# Cognitive Aging and Flexibility in Canids

by

Joshua Lazar Van Bourg

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2022 by the
Graduate Supervisory Committee:

Clive. D. L. Wynne, Chair
Ian C. Gilby
C. Athena Aktipis
Julie K. Young

ARIZONA STATE UNIVERSITY

May 2022

# ABSTRACT

This dissertation describes a series of four studies on cognitive aging, working memory, and cognitive flexibility in dogs (*Canis lupus familiaris*) and their wild relatives. In Chapters 2 and 3, I designed assessments for age-related cognitive deficits in pet dogs which can be deployed rapidly using inexpensive and accessible materials. These novel tests can be easily implemented by owners, veterinarians, and clinicians and therefore, may improve care for elderly dogs by aiding in the diagnosis of dementia. In addition, these widely deployable tests may facilitate the use of dementia in pet dogs as a naturally occurring model of Alzheimer's Disease in humans.

In Chapters 4 and 5, I modified one of these tests to demonstrate for the first time that coyotes (*Canis latrans)* and wolves (*Canis lupus lupus*) develop age-related deficits in cognitive flexibility. This was an important first step towards differentiating between the genetic and environmental components of dementia in dogs and in turn, humans. Unexpectedly, I also detected cognitive deficits in young, adult dogs and wolves but not coyotes. These finding add to a recent shift in understanding cognitive development in dogs which may improve cognitive aging tests as well as training, care, and use of working and pet dogs. These findings also suggest that the ecology of coyotes may select for flexibility earlier in development.

In Chapter 5, I piloted the use of the same cognitive flexibility test for red and gray foxes so that future studies may test for lifespan changes in the cognition of small-bodied captive canids. More broadly, this paradigm may accommodate physical and behavioral differences between diverse pet and captive animals. In Chapters 4 and 5, I examined which ecological traits drive the evolution of behavioral flexibility and in turn, species resilience. I found that wolves displayed less flexibility than dogs and coyotes suggesting that species which do not rely heavily on unstable resources may be ill-equipped to cope with human habitat modification. Ultimately, this comparative work may help conservation practitioners to identify and protect species that cannot cope with rapid and unnatural environmental change.

# DEDICATION

To my soon to be wife Annie,

for supporting me, encouraging me, and grounding me.

I could not have finished the marathon without you.

To my family,

for nurturing my interests in animals and nature,

and always encouraging me to do what I love.

To my dog Koda,

for being a willing test dummy, eager distraction, and constant companion.

You’ve taught me more than all the books combined.

# ACKNOWLEDGEMENTS

I would like to thank the dogs and owners who participated in these studies for their enthusiasm, patience, and interest in canine cognition research. I would also like to thank the USDA/ APHIS/ NWRC Predator Research Center, Wolf Park, the Wolf Conservation Center, and the Arizona Animal Welfare League for allowing me to work with their animals. In particular, I would like to thank the animal care and research staff at these facilities for accommodating and assisting with my studies.

I cannot overstate my gratitude to my advisor and mentor, Dr. Clive Wynne for his expertise, insight, and guidance as well as his confidence, wisdom, humor, and friendship. I would also like to thank Dr. Julie Young for guiding me through my first study on wild canids and for broadening my understanding of applied conservation and carnivore ecology. I am also grateful to Dr. Athena Aktipis and Dr. Ian Gilby for nurturing my interests in and understanding of cooperation and sociobiology.

I would like to thank my fellow students in the Canine Science Collaboratory and the Behavioral Neuroscience program for their feedback and friendship. In particular, I would like to thank my collaborators and coauthors Rachel Gilchrist and Lisa Gunter for their many contributions to our studies and for teaching me the ropes as a new student.

I would also like to thank the many research assistants and field technicians who were vital to the completion of this dissertation and other studies I completed during my PhD. In particular, I would like to thank Jordan Patterson, Victoria Quintana, Raneem Alkhalifah, Vlada Markov, Jaqueline Morton, and Eric Johansson.

Finally, I would like to thank the Arizona Alzheimer's Disease Core Center for supporting my research on cognitive aging in dogs, the Arizona chapter of Sigma Xi for supporting my research on coyotes, and the many donors to the Canine Science Collaboratory for supporting all of these projects.

TABLE OF CONTENTS

CHAPTER Page

APPENDIX

## LIST OF TABLES

Table Page

## LIST OF FIGURES

Figure Page

# CHAPTER 1

# GENERAL INTRODUCTION

This dissertation includes a series of four studies on cognitive aging, working memory, and cognitive flexibility in dogs and wild canids. The first two studies focused on designing assessments for cognitive aging in pet dogs that can be easily deployed by owners, veterinarians, and clinicians. The first study focused on short-term memory and the second study focused on cognitive flexibility.

In the third and fourth studies of this dissertation, I modified the cognitive flexibility assessment for pet dogs to test whether elderly coyotes and wolves also develop cognitive deficits. Because these canines are closely related to dogs, these studies may help to understand the genetic basis of dementia in dogs.

Another primary focus of this dissertation was to examine which ecological traits drive the evolution of behavioral flexibility. Because behavioral flexibility allows species to cope with anthropogenic habitat modification, these studies also aimed to better understand how species ecology and cognition can be used to predict species resilience. The first section of the chapter overviews our current knowledge about lifespan cognitive changes in dogs and how these changes mirror those that occur in humans. This section then explains how this understanding can be improved by designing more efficient cognitive tests for pet dogs and by studying lifespan cognitive changes in wild canids. The second section of this chapter overviews evolutionary theory on the ecological drivers of behavioral flexibility and its relationship to species resilience. This section then describes relevant ecological differences between the canid species studied in this dissertation and illustrates how these differences can be used to test hypotheses about the relationships between ecology, behavioral flexibility, and species resilience.

## Lifespan Cognitive Changes

### Cognitive Aging in Dogs and Humans

Millions of dogs (*Canis lupus familiaris*) develop age-related cognitive deficits (ARCDs), a form of dementia that mirrors Alzheimer's disease (AD) in humans (Head, 2001; Salvin et al., 2010). Like AD in humans, these ARCDs are correlated with naturally occurring deposits of neuritic plaques composed of beta-amyloid proteins, particularly in the prefrontal cortex (Head, 2001; Vite and Head, 2014). Although a variety of cognitive functions may be affected, these plaques and the neurotrophy they cause are most strongly associated with deficits in working memory, reversal learning and other executive functions, both in dogs (Adams et al., 2000a, Head, 2013) and in humans (Kensinger et al., 2003; Simone and Baylis, 1997). Cognitive senescence in dogs is also characterized by a suite of behavioral changes commonly referred to in veterinary settings as cognitive dysfunction syndrome (CDS; Landsberg et al., 2012; Szabó et al., 2016). These behavioral symptoms also align with the behavioral symptoms of AD in humans (Rofina et al., 2006). For example, both ARCDs in dogs and AD in humans are associated with anxiety, agitation, and sleep disruption (Landsberg et al., 2011, Reisberg et al., 1987). Thus, in terms of cognition, physiology, and behavior, ARCDs in dogs provide a strong, naturally occurring model of AD in humans.

The close associations formed between dogs and humans also provides a unique opportunity to study the environmental components of AD and ARCDs (Kaeberlein et al., 2016). Dogs and humans overlap heavily in their exposure to pollutants and pathogens, as well as their activity levels and patterns (Cotman and Head, 2008; MoMozawa, 2019). Thus, research on elderly dogs living with elderly owners may help to identify environmental predictors of AD and ARCDs.

Estimates of the incidence of CDS vary considerably (e.g., Osella et al., 2007; Azkona et al., 2009), likely due to a lack of uniform diagnostic criteria (Szabó et al., 2016). Nonetheless, given that there are approximately 80 million pet dogs in the US (American Veterinary Medical Association, 2018), by the most conservative estimates, one million dogs in the US will develop CDS each year. Like AD, these numbers are expected to grow as average life expectancies and adoption rates of older dogs increase (Inoue et al., 2018; Rowan and Kartal, 2018). Thus, CDS in dogs is not only a strong model for AD research, but also a major veterinary challenge which is personally relevant to millions of pet owners.

Research on ARCDs in dogs has primarily utilized lab-housed colony beagles. As a result, established tests for ARCDs were not designed with the challenges of widespread deployment in mind. These tests generally require weeks or months of daily testing, expensive custom-made machinery for automated testing, and highly controlled settings free from the distractions that characterize homes and veterinary clinics (e.g., Adams et al., 2000b; Milgram et al., 1999). Although focusing on colony beagles provides many benefits, beagles represent only a small subset of the dog population which is diverse both genetically and in the environments in which they live. Thus, designing new tests for cognitive aging in pet dogs which are practical and convenient for clinicians and owners may facilitate research on CDS in dogs and AD in humans.

Towards this goal of creating a widely deployable test for ARCDs in pet dogs, I conducted two studies (Chapters 2 and 3) in which I designed and deployed three rapid cognitive assessments using inexpensive and readily available materials. These tests measured cognitive functions which are particularly susceptible to decline in old age, both in dogs and in human. For each of these assessments, I tested the ability of the dog to remember and follow simple rules in order to determine which of two identical boxes contained a hidden treat. In the first study, I used transformed staircase procedures to efficiently measure the duration and accuracy of the dog's short-term memory. In the second study, I used a serial reversal learning procedure to rapidly assess behavioral inhibition, perseveration, and cognitive flexibility. For each test, I then examined relationships between age and cognitive performance.

