

Behavioral Diversity in Mixed-Species Flocks

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Scott Alan Benson

December 2024

ACKNOWLEDGEMENTS

So much of the road to completing this dissertation was paved in unexpected ways by unexpected people. Of course, I want to thank my parents, Beverly and Dwayne, who always managed to give just the right things at just the right time even when they couldn't give to themselves. Mom gave me such a leg up tutoring me on division and multiplication of all things before I even started kindergarten that I never felt like school was a burden or barrier in my life. Dwayne, who raised me as his own, overcame so many of his personal demons just to be in my life when I needed him. He taught me to be confident in every crowd but to lead with humility wherever I went.

Thank you to Dr. Christian Sawyer and Dr. Seth Swihart: you saved me from myself in high school. Being smart isn't enough. Coasting gets you nowhere good. Without your intervention I never would have developed the skills I needed for an academic career. To Anna Petroff, Dr. Hartman, and Dr. Brinthaup, I can't think of my undergraduate years without thinking of my favorite teachers; thank you for seeing me and treating me as though someday being your peer was realistic.

Thank you to the countless administrators, mechanics, loan officers, and anyone else in the world who helped me to push this far despite my abysmal financial situation over the years. Each of you were windows whenever a door was slammed in my face, and I wish I knew the names of every one of you going back nearly 20 years. An immense thank you to Connie Ogle at the University of Tennessee; I hadn't the slightest clue how graduate school worked and without your help I would have flunked out on any number of technicalities.

Thank you to Dr. Todd Freeberg. I couldn't have asked for a better mentor, and the longer I spent in your lab the more I appreciated how lucky I was to be there. Where other graduate students feel overworked, overwhelmed, and sometimes can't stick it out, I instead had a mentor that exuded support and insight and understanding. You once texted me the day after I told you I had a migraine just to check if I'd want you to feed the birds for me, and that gesture perfectly captures why I'm graduating where others don't. And you're just a cool dude.

Thank you to all my lab mates, past and present. Dr. Brittany Coppinger and Dr. Steven Kyle were great mentors in their own right, always available day or night (or after graduation) to answer some methodological question. Steven, you tried to warn me banding entire flocks wouldn't be feasible and I didn't listen; I'll never doubt your wisdom again. Thank you to Heather Brooks, Eric Frazier, Harry Pepper, Dr. Jeff Kelly, Dr. Hwayoung Jung, Colton Adams, Ryan Risner, and Zaharia Selman for all your time spent reading and commenting on every academic word I wrote for the last 6 years – I'll repay the favor any time you need in perpetuity.

Thank you to my doctoral committee, Dr. Gordon Burghardt, Dr. Elizabeth Derryberry, Dr. Garriy Shteynberg, and Dr. David Buehler. I'm a man with big ideas and it's only with your invaluable feedback that those ideas were able to take the shape of this dissertation.

But above all I want to thank my wife and partner of nearly 20 years, Renée Benson. Anything anyone else has done to support me, encourage me, see me, and make me a better, smarter, stronger, or wiser person you have done a thousand times more. Any risk I took in life was no risk at all, because at the end of the day I would still have you and that's everything to me.

ABSTRACT

Mixed-species groups are known to provide foraging and anti-predator benefits to group members. These benefits may be due to an increase in behavioral diversity unique to mixed-species groups. This dissertation examines boldness behavioral differences among individuals in mixed-species flocks of Carolina chickadees (*Poecile carolinensis*) and tufted titmice (*Baeolophus bicolor*). The studies in this dissertation applied the Behavioral Syndromes framework (personality-like components of correlated behaviors), principal component analysis, cluster analysis, and multiple-response permutation procedures to examine group differences in behavioral variability between species and flock composition conditions (mixed- vs single-species flocks). These identified behavioral types became predictors for individual behavior involving a novel foraging task, an approach that differs with more typical single-test predictive models common in the literature. Finally, to test for differences in social-context versus isolated-context behavioral testing, a Novel Environment Test was given to two different sets of individual tufted titmice, one set surrounded by a familiar flock and one set removed from theirs. Results indicated that mixed-species groups do show greater variability in their boldness principal component scores, with chickadees in mixed-species groups showing substantially more variation in their boldness behaviors than either titmice or single-species chickadee flocks. These same principal component scores out-performed all single-test predictors of novel foraging behaviors, finding that the two species differed slightly in the types of neophilia that correlated positively with novel foraging success. Between-group comparisons confirmed that the Novel Environment test depended on social/asocial context, possibly due to effects of sudden isolation and social inhibition. These findings confirm the expected condition of increased behavioral diversity in mixed-species flocks and effects of isolated testing, helping to explain the apparent diversity effect in novel foraging efficiency found in previous studies.

TABLE OF CONTENTS

Chapter 1: Introduction	1
Behavioral syndromes and individual differences.....	1
Methodological considerations.....	3
Isolation vs social context.....	3
Repeatability	4
Matching measurements to constructs	5
Study system.....	6
Hypotheses.....	6
H1: Behavioral syndromes correlate with individual novel foraging success	6
H2: Behavioral syndromes differ between species in a mixed-species flock.....	7
H3: Performance on a novel environment test differs between social and asocial contexts	7
Methods	7
Subjects, capture, and housing.....	7
Aviary testing procedure.....	9
Statistical analyses	Error! Bookmark not defined.
Behavioral Syndromes:	Error! Bookmark not defined.
Conclusion	Error! Bookmark not defined.
Dissertation outline.....	Error! Bookmark not defined.
Chapter 2: Do behavioral syndromes predict behavior better than individual tests?	15
Introduction	15
Methods	17
Subjects and housing.....	17
Materials and procedures	17
Novel feeder test	17
Mealworm test	18
Novel object test	18
Predator tests	19
Novel environment test	19
Statistical analysis	20
Results	21
Principal component analysis	21
Component 1 – Neophobia	21
Component 2 – Non-exploration	24

Component 3 – Neophilic threat discrimination	24
Component 4 – Social information preference	25
Component 5 – Neophobic exploration	25
Model comparisons	25
Identified predictors	30
Discussion.....	30
Chapter 3: Behavioral syndromes in experimental mixed-species flocks	33
Introduction	33
Methods	36
Subjects and housing.....	36
Materials and procedures	36
Behavioral syndromes testing	37
Statistical analysis.....	37
Results	39
Cluster analysis	41
MRPP	41
Discussion.....	41
Chapter 4: Asocial testing of social species.....	46
Introduction	46
Methods	47
Subjects and housing.....	47
Materials	47
Procedure	47
Results	48
Discussion.....	49
Chapter 5: Conclusion.....	52
Summary.....	52
Broader considerations and future directions	53
Producer-scrounger dynamics.....	53
Behavior diffusion	54
Cognitive aptitude	55
Conclusion	55
References.....	56
Vita.....	66

List of Figures

Figure 1: UTFRREC map of feeder locations	8
Figure 2: Test stimuli for RFID measures	12
Figure 3: Photo of the novel environment test cage.	13
Figure 4: Novel feeder approach order by ranked predictor variables.	28
Figure 5: Novel feeder approach order by principal component ranked scores	29
Figure 6: Order latency by flock composition and species	40
Figure 7: Dendrogram from hierarchical cluster analysis.....	42
Figure 8: Histogram and density of principal component scores by species	43
Figure 9: NET scores in asocial vs social contexts.....	50

LIST OF TABLES

Table 1: Summary of behavioral tests and measures.....	10
Table 2: Summary statistics of test measures by species.....	22
Table 3: Top-contributing variables to selected components	23
Table 4: Summary of predictors from ordinal logistic regression models.....	26
Table 5: McFadden's Adjusted Pseudo-R2 per predictive model	27

CHAPTER 1: INTRODUCTION

While sociality in the animal kingdom has been the focus of many decades of research, interest in mixed-species groups has only accelerated recently. A landmark review of over forty years of mixed-species groups research culminated in “Mixed company: a framework for understanding the composition and organization of mixed-species animal groups” by fifteen preeminent researchers in that field (Goodale et al., 2020). The framework includes classifications for different interspecific relationships in groups, theorizes on the evolutionary catalysts for mixed-species living, and dedicates significant effort to classifying the types of benefits that can accrue in mixed-species groups. However, recent discoveries about potential benefits of MSGs to solving novel problems in a foraging context (Freeberg et al., 2024; Freeberg et al., 2017) are considered outside the scope of the framework posed by Goodale et al. (2020). Questions of behavioral syndromes, consistent individual differences in correlated behaviors akin to personality traits, are also explicitly excluded from consideration in the final framework:

“Individuals of different sexes, ages, body conditions, or personality types (e.g. axes of boldness, exploration and aggressiveness) may accrue or provide different benefits or costs within any group, and particularly in mixed-species groups... However, understanding all interactions within groups is beyond our scope... In the future, empirical tests of the relative extent of intraspecific versus interspecific variation in behaviour or benefit provisioning are needed under different conditions to test their relative importance in structuring MSGs.”

(Goodale et al., 2020, p. 5)

The purpose of this dissertation is to explore the cognitive and behavioral dynamics of mixed-species groups. Two primary aims are to better understand 1) how behavioral syndromes are defined and distributed across closely related species, and 2) how these behavioral syndromes correlate with variation in response to novel foraging contexts. Two secondary aims of this dissertation are 1) to test the efficacy of behavioral tests given to individuals of highly social species in social contexts rather than in isolation, and 2) to test a novel approach for assessing the consistency of behavioral syndromes, such that multiple tests within an identified behavioral set replace repeated testing of individual behaviors.

Behavioral syndromes and individual differences

Behavioral syndromes are suites of correlated behaviors within a species, similar to personality traits (Sih, Bell, Johnson, et al., 2004). Just as human personality research works to identify the dimensions of behavior across which individuals tend to differ, behavioral syndromes research focuses on patterns of behavior observed as a phenotype that may or may not have a direct opposite in other phenotypes within a species. For example, a “bold” behavioral syndrome in a songbird may include high scores and positive correlations between tests of exploration, anti-predator, and dominance behaviors. A “shy” behavioral syndrome, rather than simply scoring low on the same correlated tests, might display a unique pattern of correlated behaviors such as negative response to novelty, scrounging more than producing, and/or slower

exploration (Bergmuller, 2010). Through the use of behavioral syndromes approach, we can understand and categorize individuals according to the types of correlated behaviors they exhibit, predict which behaviors may be correlated in closely related species, and understand the true diversity of behaviors that can be reasonably expected in a species with one or more common behavioral syndromes (Carter et al., 2013; Sih, Bell, & Johnson, 2004).

The ontogeny of behavioral syndromes is explained by several theories, but the evidence available is mixed. One explanation is *state-dependent feedback loops*, where an initial difference in states produces a behavioral difference between two individuals, which in turn changes the state for each individual, which informs subsequent behaviors, and so on (Mathot et al., 2021; Sih et al., 2015). For example, if two birds in a flock have two different levels of hunger but equal risk of predation, the hungrier bird will likely decide to take that risk first. The presence of the hungry bird diffuses predation risk for the second bird, so that bird forages, too. But since the second bird was not as hungry, it is satiated first and leaves the area, increasing predation risk for the remaining bird. The hungrier bird may have foraged enough to make the predation risk too high to tolerate, but again finds itself less satiated than the second bird as it leaves the area. This hungrier bird will consistently perform as the risk-taker in the dyad, driven by a difference in initial state rather than any genuine difference in phenotype (Sih et al., 2015).

Recent research has struggled to find clear and consistent support for the state-dependent feedback hypothesis (Macgregor et al., 2021; Naslund, 2021; Sakai, 2020). Some of the only support for the hypothesis came from studies of competition for territory and resources (Fokkema et al., 2021). However, a critical aspect of the idea, as proposed in Sih et al. (2015), is the social context for this feedback. A trend in the literature seems to be that state-behavior feedbacks were not detected when the methodology involved tests in isolation, but state-behavior feedbacks were present when individuals were tested in a more natural social context (Sih et al., 2015). This may be because interaction from individuals of different states contributed to the maintenance or exacerbation of the differences in behavior between those individuals.

An alternative view is the “*pace-of-life*” syndromes, also called growth-mortality tradeoffs. This idea focuses on phenotypic trade-offs between increased chances of survival against increased chances of reproduction (Stamps, 2007). Rather than an initial state difference influencing behavior alone, this state difference produces a direct change in phenotype that is often irreversible, such as earlier sexual development or a change in overall body size (Brommer & Class, 2017). This is conceptually similar to *r/K* selection theory (Pianka, 1970).

The straightforward view put forward by Sih et al. (2015) captures an intuitive relationship between state and motivation, while the strength of the pace-of-life view is in explaining apparently contradictory behaviors across their respective contexts. For example, a “bold” individual could be assumed to be hungrier or more desperate for a mating opportunity in the state-dependent feedback view, taking big risks in these contexts; alternatively, proneness or aversion to risk could be viewed as a calculation of reproducing at all costs today or surviving for better opportunities tomorrow, which better captures situations where an individual is bold only in some contexts and not others. However, a recent study testing both the pace-of-life and state-feedback hypotheses failed to find support for either (Petelle et al., 2019). The authors tested yellow-bellied marmots (*Marmota flaviventris*) on a variety of behaviors over the course of their active season and found that behavior did not seem to correlate strongly with changes in age and body weight in any pattern predicted by either theory. While it remains likely that both

theories offer some explanatory power, a prudent approach may be to continue to focus on establishing the structure of each study species' behavioral syndromes, as the ontogeny of those behavioral syndromes may differ across taxa and only be apparent with taxonomic comparisons at a later date.

Regardless of the ontogenetic cause of behavioral syndromes, their existence does allow for questions of adaptability and evolution. Dochtermann and Dingemanse (2013) conducted an analysis on available data of known genetic variance across several species and concluded that behavioral syndromes create massive constraints on evolutionary responses, more so even than life-histories or morphologies of those species. This finding underscores the importance of not only learning what we can about species-typical behaviors, but how those behaviors are distributed (as syndromes) across the population.

Methodological considerations

There are several methodological considerations relevant to behavioral syndromes, and to each topic associated with the mixed-species group context. This section will enumerate these considerations as they relate to the current dissertation and discuss the merit of the approaches used in this dissertation in the context of the current literature.

Isolation vs social context

Nearly all behavioral assays in the current literature are given to a single individual at a time in an asocial context, unless the test is specifically one that involves interaction with other individuals (Bell et al., 2009). Of all the controls thought of as beneficial and necessary in laboratory settings, this asocial context test may serve only to reduce ecological validity of the tests for many social species.

Among the few examples of testing whole groups at once, important revelations for animal personality have come to light. Olafsdottir and Magellan (2016) compared scores on behavioral assays across isolated and social contexts for a sample of three-spined sticklebacks (*Gasterosteus aculeatus*), finding that individuals that were identified as shy in the isolated tests became much bolder when tested in social contexts. If the study of stickleback populations relied only on data from these tests in asocial contexts, then the expected distribution of personality traits in the wild would appear to be wildly inaccurate, with the true range in ecologically valid contexts being much more restricted and bolder overall. It is possible that isolation itself is a confound, a stressor that differentially affects individuals of different species or personality types.

Aside from potentially changing or misjudging the behavior of a given sample through isolated testing, social context testing can reveal aspects of the natural behavioral repertoire that tests in isolation miss. By testing small groups, Zandberg et al. (2017) were able to show that personality traits can predict problem-solving behaviors in great tits. Personality traits can predict which members of the group may act first in response to novelty, but an important effect of this social testing was to change *what* the second individual to act was reacting to. The situation was no longer entirely novel, but was mostly novel plus the information produced by the first to act, and so on for the third to act, etc. An asocial context test would focus only on

latency to respond to a novel stimulus, which is only arguably ecologically relevant for the individual who habitually approaches novelty before all others.

If tests of behavior in captivity are to depict behavior in the wild with accuracy, then our methods ought to reflect this for those research questions that are concerned with high ecological validity. In the realm of animal personality, where the extreme ends of each personality trait are determined through comparison to other individuals, where the second to act necessitates a first, I will explore behavioral syndromes in social contexts using presentations of behavioral assays to groups.

Repeatability

The study of animal personality involves *consistent* individual differences in behavior. Repeatability across contexts is the simplest way to distinguish trait differences from state differences in behavior between individuals and has been the generally accepted approach for many years (Carter et al., 2013; Griffin et al., 2015; Huang et al., 2018). However, with behavioral syndromes growing as an area of study under the animal personality umbrella, extensive repetition of tests may not be the only, or optimal, methodological option for identifying behavioral syndromes within species.

Establishing the repeatability of behavior through repeated testing is straight-forward, but how many repetitions are enough? A meta-analysis by Bell et al. (2009) compiled over 700 repeatability scores across 114 studies and nearly as many species. While repeatability tended to differ across sex and context, a major finding of the meta-analysis was that testing an individual more than twice offers little added benefit for finding how repeatable behaviors tend to be within a species. This is a surprising finding, especially as the typical approach in the research analyzed included an average of 4.41 repeated measures per individual. Unintuitively, repeating the measure just once is not in aggregate of a single repetition testing the repeatability of behavior; repetitions across the entire sample contribute to the same estimate of repeatability such that average rate of consistency across individuals within a species make the repeatability detectable even where individuals might show only moderately repeatable behaviors. Contrary to the assumptions of common practice, unreliable behaviors can be detected with just two iterations of testing (Bell et al., 2009; Cauchoux et al., 2018).

Besides needing fewer repetitions to establish repeatability than we may have thought, a recent study by Huang et al. (2018) employed two other techniques that may streamline behavioral syndrome research. First, instead of collecting repeated behavioral measures from every individual, these researchers only gave repeated measures to 20 of the 100 Northern cardinals (*Cardinalis cardinalis*) they tested. The reasoning was that if a behavior on a particular assay is repeatable for a species, we need not test every individual of that species for repeatability. For the 14 measures of boldness behaviors, repeatability ranged from .35 to .99. The second technique, more specific to behavioral syndromes research than to animal personality as a whole, involved relying on the pattern of correlations between behaviors in order to find consistent individual differences in behavioral syndromes. Due to the nature of behavioral syndromes, where one behavior may predict several others and vice versa, testing multiple dimensions of a given behavioral syndrome is tantamount to repeated tests of that behavioral syndrome. In other words, because within-individual differences on a suite of behavioral assays are partially constrained by non-independence of behaviors, repeatability of the behavioral syndrome as a whole can be tested with *different* assays of behavior in much the same way that

repeatability of individual traits can be tested with *repeated* assays. The studies in the current dissertation follow from these findings. Only a subset of each study species were subjected to repeated measures, and the greater emphasis in the analysis was on behavior syndromes reinforcing themselves through correlated behaviors rather than repeated behaviors.

Matching measurements to constructs

The history of animal behavior and personality research has produced an array of behavioral assays and trait labels, each having different interpretations at different periods, in different contexts, or for different species. This variance in interpretations is called a “jingle-jangle” problem (Carter et al., 2013; Gosling, 2001). “Jingle” problems are those where a single trait label or test might actually be exploring any number of component behaviors; behavioral tests of “boldness” and “exploration” could be inadvertently measuring motivations or activity levels (Cussen, 2017), or even confusing exploration and boldness behaviors if the two are interpreted differently. A “jangle” problem is the reverse, where multiple tests and labels exist for a single underlying component of behavior; testing and labeling reactions to novel objects, novel environments, and novel foods separately may all be tests of a single trait of neophobia/neophilia.

Jingle-jangle problems pose a particular threat to correlations in behavioral syndromes. Many behavioral assays have been alternately used in personality and cognition studies, with different interpretations for the measures (Audet & Lefebvre, 2017; Dingemanse et al., 2010; Krasheninnikova et al., 2020; Olafsdottir & Magellan, 2016; Stamps & Biro, 2016; Tello-Ramos et al., 2019). Even with the knowledge that cognition and personality can often interact to produce consistent individual differences, it is still rare to attempt to control for one when testing the other. For example, Cauchoix et al. (2018) cautioned that the repeatability of cognitive performance in their meta-analysis did not control for differences in repeatability due to personality or dexterity. This could look like an obvious oversight in the field, since several studies focus specifically on the correlation between performance on cognitive tests and personality traits (Guillette et al., 2015; Sih & Del Giudice, 2012). For example, even a subtle difference in initial behavior can change the likelihood or frequency of encountering certain types of information, which may create apparent correlations between personality and cognition (Guillette et al., 2017). Even if it becomes clear that consistent differences in cognition can be operationalized as differences in apparent strategy, differences in aptitude or general intelligence (or experience, or preference) should also be taken into account (Rowe & Healy, 2014).

The problem may sometimes be even broader, in that perhaps *any* measure of “boldness” as a construct does not, in itself, prove that boldness exists as a concrete function in the tested species. This issue, *reification*, is a many-decades-old caveat in social psychology research (Berger & Pullberg, 1965). Reification is a logical fallacy that can allow behavioral measures and the language applied to those measures by researchers to be understood as existing in a concrete sense. In reality, our constructs (like “boldness”) are shorthand descriptions of behavioral patterns we detect through those behavioral measures (Adams & Markus, 2001; Machery, 2014).

