levels of corticosterone in seabird chicks. *Hormones and Behavior*, *43*(1), 140–149.

Kitaysky, Alexander S., Kitaiskaia, E. V., Piatt, J. F., & Wingfield, J. C. (2006). A mechanistic link between chick diet and decline in seabirds? *Proceedings of the Royal Society B: Biological Sciences*, *273*(1585), 445–450.

Krebs, J. R. (1990). Food-storing birds: adaptive specialization in brain and behaviour? *Phil. Trans. R. Soc. Lond. B*, *329*, 153–160.

Lange, H., Walker, L., Orell, M., & Smulders, T. V. (2021). Seasonal changes in the hippocampal formation of hoarding and non-hoarding tits. *Learning and Behavior*.

Madden, J. R., Langley, E. J. G., Whiteside, M. A., Beardsworth, C. E., & van Horik, J. O. (2018). The quick are the dead: pheasants that are slow to reverse a learned association survive for longer in the wild. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756), 20170297.

Malleret, G., Alarcon, J. M., Martel, G., Takizawa, S., Vronskaya, S., Yin, D., Chen, I. Z., Kandel, E. R., & Shumyatsky, G. P. (2010). Bidirectional regulation of hippocampal long-term synaptic plasticity and its influence on opposing forms of memory. *Journal of Neuroscience*, *30*(10), 3813–3825.

Nilsson, S. R. O., Alsiö, J., Somerville, E. M., & Clifton, P. G. (2015). The rat's not for turning: Dissociating the psychological components of cognitive inflexibility. *Neuroscience and Biobehavioral Reviews*, *56*, 1–14.

Osborne, O. E., Hara, P. D. O., Whelan, S., Zandbergen, P., Hatch, S. A., & Elliott, K. H. (2020). Breeding seabirds increase foraging range in response to an extreme marine heatwave. *Marine Ecology Progress Series*, *646*, 161–173.

Poirier, M. A., Kozlovsky, D. Y., Morand-Ferron, J., & Careau, V. (2020). How general is cognitive ability in non-human animals? A meta-analytical and multi-level reanalysis approach. *Proceedings of the Royal Society B: Biological Sciences*, *287*, 20201853.

Pravosudov, V. V., & Clayton, N. S. (2002). A test of the adaptive specialization hypothesis: population differences in caching, memory, and the hippocampus in black-capped chickadees (*Poecile atricapilla*). *Behavioral Neuroscience*, *116*(4), 515–522.

Pravosudov, V. V., & Roth II, T. C. (2013). Cognitive ecology of food hoarding: the evolution of spatial memory and the hippocampus. *Annual Review of Ecology, Evolution, and Systematics*, *44*, 173–193.

Roth II, T. C., LaDage, L. D., & Pravosudov, V. V. (2010). Learning capabilities enhanced in harsh environments: a common garden approach. *Proceedings of the Royal Society B*, *277*, 3187–3193.

Shaw, R. C., Boogert, N. J., Clayton, N. S., & Burns, K. C. (2015). Wild psychometrics: Evidence for "general" cognitive performance in wild New Zealand robins, *Petroica longipes*. *Animal Behaviour*, *109*, 101–111.

Sonnenberg, B. R., Branch, C. L., Pitera, A. M., Bridge, E., & Pravosudov, V. V. (2019). Natural selection and spatial cognition in wild food-caching mountain chickadees. *Current Biology*, *29*, 1–7.

Tebbich, S., & Teschke, I. (2014). Coping with uncertainty: woodpecker finches (*Cactospiza pallida*) from an unpredictable habitat are more flexible than birds from a stable habitat. *PLoS ONE*, *9*(3).

Tello-Ramos, M. C., Branch, C. L., Kozlovsky, D. Y., Pitera, A. M., & Pravosudov, V. V. (2018). Spatial memory and cognitive flexibility trade-offs: to be or not to be flexible, that is the question. *Animal Behaviour*, 1–8.

Troisi, C. A., Cooke, A. C., Davidson, G. L., de la Hera, I., Reichert, M. S., & Quinn, J. L. (2021). No evidence for cross-contextual consistency in spatial cognition or behavioral flexibility in a passerine. *Animal Behavior and Cognition*, *8*(3), 446–461.

van Horik, J. O., Langley, E. J. G., Whiteside, M. A. & Madden, J. R. (2017). Differential participation in cognitive tests is driven by personality, sex, body condition and experience. *Behav. Processes, 134*, 22–30.

Wilson, A. J. (2018). How should we interpret estimates of individual repeatability? *Evol. Lett., 2*, 4–8.

Appendix

Appendix 1: *List of co-authored articles not utilized as chapters in the thesis.*

Morand-Ferron, J., **Hermer, E.,** Jones, T. B. & Thompson, M. J. (2019). Environmental variability, the value of information, and learning in winter residents. *Animal Behavior, 147*, 137–145.

Cauchoix, M., Chow, P. K. Y., van Horik, J. O., Atance, C. M., Barbeau, E. J., Barragan-Jason, G., Bize, P., Boussard, A., Buechel, S. D., Cabirol, A., Cauchard, L., Claidière, N., Dalesman, S., Devaud, J.M., Didic, M., Doligez, B., Fagot, J., Fichtel, C., Henke-von der Malsburg, J., **Hermer, E.**, Huber, L., Huebner, F., Kappeler, P. M., Klein, S., Langbein, J., Langley, E. J. G., Lea, S. E. G., Lihoreau, M., Lovlie, H., Matzel, L. D., Nakagawa, S., Nawroth, C., Oesterwind, S., Sauce, B., Smith, E. A., Sorato, E., Tebbich, S., Wallis, L. J., Whiteside, M. A., Wilkinson, A., Chaine, A. S., Morand-Ferron, J. (2018). The repeatability of cognitive performance: a meta-analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences., 373*, 20170281.

Cauchoix, M., **Hermer, E**., Chaine, A. S. & Morand-Ferron, J. (2017). Cognition in the field: comparison of reversal learning performance in captive and wild passerines. *Scientific Reports, 7*, 1–10.