## Cognitive Aging in Other Canids

Although ARCDs in dogs are well documented, it is unknown whether other canids develop AD-like cognitive deficits. Thus, testing for dementia in closely related wild canines as well as more distantly related canids may help to differentiate between the genetic and environmental components of dementia in dogs and in turn, of AD in humans. Moreover, if other canid species are equally prone to the same cognitive, behavioral, and neurodegenerative changes, the shared dog-human environment may not be the driving force behind the co-occurrence of ARCDs in dogs and AD in humans. Although pet dogs are an ideal population for studying environmental components of AD, there is no control population of dogs which can survive without human-created resources. In contrast, wild canids such as coyotes (*Canis latrans*) and red foxes (*Vulpes vulpes*) are exposed to a range of anthropogenic pollutants and materials in urban, suburban, and rural habitats, but these species also thrive in undisturbed ecosystems. Thus, comparing cognitive aging between populations of wild canids may help to identify associations between synthetic materials, ARCDs, and AD.

To lay the foundation for such comparisons, I modified the serial reversal learning assessment that I designed for pet dogs in order to test for age-related cognitive flexibility deficits in captive coyotes (Chapter 4) and wolves (Chapter 5), the closest living relatives to dogs (Vilà and Jennifer, 2012). To the same end, I also piloted the use of these tests for smaller wild canids (Chapter 5).

## Cognitive Development

In addition to investigations into senescence, in each experiment I also examined cognitive changes early in adulthood by testing for curvilinear relationships between age and each measurement of performance. Although the development of social cognition in canids has been well studied, little is known about other aspects of their cognitive development (Bray et al., 2021). While it is documented that some cognitive abilities continue to develop beyond adolescence in humans (Gathercole et al., 2004; Pickering, 2001), studies on dogs have only recently begun to identify improvements in executive functions and sensorimotor abilities in young adults (Wallis et al., 2014; Watowich et al., 2020). Understanding these developmental changes may help owners and handlers to better care for and train both pet and working dogs. In addition, clarifying when dogs reach peak cognitive performance should improve the sensitivity of cognitive tests for ARCDs. Moreover, studies have historically used performance during early adulthood as a baseline, which may underestimate the extent to which performance declines in old age if cognitive abilities peak in midlife rather than early adulthood.

# Cognitive Flexibility

## Evolution of Cognitive Flexibility

Behavioral plasticity, which can arise through a variety of processes including ontogenetic expression, genetic diversity, and learning, allows individuals and populations to adapt to environmental changes. Cognitive flexibility, which involves basic learning processes and more complex cognitive functions like inhibition, insight, and creativity, is one of the primary pathways to behavioral plasticity. Thus, cognitive flexibility may enable species to invade human habitats and to cope with anthropogenic habitat modification (Snell-Rood, 2013; Wright et al., 2010). In turn, identifying the evolutionary drivers and ecological predictors of cognitive flexibility may help practitioners and policy makers to identify and protect species at risk of population decline.

Studies on the evolution of behavioral flexibility have primarily focused on primates and thus the extent to which similar evolutionary processes occurred in other taxa is unclear (MacLean et al., 2014). Some of these studies suggest that cognitive complexity evolved to allow primates to effectively forage for different fruits in response to frequent changes in quality and availability (Barton, 1996; Milton, 1981). More broadly, species that rely on patchy resources, utilize diverse behaviors to acquire these resources, and innovate new feeding strategies should evolve greater cognitive flexibility than species that do not experience these selection pressures (Day et al., 1999; Reader et al., 2011; Zuberbühler and Janmaat, 2010). In addition, species with greater dietary breadth may be able to meet the metabolic demands of growing and maintaining larger brains, which allow for more complex cognition (Aiello and Wheeler, 1995; Fish and

Lockwood, 2003). Furthermore, species that navigate complex habitats and respond appropriately to dynamic environments should exhibit greater flexibility than should species which live in stable environments with stable resources (Jones, 2005; Shettleworth, 1998).

Behavioral flexibility is also thought to evolve in social species, particularly those characterized by fission–fusion group dynamics (Amici et al., 2008), complex social interactions (Easton, 2005), hierarchical group organization (Amici et al., 2009), and large group sizes (Dunbar, 1998). Balancing individual and group needs through negotiations between individual actors or groups of actors requires constant modification to, shifts in, and inhibition of behavior (Byrne and Whiten, 1988; Easton, 2005). Similarly, accounting for the actions of others during coordinated activities requires rapid and reversible behavioral adjustments. Thus, flexibility may also evolve in species which cooperate in pair-bonding, rearing young, hunting, vigilance, or defense (for discussion and review, see Marshall-Pescini et al., 2015).

## Cognitive Flexibility, Ecology, and Species Resilience in Canids

Canid species vary greatly in most of these variables. Importantly, some canids have been highly successful in adapting to human-modified ecosystems while others have experienced extensive population declines due to human activities. Thus, canids are well-suited for studying the relationships between social ecology, diet, cognitive flexibility, and species resilience.

Whereas coyotes, dogs, and red foxes thrive in virtually every type of human habitat (Hill et al., 1987; Kuijper et al., 2019; Schipper et al., 2008), stable populations of gray wolves (*Canis lupus lupus*) are generally restricted to areas with low human densities (Massolo and Meriggi, 1998). The association between gray foxes (*Urocyon cinereoargenteus*) and human habitats is poorly understood (Allen et al., 2021) but they have fared better than wolves and worse than coyotes, dogs, and red foxes.

Although most canids will scavenge and forage opportunistically, wolves (Mech, 1970) and red foxes (Jędrzejewski and Jędrzejewski, 1992) rely more heavily on hunting than do coyotes (Andelt et al., 1987), dogs (Coppinger and Coppinger, 2002), and gray foxes (Hockman and Chapman, 1983). However, dogs have evolved to be entirely dependent on scavenging and begging (Marshall-Pescini et al., 2017a) whereas coyotes and gray foxes are also adept hunters.

Coyotes have invaded almost every type of habitat in the Americas but were historically restricted to arid and semi-arid open habitats in North America (Hody and Kays, 2018). Red foxes inhabit diverse temperate and subarctic habitats but require cooler climates than do coyotes (Lloyd, 1980). Wolves can also inhabit a variety of habitats (Mech and Boitani, 2007) but are generally more restricted by forest cover, prey abundance, and human presence (Massolo and Meriggi, 1998) than are coyotes or red foxes. Gray foxes primarily inhabit temperate and subtropical woodlands in North and Central America (Allen et al., 2021). Finally, dogs can survive in a wide range of climates but only in or around human habitats (Paul et al., 2016; Sen Majumder et al., 2016).

Although social structure is variable in dogs, they can form relatively large, complex, and fluid social groups (Bonanni and Cafazzo, 2014; Cafazzo et al., 2010; Marshall-Pescini et al., 2017a). In contrast, wolf packs are usually composed of a single breeding pair, their adult and sub-adult offspring, and the most recent litter (Packard, 2003). In certain conditions coyotes can be relatively solitary (Bekoff and Wells, 1980), but they generally form family groups consisting of a breeding pair, their most recent litter, and in some cases yearlings from the previous litter (Bowen, 1981; Hennessy et al., 2012). Red foxes form monogamous breeding pairs (Iossa et al., 2008) which sometimes associate with adult subordinates (Moehlman, 1989). Dominant males may also engage in polygyny (Moehlman, 1989). Gray foxes seasonally form small family units of monogamous breeding pairs and their young but are solitary for most of the year (Nicholson et al., 1985, Tucker et al., 1993, Chamberlain and Leopold, 2002). Cooperation is a central component of gray wolf social, reproductive, and feeding ecology (MacNulty et al., 2014; Mech et al., 2021). To a lesser degree, coyotes also cooperate in rearing young, defending territory and hunting (Bowen, 1981; Rathbun et al., 1980; Lord et al., 2013; Thornton et al., 2018). In contrast, red foxes, gray foxes, and free-ranging dogs hunt and forage alone (Moehlman, 1986; 1989; Sen Majumder et al., 2014). Allomaternal care is rare in dogs (Pal, 2005) and red foxes (Moehlman, 1986), and has not to our knowledge been observed in gray foxes. Free-ranging dogs cooperate in territorial defense (Pal, 2015) and owned dogs (particularly working breeds) cooperate with humans, but this may require training (Range and Virányi, 2015).

These ecological differences lead to a number of predictions about which canids should display more behavioral flexibility. For example, if cooperation is a strong evolutionary driver of flexibility, wolves should outperform other canines in cognitive flexibility tests. And if behavioral flexibility has helped dogs and coyotes to adapt to human habitats, while a lack of flexibility has hindered the ability of wolves to cope with anthropogenic habitat modification, these differences should be detectable with cognitive flexibility assessments.

By using the same serial reversal learning assessment to test for age-related cognitive deficits in dogs, coyotes, and wolves, I was able to measure and compare species typical cognitive flexibility in these canids. Thus, these cognitive aging assessments also allowed me to test hypotheses about the ecological traits driving the evolution of cognitive flexibility and species resilience. In addition, by piloting the use of this test for red and gray foxes I was able to assess whether future studies can add these and other small-bodied canids to this comparative framework.

# CHAPTER 2

# ADAPTIVE SPATIAL WORKING MEMORY ASSESSMENTS FOR AGING PET DOGS

**Keywords**

·dog ·cognitive decline ·aging ·development ·spatial working memory ·staircase methods

**Declarations**

Not applicable

**Funding**

This research was funded by the Arizona Alzheimer's Disease Center

**Conflicts of interest/Competing interests**

Not applicable

**Ethics approval**

This research was approved by the Arizona State University Institutional Animal Care and Use Committee (Protocol Number: 16-1490R)

**Consent**

Informed consent was received from the owners of all dogs that participated in this study and all participation was voluntary

**Data and code availability**

All data and statistical code are available from the corresponding author on reasonable request

**Authors' contributions**

All authors contributed to the study conception and design. JVB and RG prepared materials and conducted the experiments. JVB conducted the analyses and wrote the manuscript. CW helped with analysis and edited the manuscript. All authors read and approved the final manuscript.