Study system

My study system is a population of naturally occurring mixed-species flocks located in the UT Forest Resources, agResearch, and Education Center (UTFRREC) in Oak Ridge, TN. The focal species are tufted titmice (*Baeolophus bicolor*) and Carolina chickadees (*Poecile carolinensis*), which are the most abundant – and the nuclear species – in these flocks. This is an ideal local study system given the wealth of related research done on these and closely related species (Cantwell et al., 2016; Cimprich & Grubb Jr, 1994; Freeberg et al., 2014; Kyle & Freeberg, 2016; Nolen & Lucas, 2009). They are easy to capture and house and participate readily in behavioral measures. The mixed-species flocks form in the overwintering period and remain largely consistent within each season, and these two species regularly forage together during this time. Tufted titmice are considerably larger than Carolina chickadees (17-26 g vs 8-12 g) and can easily supplant chickadees during foraging, although differences in actual aggression aren't well studied. Some larger species serve as satellites and occasionally appear alongside titmice and chickadees during foraging; white-breasted nuthatches (*Sitta carolinensis*) most frequently, hairy woodpeckers (*Leuconotopicus villosus*), downy woodpeckers (*Picoides pubescens*), and red-bellied woodpeckers (*Melanerpes carolinus*) less frequently. Satellite species are usually far outnumbered by the focal nuclear species and their relative influence should be minimal (Goodale et al., 2020).

Hypotheses

Through a series of behavioral tests within large outdoor aviaries I tested the influence of behavioral syndromes in the social context of a mixed-species flock. Where previous research examined effects of either species composition or personality in wild bird flocks, I explored how personality composition interacted with species composition to predict how natural flocks react to threat, novelty, and foraging puzzles. Behavioral tests included a novel feeder test (exposure to food in a treadle trap), a novel object test (food near a red plastic cup), two predator tests (food near rubber snake and plastic hawk), and a novel environment test (movement tracked in an unfamiliar cage).

My dissertation research program tested three hypotheses. First, behavioral type differences between individuals can predict which individuals succeed with a novel foraging task. Second, behavior in mixed-species groups varies both within and between species. Finally, behavioral testing in isolation confounds the test measure with the additional stressor of isolation. I predicted that tufted titmice and Carolina chickadees, despite being closely related species with similar ecologies, would each contribute unique patterns of behavior to mixed-species flocks, that these patterns of behavior would predict interactions with a novel foraging task for each individual, and that tufted titmice tested in social contexts would be generally bolder on a novel environment test than titmice tested in isolation.

H1: Behavioral syndromes correlate with individual novel foraging success

Which individuals in a group are the quickest to attempt to forage in a novel situation? Previous studies have found associations between novel foraging behaviors and neophilia, fast-exploring behavioral types, and foraging under predation threat (Jenkins et al., 2023; Stöwe et al., 2006; Takola et al., 2021). Dimensions of individual differences identified by behavioral

syndromes analysis provide a way to partition out the potential multivariate correlations between all aspects of boldness and novel foraging. H_0 : Finding no correlation between behavioral syndrome components and the novel foraging tasks would suggest that novel foraging behaviors are themselves a distinct dimension of behavior from all other components, and that future behavioral syndrome analysis on these species ought to include novel foraging test measures as an additional component. See Chapter 2 for more information on these ideas.

H2: Behavioral syndromes differ between species in a mixed-species flock

As outlined in Goodale et al. (2020), mixed-species groups often offer additional benefits to group members above what is common for single-species group living. Mixed-species groups of closely related species often share similar ecologies and enjoy “supplementary” benefits (Goodale et al., 2020), where the normal foraging efficiency and anti-predator benefits of group living are present but with reduced competition for mates and specific food types. To better understand why mixed-species flocks of closely related species solve novel foraging problems best in the most diverse flocks (Freeberg et al., 2017), I propose that species differences in behavior play a role. To address this possibility, I first attempt to show frequencies of different behavioral syndromes differing between species (i.e., titmice constrained to positively correlated exploration and anti-predator behavior whereas chickadees are constrained to positively correlated neophilia and time spent foraging), whereby a species-diverse flock is also a more behaviorally-diverse flock. H_0 : Finding no difference in behavior between these two closely related species would suggest that some aspect of MSGs may demand or reinforce similarity in behavioral syndromes. Chapter 3 describes the tests of these ideas.

H3: Performance on a novel environment test differs between social and asocial contexts

To address whether testing highly gregarious species in isolation poses ecological validity concerns, two different versions of the Novel Environment Test were given to separate subgroups of tufted titmice. I predicted that the test given in isolation would reflect more anxious behaviors than the test given in the presence of familiar flockmates. H_0 : If no difference is found between groups, then isolated testing may not be problematic for measuring individual differences in titmice. Chapter 4 describes the test of this hypothesis.

Methods

Subjects, capture, and housing

All tests focused on wild mixed-species flocks of Carolina chickadees and tufted titmice at UTFRREC. Over 30 feeder platforms were situated throughout the 2,200 acre area of UTFRREC, separated by > 400 m so that we can be reasonably sure that activity at each feeder represented a single overwintering flock (Bartmess-Levasseur et al., 2010). All individuals recorded for this study were caught at these feeder sites with mist nets and banded with colored plastic bands with RFID pit tags inside. Some individuals caught at these sites were housed in one of six outdoor aviaries on-site at UTFRREC for these captive experiments (see Figure 1). Housed birds were always kept in the same aviary as two other birds caught the same day at the same site, creating an experimental flock size of three individuals familiar with one another.

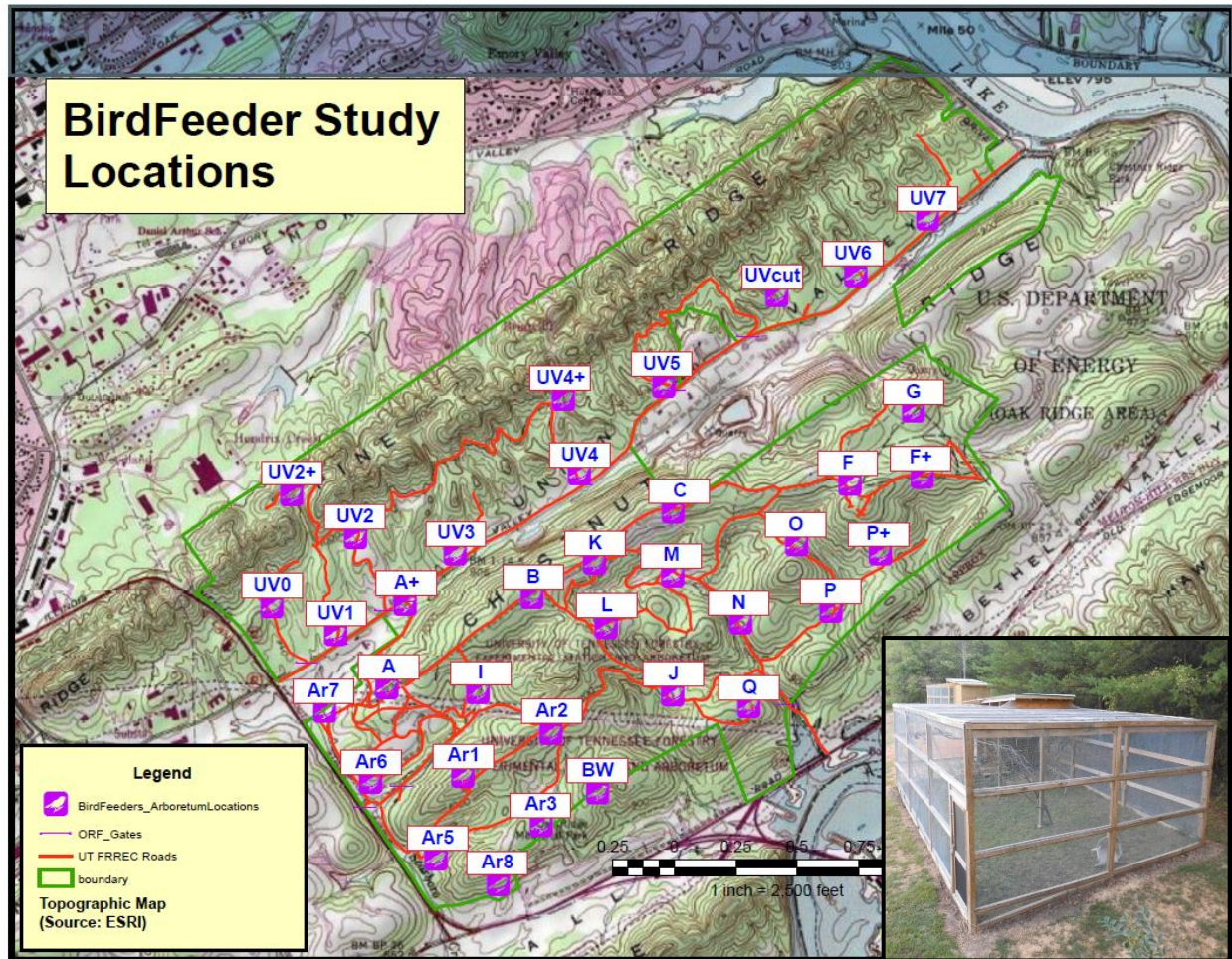


Figure 1: UTFRREC map of feeder locations. Total park area approximately 2,200 acres. *Inset:* Photo of two of the six aviaries used to house flocks temporarily (second aviary in the top-left background of photo). All aviaries are located in the area between sites “A” and “I”.

Each outdoor aviary measures approximately 6m x 9m x 3.5m in outdoor space and a small indoor area accessible to housed birds as shelter from inclement weather. Each aviary was provided with fresh seed and water *ad libitum* in designated dishes every day. Seed dishes were small clear glass candle holders (~65 mm in diameter) with RFID antennas taped to the rim, such that birds needed to perch on the antenna to reach seed in the dish. Birds housed in aviaries were allowed approximately five days to acclimate to the aviary before any testing began. Except for the time spent setting up tests and restocking food and water, birds in aviaries were not purposefully acclimated to human presence. At the end of data collection in the aviaries, birds were recaptured using treadle traps and were released back at their original site of capture. Including the 5-day acclimation period and all testing, most flocks were housed in the aviaries for no more than 2 weeks, though inclement weather sometimes led to longer delays to release the birds.

Sample size varied between studies and hypotheses based on the test data reliably collected from each individual. In total, 51 flocks (153 individuals) were housed in aviaries during the course of the research. Not all individuals participated in all tests; for example, a subset of flocks participated in the social-context Novel Environment Test instead of the typical asocial-context version of that test. Additionally, one data logger unit was stolen from an aviary resulting in incomplete data for several flocks. Also, all data loggers were eventually found to keep poor time when the battery power was below 50%, leading to timestamp discrepancies in collected data for many of the RFID-based behavioral tests. Any scheduled test whose start and end times could not be realigned with the recorded data were dropped.

Artificial flock compositions consisted of all species combinations (i.e., 3 titmice, 2 titmice and 1 chickadee, and 1 titmice and 2 chickadees, or 3 chickadees). The artificial flock size was smaller than the wild flock size, but the limited size of the aviaries kept population density high (Hetrick & Sieving, 2011).

Aviary testing procedure

Effort was made to ensure artificial flocks housed in the aviaries were caught on the same day and time to limit total acclimation time for the flock and to increase the confidence that individuals housed together in artificial flocks are familiar with each other (Coppinger et al., 2018). Some artificial flocks formed during winter 2021-22 were reformed with other artificial flocks caught at the same site to make up for lost and deceased individuals in those flocks. Though loss and death were rare, if they did occur they always occurred during acclimation periods and did not otherwise interfere with tests.

Behavioral tests (see Table 1) followed the acclimation period. The presented tests were given in sequence of least to most threatening stimuli (with the exception of the novel environment test given immediately before release from the aviary), so the order of tests was not counterbalanced. Tests typically began between 10am and 2pm each day, following from pilot data collected in winter 2020-21; activity drops sharply after 4pm for this population. Except for the Novel Environment Tests, all tests involved presentations of stimuli on the same platform as the seed dishes within the aviary. Only two seed dishes were available on the platform and both were equipped with RFID antennas. These antennas collected timed presence/absence data of the unique ID of each individual's RFID band. All of the following tests were designed to capture visits to seed dishes in the presence of various stimuli, or in the case of the puzzle feeders, to

Table 1: Summary of behavioral tests and measures

<i>Test</i>	<i>Individual Variables</i>
Novel feeder (48hrs)	Order of first approach Order latency ^a Visit rate
Mealworms (24hrs/15 worms)	Order of first approach Order latency ^a # Visits before empty
Novel Object (1hr)	Order of first approach Order latency ^a Visit rate
Predator models (2hrs snake) (2hrs hawk)	Order of first approach Order latency ^a Visit rate
Novel Environment	Number of movements Maximum time in one place Seeds taken

^a For 1st to approach, time to first approach. For nth to approach, first approach timed starting from (n – 1)th's first approach. Remaining test time given for no approach.

record exploration and solutions of the puzzle. A subset of trials (approx. 15% of the final data) were video recorded to verify the accuracy of the RFID antennas.

Statistical analyses

Behavioral Syndromes: Each behavioral test includes several measures (see Table 1, Figure 2, and Figure 3). Using principal component analysis (PCA) with these measures, the number of points of comparison between individuals can be reduced to a small number of principal components (depending on the number of eigenvalues >1). This reduction allowed smaller samples to retain statistical power by reducing the number of terms necessary in later models. This process applied to the mealworms test, novel object test, both predator tests (pooling the scores of both together), and the novel environment test using the measures listed in Table 1 for each test. This analysis pooled the scores of both species together. A subsequent cluster analysis determined whether species differences in behavior were detectable in the data by assessing the species composition of those calculated clusters.

Conclusion

This dissertation explores the unique aspects of diversity found in mixed-species flocks of birds. Through the below studies, I demonstrate the utility of the behavioral syndromes framework for answering mixed-species group questions and report unexpected dynamics of behavioral diffusion and isolation.

Dissertation outline

- **Chapter 1: Introduction** – Here I reviewed the current literature on behavioral syndromes and defined/discussed relevant features of behavioral flexibility. I analyzed testing methods for behavior and problem-solving and discussed their appropriateness for interpreting behavioral syndromes as they relate to personality, particularly in the context of social groups. How these concepts apply to mixed-species groups were presented as an open question.
- **Chapter 2: Behavioral syndromes as predictors of flexible foraging strategy** – Hypothesis 1 (Behavioral syndromes correlate with individual novel foraging success) was tested. Research question: *Which individuals in a flock are most likely to exploit a novel foraging task?* Analysis used behavioral syndromes data from Chapter 2 and two Novel Feeder tests to determine if exploitation of a novel foraging task is more common for some personalities than others. Individual test variables were also considered as predictors, and strengths of the two forms of predictive model were compared. A null result could warrant incorporating the Novel Feeder behavior into a more comprehensive PCA/HCA behavioral syndromes analysis.
- **Chapter 3: Behavioral syndromes in artificial mixed-species flocks** – Hypothesis 2 (Behavioral syndromes differ between species in a mixed-species flock) was tested in pursuit of the research question: *Does species diversity add to the behavioral diversity of an mixed-species group?* Analyses included PCA and Cluster Analysis to identify behavioral syndromes in pooled species; Clusters (behavioral syndromes) primarily ascribed to one species support the hypothesis.

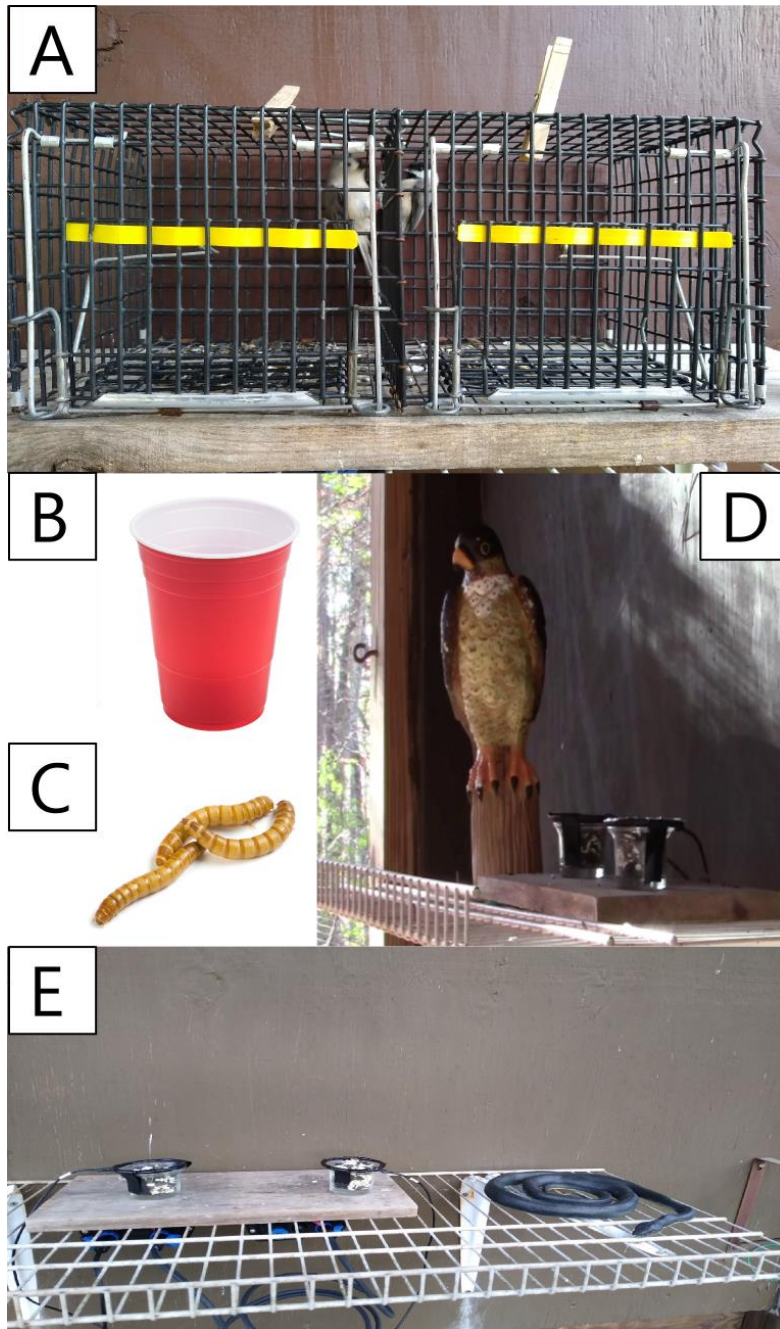


Figure 2: Test stimuli for RFID measures. (A) Treadle trap with one RFID food dish inside for Novel Feeder test. (B) Red plastic cup for Novel Object test. (C) Mealworm stock image, representing Mealworm test where mealworms were placed inside one of the two RFID food dishes. (D) Plastic hawk model approximating a Cooper's hawk or sharp-shinned hawk for the Predator tests. (E) Rubber snake model approximating a gray rat snake for the Predator tests. (D) and (E) represent the typical presentation distance from RFID food dishes for Predator and Novel Object tests.



Figure 3: Photo of the novel environment test cage.

- **Chapter 4: Asocial testing of social species** – Hypothesis 3 was tested using a between-groups design separate from the previous analyses. Comparing scores on novel environment tests between isolated and social context groups, this study addressed whether isolation significantly impact foraging behavior, activity levels, or freezing behaviors. Tested only with tufted titmice, three different aspects of novel environment behavior were compared across groups.
- **Chapter 5: Conclusion** – Do mixed-species flocks of closely related species benefit from more behavioral diversity? Findings from the previous chapters were summarized in answering these questions and in encouraging future research on a broader question: How important is diversity in social groups when challenged by changing environments?

CHAPTER 2: DO BEHAVIORAL SYNDROMES PREDICT BEHAVIOR BETTER THAN INDIVIDUAL TESTS?

Introduction

Research on novel feeding behaviors, particularly in birds, has long been a key aspect of studying animal intelligence, social learning, and adaptability (Lefebvre & Giraldeau, 1996; Lefebvre et al., 1997). The term “behavioral flexibility” is ascribed to such famous examples of foraging innovation as removing the caps from milk bottles by birds in the UK many decades ago (Aplin, Sheldon, et al., 2013; Fisher & Hinde, 1949). Somewhere in the cross-section of intelligence, social learning, and behavioral flexibility we might have the tools to predict which individuals, which populations, and which species are best poised to adapt behaviorally to environmental changes, particularly those with anthropogenic causes.

However, Audet and Lefebvre (2017) worry that the concept of behavioral flexibility is often being misused in ecology and animal behavior research, due in particular to what they call a “jingle-jangle” problem. Jingle-jangle problems are those where a single trait is measured by several different tests, but in the hands of other researchers those same tests are used to measure *different* traits. For example, presenting novel objects to a given individual and recording the reaction has been used as a measure of neophilia/neophobia under the umbrella animal personality term “boldness” (Takola et al., 2021), but neophilia is also sometimes measured with open-field or novel environment tests, which in turn are used to measure exploratory behavior (Perals et al., 2017). In fact, Audet and Lefebvre (2017) argue reversal-learning tasks are more conceptually accurate tests of behavioral flexibility than tests of innovation.

Even if I were to avoid the term “behavioral flexibility” altogether in this current work on predicting novel foraging behavior, the authors’ point on jingle-jangle problems remains relevant. What would I mean if I were to say the boldest individuals were the most likely to forage in a given context? Belgrad and Griffen (2017) examined how boldness and social context interacted to predict foraging in the presence of predator cues in mud crabs (*Scylla serrata*); Jenkins et al. (2023) referred to responses to predator cues as “risk-taking behavior” and correlated them to other “boldness” behaviors like exploration, and activity level in black-capped chickadees (*Poecile atricapillus*); Herborn et al. (2014) predicted the behavioral flexibility of blue tits (*Cyanistes caeruleus*) with age and broad personality variables like “neophobia” and “exploration”; Kurvers et al. (2012) measured “boldness” in barnacle geese (*Branta leucopsis*) using only novel object presentations; and Olafsdottir and Magellan (2016) used latency to leave a refuge as their “boldness” score when trying to predict foraging speed and accuracy in threespine sticklebacks (*Gasterosteus aculeatus*). Of course, this is no criticism of any one lab or paper, but instead meant to explain the varying levels of success each line of research had while attempting to explain foraging behaviors as a predictable aspect of boldness.

Clearly, from the above examples, we do have the ability to predict several aspects of foraging behaviors from personality research on our study species. The problem, from my perspective, is an eagerness to discuss test measures as familiar personality constructs. The results of the same Jenkins et al. (2023) paper defied a traditional understanding of “boldness” behaviors; the chickadees with the highest scores on activity level and exploratory behavior were the least “bold” in the presence of a predator. In the Herborn et al. (2014) paper, neophobic behaviors had a negative correlation with exploratory behaviors, and both correlated positively with behavioral flexibility. Ultimately, there was no coherent way to say “bolder birds innovate more” or “neophobia predicts novel foraging strategies,” particularly when the constituent behaviors are sparsely measured or not understood in the context of larger behavior patterns.

There are tools in animal personality research that can help us to describe behavior patterns quantitatively and solely in terms of the correlations between raw test variables rather than our interpretation of them. The behavioral syndromes framework (Sih, Bell, Johnson, et al., 2004) recommends compiling multiple variables from each of several behavioral tests and allowing principal component analysis to tell us what the key dimensions of individual differences are in a study species. Mertes et al. (2022) reduced the behavioral test variables of scimitar-horned oryx (*Oryx dammah*) into two principal components before describing these components as “vigilance or caution” and “changing social context” respectively, based not on a single test or a single variable, but from correlated variables across tests that themselves suggested the pattern of differences between individuals. Predictions of which personalities or behavioral types in a population are most likely to exploit novel resources in a changing environment can be understood from the natural patterns of measured behavior rather than our preconceived definitions of whether it is the “bold” or “fast-exploring” or “neophilic” individuals that do the work of adapting to those changes.

The current study aimed to demonstrate the utility of a behavioral syndromes framework to understanding this novel foraging question, but also to compare it quantitatively to single-test methods for correlating behaviors under the umbrella of “boldness” to novel foraging behaviors. Tufted titmice (*Baeolophus bicolor*) and Carolina chickadees (*Poecile carolinensis*) serve as nuclear members (Jones & Sieving, 2019) of mixed-species flocks during the overwintering period (approx. mid-October to early March each year in Eastern Tennessee). Titmice are larger and generally dominant over chickadees in foraging contexts (Cimprich & Grubb Jr, 1994). Using flocks of varying species composition of tufted titmice and Carolina chickadees I examined two different kinds of predictors: raw scores and scores on 5 principal components of behavior in these mixed-species flocks. First, using 12 variables from 4 behavioral tests, I tested if performance on any single test was a good predictor of Novel Feeder Test performance for both chickadees and titmice. Second, using the Principal Components derived from the same 12 variables, I tested if any components were a good predictor of Novel Feeder performance. Finally, I examined the relative strengths, parsimony, and utility of each technique. Finding no correlation between behavioral syndrome components and the novel foraging tasks would suggest that novel foraging behaviors are themselves a distinct dimension of behavior from all other components, and that future behavioral syndrome analysis on these species ought to include novel foraging test measures as an additional component.

Methods

Subjects and housing

Wild tufted titmice and Carolina chickadees were captured at feeder platforms at the University of Tennessee Forestry Resources AgResearch and Education Center (UTFRREC) in Oak Ridge, TN. Each platform was > 400 m away from its neighbor to ensure that each platform represented distinct overwintering flocks (Bartmess-Levasseur et al., 2010). Capture protocol involved the use of mist nets to gather flocks of 3 individuals (species composition varied, but always limited to tufted titmice and Carolina chickadees) that were housed in large outdoor aviaries (6 m × 9 m and 3.5 m high) as experimental but familiar flocks. While in the aviary, flocks had access to ad libitum food and water and were held for a period of 12-18 days depending on weather and protocol of their initial study prior to participation in the current study. This time period included a 4-5 day acclimation period.

Materials and procedures

RFID dataloggers and receivers were created by Eccel Technology Ltd (Groby, UK). RFID antennas were attached to food dishes with electrical tape to capture visits to each dish. Each bird caught for the study was banded with a plastic PIT tag with a unique ID. To extend battery life, each receiver was configured to collect data only between the hours of 7am and 6pm Eastern Standard Time. Pilot data indicated this daytime timeframe captured greater than 95% of the visits to the seed dishes during aviary housing.

Novel feeder test

The Novel Feeder Test was a choice test between an open, unobstructed seed dish filled with ad libitum seed or an identical dish placed with a locked-open treadle trap just a few centimeters away. Testing generally began between 10 am and 1 pm on the 5th day of aviary housing – as weather permitted – and ran for 48 hrs. Seed dishes were large enough to accommodate the 3-bird flocks for the full 48 hrs but fresh water and additional food was added the day after testing started regardless, in part to prevent one dish from becoming noticeably more depleted than the other during the test. There was no other researcher disturbance to the flock during the test.

RFID visits during the test were timestamped including the date and uniquely attributable to each bird and the antenna activated. While two birds could activate each of the two antennas at the same time, only one bird (typically the first to arrive) could activate a single antenna at a time. The small diameters of the round antennas (60mm) made this kind of co-occurrence difficult and unlikely.

Variables deduced from the RFID data were:

- *Order of approach* – The rank order of the birds in each flock to approach during the test period, with respect to each individual's first approach only. If one or more birds in a given flock did not approach during the test period, they were

scored as a 4 rather than a 1, 2, or 3 so that late approachers were not considered equal to non-approachers in the data.

- *Order latency* – The time in seconds since the start of the test (for the first to approach) or the time in seconds since the previously ranked bird's first approach. For example, if the first bird visited 3 minutes after the start of the test and the second bird first visited 5 minutes after the start of the test, the birds were scored 180 s and 120 s respectively. This serves two purposes: 1) avoid redundancy with Order of Approach and 2) to account for the qualitative difference in the test between a stimulus no familiar flock mate had approached to one that one (or two) flockmates had safely approached. The interpretation of this variable is then different for the first to approach (a measure of boldness) than for subsequent birds (a measure of boldness interacting with some local enhancement or other social learning). Any bird that did not approach during the test period was scored for the maximum length of the test (i.e., 172,800 seconds).
- *Visit rate* – A ratio between the number of visits during the test period and the length of the test itself. Between-species differences in metabolism should mean that virtually every titmouse has a higher visit rate than every chickadee, but within-species differences in activity level has been linked to both personality (Jenkins et al., 2023) and social context (Bibi et al., 2022).

Mealworm test

Two food dishes were placed within 20cm of each other, one contained the same seed mix birds had been given since entering the aviaries, and the other held 15 medium sized live mealworms. While this test was initially included as a novel food test, since mealworms were not thought common in the range of the tested population, pilot tests suggested that both titmice and chickadees immediately recognized mealworms as a highly desirable food source, although individuals varied widely in how quickly or aggressively they sought out the limited resource. This behavioral test was included in the data analysis with attention paid toward the behaviors expected between dominant and subordinate individuals competing for the limited numbers of mealworms. While pilot testing showed that one or two individuals tended to eat a disproportionate amount of mealworms in each flock, supplanting was not common. When they did occur, supplanting behaviors were recorded as a change in the detected RFID at the mealworm dish in less than 2 seconds. As only 15 mealworms were added to the dish, only the first 15 unique visits to the dish were analyzed. Order of approach, order latency, and visit rate were otherwise calculated in the same way as the Novel Feeder test.

The test ran for a maximum of 24hrs and typically started immediately after the end of the Novel Feeder test, as weather conditions permitted. Any remaining mealworms at the end of the test were discarded outside of the aviary.

Novel object test

The novel object test was the presentation of an upside-down disposable red plastic cup placed within 20 cm of one of the seed dishes (which were in turn no more than 20cm from each other). The presentation lasted 1hr for all flocks. Picnicking is not allowed in any area of the

UTFRREC, and so there is virtually no chance any of the birds in the study had encountered anything like this object before. No bright red plastic objects were used anywhere in the UTFRREC prior to the start of data collection, and so I am confident that the object represented a novel and easily replicable presentation. Brightly colored objects have been used before to induce neophobia in closely related species (Vrublevska et al., 2015). To account for the light weight of the cup in potentially windy conditions, a small rock was placed on the top of the cup and, when possible, a wooden clothespin was attached at the bottom of the cup across the metal grating the seed dishes were placed on.

Predator tests

Two different predator tests were conducted to account for differences between threats posed by airborne and terrestrial predators, such as a differences in calling and vigilance behavior between the two types of predation risks (Sieving et al., 2010; Suzuki, 2012). The first was a plastic model of a Cooper's hawk (*Accipiter cooperii*), the second was a rubber coiled snake that looked superficially similar to gray rat snakes (*Pantherophis spiloides*) and eastern black kingsnakes (*Lampropeltis nigra*) (Cantwell et al., 2016). The order of presentation was randomized between flocks. Predator models were placed within 20 cm of one of the food dishes, similarly to the Novel Object presentation. Presentations lasted for 2 hrs, and in consideration of the extra stress of being in an enclosed space with a supposed predator, no other tests were conducted on days when there was a predator presentation. Each predator presentation typically occurred with one day of the other.

Scoring of both predator tests followed the same focal variables as the Novel Object test, including preference for the nearer or farther food dish during the test.

Novel environment test

The novel environment was a small white bird cage (approx. 0.8 m x 0.8 m x 1.2 m) with two cross-length perches and two seed dishes stocked with black oil sunflower seed (see Figure 3). This size is similar to small "flight cage" novel environment tests (Katsis et al., 2024) rather than whole-aviary novel environment tests (Kerman et al., 2018; Tuliozi et al., 2018). All trials were video recorded using a Canon VIXIA HF R800 camera set on a nearby tripod such that the perspective angle allowed easy viewing of all 6 inside faces of the cage. The cage itself was set on a folding table approximately 0.5 m off the ground, depending on the terrain.

The same video camera was also used to film a subset of all of the above tests (16 or 15.1% of the total data set, including pilot data) of the RFID-based tests to ensure the reliability of the RFID receiver data. The calculated F_1 score was .813, indicating good precision and recall of the RFID data (Taha & Hanbury, 2015).

All Novel Environment Test videos were coded in BORIS (Friard & Gamba, 2016). Trials were marked continuously for 10 min, regardless of the length of the actual video. The beginning of the 10 min trial was no earlier than 30 s after the bird entered the cage on video to allow the researcher to leave the area, and the end of the trial was no later than 30 s from the end of the video, to account for the re-approach of the researcher. Eleven different behaviors were marked according to where the bird moved within the cage and whether it foraged (i.e., top,

bottom, left, right, front, back, upper perch, bottom perch, left bowl, right bowl, seed taken). Activity level was calculated alternately as the total number of moves during the test and the longest continuous perch during the test. Whether or not any seeds were taken was calculated as a simple binary, as many birds did not stop to forage during the 10 min trial.

I coded all video trials, and reliability was assessed on a subset of trials (12 trials or 18.2%) by a trained research assistant. Interrater reliability measured by Cohen's Kappa was 0.782, indicating substantial agreement on Novel Environment Test trials.

Statistical analysis

The behavior of interest in this study was the Order of Approach to the Novel Feeder Test, since this should have best demonstrated whether initiating novel foraging behavior was predictable. Because this response variable was ordinal, each model specified below was an ordinal logistic regression model using the “ordinal” package in R (Christensen, 2023; R Core Development Team, 2016). There were two different levels of analysis: Principal Component, and individual tests.

Scores for titmice and chickadees were pooled for principal component analysis (PCA) without regard to species as a variable so as not to bias the model toward amplifying species differences. A total of 12 variables went into the PCA (order of approach, order latency and visit rate for “Novel Object,” “Mealworm,” and “Predator” tests, number of moves, longest freeze, and number of seeds taken for “Novel Environment” tests). Parallel analysis with the “paran” package in R (Dinno, 2018) allowed predictions for the optimal number of components to keep adjusting for any bias introduced by the imputation of missing values.

In lieu of repeated tests, I calculated consistency in individual differences along the identified principal component dimensions (Huang et al., 2018). Each variable either loaded positively or negatively on each principal component, such that the relationship between any two variables in a principal component could be considered to either negatively or positively correlate with one another. By scaling test measures to zero and comparing each pair of variables to their expected relationship according to contributor loadings, I calculated how consistently each bird conformed to the expected relationship of variables in each principal component relative to the number of variables contributing to that component. For example, if the number of visits to the mealworm test loaded positively and the longest freeze on the novel environment test loaded negatively on the same component, a consistent set of scores for a given individual could be above-average mealworm visits and below average longest freeze. For each principal component, only those test variables with above-average contribution to that component were considered for consistency analysis, since these are the variables that present the strongest signal of the underlying correlational structure of the behavioral data. Permuting 1000 random samples of the data allowed for comparison of the mean of these consistency calculations to determine whether the mean consistency was significantly different from random. Significant consistency on a given principal component acted here as a substitute to repeated testing on each test.

The test-specific models were not run with species data pooled together, although it should be noted that the tests were given in semi-naturalistic flocks of varying species diversity –

though these are raw individual scores, they may not be fully independent of heterospecific influence (see Chapter 3, Chapter 4). For each species, a model for each of four different tests included that test's three measured variables as predictors for Order of Approach to the Novel Feeder.

For any models that found significant predictors of Novel Feeder behavior, the McFadden's adjusted pseudo- R^2 statistics for those models were compared (Mcfadden, 1974). This metric works as a measure of model fit specific to logistic regression models; this adjusted pseudo- R^2 cannot be interpreted directly as a percent of variance explained but it can still allow for comparisons between models and for an estimate of the goodness of fit for each model. Adjusted pseudo- R^2 is generally lower than comparable values in linear regression models, and good fit is generally between 0.2 and 0.4 (Mcfadden, 1974). Note that this metric is intended only to compare models regressed on the same response variable data, so comparing models across different subsets of the data (i.e., different n 's) is not advisable.

Significance was determined without adjustment for multiple comparisons; the primary purpose of these models was broad comparisons of fit between categories of predictors rather than the predictors themselves.

Results

Sample sizes depended on the model and species specification. A total of 66 birds across 22 artificial flocks were included in the principal component analysis: 35 were tufted titmice and 31 were Carolina chickadees. A subset of 45 individuals was included in the subsequent model comparison analysis (19 Carolina chickadees and 26 tufted titmice), representing those flocks with complete data for the novel feeder test. See Table 2 for summary statistics of each test measure.

Principal component analysis

Principal component analysis produced 5 components with eigenvalues greater than 1, explaining a combined 76.38% of the variance in behavioral differences. The appropriate number of retained components was confirmed with Horn's parallel analysis for component retention (Dinno 2018; Horn 1965), which adjusted eigenvalues to account for potential bias introduced by imputed datasets. The labels for each of the components below are brief interpretations of the contributions and loadings as they relate to the test behaviors. Other interpretations may be possible as these are written with the highest or lowest scorers for the given component in mind. See Table 3 for a summary of the top contributing variables to each principal component.

Component 1 – Neophobia

Four of the measured variables had above-average contributions for the first principal component: Novel Object Visit Rate (22.93%, loading -.479), Mealworm Order of Approach (17.73%, loading .421) and Visit Rate (24.95%, loading -.499), and Novel Environment Seeds

Table 2: Summary statistics of test measures by species

		CACH		ETTI	
		Mean(SD)	Range	Mean(SD)	Range
Novel Object	Order of Approach	1.71(.85)	1 - 3	1.90(.81)	1 - 3
	Order Latency (s)	630.6(741.35)	27 - 2,978	561.8(504.90)	77 - 2,304
	Visit Rate (per hr)	7.28(4.91)	1 - 15	3.84(1.61)	1 - 7
Mealworm	Order of Approach	1.78(.88)	1 - 3	2.05(.79)	1 - 3
	Order Latency (s)	3,707.5(14,560.89)	21 - 62,039	718.2(758.40)	4 - 2,523
	Visit Rate (per hr)	6.01(3.83)	.25 - 13	4.49(2.38)	.75 - 9
Predator tests	Order of Approach	1.70(.68)	1 - 3	2.04(.80)	1 - 3
	Order Latency (s)	748.9(865.67)	25 - 3,122	1,365.5(1,606.64)	9 - 5,264
	Visit Rate (per hr)	4.49(4.04)	.50 - 13	3.35(2.13)	.50 - 7.5
Novel Environment	Longest Freeze (s)	49.90(50.06)	9 - 170	52.23(54.72)	8 - 190
	Seeds Taken	7.11(7.43)	0 - 23	2.42(3.17)	0 - 11
	Moves	257.4(108.47)	34 - 449	168.2(70.45)	21 - 267

Table 3: Top-contributing variables to selected components

Component	Variable	Contribution (%)	Loading
PC1: “Neophobia” 22.21%	Novel Object Visit Rate	22.93	-0.479
	Mealworm Order of Approach	17.73	0.421
	Mealworm Visit Rate	24.95	-0.499
	Novel Environment Seeds Taken	9.91	-0.315
PC2: “Non-exploration” 18.91%	Novel Environment Moves	33.43	-0.578
	Novel Environment Freeze Time	29.64	0.544
	Predator Visit Rate	19.2	0.438
PC3: “Neophilic Threat Discrimination” 13.90%	Predator Order Latency	33.3	0.577
	Predator Visit Rate	19.2	-0.372
	Novel Object Order Latency	21.16	-0.46
PC4: “Social Information Preference” 11.71%	Predator Order of Approach	26.34	0.513
	Novel Object Order of Approach	16.22	0.403
	Novel Object Order Latency	22.22	0.471
	Mealworm Order Latency	18.58	-0.431
PC5: “Neophobic Exploration” 9.65%	Novel Environment Seeds Taken	32.53	0.57
	Mealworm Order Latency	29.47	0.543
	Novel Object Order of Approach	14.2	0.377

Taken (9.91%, loading -.315). This pattern of scores was consistent with neophobia; high scoring individuals on this component foraged less in the presence of a novel object (Guido et al., 2017) or in a novel environment (Tuliozi et al., 2021), and were slower and less likely adopt mealworms into their aviary diet (Tebbich et al., 2016). Up to 22.21% of the differences between individuals can be attributed to differences along this component.

The results of the permutation test indicated that the observed mean consistency score of 60.94% (mean = 3.656 out of 6 comparisons) for Principal Component 1 was significantly higher than what would be expected by chance alone ($p < .001$). This strong signal suggests that Principal Component 1 captures meaningful patterns of individual differences in behavior that are more consistent than a random dataset.

Component 2 – Non-exploration

Three of the measured variables had above-average contributions for the second principal component: Novel Environment Moves (33.43%, loading -.578) and Freeze Time (29.64%, loading .544), and Predator Test Visit Rate (19.20%, loading .438). Individuals who scored highest on this component were those who moved the least and spent the most longest period staying still during the Novel Environment Test, and foraged frequently in the presence of predator models. The strongest signal in this component was a lack of activity in a novel environment, but the high foraging rate in the presence of predators contradicts earlier positive correlations found between predator avoidance and non-exploration in house sparrows (Tuliozi et al., 2021) suggesting this component may not be primarily attributable to fear or boldness. Up to 18.91% of the differences between individuals can be attributed to differences along this component.

The results of the permutation test indicated that the observed mean consistency score of 69.79% (mean = 2.094 out of 3 comparisons) for Principal Component 2 was significantly higher than what would be expected by chance alone ($p < .001$).

Component 3 – Neophilic threat discrimination

Three of the measured variables had above-average contributions for the third principal component: Predator Order Latency (33.30%, loading .577) and Visit Rate (19.20%, loading -.372), and Novel Object Order Latency (21.16%, loading -.460). This combination of early approaches to the novel object with late and rare approaches to predator models suggested a strong discrimination between threatening and neutral stimuli in the highest scoring individuals (Blumstein et al., 2006; Liu et al., 2024). Low-scoring individuals were not rare in the dataset, suggesting that some individuals were so neophobic that they were more likely to approach predator models than novel objects. Up to 13.90% of the differences between individuals can be attributed to differences along this component.

The results of the permutation test indicated that the observed mean consistency score of 57.29% (mean = 1.719 out of 3 comparisons) for Principal Component 3 was significantly higher than what would be expected by chance alone ($p = .037$).

Component 4 – Social information preference

Four of the measured variables had above-average contributions for the fourth principal component: Predator Test Order of Approach (26.34%, loading .513), Novel Object Order of Approach (16.22%, loading .403) and Order Latency (22.22%, loading .471), and Mealworm Order Latency (18.58%, loading -.431). The highest scoring individuals were those who were last to in their flocks to approach predator models and the novel object, but were quick to eat mealworms after a flockmate had eaten one first. This pattern could be consistent with an indiscriminate shyness toward predators and novel objects but not novel foods, however the orthogonal nature of principal components means that non-discrimination of threatening stimuli is more likely represented by a score near 0 on Component 3. This pattern would also be consistent with a preference for socially gathered rather than individually gathered information (Arvidsson & Matthysen, 2016; Mamuneas et al., 2015). Up to 11.71% of the differences between individuals can be attributed to differences along this component.

The results of the permutation test indicated that the observed mean consistency score of 56.25% (mean = 3.375 out of 6 comparisons) for Principal Component 4 was significantly higher than what would be expected by chance alone ($p = .015$).

Component 5 – Neophobic exploration

Three of the measured variables had above-average contributions for the fifth principal component: Novel Environment Seeds Taken (32.53%, loading .570), Mealworm Order Latency (29.47%, loading .543), and Novel Object Order of Approach (14.20%, loading .377). Higher scoring individuals on this component were those that were quick to forage in a novel environment but not in the presence of novel objects or food. Up to 9.65% of the differences between individuals can be attributed to differences along this component.