## Abstract

Assessments for spatial working memory (SWM) in pet dogs that can detect age-related cognitive deficits in a single session may aid in diagnosing canine dementia and may facilitate translational research on Alzheimer's disease in humans. Adaptive testing procedures are widely used in single-session assessments for humans with diverse cognitive abilities. In this study, we designed and deployed two up-down staircase assessments for SWM in which 26 pet dogs were required to recall the location of a treat hidden behind one of two identical boxes following delays of variable length. In the first experiment, performance tended to decline with age, but few dogs completed the test (n =10). However, all of the dogs that participated in the second experiment (n = 24) completed the assessment and provided reliable evidence of learning and retaining the task. Delay length and age significantly predicted performance supporting the validity of this assessment. The relationships between age and performance were described by inverted U-shaped functions as both old and young dogs displayed deficits in weighted cumulative-scores and trial-by-trial performance. Thus, SWM in pet dogs may develop until mid-life and decline thereafter. Exploratory analyses of non-mnemonic fixation strategies, sustained engagement, inhibitory control, and potential improvements for future SWM assessments which adopt this paradigm are also discussed.

As in humans, the cognitive abilities of dogs may decline with age. For some, these impairments are minor and only present late in life (Chapagain et al., 2018; Head, 2001). For others, cognitive decline begins early in life and progresses more rapidly, ultimately leading to severe impairments (Adams et al., 2000a; Head, 2001). As with Alzheimer's Disease (AD) in humans (Kensinger et al., 2003; Simone and Baylis, 1997), age-related cognitive deficits (ARCD) in dogs are characterized by the deterioration of recent memory and other executive functions including selective attention, behavioral inhibition, and concept learning (Adams et al., 2000a; Head, 2013). Although the pathogenesis of ARCD is unclear, both ARCD and AD are correlated with the development of neuritic beta-amyloid plaques (Head, 1998; Vite and Head, 2014). In clinical settings, the development of behavioral impairments in elderly pet dogs is often referred to as Cognitive Dysfunction Syndrome (CDS) (Landsberg et al., 2012; Szabó et al., 2016). As with AD, CDS is characterized by increased anxiety, nocturnal activity, agitation, and disorientation (Madari et al., 2015; Landsberg et al., 2012; Fast et al., 2013).

This cognitive, behavioral, and physiological overlap suggests that dogs may provide a strong animal model for translational AD research as well as a naturally occurring alternative to transgenic rodent and invertebrate models (Araujo et al., 2017). In addition, dogs are highly tractable, widely accessible, and share a number of anatomical similarities with humans but have shorter lifespans (Gilmore and Greer, 2015; Head, 2013; Kaeberlein et al., 2016). Furthermore, research on ARCD in pet dogs living with elderly owners may provide insight into the environmental correlates of AD (Kaeberlein et al., 2016).

ARCD in dogs are domain specific and are more readily detected with complex tasks that require coordination of multiple cognitive functions (i.e., executive functions). For example, older beagles are slower to acquire, and commit more errors in, tests for egocentric spatial learning, oddity and size discrimination, and reversal learning but do not display deficits in simple visual discrimination or procedural learning tasks (see Head 2013 for review). Different studies have produced different estimations for the onset age of dementia in dogs, but most suggest that neurological, cognitive and behavioral changes begin in midlife (for review, see Chapagain et al. ,2018; Szabó et al., 2016). Importantly, sensorimotor abilities and some executive functions including attention and working memory may continue to develop after adolescence in dogs (Wallis et al., 2014; Watowich et al., 2020). Tests for spatial working memory (SWM), “the process of maintaining a limited amount of spatial information in an active representation for a short period of time so that it is available for use” (Adams et al., 2000b, p. 48), are perhaps the most useful assessments for ARCD in dogs (Head et al., 2013). Moreover, SWM deficits are correlated with cortical atrophy (Rofina et al., 2006), reduced neurogenesis (Siwak-Tapp et al., 2007), and old age in colony beagles (Adams et al. ,2000b; Chan et al., 2002; Tapp et al., 2003b; Studzinski et al., 2006).

Research on ARCD in pet dogs is currently limited by a lack of reliable cognitive assessments which can be administered in a single session using cheap and accessible materials and without requiring extensive behavioral shaping (Piotti et al., 2017). SWM assessments for colony beagles implement stringent training criteria and yield reliable performance estimates but require multiple shaping, training, and testing sessions. Indeed, established tests for SWM require weeks or months of daily sessions (e.g., Adams et al., 2000b; Milgram et al., 1999). In contrast, existing SWM assessments for pet dogs (González-Martínez et al., 2013; Piotti et al., 2017) can be completed in a single session but do not implement training criteria or other important controls. Ultimately, valid and reliable single-session assessments for SWM are needed to better understand ARCD in pet dogs.

Such assessments may help to clarify the relationships between age-related cognitive deficits, behavioral changes, and neurological changes in pet dogs. Used in combination with physiological tests and questionnaires, cognitive assessments may also aid in the diagnosis of CDS (Wallis et al., 2016), thereby facilitating veterinary care for millions of pet dogs (for prevalence estimates, see Azkona et al., 2009; Salvin, 2010). Moreover, physicians use a number of tools to diagnose AD including behavioral questionnaires, psychometric assessments, physiological assays, and neurological imaging. In contrast, veterinarians primarily diagnose CDS using only owner reports and questionnaires, the validity and consistency of which remain controversial (Szabó et al., 2016; 2018).

Adaptive testing procedures (ATPs) accommodate a wide range of abilities within a single assessment by using a response-dependent testing strategy to tailor each test to the individual test-taker (Wainer and Lewis, 1990). Relative to traditional tests which present tasks spanning a broad range of difficulty to all test-takers, ATPs improve efficiency by omitting tasks that are far from an individual's performance threshold - tasks that are too challenging or too simple (Marinissen et al., 2010; Watson and Fitzhugh, 1990). ATPs are primarily used to estimate maximal ability, peak performance, or cumulative performance. For example, the Graduate Management Admission Test®, a computerized adaptive test that determines the difficulty of future questions based on the outcomes and difficulty of preceding questions, produces a weighted cumulative score which awards more points for correct answers to difficult questions.

Up-down staircase methods, in which an individual may advance or regress through a sequential experimental design, have been widely used to measure performance thresholds in humans and non-human animals (Cornsweet, 1962; Ehrenstein and Ehrenstein, 1999). For example, staircase procedures are commonly used in hearing tests to rapidly determine the highest frequency that is audible to the test-taker. In each trial, the proctor plays a recording of a short tone and the test-taker indicates whether she hears this tone. When this occurs, the proctor plays a tone of a higher frequency in the following trial. Otherwise, the proctor plays a tone of lower frequency in the following trial.

In staircase assessments, test-takers oscillate between completing tasks which fall within the range of their abilities and failing tasks which exceed their abilities. Thus, staircase methods can reliably estimate performance thresholds without stringent progression criteria (i.e., criteria for "stepping-up" in the staircase). Given that task difficulty is additive, staircase procedures can also rapidly estimate ability by measuring performance within ranges of task difficulty.

The purpose of this study was to design an adaptive test for SWM which can reliably and rapidly detect ARCD in pet dogs. In two experiments, we designed and deployed assessments in which 26 pet dogs were required to recall the location of a treat hidden behind one of two boxes following a delay (retention period) of variable length. The first assessment used a transformed staircase procedure (Kaernbach, 1991) which required the dog to make consecutive correct choices to progress into trials with longer delays and implemented zero-delay correction trials following incorrect choices. To address an apparent ceiling effect and to reduce the rate of fail-out, the second assessment used a transformed and weighted staircase procedure (Kaernbach, 1991) in which the magnitude of change in the delay was also performance-dependent (i.e., step sizes varied).

To test the validity of the modified assessment in the second experiment, we examined whether each dog learned and followed the rules of the task, age predicted performance, and performance was delay-dependent. We then conducted supplementary analyses of the dog's head and body orientations to explore the processes through which dogs stored and maintained information in SWM and to assess whether the dog's ability to remain engaged with the task could account for age effects observed in the primary analyses. Finally, we conducted two exploratory analyses to inform the design of future assessments which could adopt this paradigm. Specifically, we examined whether this experimental design may be used to detect age-related deficits in inhibitory control, and whether a shorter, fixed-delay procedure may be used to detect age-related SWM deficits.

## Method

### Subjects

Twenty-six pet dogs (13 of each sex) of various breeds and ages (17 - 181 months, mean = 89.82) were tested (see Appendix C-1). Given that the age estimation of an adult dog may not be reliable, only dogs seen by a veterinarian before reaching one year of age were included in this study. All owners volunteered their dogs to participate.

### Materials and Layout

One-meter tall collapsible, plastic fencing was used to create a 5 by 5m testing area and a 2 by 1m holding area in an adjacent room (Fig. 2-1). A 1 by 1m doorframe-mounted dog gate was used to restrict the dog to the holding area without blocking visibility into the testing area. To prevent visibility, a 1.5 x 1.0 x 0.04m foam board was placed against the dog gate. Throughout each session, a demonstrator (E1) remained in the testing area, an assistant (E2) and the owner remained in the holding area, and the dog moved between the two areas. To minimize distraction while preventing separation anxiety, the owner sat in a chair opposite the testing room facing away from the experimental area, reading silently and ignoring the dog. To control access and visibility into the testing room while minimizing interactions with the dog, E2 sat in a chair directly adjacent to the doorway facing the testing room.