The results of the permutation test indicated that the observed mean consistency score of 48.96% (mean = 1.469 out of 3 comparisons) for Principal Component 5 was not significantly higher than what would be expected by chance alone ($p = .693$). This low consistency may be due to Order of Approach appearing twice in the three top contributing measures for Component 5, which was the measure most sensitive to flockmate activity and thus not itself expected to be a highly repeatable measure for individuals.

Model comparisons

Out of 11 different ordinal logistic regression models to predict the Order of Approach to a Novel Feeder (NF-Order), only 6 models identified a total of 7 significant predictors out of 17 unique candidate predictors (see Table 4). McFadden's Adjusted Pseudo- R^2 for each model is given in Table 5 for comparison of model fit. Figures 4 and 5 use locally estimated scatterplot smoothing (Cleveland et al., 1990) to visualize local trends along rank-ordered test variables.

All three models based on Principal Components identified significant predictors (see Table 4 & Figure 5), and of all comparable species-specific models the Principal Components models (CACH: McFadden's $R_{adj}^2 = 0.487$; ETTI: McFadden's $R_{adj}^2 = 0.172$) fit better than any of the single-test models (CACH: McFadden's $R_{adj}^2 \leq 0.223$; ETTI: McFadden's $R_{adj}^2 = 0.082$), see Table 5.

Table 4: Summary of predictors from ordinal logistic regression models

Predictors	Models	Significance (N)		
		Pooled (n = 45)	ETTI (n = 26)	CACH (n = 19)
PCA				
Comp.1	<i>NF ~ Comp.1 + ... Comp.5</i>	.785**	.995*	1.135*
Comp.2		-.138	.454	-.232
Comp.3		.658**	.303	1.761*
Comp.4		.358	.744	.508
Comp.5		-.610	-1.380*	-.288
Novel Object Test				
Approach Order	<i>NF ~ AppOrder + OrderLat</i>		.958	1.900*
Order Latency	<i>+ VisitR</i>		-1.105*	-.873.
Visit Rate			-.494*	-.162
Mealworm Test				
Approach Order	<i>NF ~ AppOrder + OrderLat + VisitR</i>		.616	.977
Order Latency			-.575*	-.228
Visit Rate			-.210	-.349
Predator Test				
Approach Order	<i>NF ~ AppOrder + OrderLat</i>		.410	.286
Order Latency	<i>+ VisitR</i>			
Visit Rate			.029	-.993
			-.025	-.390.
Novel Environment Test				
Total Moves	<i>NF ~ Moves + Freeze +</i>		-3.389.	.805
Longest Freeze	<i>Seeds</i>		-.032.	.018
Seeds Taken			.060	-.182

Note: Significance based on unadjusted p-values. * = $p < .05$, ** = $p < .01$, . = $p < .10$. NF = Novel Feeder Order of Approach. “Comp” refers to scores on the corresponding principal component.

Table 5: McFadden's Adjusted Pseudo-R2 per predictive model

Model	R2		
	Pooled	CACH	ETTI
PCA			
Pooled†	0.1004*		
CACH		0.4871*	
ETTI			0.1718*
Novel Object Test			
CACH		0.2235*	
ETTI			0.0234*
Mealworm Test			
CACH		0.1119	
ETTI			-0.0366*
Predator Test			
CACH		0.1138	
ETTI			0.0188
Novel Environment Test			
CACH		0.0050	
ETTI			-0.0459

Note: * = one or more significant predictors, † = not comparable to other models

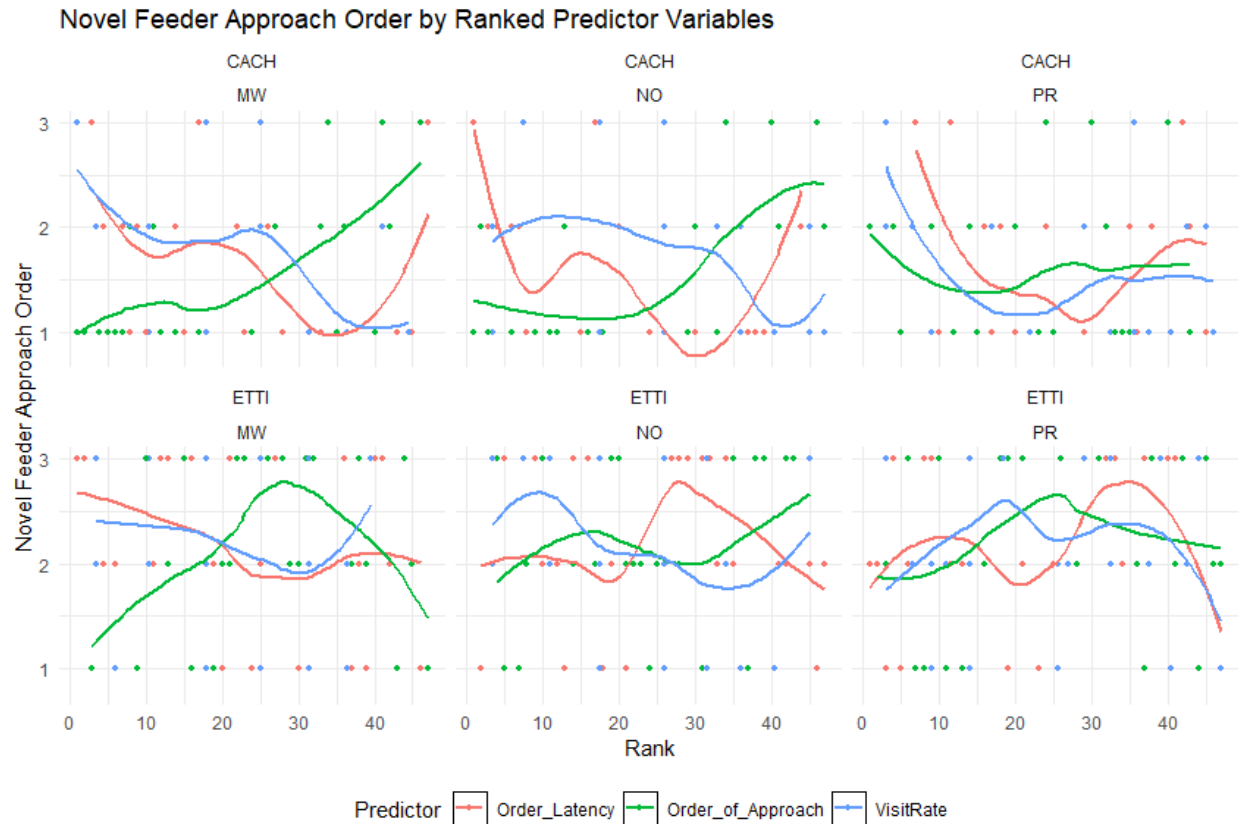


Figure 4: Novel feeder approach order by ranked predictor variables.

Rank-ordered scores on Order Latency, Order of Approach, and Visit Rate each for the Mealworm test, Novel Object test, and combined Predator test scores presented with individual's Order of Approach to the Novel Feeder test. Each point represents one individual's rank on one test measure and their corresponding Novel Feeder Order of Approach. LOESS regression lines illustrate estimated non-parametric relationships between predictor measures and Novel Feeder Order of Approach. Species are presented separately.

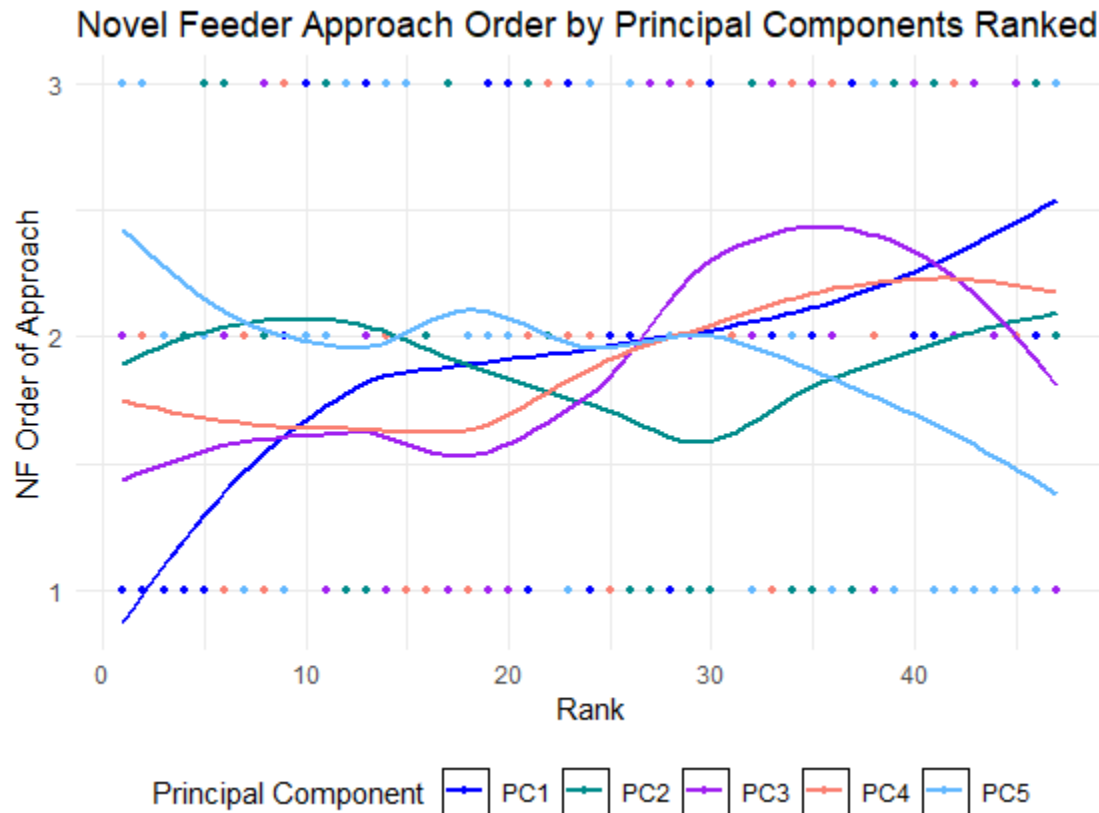


Figure 5: Novel feeder approach order by principal component ranked scores

Rank-ordered scores on the five selected principal components presented with individual's Order of Approach to the Novel Feeder test. Each point represents one individual's rank on one test measure and their corresponding Novel Feeder Order of Approach. LOESS regression lines illustrate estimated non-parametric relationships between each principal components and Novel Feeder Order of Approach.

Identified predictors

In both the pooled and CACH-only Principal Component models, the “Neophobia” (PC1; Pooled: Est = .785, $p < .01$; CACH-only: Est = 1.135, $p < .05$) and “Neophilic Threat Discrimination” (PC3; Pooled: Est = .658, $p < .01$; CACH-only: Est = 1.761, $p = .02$) components were positively associated with NF-Order. In the ETTI-only Principal Component models, “Neophobia” (PC1, Est = .995, $p = .03$) was also positively correlated with NF-Order and “Neophobic Exploration” was negatively correlated with NF-Order (PC5, Est = -1.380, $p = .03$).

Each of the three Novel Object Test variables were significant in different species models. Order Latency (Est = -1.105, $p = .04$) and Visit Rate (Est = -.494, $p = .02$) were significant predictors in the ETTI-only model, and Order of Approach (Est = 1.900, $p = .02$) was significant in the CACH-only model. Order Latency was also marginally significant in the CACH-only model (Est = -.873, $p = .08$).

For the Mealworm Test models, only Order Latency (Est = -.575, $p = .04$) was significant for the ETTI-only model. None of the test variables were significant in the CACH-only model. None of the variables were significant in either the Predator Test models or the Novel Environment Test models, although Visit Rate (CACH-only, Est = -.390, $p = .06$), Longest Freeze (ETTI-only, Est = -.032, $p = .08$), and Total Moves (ETTI-only, Est = -3.389, $p = .06$) were marginally significant, depending on the species-specific model.

Discussion

Using the Adjusted R^2 for a series of generalized linear models, I showed that my proposed method for predicting novel foraging behavior with behavioral syndromes classifications is not necessarily more effective at predicting novel foraging than with scores from competing single-test models. While the PCA-based models had better fit statistics, none of the models presented here were significant after adjusting for multiple comparisons. This result supports the current trend in the literature to use a single behavioral variable or several variables from a single behavioral test to predict behavior on one other test (in this case, novel foraging).

The best-fitting models for each species were those that used Principal Components as predictors. This result is not too surprising; the principal components capture the orthogonal variance of all 12 individual test measures, and so simultaneously contain the most behavioral data for each individual of any model and avoid any risk of multicollinearity between terms. McFadden’s adjusted pseudo- R^2 does penalize models for the addition of predictor terms that do not meaningfully contribute to the predictive power of the model. That the principal component model was still the best-fitting model, even including insignificant components, suggests that the strength of the model actually comes from the relationship between Novel Feeder behavior and the personality dimensions described by each of the significant principal components.

In my mixed-species study system, it is interesting to see the minor differences in predictors between species. Both species, perhaps unsurprisingly, were slower to approach the Novel Feeder when they scored high on the “Neophobia” principal component term. However,

titmice were also predicted to approach later when they scored especially low on the “Neophobic Exploration” component; this suggests that the appraisal of a novel feeder as more like a novel object or a novel environment might differ between titmice.

Note that while the order of stimuli presentations was designed to steadily increase threat throughout the experiment, some individuals habituated to predator models in under 1 minute, or within 1 minute from seeing a flockmate approach the predator stimuli. This may be the result of habituation toward the aviary environment and testing procedures; all stimuli were placed in, over, or very near the seed dishes in every RFID-measured test. This kind of environment-procedure habituation may have been first tested in angel fish (*Pterophyllum eimecki*) in an experiment by Schleidt et al. (1983). However, most individuals did not habituate so quickly and some birds took well over an hour to approach the predator stimuli even after flockmates had done so safely. While future aviary studies should attempt to disrupt the stereotyped presentation procedures, the wide range of scores in the current study suggests stark individual differences in this habituation.

The unfortunate drawback of using principal components derived from a variety of different test scores (beside the additional testing effort) is that it may not meaningfully address the “jingle-jangle” problem as described by Carter et al. (2013). Often the loadings of the contributing variables conflict with our typical expectations for a pattern of behavior. For example, Component 3 was primarily weighted by a mix of slow approaches to a predator model and fast approaches to a novel object – both behaviors considered “bold” depending on the study (Belgrad & Griffen, 2017; Kerman et al., 2018; Takola et al., 2021) – which is made only more complicated by the positive association with late approaches to the Novel Feeder in Carolina chickadees. The labeling of principal components as behavioral syndromes constructs may also cause reification problems, where the pattern of measured behaviors is thought to exist in a functional and concrete sense instead of simply serving as a label to observed correlations in the data (Berger & Pullberg, 1965; Machery, 2014). It is important to remember that part of the appeal of the behavioral syndromes framework is the mechanism-agnostic approach to analyzing these behavioral correlations when exploring individual differences (Sih, Bell, & Johnson, 2004); labels here serve only as a shorthand reminder of the shape of the PCA data and do not presuppose or require the existence of those principal components as concrete functions in either species.

Attempting to label each principal component requires careful consideration of the direction of the correlation between the top contributing variables, although some variables may be categorized under several different labels depending on the researcher’s area of interest (Audet & Lefebvre, 2017). Principal component 4, dubbed “Social Information Preference,” is a good example of a difficult-to-classify dimension of individual differences due to seemingly contradictory relationships between the top contributing variables. Late approach orders to predators and novel objects combined with low order latency during the mealworm presentation could indicate a tendency for some birds to only attend to social foraging cues about food patch quality, rather than for information about safety. This would be consistent with the kinds of information-gathering preferences described by Stamps and Biro (2016) and Arvidsson and

Matthysen (2016) but should not be taken as sufficient evidence that social learning occurred in the data. Low scoring individuals are more easily thought of as “opposite” the high scoring individuals in this case, however; low scorers were most likely to be first to approach novel objects and predator models, not waiting for social cues regarding their safety, and also to follow others to the limited number of mealworms more slowly as a (potential) result of a low preference for social information.

Predicting novel feeder behavior was secondary to comparing the strength of the methods of prediction, but it should be reiterated that the predictors found significant were only significant according to p-values that were not adjusted for multiple comparisons. While some trends in the data presented themselves, the behavior of both species was difficult to predict even with the best-fitting model, the Principal Component model. Whereas titmice had more significant predictors across their single-test models, the fit for these models was generally poor (McFadden’s $R_{adj}^2 < .05$). Conversely, chickadees had several single-test models with moderate to good fit (McFadden’s $R_{adj}^2 > .10$), but only Novel Object Order of Approach was significantly correlated with Novel Feeder Order of Approach. Several of the models included one or more predictors in the range of “marginally significant” and had fewer than 30 observations, so these interpretations may be hindered by small species-specific sample sizes.

Future studies with larger sample sizes, besides having greater statistical power, would be able to test an even higher-order level of analysis as a predictor: behavioral syndrome *clusters*. The current dataset was found to be divisible into two clusters representing distinct patterns of high and low scores across the five principal component behavioral dimensions, although the sizes of each cluster were so unbalanced as to make meaningful analysis in the context of the current study impossible. Using clusters instead of principal components could expand further on the advantages gained by replacing single-measure predictors with principal components; replacing single components with correlated relationships between components observed in the population would essentially be predicting behavior using a category more akin to a personality type.

CHAPTER 3: BEHAVIORAL SYNDROMES IN EXPERIMENTAL MIXED-SPECIES FLOCKS

Introduction

In well-studied species, patterns of individual differences in behavior are often associated with survivorship, dominance hierarchies, or personality structures (Dall et al., 2004; De Azevedo & Young, 2021; Dingemanse & Reale, 2005). Social contexts can also significantly impact the observed behaviors of individuals relative to conspecifics (Nelson et al., 2008; Von Merten et al., 2017). Because personality dimensions are typically thought of as repeatable differences in particular behaviors, and some of the variation in these behaviors is most often restricted by a natural social context and the presence of conspecifics, it can be difficult to determine the structure of personalities in the natural setting for highly social species. Are we classifying “personalities” based on underlying individual differences or are we mistaking socially-influenced restrictions on these behaviors for the true pattern of variation in a population? The question has been raised in both the “person-situation debate” and “interactionism” starting with human personality research several decades ago (Furnham & Jaspars, 1983; Houts et al., 1986). Neither is the question new among animal behavior studies, or even flocks of birds specifically (Van Oers et al., 2005).

Allowing for context-specific repeatability in those individual behavior differences isn’t impossible, but it does imply that the ecological validity of personality or behavioral type labels are somewhat limited to the most common contexts encountered by a population (Beekman & Jordan, 2017; Sih, 2017). We might then be tempted to run completely in the other direction away from whole-organism behavioral classifications like “personality” and instead study each behavior across contexts, as with the reaction norms model popularized by Dingemanse et al. (2010). Subsequent research has reinforced that context is key to understanding inter- and intra-individual differences in behavior in a way that accounts for differences in behavioral plasticity. For example, some studies have shown that “bold” individuals don’t change their behaviors much between social and asocial contexts (e.g., *Parus major*, (Van Oers et al., 2005); *Parus cinereus* (Bibi et al., 2022), while others have shown the effect of social context depends on the behavioral phenotypes of nearby group members (e.g., *Danio rerio*, (Guayasamin et al., 2017); *Taeniopygia guttata* (Kerman et al., 2018; Mainwaring et al., 2011).

It is also possible to examine suites of correlated behaviors in a given species, including the differences in plasticity between these contexts, in a way that gives some parallel to personality dimensions. The Behavioral Syndromes approach (Sih, Bell, Johnson, et al., 2004) allows us to find differences in the patterns of correlated behaviors between groups in a population. Typically employing principal component analysis or factor analysis approaches, identified principal components or factors represent the most prominent dimensions of individual differences in a studied population, such that higher or lower scores on these dimensions are the result of patterns of correlated behavioral variables. This framework has been applied to such wide-ranging topics as regional variation in behavior in the delicate skink, *Lampropholis*

delicata (Michelangeli et al., 2018), differences in social behavior among the presumed asocial cane toad, *Rhinella marina* (Gonzalez-Bernal et al., 2014), decision-making speed and accuracy of three-spined sticklebacks, *Gasterosteus aculeatus* (Mamuneas et al., 2015), and even variation in “boldness” of across different colonies of ants, *Aphaenogaster senilis* (Blight et al., 2016).

Behavioral syndromes also allow us to make comparisons across species, especially where we might expect some behavioral overlap between two or more species (Sih, Bell, & Johnson, 2004). While this was always an intended use of the framework, it is only rarely and more recently that it has been applied in this way. Two different labs studying several fish species have separately used Behavioral Syndromes to compare two species of mosquito fish, *Gambusia affinis* & *Gambusia holbrooki* (Michelangeli et al., 2020) and two species of fugu, *Takifugu rubripes* & *Takifugu niphobles* (Hosoya et al., 2015). The ability to compare across species also makes it an ideal method for studying the behavioral diversity of mixed-species groups.

Mixed-species groups (MSGs) provide many of the same important benefits of living in a single-species social group (e.g., increased foraging efficiency, lower predation risk) but with reduced intra-specific competition (e.g., no mate competition, foraging niche differences) (Goodale et al., 2020). Some MSGs formed by taxonomically diverse species may provide asymmetric benefits to each other, such as wrasses, *Halichores hortulanus*, who receive a disproportionate benefit when hunting alongside goatfish, *Parupeneus barberinus* (Lukoschek & McCormick, 2002). MSGs made up of closely related species more commonly provide “supplementary” benefits, meaning they function very similarly to a single-species group – providing basic and fairly equal benefits toward finding and evaluating the quality of food patches and remaining vigilant for many of the same predators (Goodale et al., 2020; Morse, 1970).

The formation and maintenance of MSGs can depend on behavior differences in the component species. For example, if the members of one species are more likely to alarm call when a predator is detected, all other members of their group (conspecifics and heterospecifics alike) stand to benefit, encouraging the formation of such a group through their behavior (Gil et al., 2017; Goodale & Kotagama, 2005). Such “sentinel species” have been identified in MSGs, including fork-tailed drongos (*Dicrurus adsimilis*) in groups of other African bird and mammal species (Flower et al., 2014), chemically signaling fish (Chivers & Smith, 1998), and tufted titmice (*Baeolophus bicolor*) in groups of other passerines in the eastern United States (Hetrick & Sieving, 2011).

While single-species social groups have received a wealth of attention from animal personality researchers, the diversity of behavioral types beyond species differences in behavior is rarely considered in mixed-species groups. However, intraspecific variation in behavior should not be ignored in mixed-species contexts, especially in groups primarily composed of closely related species with similar ecologies. For example, across Europe, the great tit (*Parus major*) is the nuclear species in large and diverse mixed-species flocks of songbirds, and their individual personalities correlate strongly to their dominance during foraging and exploratory behaviors (Bibi et al., 2019, 2023). This is the same social environment a mixed-species group might form,

and the behavior of that group is at least in part influenced by these personality differences among the great tits in the flock. Even within non-cooperative interspecific interactions personality can play an important role: the overlapping ranges and cohabitation of various stickleback species seems to be constrained by the relative boldness of individuals within and between species (Webster et al., 2009).

Behaviorally diverse groups may also be better suited to solving novel problems in changing environments, as the additive plasticity between members makes producing a favorable solution more likely. Freeberg et al. (2017) found evidence to suggest that individuals in more diverse mixed-species flocks of songbirds were more likely to take seed from a novel feeder in a natural environment. It is possible that individuals were responding to the presence of the species they attend to rather than the novel feeder problem itself, as when Martinez et al. (2018) found that ant-following species of birds in the Amazon relied on the abundance and species diversity of other bird species present to determine the value of a foraging patch. It is frequently assumed to be beneficial to gather redundant information from multiple species, a unique advantage of mixed-species living (Goodale & Magrath, 2024). However, within-species differences in the preference for socially or individually gathered information could dictate whether a given individual enjoys such a benefit (Arvidsson & Matthysen, 2016; Stamps, 2007; Stamps & Biro, 2016).

The current study examines behavioral differences in tufted titmice and Carolina chickadees (*Poecile carolinensis*), two nuclear species in mixed-species flocks that form over the winter season in the eastern United States. Both species are in the family Paridae, share functionally and acoustically similar alarm calls (Krams et al., 2012), have similar diets, and are prey to the same predators. Tufted titmice are larger, heavier, and are generally dominant over Carolina chickadees in foraging contexts (Cimprich & Grubb Jr, 1994). Individual differences in boldness behaviors are not as well studied in these two species as they are in their relatives such as black-capped chickadees (*Poecile atricapillus*) (Jenkins et al., 2023) and great tits (Aplin, Farine, et al., 2013; Bibi et al., 2019; Vrublevska et al., 2015). Using a suite of behavioral assays related to boldness behaviors, principal components analysis (PCA), and hierarchical cluster analysis (HCA), I determined the most common bold/shy behavioral syndromes for this study system in semi-natural captive environments. To test whether behavioral diversity was significantly increased by species diversity, I measured the species composition of each identified cluster; if one or more clusters were primarily composed of just one of the two species, it would indicate that this cluster of behaviors was based upon unique contributions by that species. Finding no difference in behavior between these two closely related species could point towards evolutionary constraints of the behavioral syndromes shared between these species (Dochtermann & Dingemanse, 2013); finding species differences only between single-species flocks might suggest that some aspect of mixed-species group living demands or reinforces similarity in behavior.

Methods

Subjects and housing

A total of 66 birds used for the current study came from simultaneous studies requiring temporary aviary housing (see Chapter 2). Wild tufted titmice and Carolina chickadees were captured at feeder platforms at the University of Tennessee Forestry Resources AgResearch and Education Center (UTFRREC) in Oak Ridge, TN. Each platform was > 400 m away from its neighbor to ensure that each platform represented distinct overwintering flocks (Bartmess-Levasseur et al., 2010). Capture protocol involved the use of both mist nets and treadle traps to gather flocks of 3 individuals (species composition varied, but was always limited to tufted titmice and Carolina chickadees) that were housed in large outdoor aviaries (6 m × 9 m and 3.5 m high) as experimental but familiar flocks. While in the aviary, flocks had access to ad libitum food and water and were held for a period of 12-18 days depending on weather and protocol of their initial study prior to participation in the current study. This time period included a 5 day acclimation period.

Materials and procedures

RFID dataloggers and receivers were obtained from Eccel Technology Ltd (Grobby, UK). RFID antennas were attached to food dishes with electrical tape to capture visits to each dish. Each bird caught for the study was banded with a plastic PIT tag with a unique ID. To extend battery life, each receiver was configured to collect data only between the hours of 7am and 6pm Eastern Standard Time, as pilot data indicated this time frame captured greater than 95% of the visits to the seed dishes during aviary housing.

The tests below were given in their listed order for all flocks until birds were released from the aviary back at their capture sites at the conclusion of the Novel Environment Test.

Variables deduced from the RFID data were:

- *Order of approach* – The rank order of the birds in each flock to approach during the test period, with respect to each individual's first approach only. If one or more birds in a given flock did not approach during the test period, they were scored as a 4 rather than a 1, 2, or 3 so that late approachers were not considered equal to non-approachers in the data.
- *Order latency* – The time in seconds since the start of the test (for the first to approach) or the time in seconds since the previously ranked bird's first approach. For example, if the first bird visited 3 minutes after the start of the test and the second bird first visited 5 minutes since the start of the test, the birds were scored 180 s and 120 s respectively. This was to account for the qualitative difference in the test between a presentation no familiar flock mate had approached to one that one (or two) flockmates had safely approached. The interpretation of this variable is then different for the first to approach (a measure of boldness) than for subsequent birds (a measure of boldness interacting with some local enhancement

or other social learning). Any bird that did not approach during the test period was scored for the maximum length of the test.

- *Visit rate* – A ratio between the number of visits during the test period and the length of the test itself. Between-species differences in metabolism should mean that virtually every titmouse has a higher visit rate than every chickadee, but within-species differences in activity level have been linked to both personality (Jenkins et al., 2023) and social context (Bibi et al., 2022).

Behavioral syndromes testing

The individuals included in this study were previously tested in Chapter 2. These assays used a combination of RFID automation similar to the Novel Feeder Test above and video-recordings. Twelve variables were derived from these tests (4 tests with 3 measures each). Detailed descriptions of each test, measure, principal component eigenvalues and loadings are available in Chapter 2. Brief descriptions of each test are below.

Novel object presentation: A red disposable plastic cup was placed several centimeters away from one of the two seed dishes for 1 hour. The same three variables were recorded as in the Novel Feeder Test.

Mealworm presentation: Fifteen medium-sized mealworms were placed in one of the two seed dishes with ad libitum seed in the other. The presentation lasted 24 hours and only the first 15 unique visits to the mealworm dish were counted for each flock. The same three variables were recorded as in the Novel Feeder Test.

Predator presentations: A plastic hawk model and a rubber snake model were presented at least 1 day apart within a few centimeters of one of the seed dishes for 2 hours each. The scores for both tests were pooled together to account for missing data for some tests for some flocks (see Chapter 2). The same three variables were recorded as in the Novel Feeder Test.

Novel environment test: Individual birds were placed alone in a small white bird cage (approximately 0.8 m x 0.8 m x 1.2 m) with two cross-length perches and two seed dishes stocked with black oil sunflower seed. Birds video recorded for 10 min to measure “Number of Movements,” “Longest Freeze,” and “Number of Seeds Taken.” This test was the only one that occurred outside the aviary and without using RFID to automatically record data.

Statistical analysis

Data from the RFID dataloggers were cleaned and analyzed primarily in R (R Core Development Team, 2016). Code for cleaning and parsing the raw data according to test schedules is available from Scott Benson on request (sbenson9@vols.utk.edu). Any data for tests whose start and end times could not be confirmed to be accurate or accounted for were removed. Flocks with fewer than 4 complete tests were dropped from the analyses (16 of 35 flocks). For the remaining flocks, missing and removed data were imputed using Multivariate Imputation by Chained Equations (MICE; Azur et al. (2011)) prior to principal component analysis. This method allowed for the underlying correlation structures of the data to be preserved and for individual variables to have specified constraints on their maximum values (i.e., no “Order of

Approach” scores below 1 or above 4 or latencies longer than the test length) while limiting the risk of overfitting found in other methods. Due to a substantial and roughly equal loss of data during the Hawk and Snake presentations in one season (~50% of the data for each test, distributed randomly across flocks for that season) from equipment theft and other equipment failures, the scores of the two tests were pooled together for all individuals, creating 3 “Predator” variables where snake and hawk variables had been planned.

Scores for titmice and chickadees were pooled for both MICE and PCA without regard to species as a variable so as not to bias the model toward amplifying species differences. Twelve variables went into the PCA (order of approach, order latency and visit rate for “Novel Object,” “Mealworm,” and “Predator” tests; number of moves, longest freeze, and number of seeds taken for “Novel Environment” tests). Parallel analysis with the “paran” package in R (Dinno, 2018) allowed predictions for the optimal number of components to keep adjusting for any bias introduced by the imputation of missing values.

Hierarchical cluster analysis conducted on the PCA results helped to distinguish distinct patterns of scores across the principal components by comparing the Euclidean distances between individuals in the principal component space. This method iteratively merges the closest clusters based on these distances, eventually forming distinct clusters that minimize within-cluster distances and maximize between-cluster distances. The purpose of the cluster analysis used here was to determine the number of common behavioral syndromes in the data as well as to identify whether any species-specific clusters of behaviors existed in the term-reduced PCA data. If, for example, a cluster primarily composed of one species existed, then we can be confident that our mixed-species PCA was able to identify where one species contributed additional behavioral variance to the flock. The goal of this approach is similar to some discriminant function analysis approaches in the behavioral literature where agnostic models are evaluated for their ability to correctly classify *a priori* groups from a set of behavioral data. An especially successful example was such an approach by Almlie and Burghardt (2006), in which their model was able to correctly classify 100% of ratsnakes (*Elaphe obsoleta*) raised in either an enriched or non-enriched environment based on their scores on a suite of behavioral tests.

A multiple-response permutation procedure (MRPP)(Zimmerman et al., 1985) was applied to the PCA scores to determine whether PCA scores differed between species, flock compositions, or as a result of an interaction between the two. MRPP uses the average multidimensional distances between individuals in the PCA space to construct random permutations and an “expected delta” representing the random closeness between individuals without a group assignment. The expected delta is then compared to the observed delta of the *a priori* group designations (i.e., species, flock composition, or a species x flock composition designation) to determine whether group differences are greater than expected by random chance (Carter & Feeney, 2012). MRPP also gives a value “A” ($1 - \text{observed delta/expected delta}$) to represent the fit of each grouping; values very near zero indicate random behavior, a value of 1 indicates identical behaviors for every group member, and realistic observed values for A in the ecological literature are considered quite strong when > 0.3 (Carter & Feeney, 2012; Mccune et al., 2002).

Results

Data for 66 birds across 22 artificial flocks were included in the final analysis; 35 were tufted titmice and 31 were Carolina chickadees. See Figure 6 for species and flock composition comparisons of Order Latency scores per test. Completeness of the data for each individual varied considerably, even after pooling the two predator tests; up to 35 included individuals did not have reliable data for the Novel Object test, due primarily to unrecoverable time and date drift of the timestamped RFID data. The data loss, while at times substantial, was

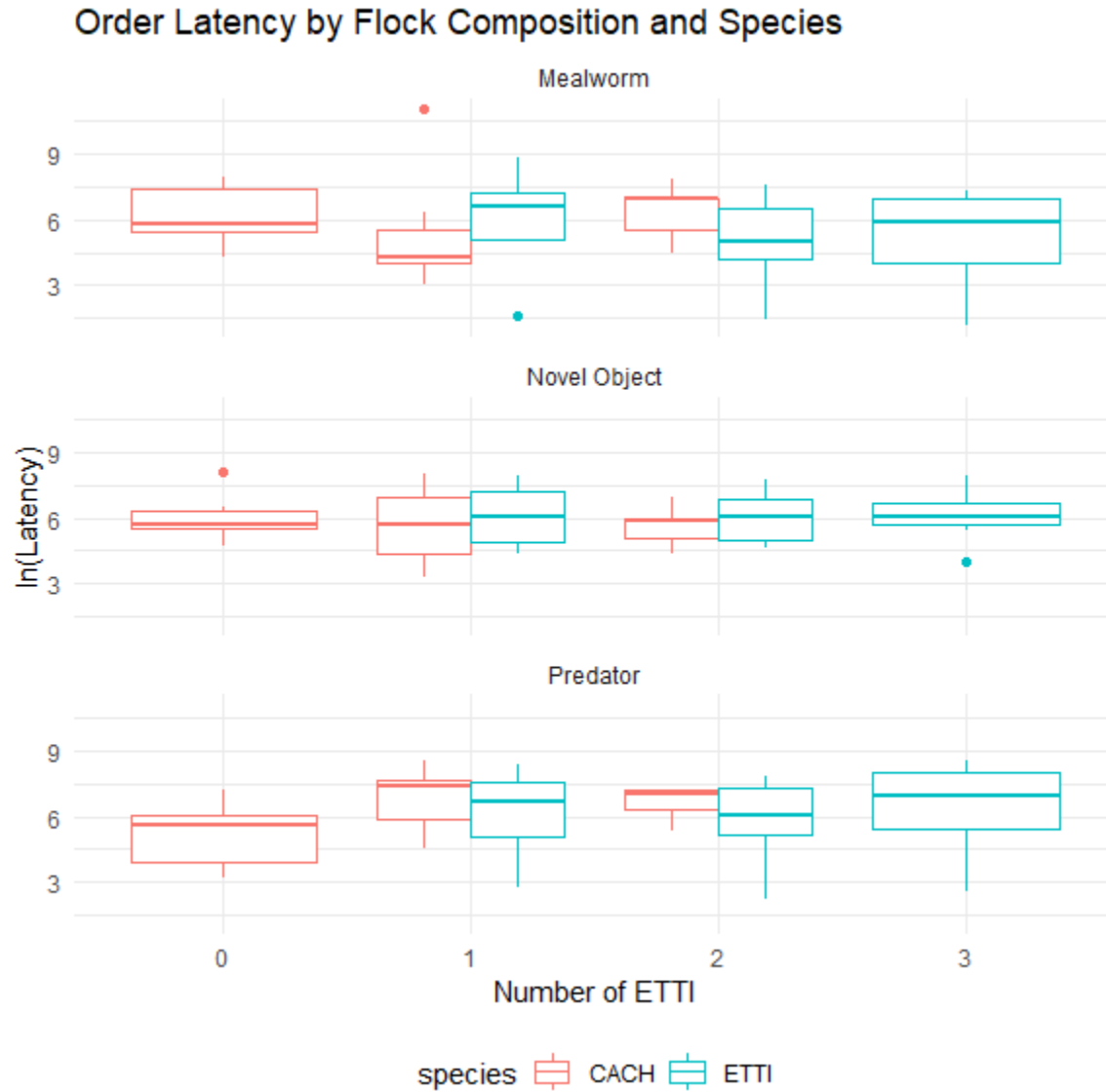


Figure 6: Order latency by flock composition and species. Latency is presented transformed with natural logarithm for the Mealworm test, Novel Object test, and the combined Predator test scores. Flock composition is represented as the number of ETTI in the three-bird flock, such that 0 and 3 are single-species flocks.

not systematic, and any individuals with fewer than 6 measured variables (i.e., 2 of the 4 complete tests) were discarded from analysis (39 of the original sample size of 105). For a detailed explanation of the PCA results, see Chapter 3 and Table 5.

Cluster analysis

The NbClust analysis (Charrad et al., 2014) found a slight plurality in favor of examining just two clusters after polling 30 different indices of optimal numbers of clusters. Clusters were derived and cluster labels assigned to individuals using hierarchical cluster analysis. See Figure 7 for the resulting dendrogram. Below are interpretations of these clusters based on their mean scores on each of the five principal components described above.

Cluster 1 included most birds (59 out of 66; 24 chickadees and all 35 titmice). Cluster 1, as a group, did not have any principal component scores that differed significantly from zero, likely due to representing most of the dataset.

Cluster 2 only included only 7 individuals, but all 7 of those individuals were chickadees, confirming that cluster analysis could separate out unique contributions to behavioral diversity of one or more species in the mixed-species data. Cluster 2, as a group, scored lower than average on “Neophobia” ($z = -3.74$, $sd = .429$), and significantly higher on “Neophobic Exploration” ($z = 1.87$, $sd = .669$).

MRPP

The multiple-response permutation procedure found significant group differences for flock species composition, with group mean Euclidean distances of 3.392, 4.753, 4.202, and 3.308 for flocks with 0, 1, 2, and 3 titmice respectively ($A = .023$, observed delta = 4.043, expected delta = 4.14, $p = 0.019$), meaning both of the mixed-species flock compositions produced more variable behaviors than either single-species composition. MRPP also found significant group differences between species, with CACH producing more variable behavior than ETTI overall (CACH mean distance = 4.758, ETTI mean distance = 3.503, $A = 0.012$, observed delta = 4.093, expected delta = 4.14, $p = .039$). See Figure 8 for comparison of species differences in PCA scores across components.

To examine interaction effects, groups were designated for MRPP using the focal individuals’ own species and whether they were tested in single- or mixed-species flocks. MRPP found significant group differences for this designation, with mean distances of 4.966 for “CACH-mixed,” 3.392 for “CACH-single,” 3.694 for “ETTI-mixed,” and 3.308 for “ETTI-single” ($A = .042$, observed delta = 3.965, expected delta = 4.14, $p < 0.001$), meaning the chickadees in mixed-species flocks were much more variable in their observed behaviors than any other species/composition combination.

Discussion

Consistent with my predictions, Carolina chickadees and tufted titmice showed significantly different distributions of behavior. The largest variation in behavior was found

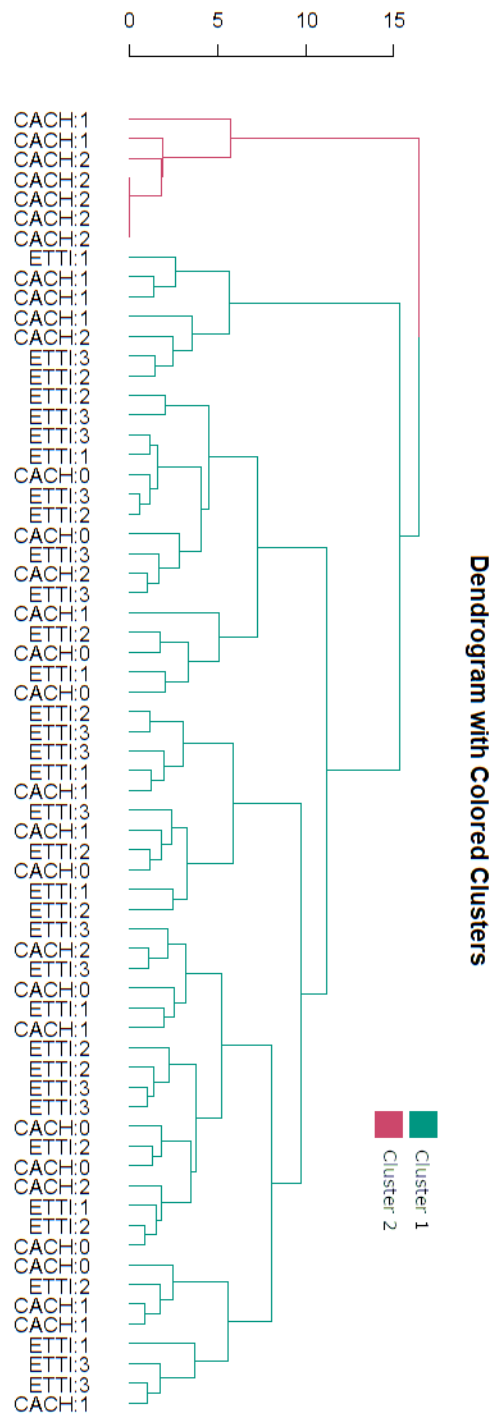


Figure 7: Dendrogram from hierarchical cluster analysis.

All birds are organized by their relative distances from each other in principal component space. Cluster 2 highlighted in red, containing only chickadees (“CACH”). Number after “:” is the number of ETTI in the flock (“0” and “3” represent single-species flocks).