Treats were hidden in one of two identical boxes placed on the ground in the testing area (Fig. 2-1). These boxes were spaced 2m apart and 2.5m from the starting location of both the dog and E1. Each box consisted of a 0.2 x 0.2 x 0.1m base, upon which the experimenter placed food treats, and a 0.1 x 0.2 x 0.2m shield, which could be configured to control both visibility and access to treats (Fig. 2-1.d). The shield consisted of a transparent plastic face housed in a wooden frame which contained a slot for an opaque plastic sheet (an occluder). The base contained a hidden compartment for smell controls: treats placed in the boxes before each session to prevent dogs from locating treats by smell.

### Procedure

To prevent satiation, dogs were fasted for four hours preceding the test. Before training, each dog was allowed to freely explore the testing and waiting areas. This acclimation period concluded when the dog ceased all exploratory behaviors and at least five minutes had elapsed.

In both training and testing, trials consisted of a demonstration by E1 and a choice by the dog. During demonstrations, E1 baited one of the boxes while the dog observed from the holding area. Starting from the far end of the testing area, E1 first attracted the dog's attention by taking one step forward while simultaneously holding up a treat and calling the dog's name. Keeping the treat raised, E1 walked towards, and then placed a treat on, one of the boxes. If the dog looked away during this demonstration, E1 paused and called to the dog until the dog redirected its attention towards E1 and the treat. After returning to the starting location, E1 turned to face the holding area and gazed at the top of the doorway. The choice portion of each trial began when E2 opened the gate while simultaneously speaking the specific release word used by the dog's owner. The dog was then given up to one minute to retrieve the treat but was only allowed to visually inspect a single box.

In between each trial, E1 returned the dog to the holding area and remained by the gate until E2 placed a foam board in front of the gate to block the dog's view of the testing area. This allowed E1 to remove an un-retrieved treat from the previous trial while maintaining the association that dogs were required to learn for this task (i.e., the box approached and touched by E1 always indicated the location of hidden food). After E1 collected a new treat and returned to the starting location, E2 indicated to E1 the next box to be baited in the upcoming trial.

Before testing, each dog completed three stages of training in which the choice portion of the trial immediately followed the demonstration. In Training Level 1 (T1) only one box was present in each trial and the treat remained visible throughout the demonstration and choice (i.e., the treat remained in view after being placed on the box platform in the demonstration). In Training Level 2 (T2) only one box was present in each trial but the occluder was added to the box shield so that the treat was not visible once placed on the box. In Training Level 3 (T3) and throughout testing, both boxes were present for each trial and treats were not visible once placed on a box.

With the box occluders in place, the dog was required to walk past the sides of the box shields in order to see which of the boxes held the treat. Thus, in testing, the first box visually inspected by the dog after entering this area of visibility was scored as the dog’s choice for that trial. Choices were coded as a correct if the inspected box corresponded to the baited box containing the treat. Choices were coded as incorrect if the inspected box corresponded to the un-baited box or if the dog inspected neither box within one minute. Choices were determined by E1 and later confirmed from videos recorded by two synchronized cameras. The north camera was used to determine when treats became visible (Fig. 2-1a) and the south camera was used to track the dog’s eyes (Fig. 2-1b)

Given that only one box was present in T1 and T2, the dog was not prevented from exploring the testing area before retrieving the treat in these training levels. However, a choice was only coded as correct if the dog walked directly to the correct box.

#### Experiment I protocol.

Training trials were presented in blocks of two and training blocks were repeated following an incorrect choice on either trial (a failed block). Correct choices on both trials in a block (a completed block) lead to a "step up" to the next level, but only if the proportion of correct choices (PCC) across all trials at the current training level was also greater than 50% (Table 2-1: after trials 2, 6, and 12, but not after trial 10). If the dog failed three blocks at the same training level, the test was ended.

During testing, a delay varying from 15 to 180s was implemented between the demonstration and choice, and the baited box followed a pseudorandomized order (see Fig. 2-2 for example of test progression). Trials at each delay were repeated in blocks of two and the length of the delay increased or decreased from block to block in increments of 15s. A completed block resulted in a step up to a longer delay, but only if PCC at the given delay was also greater than 50%. If PCC at the delay was equal to or less than 50% following a completed block (e.g., Fig. 2-2: trials 41-42), the delay was repeated (e.g., Fig. 2-2: trials 43-44).

Following a failed delay block, the dog was given a block of two trials in which a delay was not implemented between the demonstration and choice (i.e., "refreshers"). When the dog failed a refresher block (e.g., Fig. 2-2: trials 17-18), a second refresher block was implemented (e.g., Fig. 2-2: trials 19-20). If the dog failed the second refresher block, the test ended. Following a successfully completed refresher block, the dog returned to the delay staircase.

The first time the dog returned to the delay staircase (e.g., Fig. 2-2: trials 27-28), the previously attempted delay (e.g., Fig. 2-2: trials 23-24) and box were repeated. If the dog failed this block of delay trials (e.g., Fig. 2-2: trials 27-28) and then completed another block of refresher trials (e.g., Fig. 2-2: trials 29-30), the delay was decreased following the second (e.g., Fig. 2-2: trials 31-32) and third (e.g., Fig. 2-2: trials 35-36) consecutive returns to the delay staircase. If the dog failed the delay block after the third consecutive return, the test ended (e.g., if the dog made an incorrect choice in Fig. 2-2: trial 35 or 36). In addition, the test ended if the dog failed three consecutive blocks at the same delay (e.g., Fig. 2-2: after trials 55-56). If none of the above criteria for ending the test were met, the session continued for up to one hour. At the end of the test, two refresher trials were implemented to confirm that the dog was not satiated and to assess whether the dog continued to choose the box most recently visited by E1 throughout the test (e.g., Fig. 2-2: trials 57-58).

This up-down staircase procedure did not require stringent progression criteria to produce reliable evidence that the dog's memory-span was at least as long as a completed delay. Given that time is additive and that the spatial-matching rule was the same in T3, refreshers, and all delays, correct choices in longer delays also functioned as correct choices in shorter delays. For example, the dog in Figure 2-2 made correct choices in 13 of 14 trials with delays of 15-60s and therefore, provided highly reliable evidence of a 15s memory-span despite attempting only two trials at the 15s delay. The same reasoning holds for delays of at least 30s and delays of at least 45s. Thus, additional trials at these shorter delays would have been unnecessary and inefficient.

The up-down procedure and greater difficulty of remembering the correct location at longer delays also mitigated potentially detrimental consequences of allowing the dog to progress beyond its true SWM threshold. As demonstrated in trials 21-22 of Figure 2-

2, a dog with a 60s memory-span was not unlikely to progress beyond the 75s delay by randomly guessing the correct box in two consecutive trials ($p = .25$ if the delay was novel and $p \leq .19$ if the dog had previously attempted the delay). However, the probability of completing consecutive blocks by random chance was low ($p = .06$ if both delays were novel and $p \leq .03$ if the dog had previously attempted blocks at these delays). Additionally, the dog was unlikely to avoid repeating a delay that was previously completed by chance. Indeed, the example dog in Fig. 2-2 was ultimately required to complete six more blocks at the 75s delay after erroneously progressing to the 90s delay. Most importantly, the test provided evidence that the dog's memory-span was at least 60s (7 of 8 correct) but less than 75s (8 of 16 correct).

### Experiment II protocol.

The protocol for Experiment 2 differed from Experiment 1 in the following respects. Both in training and testing, trials were not necessarily implemented in blocks of two. In training, the baited box alternated after a correct choice. An incorrect choice immediately resulted in a correction trial in which the training level and baited box were repeated. Following a correct choice on a correction trial, the training step was repeated at the other box (Table 2-2: trials 4-6). If the dog chose incorrectly on a correction trial, that trial was repeated. Two successive incorrect choices on correction trials resulted in a step down to a correction trial on the previous training level (Table 2-2: trials 7-10). If at any point a dog did not retrieve food in four successive trials, the session was ended.

During the test, a delay of 15-300s was implemented between the demonstration and choice. From 0 to 59s, delays increased by 15s; from 60 to 179s, delays increased by 30s; and from 180s on, delays increased by 60s (Fig. 2-3: dog A, trials 1-20). An incorrect choice on any delay trial immediately resulted in a zero-delay refresher trial (e.g., Fig. 2-3: dog B, trial 23). After one correct choice on a refresher trial, the dog was given a second refresher trial on the other box. Following an incorrect choice on a refresher trial, the dog was given a correction trial which consisted of another refresher trial on the same box. Correction trials were repeated until a correct choice was made, after which another refresher at the other box was implemented. The dog returned to the delay staircase following consecutive correct choices if PCC in refresher trials was above 50%.

The first time a dog returned to the delay staircase, the previously attempted delay and box were repeated. However, in successive returns, the previously attempted box was repeated but the delay decreased. If the failed delay was between 0 and 180s, the delay decreased by 15s (Fig. 2-3: dog B, trials 40-55). If the failed delay was more than 180s or, the delay decreased by 30s (Fig. 2-3: dog A, trials 25-33).

After failing the first block of 15s delay trials, one dog (Greta) developed a strong side bias and failed to return to the delay staircase despite 10 refresher and correction trials. Therefore, she was given a counter-training procedure which required four consecutive correct choices in both T2 and T3.

### Analysis Overview

Due to a low rate of test completion in Experiment I, we used only descriptive statistics and graphical analyses of test outcomes to inform the design of the second assessment. Statistical analyses for Experiment II are outlined in Table 2-3. Tests of assumptions and case-wise diagnostics for these analyses are described in Appendix C-2.

All data were analyzed in R version 3.4.1. Generalized linear mixed models (GLMMs) were constructed and tested using the package “lme4” (Bates et al., 2015b).

## Body Size and Sex Effects

The prevalence of CDS may be similar in breeds which differ greatly in size (Salvin et al., 2010; 2012). In addition, Watowich et al., (2020) recently found that lifespan changes in cognitive performance did not depend on the average lifespan of a dog’s breed. However, lifespan is inversely related to body size in dogs and physiological deterioration may progress more rapidly in large breeds (Kraus et al., 2013). Thus, to account for possible effects of body size, all statistical analyses were repeated with age measured in months (AGE) and as a proportion of the dog’s size-predicted lifespan (PSL). Size-predicted lifespan was calculated using the following equation from Greer et al. (2007): predicted lifespan (years) = 13.62 + (0.0702 x height (inches)) - (0.0538 x weight (pounds)). The relative strength and significance of predictors did not change in any analysis when age was treated as PSL. Therefore, only analyses of AGE are reported.

Sex may affect spatial learning and memory in cognitive aging tests for dogs (e.g., Mongillo et al., 2017). Thus, all analyses were first conducted with sex included as a covariate (see Appendix C-3). Sex did not predict performance in any analysis and lead to issues of non-convergence in analyses which used GLMMs. In addition, accounting for sex did not lead to any meaningful change in the outcome of an analysis.

## Experiment II Measurements

### Proportion of correct choices.

The validity of this assessment was contingent on evidence that each dog attempted to search for food at the box visited by E1 during the most recent demonstration (i.e., the correct box). Moreover, incorrect choices in delay trials could not be attributed to SWM deficits if the dog failed to follow this spatial-matching rule. Thus, for each dog, PCC in zero-delay trials (the third step of training and refreshers) and PCC in delay trials were used to examine whether the dog learned the spatial-matching rule and chose accordingly throughout the test.

#### Cumulative memory score.

To measure overall performance, the sum of the delays after which the dog chose the correct box was used as a weighted Cumulative Memory Score (CMS). Only the first 18 delay trials (the fewest completed by any dog) were included in this score. Thus, CMS measured the total amount of time in the first 18 delay trials that the dog remembered the location of the treat.

#### Orientation behaviors.

To investigate whether the dog used non-mnemonic orientation strategies to locate the treat, a coder recorded from video whether the dog's head and body (coded separately) remained oriented towards the correct box for the entirety of the delay. Orientation was scored as correct when the long axis of the head or body pointed towards the correct box, between the midpoint and upright of the doorway (Fig. 2-4). To measure initial and sustained attention, the coder scored whether the dog's head and body (coded separately) remained oriented toward the testing area for the entirety of the demonstration and the delay (coded separately). Orientation was scored as forward when the long axis of the head or body pointed towards the testing area between the uprights of the doorway (Fig. 2-4). To assess interrater reliability, 25% of the videos were scored by a second coder. For each orientation behavior, coders agreed on at least 98% of trials. Due to video inadequacies, orientations could not be coded for 11 (2.5%) trials.

#### Box preference (side bias).

Perseveration rather than inadequate memory span may result in incorrect choices. The dog may incorrectly choose its preferred box despite correctly recalling that E1 visited the other box in the most recent demonstration. In the first experiment, the dog completed the same number of trials at each box. Therefore, the number of errors the dog made at each box was used to assess whether it was hindered by a perseverative tendency to choose its preferred box. In the second experiment, correction trials were implemented to combat the development of a side bias. Thus, dogs were not required to complete the same number of trials at each box. As a result, the magnitude of the dog's box preference was used to measure his perseverative tendency. Specifically, side bias was calculated as the absolute value of the difference between PCC on Box 1 and PCC on Box 2.

## Results

### Experiment I

The completion rate in the first experiment was too low to test whether age significantly predicted performance. Thus, to determine which procedural modifications were needed to allow dogs of all ages to complete the test, the longest delay attempted (i.e., the highest step reached by the dog) and the reason for ending the test were plotted as a function of age (Fig. 2-5). To assess whether such modifications were likely to produce a viable test for ARCD, a regression line was fitted to the data of dogs that completed the test.

Ten of the twenty-six dogs did not pass training. Three of these dogs stopped participating, three never participated, and four made incorrect choices in three consecutive blocks at T3. Two dogs stopped participating after the first incorrect choice in a delay trial and three made incorrect choices in consecutive blocks of refresher trials. One test was ended prematurely due to an experimenter error. Of the seven dogs that participated fully but failed to complete the test, five provided evidence of a strong side bias; each of these dogs made at least four times as many errors when the preferred box was not baited. Collectively, these dogs made 28 errors when the non-preferred box was baited and three errors when the preferred box was baited.

Among the ten dogs that completed the test (AGE: $\bar{x} \pm s = 97.7 \pm 33.0$ months), the longest delay attempted decreased rapidly with age (delay = 138.3 - 1.13 * centered AGE). However, no dogs younger than four years and older than twelve years completed the test.

## Experiment II – Assessment Validity

### Test completion.

Of the 25 dogs that returned, 23 completed the test. The two dogs that did not complete the test never participated in either experiment. These dogs were among the oldest tested (AGE = 163 and 134 months). The other dog that did not participate in Experiment I acquired and retained the spatial-matching rule in Experiment II and completed this modified assessment but still did not make a choice in 25% of all trials (AGE = 40 months). This dog was only included in analyses of PCC and side bias, which did not include no-choice trials. Of the remaining 22 dogs, 18 participated in all trials and four participated in at least 90% of trials.

### Proportion of correct choices.

For each dog that participated, two binomial tests were conducted to assess whether PCC in zero-delay and delay trials differed from chance ($E(\text{PCC}) = 0.5$). PCC was significantly above chance in zero-delay trials for 20 of the 23 dogs (Fig. 2-6). The other three dogs were each far above chance in delay trials ($p < .0001$; Appendix C-4). PCC was significantly above chance in delay trials for 19 of the 23 dogs (Fig. 2-6). The other four dogs were each far above chance in zero-delay trials ($p < .0005$; Appendix C-4).

### Cumulative memory score.

A regression analysis of CMS by centered AGE (AGE.CENT) was conducted to assess whether age predicted cumulative performance. However, an inverted U-shaped relationship was observed between AGE.CENT and CMS. Thus, a coefficient for the predictor AGE.CENT$^2$ was also included in the regression equation. An F-test was used to test the overall prediction of the model and two-tailed t tests were used to test the significance of the predictors. AGE.CENT and AGE.CENT$^2$ yielded significant overall prediction of CMS, $F(2,19) = 7.24$, $p = .005$; $r^2 = 0.43$. The linear effect of AGE.CENT was not significant, $b_1 = -0.40$, $SE_{b1} = 1.65$; $t(19) = -0.24$, $p = .81$, but AGE.CENT$^2$ significantly predicted CMS, $b_2 = -0.12$, $SE_{b2} = 0.03$; $t(19) = -3.73$, $p = .001$. Memory scores ($\bar{y} = 967.5$, $SE_y = 92.41$) were higher in middle-aged dogs than in young and old dogs (Fig. 2-7).

### Test-trial outcome.

To assess whether performance during the test (i.e., in delay and refresher trials) was age- and delay-dependent, a binomial regression analysis of test-trial outcome (correct or incorrect choice) was conducted using a GLMM with a logit-link function (Table 2-4). A hypothesis-driven (full) model was first constructed with random subject slopes for delay and trial, random intercepts for subjects, and fixed effects for delay, trial number, AGE, $\text{AGE}^2$, delay*AGE, delay*$\text{AGE}^2$, trial*AGE, and trial*$\text{AGE}^2$. The overall fit of this model was then compared to that of the maximal converging (final) model (tolerance = 0.001) using Akaike and Bayesian Information Criteria (AIC and BIC). To aid in convergence, continuous predictors were rescaled to *Z* scores. Profiled confidence intervals of fixed-effect estimates and of random-effect variances were calculated using parametric bootstraps. In addition, likelihood ratio tests of competing models ($\alpha = .05$) were used to test the significance of interactions and random effects (this was not possible for the main effects).

The full model did not converge due in part to overfitting with interaction terms and random slopes. Furthermore, the random slopes for trial and delay were perfectly correlated with the random intercepts. The final model included only the main effects and random intercepts. Removing the interactions and random slopes improved overall fit (decreased AIC and BIC) but did not change whether the profiled confidence intervals of the remaining predictors overlapped with zero (i.e., the same predictors were significant). Both models are summarized in Table 2-4.

In the final model, the probability of choosing the correct box decreased significantly with delay length, $X^2(1) = 20.89$, $p < .0001$, but not trial number, $X^2(1) = 1.29$, $p = .26$. The effects of AGE, $X^2(1) = 11.68$, $p = .0006$, and $\text{AGE}^2$ were significant, $X^2(1) = 10.54$, $p = .001$. Middle-aged dogs chose correctly more often than young and old dogs. The random intercept for study subject was not significant, $X^2(1) = 2.25$, $p = .13$.

## Experiment II – Exploratory Analyses

### Head and body orientations.

Because they did not apply to refresher trials, delay orientations could not be treated as covariates in the previously described GLMMs, which estimated the outcome of either a refresher or a delay trial. Therefore, separate exploratory analyses were used to examine whether head and body orientations predicted the outcome of a delay trial. However, the previous analysis indicated that this reduced set of data could not support a model with more than a limited set of predictors (i.e., a model with coefficients for all six orientation behaviors, AGE, AGE$^2$, delay, trial, the interactions of these fixed effects, and random effects would be overfit). Thus, the predictive value (i.e., utility) of each orientation behavior was first estimated using a GLMM of trial outcome with a binomial (logit) error distribution, fixed effects for each orientation (orientation maintained or broken) and random intercepts for study subjects (see Table 2-5).