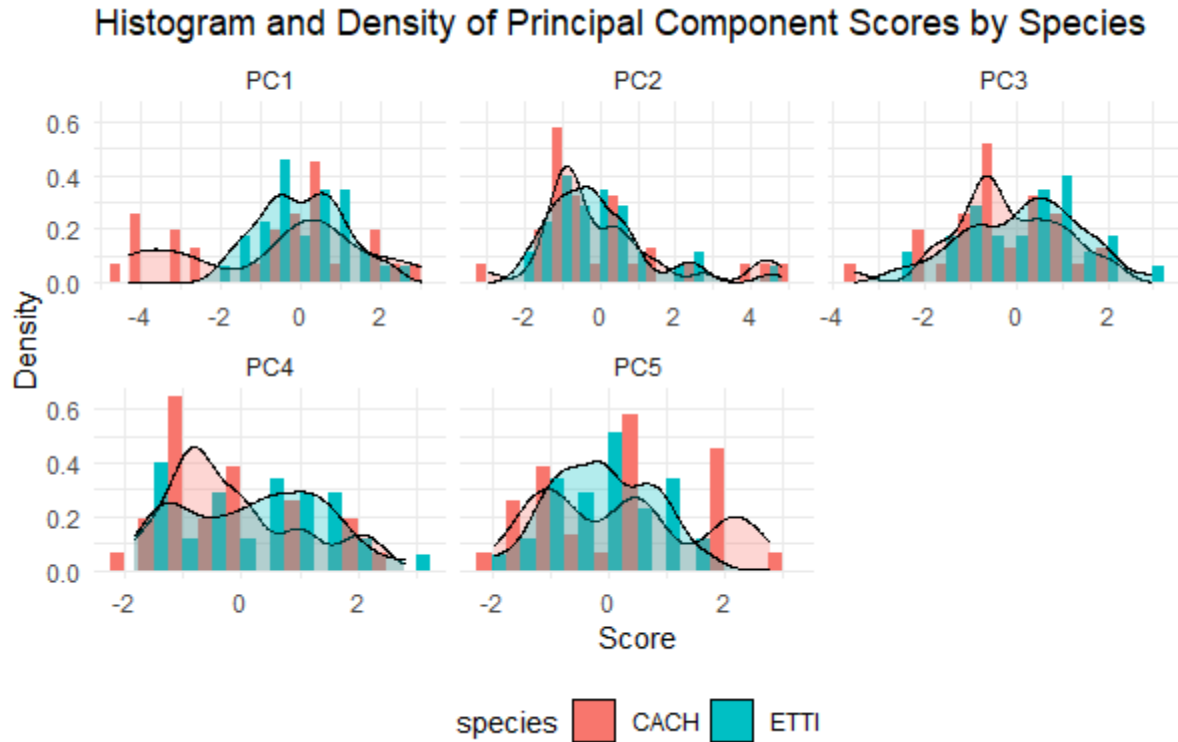


Figure 8: Histogram and density of principal component scores by species. Species differences in scores on the five retained principal components are represented by non-overlapping density curves on each graph. PCs represent “Neophobia,” “Non-exploration,” “Neophilic Threat Discrimination,” “Social Information Preference,” and “Neophobic Exploration” respectively, per the descriptions in Chapter 2.

among chickadees in mixed-species flocks. Of the two clusters identified (or more accurately, the one cluster that was separated from the bulk of the sample), one of them was made up entirely of chickadees. This cluster of chickadees differed from the rest of the sample in having lower scores in the Neophobia principal component and higher scores in the Neophobic-Exploration component. Interestingly, it was observed that chickadees assigned to the smaller cluster were always in mixed-species flocks rather than chickadee-only flocks, although with only seven individuals in the cluster and less than half the flocks being single-species flocks more data are needed to confirm this result. This finding was corroborated by the results of the three MRPPs. Chickadees (in the species comparison) and mixed-species groups (in the flock composition comparison) produced more variable behaviors than titmice and single-species groups, respectively. When comparing combinations of species and flock composition, the mean Euclidean distance (e.g., behavioral diversity) was the highest between chickadees in mixed-species groups and lowest for titmice in single-species groups, suggesting that it was largely chickadees' responses to flock composition that produced the observed effects on behavioral variability of these groups.

Presumably, a large sample size and more test variables would allow for more clusters to be reliably identified in the data. With more granularity in the clusters, we might be able to describe under which conditions of flock composition some behavioral syndromes emerge. Future research in this area should assess whether shifts in behavioral syndromes are more or less adaptive to the social niche (Montiglio et al., 2013) over the ecological niche (Fokkema et al., 2021; Sih et al., 2015), and whether any difference in tendency of reactivity between individuals could be associated with their preference for social versus individual information gathering (Arvidsson & Matthysen, 2016; Mccune et al., 2018; Stamps & Biro, 2016). It may also be that typical chickadee foraging behaviors are disrupted in the presence of dominant and competing titmice; an earlier study in which chickadees in flocks with all titmice removed foraged in niches more typically occupied by titmice (Cimprich & Grubb Jr, 1994).

Identifying behavioral syndromes with the three-step technique of PCA → HCA → MRPP helped to conceptually cover some of the limitations of whole-group testing. Several of the test measures used for the PCA were not independent measures (i.e., ranking approach order in a flock). Looking at individual scores on the principal component alone may reduce the number of terms needed to describe an individual, but the interdependent nature of the data limits the number of “top” spots in each category, even where true boldness between two individuals might not actually be categorically different. This interdependence issue was made apparent when trying to apply a consistency measure to Principal Component 5, for which the majority of the component score was made up of this ranked Order of Approach data (see Table 5). Through the clustering of individuals across the 5-dimensional PCA space, there were several avenues to membership in each behavioral syndromes category. Without clustering, two individuals with very similar behaviors necessarily have different orders of approach on each test if they belong to the same flock, which would result in substantial differences on any principal component with heavy loading for order of approach, despite the individuals scoring similarly in every measure. Instead, clustering considers similarity in scores across all five components, allowing for slight

deviations – like what a slip in rank against an equal might cause – without changing the classification of an individual's behavioral syndrome.

CHAPTER 4: ASOCIAL TESTING OF SOCIAL SPECIES

“Exposure to novelty and the sudden interruption or denial of social contact are both regarded as frightening and/or stressful procedures.”

-Jones and Merry (1988)

Introduction

Behavior assays in many areas of animal behavior research are most often given in tightly controlled conditions and in isolation, limiting outside factors to best understand any individual differences in behavior. In many social species, however, the condition of isolation is alien and can serve as a stressor and confound during these behavioral tests (Bibi et al., 2022; Jones & Merry, 1988). Too little attention is paid to the possibility that any significant individual differences found may be (at least in part) attributable to differences in response to sudden and/or prolonged isolation.

Isolation is a known stressor among well-studied groups. Mice and rats housed in isolation are found to have increased metabolic rates, experience weight loss, show signs of depression, and even suffer reduced neuroplasticity and permanent impairments in memory function (Liu et al., 2020; Muta et al., 2023; Potrebic et al., 2022). Similarly, European starlings (*Sturnus vulgaris*) lost weight, raised their activity levels, and showed increased testosterone and corticosterone levels after a period of isolation (Apfelbeck & Raess, 2008). Chronically isolated zebrafish (*Danio rerio*) showed similar signs of depression and reduced whole-brain serotonin levels (Shams et al., 2015). Corticosterone levels in zebra finches (*Taeniopygia guttata*) were significantly higher after just 10 minutes in isolation (Banerjee & Adkins-Regan, 2011), and after being isolated overnight calling rates were lower and the frequencies (Hz) of those calls were higher, a sign of increased stress (Perez et al., 2012). Isolated chicks (*Gallus gallus*) recovering from a painful injection produced more frequent distress calls when isolated compared to when isolated with a mirror nearby, suggesting even the simulated presence of conspecifics is less stressful than complete isolation (Feltenstein et al., 2002).

Even where we might be comfortable with data from an individual being tested in stressful situations to examine individual differences in reactivity and compensatory behaviors, we are limited in the external validity of these tests if what we are interested in is the natural behavior of individuals that vary in personality. Testing individuals in isolation may yield an entirely different pattern of behavior than testing in a natural social context. The current research on the subject underscores a lack of predictability in which species or individuals will change, and how much, and in what direction between isolated and social contexts. Zebra finches given novel object tests in isolation as well as when tested in familiar and unfamiliar flocks were significantly bolder in isolation; interestingly, when ranked for boldness prior to the test, all zebra finches remained in their relative ranks for boldness across all contexts, with no significant effect of sex or the familiarity of flockmates (Kerman et al., 2018).

In this case we see a flat, main effect of isolation increasing boldness in a species, but such a clean effect may be rare in the animal kingdom. For example, mud crabs (*Panopeus herbstii*) complicates this picture and illustrates the layered problems isolated personality testing can produce. This species organizes itself in denser social groups in more favorable habitats and is less social in less favorable ones. Mud crabs in larger social groups were generally bolder in

the presence of chemical predator cues than those in smaller groups while shyer mud crabs were the most plastic in their behavior between contexts, becoming significantly bolder when moving from smaller to larger congregations (Belgrad & Griffen, 2017). This leaves us with a species that will test shyer in isolation than in the wild, and whose actual ecological boldness depends on an interaction between habitat quality, group size, and behavioral phenotype. Of all the species tested in isolation under the banner of animal personality research, how many are zebra finches and how many are mud crabs?

The scope of the current study is not to answer that question comprehensively, but to exercise the same due diligence recommended for all study species whose personality is of ecological interest to their lab. Using a novel environment test on tufted titmice (*Baeolophus bicolor*) in eastern Tennessee, I compared the mean scores of activity level, foraging rate, and freezing behaviors between isolated and socially tested groups. Based on trends in the literature, I expected that the isolation plus novel environment condition would produce more anxious behaviors than the novel environment condition alone; I predicted higher rates of movement, lower foraging rates, and longer freezes in the Asocial context group compared to the Social context group.

Methods

Subjects and housing

A total of 53 (26 in Asocial group, 27 in Social group) tufted titmice used for the current study were a subset from the total data collected for Chapters 2-3. Wild birds were captured at feeder platforms at the University of Tennessee Forestry Resources, AgResearch, and Education Center (UTFRREC) in Oak Ridge, TN. Each platform was > 400 m away from its neighbor to ensure that each platform represented a distinct overwintering flocks (Bartmess-Levasseur et al., 2010). The capture protocol involved the use of both mist nets and treadle traps to gather flocks of 3 individuals (species composition varied but was always limited to tufted titmice and Carolina chickadees) to house in large outdoor aviaries (6 m × 9 m and 3.5 m high) as experimental but familiar flocks. While in the aviary, flocks had access to ad libitum food and water and were held for a period of 12-18 days depending on weather and protocol of their initial study prior to participation in the current study.

Materials

The novel environment was a small white bird cage (approximately 0.8 m x 0.8 m x 1.2 m) with two cross-length perches and two seed dishes stocked with black oil sunflower seed (see Figure 3). All trials were video recorded using a Canon VIXIA HF R800 camera set on a nearby tripod such that the perspective angle allowed easy viewing of all 6 inside faces of the cage. The cage itself was set on a folding table approximately 0.5 m off the ground, depending on the terrain.

Procedure

At the conclusion of each flock's previous study participation, an open treadle trap was introduced to the aviary to allow each bird to become acclimated to taking seed from only within

the trap. The next day the trap was set and the first bird captured was promptly removed and placed in a small cotton bag in preparation for the Social Context version of the novel environment test (sNET). The table, cage, and video camera were brought in to the aviary in the presence of the remaining two flockmates and set up to record the trial. The bagged bird was removed from the bag, the bird's flock ID and band ID were spoken aloud to the camera before being placed in the white cage. I immediately left the area and set a timer for 11 minutes, after which time I returned, ended the recording, and removed the bird from the cage to be released back at its respective home field site. The treadle trap was then reset to capture the remaining birds in the aviary for release.

Subsequent birds caught received the Isolated Context version of the test (iNET) prior to release at their site of capture. After being bagged from the treadle trap, each bird was taken approximately 50 meters from their aviary and remaining flockmate (if any) to enter the novel environment cage in an open area away from other birds, and > 400 m away from their own wild flockmates. The procedure for the iNET was otherwise identical.

All behavior was coded in BORIS (Friard & Gamba, 2016). Trials were marked continuously for 10 min, regardless of the length of the actual video. The beginning of the 10-min trial was no earlier than 30 s after the bird entered the cage on video to allow the researcher to leave the area, and the end of the trial was no later than 30 sec from the end of the video to account for the re-approach of the researcher. Ten different locations were marked according to where the bird perched within the cage (i.e., top, bottom, left, right, front, back, upper perch, bottom perch, left bowl, right bowl). The total moves during the test were counted as changes between perch locations (either by flying, hopping, or climbing). The total number of seeds taken from either bowl was also coded during the trial. The longest freeze was determined at the end of the trial, measuring the longest period in seconds that the bird was not moving from its current perch.

I coded all video trials, and reliability was assessed on a subset of trials (12 trials or 18.2%) by a trained research assistant naïve to the study's predictions. Interrater reliability measured by Cohen's Kappa was 0.782, indicating substantial agreement on Novel Environment Test trials.

Statistical Analysis – Hypothesis testing originally used MANOVA with the 3 measured test variables as the dependent term against test context as the predictor. However, the test variables violated assumptions of normality and homogeneity of covariance. Instead, each variable was tested in Generalized Linear Models with Test Context as the predictor variable. Overdispersion checks rejected a Poisson distribution for all three variables, so a negative binomial distribution was used instead.

Results

The negative binomial generalized linear models showed that Test Context significantly predicted Maximum Freeze Duration ($z = -2.04$, $p = 0.0413$) and Seeds Taken ($z = -2.229$, $p = 0.0258$), but did not predict Total Moves ($z = 1.291$, $p = 0.197$). The Social context group on

average froze for approximately 62.6% as much time as the Asocial group (39.5s vs 63.2s). The Social context group also took 44.4% as many seeds on average as the Asocial group (1.56 vs 3.50; see Figure 9).

Discussion

Results from the GLMs for each of the test variables contradicted my expectations. Tufted titmice tested in asocial contexts in a novel environment were more likely to forage than titmice tested within the context of their aviary flock. Isolated titmice also had longer maximum freeze durations than titmice tested in their social group (consistent with expectations), but the higher maximum freeze duration did not seem to significantly change the measured number of movements during the test period. In the case of tufted titmice, sudden and short-term isolation did not have the presumed effect but did produce significantly different behaviors in comparison to socially-tested titmice.

These results are somewhat consistent with results discussed in zebra finches (Kerman et al., 2018), where isolated individuals appeared to have higher levels of “boldness.” While those authors credit “Social Inhibition” or an expected reduction in boldness behaviors that may be attributable to the social signal of who in a group has and has not engaged with a task yet (Dugatkin & Alfieri, 2003; Mainwaring et al., 2011; Webster & Ward, 2011), much of the supporting evidence focuses on measures of neophobia through novel objects and novel feeding opportunities (Kerman et al., 2018; Van Oers et al., 2005) and more rarely on exploratory behavior (Mainwaring et al., 2011). Where individuals in a group context might wait to take advantage of social cues and socially generated information to inform their behavior around novel objects, foods, and environments (Arvidsson & Matthysen, 2016; Stamps & Biro, 2016), it is still generally expected that a benefit of group contexts is a reduced need for individual investment in vigilance (Lima, 1987). The combination of increased foraging and reduced time spent freezing in the above study instead seems to suggest a reduced investment overall in vigilance in favor of increased foraging.

The similar amount of movement between groups may be easier to explain, if we assume that the sudden isolation did not lead to an anxious reaction. If each movement within the cage provides a new vantage point and opportunity to explore the limits and contents of the cage, then the total movements around the cage might be about gathering information about the new surroundings. Information gathering strategies can vary widely across contexts and individuals, but the evidence that context did not change the average number of movements may suggest that isolation does not significantly affect each individual’s preference for information gathering (Arvidsson & Matthysen, 2016). While many boldness-related behaviors in this and other study systems have been shown to vary according to flock species and behavioral composition (Guayasamin et al., 2017), this lack of difference between groups in movement may predict that information gathering preference is a distinct and parallel dimension of individual difference in titmice. More research in this area is needed, but uncoupling isolation behaviors from the “boldness” umbrella may help understand the contradictory effects of isolation on boldness behaviors in the literature.

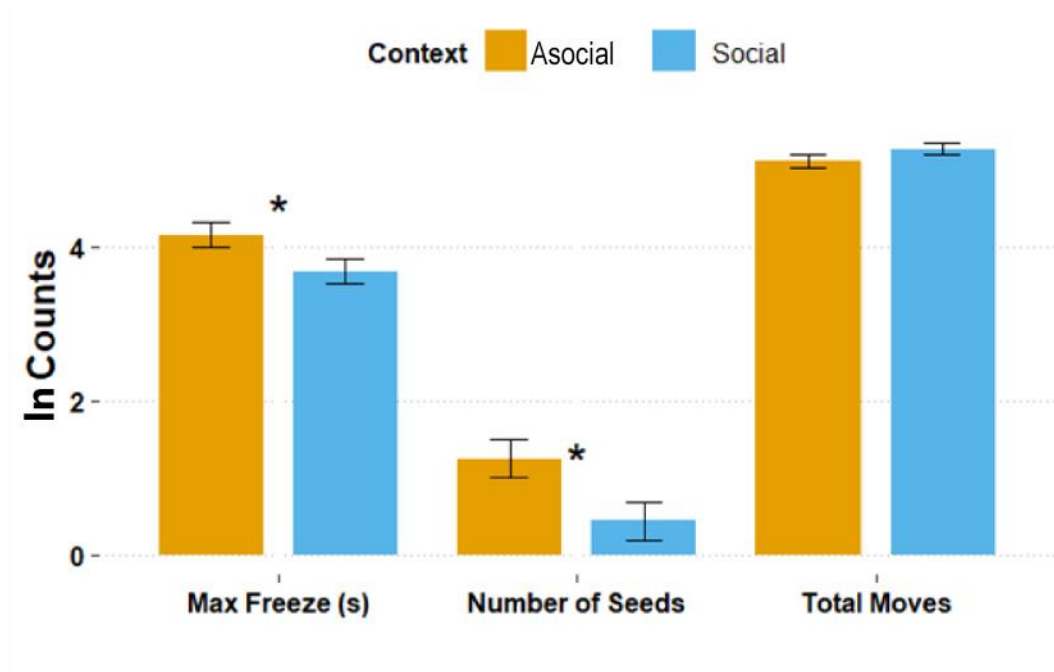


Figure 9: NET scores in asocial vs social contexts. $\ln(\text{Counts})$ of each measured variable between Asocial and Social contexts. Asterisks (*) mark the statistically significant between-context differences in scores on Maximum Freeze in seconds and the Number of Seeds taken during the test. Total Moves during the test was not significantly different between contexts.

The results of this study speak to the broader implications of isolation and asocial testing of social species as a potential confound in animal behavior research. Kimball and Lattin (2023) suggest 7 “sins” for tests of neophobia but do not include considerations for asocial versus social contexts, even among highly social species. Even where individuals may strongly prefer their own information or strategies (Arvidsson & Matthysen, 2016), the presence of variability on this dimension within a species can reduce the ecological validity of tests that remove the option to rely on group members.

CHAPTER 5: CONCLUSION

Summary

Mixed-species groups appear to provide unique benefits above and beyond what single-species groups tend to offer (Goodale & Magrath, 2024; Goodale et al., 2020; Morse, 1977). The goal of this dissertation was to explore how mixed-species groups might be better equipped for the shifting opportunities and threats in changing environments by examining how behavioral diversity within species-diverse flocks contributed to solving novel foraging problems. The body of work presented here first established that such an expected increase in behavioral diversity was present in a mixed-species flock of birds, then examined how that behavioral variability might predict individual responses to a novel foraging task. Several methodological questions about how best to test for these effects were answered throughout each chapter.

Chapter 2 provided evidence that models using principal component scores were better fitting than models using than any individual test when predicting novel feeder behavior. Even where examples in the literature applied principal component analysis to their predictive tests, there was a tendency to do so to reduce a single test to a single score, rather than allowing variables from disparate tests to correlate freely. However, the success of this approach likely depended on the true relationship between the identified components and the novel feeder test, and two alternatives were possible: 1) no components were good predictors but one individual test was, meaning that the relationship between that test and the novel feeder behavior was suppressed in the principal component space, or 2) no components and no tests were good predictors of novel feeder behavior, meaning that it is likely foraging in contexts involving novel stimuli is its own dimension of behavioral variation uncorrelated to the others and a good candidate for a sixth component in future principal component analysis that includes those novel feeder measures.

Chapter 2 also attempted a new kind of repeatability measure unique to behavioral syndromes, adapted from recommendations from Huang et al. (2018). Repeatability is considered an essential measure for animal behavior studies (Bell et al., 2009). By looking at the relationships between the top-contributing variables of each behavioral syndromes dimension (principal component), a consistency score was assigned to individuals based on whether the relationship between their individual test scores conformed to the expected relationship given by that behavioral dimension. In effect, the consistency score tested how “repeatable” a behavioral syndrome was across multiple tests of that behavioral type. By averaging the consistency scores across the sample, I showed that 4 of the 5 identified behavioral dimensions were significantly consistent, analogous to an individual test being highly repeatable.

Chapter 3 described the unique contribution by Carolina chickadees to small mixed-species flocks with tufted titmice, supporting the hypothesis that species diversity should lead to greater behavioral diversity, even among as ecologically similar and closely related species as tufted titmice and Carolina chickadees. Five dimensions of behavioral variability were described

for this study system, and a small cluster of chickadees unique to mixed-species flocks scored especially low in “Neophobia” and “Neophobic Exploration.”

Cluster analysis was shown to be a useful support tool for hypothesis testing around behavioral differences. The Behavioral Syndromes framework (Sih, Bell, Johnson, et al., 2004) favors principal component analysis approaches, but the resulting dimensions of inter-individual variability are themselves only descriptive; some model or prediction or correlation must be made after individual scores are known, and “individuals differ from one another” is hardly a scientifically interesting statement. By applying cluster analysis to the principal component space (calculated without species information), finding a cluster made up entirely of one species served as corroboration of a unique behavioral contribution of one species to an otherwise largely similar group.