Only body orientated forwards during the delay (B.F.Dl) significantly predicted trial outcome, $X^2(1) = 8.46$, $p = .004$. Therefore, to explore whether B.F.Dl and, in turn, sustained engagement might account for the age effects observed in the primary analyses, another GLMM of delay-trial outcome (logit error distribution) was constructed with fixed effects for B.F.Dl, AGE, AGE$^2$, delay, B.F.Dl*AGE, and B.F.Dl*AGE$^2$. Subject intercepts and slopes for B.F.Dl were included as random effects. The final model is outlined in Table 2-6[1]. Confidence intervals for the conditional effects of Age and Age$^2$

---

[1] Trial number was not included as a predictor in this model given that the omission of refresher trials created highly discontinuous and individualized trial number data (i.e., numerous singularities). In addition, trial number yielded no prediction of trial outcome in the previous models. Likewise, random subject slopes

on delay-trial outcome when the dog did (B.F.Dl = 1) and did not (B.F.Dl = 0) maintain orientation were simulated using the package “interplot” (Solt and Hu 2019), following the method of Brambor et al. (2006).

Dogs that remained oriented towards the testing room during the delay were more likely to choose the correct box, 95% CI of β [0.11, 1.26]. The interactions between B.F.Dl and AGE, $X^2(1) = 0.01$, $p = .94$, and between B.F.Dl and AGE$^2$, $X^2(1) = 0.02$, $p = .88$, were not significant. Middle-aged dogs chose correctly more often than young and old dogs regardless of B.F.Dl. The conditional effect of AGE was significant in trials in which the dog remained oriented forwards, β = 1.77, 95% CI: [0.46, 3.08], and in trials in which the dog turned away from the testing room, β = 1.86, 95% CI: [0.38, 3.35]. Likewise, the conditional effect of AGE$^2$ was significant whether the dog did, β = -1.84, 95% CI: [-3.18, -0.52], or did not, β = -1.71, [-3.08, -0.37], maintain forward body orientation during the delay. The probability of choosing the correct box tended to decrease with delay length, but this trend was not significant, $X^2(1) = 3.49$, $p = .06$. Random subject intercepts, $X^2(1) = 0.19$, $p = .91$, and slopes for B.F.Dl, $X^2(1) = 0.92$, $p = .63$, were not significant.

**Perseveration.**

To investigate whether inhibitory control may provide an alternative focus for future cognitive aging assessments using this paradigm, a regression analysis of side bias by AGE.CENT and AGE.CENT$^2$ was conducted. Trials in which the dog failed to make a

---

for delay, and coefficients for the interactions between delay and age were also omitted because these effects lead to singularities and non-convergence but did not predict trial outcome, improve model fit, or change the prediction of other effects in the previous models.

choice were excluded from this analysis. AGE.CENT and AGE.CENT$^2$ yielded significant overall prediction of side bias, $F(2,20) = 5.37$, $p = .01$; $r = 0.35$. The coefficients for AGE.CENT, $b_1 = 0.0008$, $SE_{b1} = 0.0004$; $t(20) = -2.15$, $p = .04$, and AGE.CENT$^2$ were also significant; $b_2 = 0.00002$, $SE_{b2} = 0.00007$; $t(20) = 2.87$, $p = .01$. Box preferences (mean = 0.11, $SE_y = 0.02$) were stronger in young and old dogs than in middle-aged dogs (Fig. 2-8).

### Delay optimization.

To identify the shortest delays that may be sensitive to ARCD in this paradigm, a separate regression analysis of PCC by AGE and AGE$^2$ was conducted for each delay that was attempted by at least 20 dogs. The F test and multiple $r^2$ of the model, and the two-tailed t tests of AGE and AGE$^2$ were then compared among delays (Table 2-7).

AGE and AGE$^2$ accounted for a significant proportion of variation in PCC at the 45s ($r^2 = 0.33$, $p = .02$) and 60s delay ($r^2 = 0.35$, $p = .02$) but not the 15s ($r^2 = 0.20$, $p = .12$) or 30s delay ($r^2 = 0.15$, $p = .22$). Two of eight old dogs ($\leq$ 8 years) displayed deficits at the 15s delay, compared to four old dogs at the 30s and 45s delays, and seven at the 60s delay. The 60s delay was also the longest delay reached by one old dog. Bivariate PCC by AGE plots for all delays are displayed in Figure 2-9.

## Discussion

### Experiment I

Although the age-related decline in performance among dogs that completed the first test suggested that this assessment may be sensitive to ARCD, the low test-completion rate indicated that this version of the assessment was not viable as a widely deployable diagnostic tool. However, this experiment did provide insights that led to effective modifications for the second assessment. The high prevalence of side biases among dogs that failed to complete the test suggested that perseverative tendencies may be strong in dogs. This led to the addition of correction trials after incorrect training and refresher trials, as well as the analysis of age-related changes in side bias in Experiment II. The finding that two dogs stopped participating after the first incorrect choice on a delay trial indicated that refresher trials may not successfully re-incentivize participation for some dogs. However, the fact that this occurred only twice in 50 tests also suggested that, for most dogs, refresher trials may help to maintain participation throughout this hour-long assessment. Finally, the finding that two dogs did not participate in either experiment suggested that insufficient motivation may present a small but unavoidable obstacle to assessments for ARCD. However, the four-fold decrease in fail-out due to non-participation in Experiment II indicated that testing-rules may be modified to greatly increase motivation and participation.

## Experiment II – Assessment Validity

### Proportion of correct choices.

Dogs that completed the second assessment provided strong evidence of learning the spatial-matching rule. Indeed, only three dogs did not significantly differ from chance in zero-delay trials and in each case, this was reflective of a small sample size (mean = 7) rather than poor performance (mean PCC = 81%). These dogs excelled in delay trials (PCC = 95%) and thus additional zero-delay trials (i.e., refreshers) were not needed to confirm acquisition of the spatial-matching rule. Moreover, the testing rules only allowed for such a small number of zero-delay trials when the dog committed few errors in either T3 or testing. Nonetheless, these non-significant results could be prevented by implementing additional refresher trials at the end of the test and this would require little additional testing time.

Evidence that dogs continued to use the spatial-matching rule to retrieve treats in delay trials was also robust. Only 4 of 23 dogs failed to choose the correct box more often than chance in delay trials but each succeeded in doing so in zero-delay trials. This indicates that incorrect choices in delay trials resulted from insufficient SWM rather than failure to apply the spatial-matching rule.

Compared to previous SWM assessments for ARCD in colony beagles, this assessment required far fewer trials to confirm that the dog learned the spatial-matching task. Whereas beagles completed up to 40 daily sessions of ten training trials in previous assessments, dogs in the present study completed an average of 9.57 training trials and 10.77 refresher trials (see Fig. 2-6). This increase in efficiency was partially enabled by adopting a less cognitively demanding task. The delayed non-matching to position (DNMP) tasks used in previous studies required additional information processing and greater behavioral inhibition relative to the simple spatial-matching task used in the present study. Additionally, completing each training step in this staircase procedure required as few as two consecutive correct choices whereas DNMP training criteria required nine of ten, or eighteen of twenty correct choices. These lenient criteria and, in turn, rapid shaping and training procedures were made possible by the inclusion of response-dependent refresher trials, which collected additional evidence of the ability to follow the spatial-matching rule from any dog that did not provide this evidence in delay trials.

To our knowledge, refresher trials are the first controls in SWM assessments for dogs which assess task retention independently from task acquisition. A dog may stop following previously learned rules due to intertrial interference, the development of ineffective strategies, or fatigue. The dog may also meet training criteria by random chance and therefore begin the test before truly learning the reward-contingencies or appropriate responses. Traditional procedures do not address these possibilities and assume that the dog retained the task based solely on evidence that the dog learned the task in training. In the present study, dogs made correct choices in over 90% of the 283 refresher trials and the only dog below 80% PCC in refresher trials (3 of 4 correct) made correct choices in 20 of 21 delay trials. Thus, refresher trials provided evidence that each dog remembered the rules of the task and in turn assessed SWM.

Although training criteria in the present study were lenient relative to assessments for colony beagles, they provided substantial improvements relative to the two previous SWM assessments for ARCD in pet dogs. González-Martínez et al. (2013) compared the tendencies of old and young dogs to locate a hidden treat in an open field test after a 60-second retention period. They scored whether the dog (1) immediately located the treat, (2) located the treat after searching, (3) failed to locate the treat, or (4) made no attempt to locate the treat. Piotti et al. (2017) compared the tendencies of old and young dogs to locate a treat hidden in one of five containers after a 30-s retention period. They analyzed the first container approached by the dog in each of five trials (one per container).

Although both studies found significant differences between young and old dogs, neither confirmed that the test assessed SWM. The single-trial procedure used by González-Martínez et al. (2013) did not allow the dog to learn the task and the repeated-trials procedure used by Piotti et al. (2017) provided no incentives for learning the task given that the dog was allowed to retrieve the treat after an incorrect choice. In addition, neither study implemented training trials and neither assessed whether individual performance differed from random chance. Thus, these studies did not demonstrate that each dog encoded and remembered the location baited by the experimenter in the most recent demonstration. Furthermore, these studies did not confirm that all dogs were motivated to retrieve treats. In contrast, González-Martínez et al. (2013) scored non-participation as the most severe memory deficit. Ultimately, the present study may provide the first assessment for ARCD in pet dogs which directly measures SWM rather than incidental or uninformed search behavior (i.e., behavior which is likely influenced by SWM but to an unknown extent).

### Delay-dependent performance.

As illustrated by these previous assessments for pet dogs, errors in trials with delays can only be attributed to memory deficits if these errors are less frequent in trials without delays. Moreover, lower performance in delay trials than in training trials indicates that an assessment is memory dependent. Demonstrating that performance decreases with increasing retention periods provides even stronger evidence of a valid memory assessment. In the present study, the probability of choosing the correct box during the test decreased significantly with delay length. Although dogs only attempted longer delays in later trials, the effect of trial number was not significant and thus the effect of delay could not be attributed to fatigue or intertrial interference. In addition, differences in the delays and number of trials attempted by dogs of different ages could not account for this finding. Ultimately, this delay-dependent performance supports the validity of the present assessment as a test for SWM.

Our findings align with previous indications that spatial-matching paradigms may be used to assess SWM in pet dogs. Fiset et al. (2003) found that the spatial recall of young pet dogs in a similar three-location test gradually declined between 10s and 60s, at which point performance remained stable and above chance for up to 240s. Hunter (1913) and Walton (1915) observed delay-dependent performance in individual dogs using two- and three-location delayed-response tests. To our knowledge, all other previous SWM tests either used non-matching to position tasks or did not test for delay dependence.

### Age-dependent performance.

The finding that older dogs were less likely to choose correctly indicates that this memory assessment was sensitive to ARCD. Moreover, the outcomes from the GLMMs of test-trial outcome indicated that differences in the delays and number of trials attempted by dogs of different ages could not account for these deficits in older dogs. However, such models of trial-by-trial performance do not provide a means of determining whether dogs that complete this assessment in the future are cognitively impaired. To be clinically deployable ARCD assessments need to produce a single, easily interpreted score which pet owners and clinicians can compare either to previous tests on the same dog, or to a performance range typical of unimpaired dogs. Thus, the finding that cumulative memory scores declined after middle-age may satisfy this important requirement of a widely deployable ARCD assessment.

This study also provided consistent evidence of memory deficits in young dogs. Each of the four dogs younger than 2.5 years was more likely to choose the incorrect box in a delay trial than each of the twelve dogs between 2.5 and 8.5 years of age. In addition, ten of the twelve middle-aged dogs achieved a higher CMS than the four young dogs. However, the age of peak performance (around 7 years) was likely overestimated given that the only dog between 2.5 and 5 years of age included in these analyses was by far the top performer in this study. This dog chose the correct box in 18 of 18 trials up to the four-minute delay, thereby achieving the only perfect CMS. Likewise, the other dog in this age range (excluded due to low participation) made correct choices in all of the delay trials in which it participated.

Given that the youngest dog tested in the present study was nearly an adult (age = 17 months), the curvilinear relationship between age and performance indicates that SWM develops slowly in pet dogs. Indeed, assessments that are sensitive to cognitive decline should also be sensitive to cognitive development. Although previous studies have primarily used decreasing linear functions to model ARCD in dogs, studies on non-human primates (Manrique and Call, 2015), humans (Craik and Bialystok, 2006; Harada et al., 2013), and dogs (Fox, 1971) suggest that curvilinear models better estimate lifespan changes in cognitive abilities. In an analysis of citizen-science data from over 4000 pet dogs, Watowich et al. (2020) found that quadratic age functions better estimated performance in a broad range of cognitive assessments. Old and young dogs displayed deficits in social communication, inhibitory control, selective attention, reasoning, and working memory. Furthermore, these quadratic trends were strongest in tests for SWM and inhibitory control. In line with these findings, Gathercole et al. (2004) and Pickering et al. (2001) found that SWM continues to develop through adolescence and even early adulthood in humans.

Methodological differences may explain why previous experimental studies (i.e., studies which did not use citizen-science data) have not observed SWM deficits in young dogs. To our knowledge, eight such studies have examined the relationship between age and SWM in dogs. Two of these studies did not report the age ranges used to group dogs (Araujo et al., 2005; Piotti et al., 2017) and five of the other six studies used unique age groups. In addition, four studies tested laboratory-housed colony beagles (Chan et al., 2002; Tapp et al., 2003b; Araujo et al., 2005; Studzinski et al., 2006) and two tested both colony beagles and shelter-sourced dogs of uncertain ages (Head et al., 1995; Adams et al., 2000b). These six studies on beagles used five different variants of the DNMP task, all of which required additional information processing and greater behavioral inhibition relative to the simple matching task used in the present study. Two previous studies have tested pet dogs but neither confirmed that each individual was capable of and was motivated to complete the task (González-Martínez et al., 2013; Piotti et al., 2017). Ultimately, additional studies controlling for these differences are needed to clarify the developmental trajectory of SWM in dogs.

Although the present study is the first to observe SWM deficits in young dogs in a controlled experiment, previous studies have found little evidence that SWM declines monotonically with age from young adulthood onwards. Only three of these eight studies compared young and middle-aged dogs and their findings were inconsistent. Head et al. (1995) found that young and middle-aged beagles and pound-sourced dogs (of uncertain age) did not significantly differ in a three-location DNMP task, but performance declined with age at some delays. However, this study did not include any dogs between three and seven years of age, and did not test for a nonmonotonic, curvilinear relationship between age and performance. Using the same task, Studzinski et al. (2006) found that puppies (< 1y) and young beagles (1-3y) performed better than middle-aged beagles (6-8y). However, only one of the thirteen middle-aged beagles was sourced from a population that was found to significantly outperform the other populations sampled in this study. In contrast, all of the puppies and one third of the young beagles came from the high performing population. González-Martínez et al. (2013) found that young and middle-aged pet dogs did not significantly differ in an open-field food-search task but it is unclear whether this assessment measured SWM.

Although four previous studies found that young beagles outperformed old beagles, only two of these findings disagree with the present study. Young beagles in Chan et al. (2002) and Tapp et al. (2003b) were between 3-5 and 3.4-6.6 years of age, respectively. In the present study, dogs within these age ranges also out-performed older dogs. This highlights how comparisons between two age groups with truncated age ranges may erroneously indicate that performance declines linearly with age. When performance peaks in midlife, age groups can be defined such that young dogs appear to perform better, worse, or the same as old dogs. Furthermore, at least three age groups are required to detect a curvilinear relationship. To avoid this confusion, we suggest that future studies on ARCD should analyze age as a continuous variable.

## Experiment II – Exploratory Analyses

### Perseveration.

The significant relationship between age and side bias indicates that this assessment may be sensitive to selective attention and inhibitory control. Selective attention refers to an individual’s ability to ignore irrelevant information in order to effectively process task-relevant information (McDowd and Oseas-Kreger, 1991). Inhibitory control requires an individual to resist the impulse to act on irrelevant, distracting information (Bray et al., 2014). Thus, if the dog formed an association between a particular box and treats, the magnitude of this box preference may reflect its ability to selectively attended to the most recent demonstration or to inhibit the behavior of choosing the preferred box. In turn, the U-shaped relationship between age and side bias may indicate that inhibitory control and selective attention in pet dogs develop slowly and decline in old age.

Most studies on inhibitory control in pet dogs have observed performance deficits and stronger side biases in older individuals (e.g., Mongillo et al., 2013; Piotti et al., 2018) but similar studies on colony beagles have produced inconsistent results (Tapp et al., 2003a; Chan et al., 2002; but see Head et al., 1998). However, only one study has, to our knowledge, explicitly examined the development of these cognitive processes in dogs. In line with the present study, Wallis et al. (2014) found that selective attention and sensorimotor coordination in Border collies peaked in midlife. In humans, tasks that require selective attention and inhibitory control are particularly difficult for adolescents (Diamond, 2013). Development of the prefrontal cortex, which may play a critical role in inhibitory control (Braver et al., 2001; Diamond, 1990; Ridderinkhof et al., 2004), is also relatively slow in humans (Bunge et al., 2002; Dempster, 1992; West, 1996). Tapp et al. (2004) found that frontal lobe volume (as a proportion of total brain volume) was larger in young colony beagles (6m - 3.9 y) than in puppies (3m). In contrast, total brain volume (as a proportion of total intracranial volume) did not differ between young dogs and puppies. Thus, neurological development also suggests that inhibitory control may develop slowly in pet dogs.

The exploratory analysis of side bias cannot rule out a number of alternative explanations. For example, when the dog failed to remember which box was baited, it may have defaulted to choosing a particular box. If so, incorrect choices would only occur when the dog failed to remember which box was baited and the preferred box happened to be incorrect in that trial. Thus, the magnitude of the dog's box preference may measure memory rather than inhibitory control, and the outcome of the side-bias analysis may provide additional evidence that middle-aged dogs were more likely to remember the location of the correct box. Alternatively, this finding may indicate that middle-aged dogs were less likely to choose a default box when they failed to remember the baited location. Ultimately, this memory assessment was not designed to measure side biases or the processes through which they develop. Therefore, future assessments using this two-location paradigm which manipulate the side baited by the experimenter in zero-delay trials may help to clarify the relationships between age, side biases, selective attention, and inhibitory control in pet dogs.

#### Head and body orientations.

The findings of the present study indicate that dogs did not orient towards baited locations to circumvent the memory requirements of this spatial-matching task. Specifically, neither head nor body orientation towards the correct box during the delay improved performance. In contrast, dogs oriented their head and body away from the correct box in over 95% and 75% of all delay trials, respectively, but still chose the correct box in nearly 80% of these trials. In line with these findings, Fiset et al. (2000)

found that the probability of a correct choice did not depend on the dog's head or body orientation immediately, 5-s, or 10-s after a reward was hidden in one of three-boxes. Given that we did not code eye-tracking, it is possible that dogs visually fixated on the correct box. However, in a four-location visual displacement task, Gagnon and Doré (1994) found that dogs interrupted their visual fixation on the target box in 96.7% of 10s and 20s delay trials. Thus, visual fixation in delays of up to 300s in the present study was unlikely.