Chapter 4 addressed a methodological concern that ran through the studies from the previous chapters: when assessing the personalities of group-living individuals, should they be measured in their groups or by themselves? While not a direct comparison of reaction norms within individuals, the between-subjects results offer some support to the claim that testing highly gregarious species in asocial contexts confounds the intended behavioral measure – and, importantly, in unpredictable directions. Ultimately, however, there remains disagreement among animal behavior researchers about whether testing behavioral syndromes or personalities in social contexts allows for a valid and independent measurement (Beekman & Jordan, 2017; Carter et al., 2013). While the data and references presented in Chapter 4 suggest that being removed from the group has an effect on behavior, Chapters 2, 3, and 4 all also provide evidence that individual behaviors are heavily influenced by other members of the group. This makes Chapter 4 less of an argument for social-context testing for the sake of accuracy and more of a rebuke of studies that don’t discuss the effect of sudden and/or prolonged isolation.

Broader considerations and future directions

Producer-scrourer dynamics

One limitation of the current design was decreased ability to assess producer-scrourer dynamics. Most of the data collection relied on readings of RFID tags on designated seed dishes or puzzle perches. While this offered the advantage of delivering low-cost, round-the-clock automated testing without the need for human coding of visual data (beyond the initial validation of the RFID method), any activity that occurred outside the range of the RFID antennas was not available for analysis. This would include a recorded individual spilling seeds out of the dish that an unrecorded scrourer could reach, although validation videos recorded no instances of this occurring.

This does not mean *nothing* can be learned about producers and scrourers in mixed-species flocks from this design. Information about the safety and value of artificial foraging patches (i.e., the feeder platforms) is both produced and scroured, so the order of approach offered information on who produces and who scrours this information. The caveat of course is to reiterate the Jingle-Jangle problem; is it fair to count an early approach as both a bold behavior and a producing behavior? Suggesting that bolder individuals produce information

while shyer individuals scrounge it is not a correlation between boldness and producing, it is a tautology. Several studies have been done that found connections between producer-scrounger strategies and behavioral syndromes (Aplin et al., 2012; Aplin & Morand-Ferron, 2017; Kurvers et al., 2010), though these studies did not depend on first-to-approach data as one of their behavioral measures.

Behavior diffusion

A follow-up to the Goodale et al. (2020) framework – which focuses on the types of benefits accrued and contributed by each member species in a mixed-species group – addresses the possibility of “interspecific information flow” (Goodale & Magrath, 2024). Interspecific information flow is not necessarily a form of social learning, but diversity as a cue may provide benefits similar to local or stimulus enhancement (Zentall & Galef, 1988). For example, playback experiments in a tropical system where several bird species follow army ants (*Eciton burchellii*) to feed on disturbed arthropods found that those bird species responded the strongest to playbacks with the higher species richness and apparent aggregation size (Martinez et al., 2018). More direct tests of information flow across species have also been carried out, such as the spread of information about the location of novel feeder sites among blue tits (*Cyanistes caeruleus*), great tits (*Parus major*), and marsh tits (*Poecile palustris*); while information seemed to flow more readily between conspecifics, heterospecific information flow was significant and species differed in the amount of information produced (i.e., foraging patches discovered) in their mixed-species flocks (Farine et al., 2015).

Even where the spread of a behavior throughout a mixed-species flock might explain how an advantage of diversity *could* affect all flock members, it would not fully explain a diversity effect on the initial latency to solve a novel foraging problem. Freeberg et al. (2024) found that for Carolina chickadees in mixed-species flocks, both the rate of success between flocks and the latency appeared to be related to the species diversity of each individual’s flock. One possibility is that mixed-species flocks show a greater range or diversity of behavioral types. Some studies have shown that, within single-species groups, behavioral variability does change depending on social contexts and the behavioral type of other group members – sometimes referred to as “relative personality” (Guayasamin et al., 2017). In Chapter 3, Carolina chickadees were shown to add to the behavioral type variability in their shared flocks with tufted titmice; cluster analysis on behavioral data sorted by principal component analysis identified a group of chickadees that tested significantly lower than the rest of the tested birds on Neophobia and a type of exploratory behavior. If certain patterns of behavior within a species (or individual) are typically only present in mixed-species contexts, it could be the additive variety in behavior that makes individuals in a mixed-species group more likely to “solve” novel foraging tasks just as accumulated perceptual differences are implicated in the increased foraging efficiency and predator detection outlined in the Goodale et al. (2020) mixed-species benefits framework.

A current study in progress will attempt to track the diffusion of behaviors through each flock according to social connection strength and species.

Cognitive aptitude

Behavioral differences between individuals can lead to differences in information gathered through experience (Carazo et al., 2014; Louison et al., 2019; Muth & Healy, 2014). Cognitive differences between those same individuals moderate how this information is used and retained. Recent research may already show how such a pathway can lead to predictable differences in some cognitive aptitudes such as learning and problem-solving.

Mazza et al. (2018) showed a speed-accuracy tradeoff in bank voles (*Myodes glareolus*) where bolder individuals were inflexible and more likely to make errors on an associative learning task. Correct answers correlated with increased decision time, and only shy individuals took their time when making decisions. While this would suggest favorability for shy/slower styles, a later study on ravens (*Corvus corax*) (Gallego-Abenza et al., 2020) showed that slower decision makers got their food stolen by flockmates who decided faster. Taken together, these studies reinforce that speed-accuracy IS a tradeoff and that the optimum strategy changes based on context.

However, consider an individual familiar with many different kinds of food and experienced with associative learning tasks. Compared to naïve individuals, we should predict fewer errors *and* shorter decision-times, especially when compared to those with similar personalities or cognitive styles. Past experience or a developed aptitude for dealing with novel problems is not easily controlled for in a wild setting. Future research in this area should consider tests of individual differences in cognitive aptitude, which is an important consideration for individual differences that receives relatively little attention outside of human research.

CONCLUSION

Taken together, the major finding of this dissertation is that effects of behavioral and/or species diversity and social context *can* be found where we would expect to find them, but the relationships themselves are unpredictable, even among closely related species. Future research should continue to explore the potential of behavioral system clustering techniques described in this dissertation to identify patterns of behavior unique to the mixed-species group context.

REFERENCES

- Adams, G., & Markus, H. R. (2001). Culture as patterns: An alternative approach to the problem of reification. *Culture & Psychology*, 7(3), 283-296.
<https://doi.org/10.1177/1354067X0173002>
- Almli, L. M., & Burghardt, G. M. (2006). Environmental enrichment alters the behavioral profile of ratsnakes (elaphe). *Journal of Applied Animal Welfare Science*, 9(2), 85-109.
https://doi.org/10.1207/s15327604jaws0902_1
- Apfelbeck, B., & Raess, M. (2008). Behavioural and hormonal effects of social isolation and neophobia in a gregarious bird species, the european starling (*sturnus vulgaris*). *Hormones and Behavior*, 54(3), 435-441.
<https://doi.org/https://doi.org/10.1016/j.yhbeh.2008.04.003>
- Aplin, L. M., Farine, D. R., Morand-Ferron, J., Cole, E. F., Cockburn, A., & Sheldon, B. C. (2013). Individual personalities predict social behaviour in wild networks of great tits (*parus major*). *Ecology Letters*, 16(11), 1365-1372. <https://doi.org/10.1111/ele.12181>
- Aplin, L. M., Farine, D. R., Morand-Ferron, J., & Sheldon, B. C. (2012). Social networks predict patch discovery in a wild population of songbirds. *Proceedings of the Royal Society B-Biological Sciences*, 279(1745), 4199-4205. <https://doi.org/10.1098/rspb.2012.1591>
- Aplin, L. M., & Morand-Ferron, J. (2017). Stable producer-scrounger dynamics in wild birds: Sociability and learning speed covary with scrounging behaviour. *Proceedings of the Royal Society B-Biological Sciences*, 284(1852), Article 20162872.
<https://doi.org/10.1098/rspb.2016.2872>
- Aplin, L. M., Sheldon, B. C., & Morand-Ferron, J. (2013). Milk bottles revisited: Social learning and individual variation in the blue tit, *cyanistes caeruleus*. *Animal Behaviour*, 85(6), 1225-1232. <https://doi.org/10.1016/j.anbehav.2013.03.009>
- Arvidsson, L. K., & Matthysen, E. (2016). Individual differences in foraging decisions: Information-gathering strategies or flexibility? *Behavioral Ecology*, 27(5), 1353-1361.
<https://doi.org/10.1093/beheco/arw054>
- Audet, J.-N., & Lefebvre, L. (2017). What's flexible in behavioral flexibility? *Behavioral Ecology*, 28(4), 943-947. <https://doi.org/10.1093/beheco/arx007>
- Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: What is it and how does it work? *Int J Methods Psychiatr Res*, 20(1), 40-49.
<https://doi.org/10.1002/mpr.329>
- Banerjee, S. B., & Adkins-Regan, E. (2011). Effect of isolation and conspecific presence in a novel environment on corticosterone concentrations in a social avian species, the zebra finch (*taeniopygia guttata*). *Hormones and Behavior*, 60(3), 233-238.
<https://doi.org/https://doi.org/10.1016/j.yhbeh.2011.05.011>
- Bartmess-Levasseur, J., Branch, C. L., Browning, S. A., Owens, J. L., & Freeberg, T. M. (2010). Predator stimuli and calling behavior of carolina chickadees (*poecile carolinensis*), tufted titmice (*baeolophus bicolor*), and white-breasted nuthatches (*sitta carolinensis*). *Behavioral Ecology and Sociobiology*, 64(7), 1187-1198. <https://doi.org/10.1007/s00265-010-0935-y>
- Beekman, M., & Jordan, L. A. (2017). Does the field of animal personality provide any new insights for behavioral ecology? *Behavioral Ecology*, 28(3), 617-623.
<https://doi.org/10.1093/beheco/arx022>

- Belgrad, B. A., & Griffen, B. D. (2017). Habitat quality mediates personality through differences in social context. *Oecologia*, 184(2), 431-440. <https://doi.org/10.1007/s00442-017-3886-4>
- Bell, A. M., Hankison, S. J., & Laskowski, K. L. (2009). The repeatability of behaviour: A meta-analysis. *Animal Behaviour*, 77(4), 771-783. <https://doi.org/10.1016/j.anbehav.2008.12.022>
- Berger, P., & Pullberg, S. (1965). Reification and the sociological critique of consciousness. *History and Theory*, 4(2), 196-211. <https://doi.org/10.2307/2504151>
- Bergmuller, R. (2010). *Animal personality and behavioural syndromes*. https://doi.org/10.1007/978-3-642-02624-9_19
- Bibi, N., Wei, Y. S., Xu, H. W., Liang, J. N., Hussain, I., Ahmad, M. S., & Wang, H. T. (2019). Personality is associated with dominance in a social feeding context in the great tit. *Behaviour*, 156(13-14), 1419-1434. <https://doi.org/10.1163/1568539x-00003572>
- Bibi, N., Yu, J. P., Gong, Y., Mo, T. S., Zubair, M., & Wang, H. T. (2022). Influence of social context on cinereous tit (*parus cinereus*) neophobic behavior. *Wilson Journal of Ornithology*, 134(2), 227-235. <https://doi.org/10.1676/21-00007>
- Blight, O., Albet Díaz-Mariblanca, G., Cerdá, X., & Boulay, R. (2016). A proactive-reactive syndrome affects group success in an ant species. *Behavioral Ecology*, 27(1), 118-125. <https://doi.org/10.1093/beheco/arv127>
- Blumstein, D. T., Holland, B. D., & Daniel, J. C. (2006). Predator discrimination and 'personality' in captive vancouver island marmots (*marmota vancouverensis*). *Animal Conservation*, 9(3), 274-282. <https://doi.org/10.1111/j.1469-1795.2006.00033.x>
- Brommer, J. E., & Class, B. (2017). Personality from the perspective of behavioral ecology. In *Personality in nonhuman animals* (pp. 73-107). https://doi.org/10.1007/978-3-319-59300-5_5
- Cantwell, L. R., Johnson, W. T., Kaschel, R. E., Love, D. J., & Freeberg, T. M. (2016). Predator-risk-sensitive foraging behavior of carolina chickadees (*poecile carolinensis*) and tufted titmice (*baeolophus bicolor*) in response to the head orientation of snake predator models. *Behavioral Ecology and Sociobiology*, 70(4), 533-539. <https://doi.org/10.1007/s00265-016-2070-x>
- Carazo, P., Noble, D. W. A., Chandrasoma, D., & Whiting, M. J. (2014). Sex and boldness explain individual differences in spatial learning in a lizard. *Proceedings of the Royal Society B-Biological Sciences*, 281(1782), 9, Article 20133275. <https://doi.org/10.1098/rspb.2013.3275>
- Carter, A. J., & Feeney, W. E. (2012). Taking a comparative approach: Analysing personality as a multivariate behavioural response across species. *PLOS ONE*, 7(7), e42440. <https://doi.org/10.1371/journal.pone.0042440>
- Carter, A. J., Feeney, W. E., Marshall, H. H., Cowlshaw, G., & Heinsohn, R. (2013). Animal personality: What are behavioural ecologists measuring? *Biological Reviews*, 88(2), 465-475. <https://doi.org/10.1111/brv.12007>
- Cauchoi, 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., Claidiere, N., Dalesman, S., Devaud, J. M., Didic, M., Doligez, B., Fagot, J., Fichtel, C., Henke-Von Der Malsburg, J., . . . Morand-Ferron, J. (2018). The repeatability of cognitive performance: A meta-analysis. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 373(1756), Article 20170281. <https://doi.org/10.1098/rstb.2017.0281>

- Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). Nbclust: An r package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, 61(6), 1-36. <https://www.jstatsoft.org/v61/i06/>
- Chivers, D. P., & Smith, R. J. F. (1998). Chemical alarm signalling in aquatic predator-prey systems: A review and prospectus. *Écoscience*, 5(3), 338-352. <https://doi.org/10.1080/11956860.1998.11682471>
- Christensen, R. H. B. (2023). *Ordinal - regression models for ordinal data*. In The Comprehensive R Archive Network (CRAN). <https://CRAN.R-project.org/package=ordinal>
- Cimprich, D. A., & Grubb Jr, T. C. (1994). Consequences for carolina chickadees of foraging with tufted titmice in winter. *Ecology*, 75(6), 1615-1625. <https://doi.org/https://doi.org/10.2307/1939622>
- Cleveland, R. B., Cleveland, W. S., & Terpenning, I. (1990). Stl: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 6(1), 3. <https://utk.idm.oclc.org/login?url=https://www.proquest.com/scholarly-journals/stl-seasonal-trend-decomposition-procedure-based/docview/1266805989/se-2?accountid=14766>
<https://libkey.io/libraries/203/openurl?genre=article&aulast=Cleveland&issn=0282423X&title=Journal+of+Official+Statistics&atitle=STL%3A+A+Seasonal-Trend+Decomposition+Procedure+Based+on+Loess&volume=6&issue=1&spage=3&date=1990&doi=&sid=ProQuest>
- Coppinger, B. A., De Launay, A. S., & Freeberg, T. M. (2018). Carolina chickadee (poecile carolinensis) calling behavior in response to threats and in flight: Flockmate familiarity matters. *Journal of Comparative Psychology*, 132(1), 16-23. <https://doi.org/10.1037/com0000090>
- Cussen, V. A. (2017). Psittacine cognition: Individual differences and sources of variation. *Behavioural Processes*, 134, 103-109. <https://doi.org/10.1016/j.beproc.2016.11.008>
- Dall, S. R. X., Houston, A. I., & McNamara, J. M. (2004). The behavioural ecology of personality: Consistent individual differences from an adaptive perspective. *Ecology Letters*, 7(8), 734-739. <https://doi.org/10.1111/j.1461-0248.2004.00618.x>
- De Azevedo, C. S., & Young, R. J. (2021). Animal personality and conservation: Basics for inspiring new research. *Animals*, 11(4), Article 1019. <https://doi.org/10.3390/ani11041019>
- Dingemanse, N. J., Kazem, A. J. N., Reale, D., & Wright, J. (2010). Behavioural reaction norms: Animal personality meets individual plasticity. *Trends in Ecology & Evolution*, 25(2), 81-89. <https://doi.org/10.1016/j.tree.2009.07.013>
- Dingemanse, N. J., & Reale, D. (2005). Natural selection and animal personality. *Behaviour*, 142, 1159-1184. <https://doi.org/10.1163/156853905774539445>
- Dinno, A. (2018). *Paran: Horn's test of principal components/factors*. In Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/package=paran>
- Dochtermann, N. A., & Dingemanse, N. J. (2013). Behavioral syndromes as evolutionary constraints. *Behavioral Ecology*, 24(4), 806-811. <https://doi.org/10.1093/beheco/art002>
- Dugatkin, L. A., & Alfieri, M. S. (2003). Boldness, behavioral inhibition and learning. *Ethology Ecology & Evolution*, 15(1), 43-49.
- Farine, D. R., Aplin, L. M., Sheldon, B. C., & Hoppitt, W. (2015). Interspecific social networks promote information transmission in wild songbirds. *Proceedings of the Royal Society B-Biological Sciences*, 282(1803), 9, Article 20142804. <https://doi.org/10.1098/rspb.2014.2804>

- Feltenstein, M., Ford, N., Freeman, K., & Sufka, K. (2002). Dissociation of stress behaviors in the chick social-separation-stress procedure. *Physiology & Behavior*, 75, 675-679, Article PII S0031-9384(02)00660-1.
- Fisher, J., & Hinde, R. A. (1949). The opening of milk bottles by birds. *British Birds*, 42((2)), 347-357.
- Flower, T. P., Gribble, M., & Ridley, A. R. (2014). Deception by flexible alarm mimicry in an african bird. *Science*, 344(6183), 513-516. <https://doi.org/10.1126/science.1249723>
- Fokkema, R. W., Korsten, P., Schmoll, T., & Wilson, A. J. (2021). Social competition as a driver of phenotype-environment correlations: Implications for ecology and evolution. *Biological Reviews*, 96(6), 2561-2572. <https://doi.org/10.1111/brv.12768>
- Freeberg, T. M., Adams, C. B., Price, C. A., & Papeş, M. (2024). Mixed-species flock sizes and compositions influence flock members' success in three field experiments with novel feeders. *Plos One*, 19(5), e0301270. <https://doi.org/10.1371/journal.pone.0301270>
- Freeberg, T. M., Eppert, S. K., Sieving, K. E., & Lucas, J. R. (2017). Diversity in mixed species groups improves success in a novel feeder test in a wild songbird community. *Scientific Reports*, 7, 9, Article 43014. <https://doi.org/10.1038/srep43014>
- Freeberg, T. M., Krama, T., Vrublevska, J., Krams, I., & Kullberg, C. (2014). Tufted titmouse (*baeolophus bicolor*) calling and risk-sensitive foraging in the face of threat. *Animal Cognition*, 17(6), 1341-1352. <https://doi.org/10.1007/s10071-014-0770-z>
- 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. <https://doi.org/https://doi.org/10.1111/2041-210X.12584>
- Furnham, A., & Jaspars, J. (1983). The evidence for interactionism in psychology - a critical analysis of the situation response inventories. *Personality and Individual Differences*, 4(6), 627-644. [https://doi.org/10.1016/0191-8869\(83\)90117-4](https://doi.org/10.1016/0191-8869(83)90117-4)
- Gallego-Abenza, M., Loretto, M. C., & Bugnyar, T. (2020). Decision time modulates social foraging success in wild common ravens, *corvus corax*. *Ethology*, 126(4), 413-422. <https://doi.org/10.1111/eth.12986>
- Gil, M. A., Emberts, Z., Jones, H., & St Mary, C. M. (2017). Social information on fear and food drives animal grouping and fitness. *Am Nat*, 189(3), 227-241. <https://doi.org/10.1086/690055>
- Gonzalez-Bernal, E., Brown, G. P., & Shine, R. (2014). Invasive cane toads: Social facilitation depends upon an individual's personality. *Plos One*, 9(7), 8, Article e102880. <https://doi.org/10.1371/journal.pone.0102880>
- Goodale, E., & Kotagama, S. W. (2005). Testing the roles of species in mixed-species bird flocks of a sri lankan rain forest. *Journal of Tropical Ecology*, 21, 669-676. <https://doi.org/10.1017/s0266467405002609>
- Goodale, E., & Magrath, R. D. (2024). Species diversity and interspecific information flow. *Biological Reviews*. <https://doi.org/10.1111/brv.13055>
- Goodale, E., Sridhar, H., Sieving, K. E., Bangal, P., Colorado, G. J., Farine, D. R., Heymann, E. W., Jones, H. H., Krams, I., Martinez, A. E., Montano-Centellas, F., Munoz, J., Srinivasan, U., Theo, A., & Shanker, K. (2020). Mixed company: A framework for understanding the composition and organization of mixed-species animal groups. *Biological Reviews*, 95(4), 889-910. <https://doi.org/10.1111/brv.12591>

- Gosling, S. D. (2001). From mice to men: What can we learn about personality from animal research? *Psychological Bulletin*, 127(1), 45-86. <https://doi.org/10.1037/0033-2909.127.1.45>
- Griffin, A. S., Guillette, L. M., & Healy, S. D. (2015). Cognition and personality: An analysis of an emerging field. *Trends in Ecology & Evolution*, 30(4), 207-214. <https://doi.org/10.1016/j.tree.2015.01.012>
- Guayasamin, O. L., Couzin, L. D., & Miller, N. Y. (2017). Behavioural plasticity across social contexts is regulated by the directionality of inter-individual differences. *Behavioural Processes*, 141, 196-204. <https://doi.org/10.1016/j.beproc.2016.10.004>
- Guido, J. M., Biondi, L. M., Vasallo, A. I., & Muzio, R. N. (2017). Neophobia is negatively related to reversal learning ability in females of a generalist bird of prey, the chimango caracara, milvago chimango. *Animal Cognition*, 20(4), 591-602. <https://doi.org/10.1007/s10071-017-1083-9>
- Guillette, L. M., Hahn, A. H., Hoeschele, M., Przyslupski, A. M., & Sturdy, C. B. (2015). Individual differences in learning speed, performance accuracy and exploratory behaviour in black-capped chickadees. *Animal Cognition*, 18(1), 165-178. <https://doi.org/10.1007/s10071-014-0787-3>
- Guillette, L. M., Naguib, M., & Griffin, A. S. (2017). Individual differences in cognition and personality. *Behavioural Processes*, 134, 1-3. <https://doi.org/10.1016/j.beproc.2016.12.001>
- Herborn, K. A., Heidinger, B. J., Alexander, L., & Arnold, K. E. (2014). Personality predicts behavioral flexibility in a fluctuating, natural environment. *Behavioral Ecology*, 25(6), 1374-1379. <https://doi.org/10.1093/beheco/aru131>
- Hetrick, S. A., & Sieving, K. E. (2011). Antipredator calls of tufted titmice and interspecific transfer of encoded threat information. *Behavioral Ecology*, 23(1), 83-92. <https://doi.org/10.1093/beheco/arr160>
- Hosoya, S., Suetake, H., Suzuki, Y., & Kikuchi, K. (2015). Genetic basis underlying behavioral correlation between fugu <i>takifugu rubripes</i> and a closely related species, <i>takifugu niphobles</i>. *Behavior Genetics*, 45, 560-572. <https://doi.org/10.1007/s10519-015-9728-4>
- Houts, A. C., Cook, T. D., & Shadish, W. R. (1986). The person-situation debate - a critical multiplist perspective. *Journal of Personality*, 54(1), 52-105. <https://doi.org/10.1111/j.1467-6494.1986.tb00390.x>
- Huang, P., Kimball, R. T., & St Mary, C. M. (2018). Does the use of a multi-trait, multi-test approach to measure animal personality yield different behavioural syndrome results? *Behaviour*, 155(2-3), 115-150. <https://doi.org/10.1163/1568539x-00003480>
- Jenkins, J. R., Pope, I. G., Dykstra, M. A., Jenkins, J. J., Dykstra, C. R., & Williams, K. A. (2023). Do standard behavioral assays predict foraging behavior of individual black-capped chickadees (<i>poecile atricapillus</i>) in response to a predator model or calls? *Journal of Comparative Psychology*, 18. <https://doi.org/10.1037/com0000364>
- Jones, H., & Sieving, K. (2019). Foraging ecology drives social information reliance in an avian eavesdropping community. *Ecology and Evolution*, 9, 11584-11597. <https://doi.org/10.1002/ece3.5561>
- Jones, R. B., & Merry, B. J. (1988). Individual or paired exposure of domestic chicks to an open-field - some behavioral and adrenocortical consequences. *Behavioural Processes*, 16(1-2), 75-86.

- Katsis, A. C., Colombelli-Négrel, D., Akçay, Ç., Common, L. K., García-Loor, J., & Kleindorfer, S. (2024). Exploration behavior differs between darwin's finch species and predicts territory defense and hatching success. *Behavioral Ecology and Sociobiology*, 78(2), 26. <https://doi.org/10.1007/s00265-024-03438-7>
- Kerman, K., Miller, L., & Sewall, K. (2018). The effect of social context on measures of boldness: Zebra finches (*Taeniopygia guttata*) are bolder when housed individually. *Behavioural Processes*, 157, 18-23. <https://doi.org/10.1016/j.beproc.2018.08.007>
- Krams, I., Krama, T., Freeberg, T. M., Kullberg, C., & Lucas, J. R. (2012). Linking social complexity and vocal complexity: A parid perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1597), 1879-1891.
- Krasheninnikova, A., Chow, P. K. Y., & Von Bayern, A. M. P. (2020). Comparative cognition: Practical shortcomings and some potential ways forward. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale*, 74(3), 160-169. <https://doi.org/10.1037/cep0000204>
- Kurvers, R., Nolet, B. A., Prins, H. H. T., Ydenberg, R. C., & Van Oers, K. (2012). Boldness affects foraging decisions in barnacle geese: An experimental approach. *Behavioral Ecology*, 23(6), 1155-1161. <https://doi.org/10.1093/beheco/ars091>
- Kurvers, R., Van Oers, K., Nolet, B. A., Jonker, R. M., Van Wieren, S. E., Prins, H. H. T., & Ydenberg, R. C. (2010). Personality predicts the use of social information. *Ecology Letters*, 13(7), 829-837. <https://doi.org/10.1111/j.1461-0248.2010.01473.x>
- Kyle, S. C., & Freeberg, T. M. (2016). Do carolina chickadees (*poecile carolinensis*) and tufted titmice (*baeolophus bicolor*) attend to the head or body orientation of a perched avian predator? *J Comp Psychol*, 130(2), 145-152. <https://doi.org/10.1037/com0000019>
- Lefebvre, L., & Giraldeau, L.-A. (1996). Chapter 6 - is social learning an adaptive specialization? In C. M. Heyes & B. G. Galef (Eds.), *Social learning in animals* (pp. 107-128). Academic Press. <https://doi.org/10.1016/B978-012273965-1/50007-8>
- Lefebvre, L., Whittle, P., Lascaris, E., & Finkelstein, A. (1997). Feeding innovations and forebrain size in birds. *Animal Behaviour*, 53(3), 549-560. <https://doi.org/10.1006/anbe.1996.0330>
- Lima, S. L. (1987). Vigilance while feeding and its relation to the risk of predation. *Journal of Theoretical Biology*, 124(3), 303-316. [https://doi.org/10.1016/S0022-5193\(87\)80118-2](https://doi.org/10.1016/S0022-5193(87)80118-2)
- Liu, N., Wang, Y., An, A. Y., Banker, C., Qian, Y.-H., & O'donnell, J. M. (2020). Single housing-induced effects on cognitive impairment and depression-like behavior in male and female mice involve neuroplasticity-related signaling. *European Journal of Neuroscience*, 52(1), 2694-2704. <https://doi.org/10.1111/ejn.14565>
- Liu, W. J., Xu, Y. C., & Rong, L. B. (2024). Predator discrimination promotes the coexistence of prey and predator. *Journal of Applied Analysis and Computation*, 14(3), 1579-1597. <https://doi.org/10.11948/20230301>
- Louison, M. J., Hage, V. M., Stein, J. A., & Suski, C. D. (2019). Quick learning, quick capture: Largemouth bass that rapidly learn an association task are more likely to be captured by recreational anglers. *Behavioral Ecology and Sociobiology*, 73(2), 13, Article 23. <https://doi.org/10.1007/s00265-019-2634-7>

- Lukoschek, V., & McCormick, M. (2002). *A review of multi-species foraging associations in fishes and their ecological significance* (Vol. 1).
- Macgregor, H. E. A., Cottage, A., & Ioannou, C. C. (2021). Suppression of personality variation in boldness during foraging in three-spined sticklebacks. *Behavioral Ecology and Sociobiology*, 75(4), 10, Article 71. <https://doi.org/10.1007/s00265-021-03007-2>
- Machery, E. (2014). Social ontology and the objection from reification. In M. Gallotti & J. Michael (Eds.), *Perspectives on social ontology and social cognition* (pp. 87-100). Springer Netherlands. https://doi.org/10.1007/978-94-017-9147-2_7
- Mainwaring, M. C., Beal, J. L., & Hartley, I. R. (2011). Zebra finches are bolder in an asocial, rather than social, context. *Behavioural Processes*, 87(2), 171-175. <https://doi.org/10.1016/j.beproc.2011.03.005>
- Mamuneas, D., Spence, A. J., Manica, A., & King, A. J. (2015). Bolder stickleback fish make faster decisions, but they are not less accurate. *Behavioral Ecology*, 26(1), 91-96. <https://doi.org/10.1093/beheco/aru160>
- Martinez, A. E., Pollock, H. S., Kelley, J. P., & Tarwater, C. E. (2018). Social information cascades influence the formation of mixed-species foraging aggregations of ant-following birds in the neotropics. *Animal Behaviour*, 135, 25-35. <https://doi.org/10.1016/j.anbehav.2017.10.024>
- Mathot, K. J., Zollinger, S. A., & Freeberg, T. M. (2021). Personality and behavioral syndromes. In D. S. Proppe (Ed.), *Songbird behavior: Implications for conservation and management in the anthropocene* (pp. 239-262). CRC Press.
- Mazza, V., Eccard, J. A., Zaccaroni, M., Jacob, J., & Dammhahn, M. (2018). The fast and the flexible: Cognitive style drives individual variation in cognition in a small mammal. *Animal Behaviour*, 137, 119-132. <https://doi.org/10.1016/j.anbehav.2018.01.011>
- Mccune, B., Grace, J. B., & Urban, D. L. (2002). *Analysis of ecological communities*. MjM Software Design.
- Mccune, K., Jablonski, P., Lee, S. I., & Ha, R. (2018). Evidence for personality conformity, not social niche specialization in social jays. *Behavioral Ecology*, 29(4), 910-917. <https://doi.org/10.1093/beheco/ary055>
- Mcfadden, D. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in econometrics*, 105-142.
- Mertes, K., Ressijac, C. A., Moraes, R. N., Hughey, L. F., Alegre, L. H. P., Horning, M., Buk, T., Harwood, A., Layman, L., Mathews, C., Vance, M., Reed, D., Stabach, J. A., & Goldenberg, S. Z. (2022). Assessing neophobia and exploration while accounting for social context: An example application in scimitar-horned oryx. *Mammalian Biology*, 102(4), 1357-1371. <https://doi.org/10.1007/s42991-022-00271-1>
- Michelangeli, M., Chapple, D. G., Goulet, C. T., Bertram, M. G., & Wong, B. B. M. (2018). Behavioral syndromes vary among geographically distinct populations in a reptile. *Behavioral Ecology*, 30(2), 393-401. <https://doi.org/10.1093/beheco/ary178>
- Michelangeli, M., Cote, J., Chapple, D., Sih, A., Brodin, T., Fogarty, S., Bertram, M., Eades, J., & Wong, B. (2020). Sex-dependent personality in two invasive species of mosquitofish. *Biological Invasions*, 22, 1353-1364. <https://doi.org/10.1007/s10530-019-02187-3>
- Montiglio, P.-O., Ferrari, C., & Réale, D. (2013). Social niche specialization under constraints: Personality, social interactions and environmental heterogeneity. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 368(1618), 20120343-20120343. <https://doi.org/10.1098/rstb.2012.0343>

- Morse, D. H. (1970). Ecological aspects of some mixed-species foraging flocks of birds. *Ecological Monographs*, 40(1), 119-+. <https://doi.org/10.2307/1942443>
- Morse, D. H. (1977). Feeding-behavior and predator avoidance in heterospecific groups. *Bioscience*, 27(5), 332-339. <https://doi.org/10.2307/1297632>
- Muta, O., Odaka, M., Fujii, Y., Fushimi, T., Sato, H., & Osakabe, N. (2023). Difference in endocrine and behavior between short-term single- and paired-housing mice in metabolic cage. *Neuroscience Letters*, 806, 137246. <https://doi.org/10.1016/j.neulet.2023.137246>
- Muth, F., & Healy, S. D. (2014). Zebra finches select nest material appropriate for a building task. *Animal Behaviour*, 90, 237-244. <https://doi.org/10.1016/j.anbehav.2014.02.008>
- Naslund, J. (2021). Unreplicable state-dependent effects on start-box emergence latency in wild-origin sticklebacks. *Ethology*, 127(8), 621-631. <https://doi.org/10.1111/eth.13169>
- Nelson, X. J., Wilson, D. R., & Evans, C. S. (2008). Behavioral syndromes in stable social groups: An artifact of external constraints? *Ethology*, 114(12), 1154-1165. <https://doi.org/10.1111/j.1439-0310.2008.01568.x>
- Nolen, M. T., & Lucas, J. R. (2009). Asymmetries in mobbing behaviour and correlated intensity during predator mobbing by nuthatches, chickadees and titmice. *Animal Behaviour*, 77(5), 1137-1146. <https://doi.org/10.1016/j.anbehav.2009.01.023>
- Olafsdottir, G. A., & Magellan, K. (2016). Interactions between boldness, foraging performance and behavioural plasticity across social contexts. *Behavioral Ecology and Sociobiology*, 70(11), 1879-1889. <https://doi.org/10.1007/s00265-016-2193-0>
- Perals, D., Griffin, A. S., Bartomeus, I., & Sol, D. (2017). Revisiting the open-field test: What does it really tell us about animal personality? *Animal Behaviour*, 123, 69-79. <https://doi.org/10.1016/j.anbehav.2016.10.006>
- Perez, E. C., Elie, J. E., Soulage, C. O., Soula, H. A., Mathevon, N., & Vignal, C. (2012). The acoustic expression of stress in a songbird: Does corticosterone drive isolation-induced modifications of zebra finch calls? *Hormones and Behavior*, 61(4), 573-581. <https://doi.org/10.1016/j.yhbeh.2012.02.004>
- Petelle, M. B., Martin, J. G. A., & Blumstein, D. T. (2019). Mixed support for state maintaining risky personality traits in yellow-bellied marmots. *Animal Behaviour*, 150, 177-188. <https://doi.org/10.1016/j.anbehav.2019.02.008>
- Pianka, E. (1970). R-selection and k-selection. *The American Naturalist*, 104, 592. <https://doi.org/10.1086/282697>
- Potrebić, M. S., Pavković, Ž. Z., Srbovan, M. M., Dmura, G. M., & Pešić, V. T. (2022). Changes in the behavior and body weight of mature, adult male wistar han rats after reduced social grouping and social isolation. *Journal of the American Association for Laboratory Animal Science*, 61(6), 615-623. <https://doi.org/10.30802/AALAS-JAALAS-22-000032>
- R Core Development Team. (2016). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rowe, C., & Healy, S. D. (2014). Measuring variation in cognition. *Behavioral Ecology*, 25(6), 1287-1292. <https://doi.org/10.1093/beheco/aru090>
- Sakai, O. (2020). Do different food amounts gradually promote personality variation throughout the life stage in a clonal gecko species? *Animal Behaviour*, 162, 47-56. <https://doi.org/10.1016/j.anbehav.2020.01.014>

- Schleidt, W., Shalter, M., & Carawan, T. (1983). The effect of spatial context on habituation to a predator model. *Zeitschrift Fur Tierpsychologie-Journal of Comparative Ethology*, 61, 67-70.
- Shams, S., Chatterjee, D., & Gerlai, R. (2015). Chronic social isolation affects thigmotaxis and whole-brain serotonin levels in adult zebrafish. *Behavioural Brain Research*, 292, 283-287. <https://doi.org/10.1016/j.bbr.2015.05.061>
- Sieving, K. E., Hetrick, S. A., & Avery, M. L. (2010). The versatility of graded acoustic measures in classification of predation threats by the tufted titmouse *baeolophus bicolor*: Exploring a mixed framework for threat communication. *Oikos*, 119(2), 264-276. <https://doi.org/10.1111/j.1600-0706.2009.17682.x>
- Sih, A. (2017). Insights for behavioral ecology from behavioral syndromes: A comment on beekman and jordan. *Behavioral Ecology*, 28(3), 627-628. <https://doi.org/10.1093/beheco/axx032>
- Sih, A., Bell, A., & Johnson, J. C. (2004). Behavioral syndromes: An ecological and evolutionary overview. *Trends in Ecology & Evolution*, 19(7), 372-378. <https://doi.org/10.1016/j.tree.2004.04.009>
- Sih, A., Bell, A. M., Johnson, J. C., & Ziemba, R. E. (2004). Behavioral syndromes: An integrative overview. *Quarterly Review of Biology*, 79(3), 241-277. <https://doi.org/10.1086/422893>
- Sih, A., & Del Giudice, M. (2012). Linking behavioural syndromes and cognition: A behavioural ecology perspective. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 367(1603), 2762-2772. <https://doi.org/10.1098/rstb.2012.0216>
- Sih, A., Mathot, K. J., Moiron, M., Montiglio, P. O., Wolf, M., & Dingemanse, N. J. (2015). Animal personality and state-behaviour feedbacks: A review and guide for empiricists. *Trends in Ecology & Evolution*, 30(1), 50-60. <https://doi.org/10.1016/j.tree.2014.11.004>
- Stamps, J. A. (2007). Growth-mortality tradeoffs and 'personality traits' in animals. *Ecology Letters*, 10(5), 355-363. <https://doi.org/10.1111/j.1461-0248.2007.01034.x>
- Stamps, J. A., & Biro, P. A. (2016). Personality and individual differences in plasticity. *Current Opinion in Behavioral Sciences*, 12, 18-23. <https://doi.org/10.1016/j.cobeha.2016.08.008>
- Stöwe, M., Bugnyar, T., Loretto, M. C., Schloegl, C., Range, F., & Kotrschal, K. (2006). Novel object exploration in ravens (*corvus corax*): Effects of social relationships. *Behavioural Processes*, 73(1), 68-75. <https://doi.org/10.1016/j.beproc.2006.03.015>
- Suzuki, T. N. (2012). Referential mobbing calls elicit different predator-searching behaviours in japanese great tits. *Animal Behaviour*, 84(1), 53-57. <https://doi.org/10.1016/j.anbehav.2012.03.030>
- Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3d medical image segmentation: Analysis, selection, and tool. *BMC Med Imaging*, 15, 29. <https://doi.org/10.1186/s12880-015-0068-x>
- Takola, E., Krause, E. T., Muller, C., & Schielzeth, H. (2021). Novelty at second glance: A critical appraisal of the novel object paradigm based on meta-analysis. *Animal Behaviour*, 180, 123-142. <https://doi.org/10.1016/j.anbehav.2021.07.018>
- Tebbich, S., Griffin, A. S., Peschl, M. F., & Sterelny, K. (2016). From mechanisms to function: An integrated framework of animal innovation. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 371(1690), 11, Article 20150195. <https://doi.org/10.1098/rstb.2015.0195>

- Tello-Ramos, M. C., Branch, C. L., Kozlovsky, D. Y., Pitera, A. M., & Pravosudov, V. V. (2019). Spatial memory and cognitive flexibility trade-offs: To be or not to be flexible, that is the question. *Animal Behaviour*, 147, 129-136. <https://doi.org/10.1016/j.anbehav.2018.02.019>
- Tuliozi, B., Camerlenghi, E., & Griggio, M. (2021). Dyadic leader-follower dynamics change across situations in captive house sparrows. *Behavioral Ecology*, 32(3), 508-517. <https://doi.org/10.1093/beheco/araa148>
- Tuliozi, B., Fracasso, G., Hoi, H., & Griggio, M. (2018). House sparrows' (*passer domesticus*) behaviour in a novel environment is modulated by social context and familiarity in a sex-specific manner. *Frontiers in Zoology*, 15, 14, Article 16. <https://doi.org/10.1186/s12983-018-0267-8>
- Van Oers, K., Klunder, M., & Drent, P. J. (2005). Context dependence of personalities: Risk-taking behavior in a social and a nonsocial situation. *Behavioral Ecology*, 16(4), 716-723. <https://doi.org/10.1093/beheco/ari045>
- Von Merten, S., Zwolak, R., & Rychlik, L. (2017). Social personality: A more social shrew species exhibits stronger differences in personality types. *Animal Behaviour*, 127, 125-134. <https://doi.org/10.1016/j.anbehav.2017.02.021>
- Vrublevska, J., Krama, T., Rantala, M. J., Mierauskas, P., Freeberg, T. M., & Krams, I. A. (2015). Personality and density affect nest defence and nest survival in the great tit. *Acta Ethologica*, 18(2), 111-120. <https://doi.org/10.1007/s10211-014-0191-7>
- Webster, M. M., & Ward, A. J. W. (2011). Personality and social context. *Biological Reviews*, 86(4), 759-773. <https://doi.org/10.1111/j.1469-185X.2010.00169.x>
- Webster, M. M., Ward, A. J. W., & Hart, P. J. B. (2009). Individual boldness affects interspecific interactions in sticklebacks. *Behavioral Ecology and Sociobiology*, 63(4), 511-520. <https://doi.org/10.1007/s00265-008-0685-2>
- Zandberg, L., Quinn, J. L., Naguib, M., & Van Oers, K. (2017). Personality-dependent differences in problem-solving performance in a social context reflect foraging strategies. *Behavioural Processes*, 134, 95-102. <https://doi.org/10.1016/j.beproc.2016.09.007>
- Zentall, T. R., & Galef, B. G., Jr. (1988). Social learning: Psychological and biological perspectives. *Social learning: psychological and biological perspectives.*, i-xii, 1-357.
- Zimmerman, G. M., Goetz, H., & Mielke, P. W. (1985). Use of an improved statistical method for group comparisons to study effects of prairie fire. *Ecology*, 66(2), 606-611. <https://doi.org/10.2307/1940409>

VITA

Scott Benson was born in Nashville, TN in 1989 to mother Beverly Parker and father Rodney Dunham. Scott attended Hillsboro High School before attending Motlow Community College, eventually transferring to Middle Tennessee State University where he graduated with a bachelor's degree in Psychology as a first-generation college student. After several years selling fabric, building car dashboards, and brokering social media ads for small businesses, he returned to academia to accept a graduate position at the University of Tennessee under Dr. Todd Freeberg. Scott is interested in social behavior and behavioral diversity in mixed-species groups of animals, and plans to extend the work to examining social learning across species barriers and the conservation outcomes associated with behavioral syndromes.