The finding that dogs were more likely to choose the correct box when they remained oriented towards the testing room during the delay indicates that the ability to remain engaged may constrain the processes through which dogs maintain locations in recent memory. Humans attend to relevant spatial locations to maintain visuospatial information in working memory (Smyth and Scholey, 1994) and thus, selective spatial attention serves as a rehearsal mechanism for SWM (Awh et al., 1998). In consequence, human observers respond faster and more accurately to stimuli within the spatial region of their current attention (Posner, 1980). In the present study, the orientation of the dog's body was used to measure its attentional space. Therefore, dogs may have ceased to actively maintain the treat location in memory when they lost interest in the test or were distracted (i.e., when they turned away from the testing room).

Such disruptions likely occur more frequently over increasing lengths of time. Thus, this increasing difficulty of remaining engaged may explain the finding that performance decreased with increasing delay. Indeed, after accounting for forward body orientation, the effect of delay on delay-trial outcome was only marginally significant. In contrast, the effect of delay on test-trial outcome was highly significant. However, this weaker prediction may be attributed to the fact that the latter model did not account for the higher performance of dogs in zero-delay trials.

These competing interpretations highlight the exploratory nature of the analyses of orientation behaviors. Moreover, an interruption to forward body orientation may be a consequence rather than a cause of a failure to remember the correct location. In addition, forward head orientation during the delay, the other hypothesized indication of sustained engagement, did not improve performance. Ultimately, additional studies are needed to clarify the relationship between head and body orientations, engagement, retention periods, and SWM.

### Delay optimization.

Future versions of this assessments may be able to detect ARCD more rapidly by using only 45s and 60s delays. Delays which are so short that they present little challenge to cognitively impaired dogs may only help to detect the most severe ARCD. Given that only two of eight old dogs ($\leq$ 8 years) displayed deficits at the 15s delay, and that age did not predict performance at the 15s or 30s delays, these delays may not be worth the additional testing time in future assessments. Given that AGE and AGE$^2$ accounted for the largest proportion of variation in performance at the 45s and 60s delays, the inclusion of longer delays in future assessments may also prolong the test unnecessarily while reducing sensitivity to ARCD.

## Conclusions

In this study, we designed an adaptive staircase assessment for SWM in pet dogs that was sensitive to ARCD. Both young and old dogs displayed performance deficits relative to middle-aged dogs indicating that SWM not only declines in old age but may develop slowly in young, adult dogs. Orientations of the dog's head and body suggested that the ability to remain engaged with the task may constrain the dog's ability to maintain locations in recent memory. However, dogs did not orient towards baited locations to circumvent the memory requirements of this task. Up-down psychometric staircase methods and, more broadly, adaptive testing procedures may be used to efficiently assess the cognitive abilities of pet dogs within a single session. However, future assessments which adopt this paradigm may be able to detect ARCD more rapidly by testing SWM in only 45s- and 60s-delay trials. Alternatively, this paradigm may be used to detect age-related deficits in selective attention and inhibitory control.

# CHAPTER 3

# A RAPID SERIAL REVERSAL LEARNING ASSESSMENT FOR AGE-RELATED COGNITIVE DEFICITS IN PET DOGS

Joshua Van Bourg[a,b,c], Lisa Gunter[a], Clive D. L. Wynne[a,d]

[a] Department of Psychology, Arizona State University, Tempe, Arizona, United States of America. PO Box 871104, Tempe, AZ 85287-1104

[b] Corresponding Author

E-mail: jvanbour@asu.edu

[c] ORCID ID: 0000-0003-0599-9602

[d] ORCID ID: 0000-0001-8408-1466

**Keywords**

Aging; Cognition; Dogs; Development; Inhibition; Reversal Learning

Abbreviations: ARCD: age-related cognitive deficits

**Funding**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Ethics approval**

This research was approved by the Arizona State University Institutional Animal Care and Use Committee (Protocol Number: 19-1681R)

**Consent**

Informed consent was received from the owners of all dogs that participated in this study and all participation was voluntary.

**Declaration of Interest**: none

## Abstract

Assessments for behavioral inhibition in pet dogs that can rapidly detect age-related cognitive deficits (ARCD) using inexpensive and accessible materials may aid in diagnosing canine dementia and may facilitate translational research on Alzheimer's disease in humans. In this study, we designed and deployed a spatial serial reversal learning test in which 80 pet dogs were required to learn which of two identical boxes contained a hidden food treat. Each time the dog chose the correct box in three consecutive trials the procedure was repeated using the other box. All dogs that completed shaping (n = 62) also completed the 30-minute assessment. Middle-aged dogs chose the correct box more often than younger and older dogs. This cognitive decline was detectable with a stand-alone score for perseveration that can be easily measured and interpreted by clinicians and dog owners. Age did not predict how frequently the dog learned the serially reversing reward contingency, but older and younger dogs displayed longer streaks of perseverative errors. Thus, ARCD in dogs may be better characterized by bouts of severe cognitive dysfunction rather than temporally consistent cognitive deficits. We  that future ARCD assessments for pet dogs should include measurements for intra-individual variability.

Some dogs spontaneously develop retrogressive neurological and cognitive deficits that resemble early-stage Alzheimer’s Disease (AD) in humans (Head ,2001). Dogs are highly tractable, widely accessible, and share a number of anatomical similarities with humans (Kaeberlein et al., 2016) but have shorter lifespans (Gilmore and Greer, 2015) and may thus provide a strong animal model for translational AD research (Araujo et al., 2017). Given that pet dogs share their environments with their owners, research on age-related cognitive deficits (ARCD) in pet dogs may also provide insight into the environmental factors that contribute to dementia in humans (Kaeberlein et al., 2016). Cognitive tests typically carried out on aging colony dogs require weeks or months of daily testing sessions and specialized laboratory equipment (e.g., Adams et al., 2000b). Thus, to facilitate veterinary care for aging dogs while expanding the utility of pet dogs as a model population for AD research, new cognitive tests must be designed which can be administered in a single session using only cheap and accessible materials (Chapagain et al., 2018).

In both humans and dogs, executive functions like inhibitory control and working memory are particularly sensitive to ARCD (Head, 2013). Behavioral inhibition requires selective attention towards task-relevant information and the suppression of irrelevant or conflicting behaviors (McDowd and Oseas-Kreger, 1991). Reversal learning tests evaluate an individual’s ability to inhibit prepotent responses to previously reinforced stimuli and to shift responses towards a previously unreinforced stimulus (Lai et al., 1995).

Previous studies on size- and object-reversal learning in colony beagles have found robust evidence that older dogs are more persistent in responding to previously rewarded objects following reversals of task contingencies, and exhibit stronger position biases (e.g., Chan et al., 2002; Tapp et al., 2003a). However, the relationship between age and inhibitory control in dogs may be task specific (Bray et al., 2014) and may differ between colony beagles and mixed-breed shelter-sourced dogs (Milgram et al., 1994).

Most studies on inhibitory control in pet dogs have observed deficits in older individuals (e.g., Mongillo et al., 2013; Wallis et al., 2014). However, only one study used a test which could be completed in a single session without specialized equipment. Piotti et al. (2018) assessed object and location reversal learning in pet dogs using a go-no-go paradigm. Dogs younger than eight years learned to approach the reinforced stimulus faster after each reversal than did older dogs. However, fewer than half of the older dogs learned the task within the cut-off of 100 trials. Thus, this task may be too difficult to rapidly assess reversal learning in older dogs.

In the present study, we designed a spatial serial reversal learning test to rapidly detect age-related cognitive decline in pet dogs. We examined whether age predicted performance after accounting for subject and test covariates. In addition, we developed an easily interpretable test score to assay cognitive impairment. In doing so, we also examined whether age-related deficits in inhibitory control were characterized by bouts of poor performance in addition to poor overall performance.

## Method

### Subjects

Eighty pet dogs (39 male) of various breeds and ages (10 - 173 months, mean = 80.64) were tested (Appendix D-1). To ensure accuracy of age reports, only dogs seen by a veterinarian before reaching one year of age were included in this study. All owners volunteered their dogs to participate.

### Materials and Procedure

The materials and layout were similar to Van Bourg et al. (2020). Additional information about the materials and layout are provided in Appendix D-2.1.

This serial reversal learning task required dogs to choose which of two identical boxes contained a hidden food treat. Between trials, E1 hid a treat behind (“baited”) one of the boxes while the dog waited with E2 in the holding room. In each trial, E2 released the dog into the testing room and the dog was allowed to search for a treat in one of the boxes. When the dog began to move away from the chosen box, E1 immediately ushered the dog back to the holding room. Thus, the dog was only allowed to retrieve the treat if it correctly chose to search the baited box.

The first box the dog oriented its head towards once the treat (or empty treat platform) was in view was scored as the dog’s choice (Appendix D-2.3). E1 was responsible for making this determination but E2 also watched the dog to provide confirmation or correction if needed. However, this was rarely the case as the dog almost always walked directly towards and brought its snout to within a few centimeters of the chosen box. In addition, all trials were verified from video recordings by a coder who was blind to the nature of the study.

E1 continued to bait the same box until the dog chose correctly in three consecutive trials. Each time the dog met this criterion, E1 repeated the procedure using the other box (a ‘reversal’). If the dog failed to choose a box or retrieve the treat within two-minutes, E2 recorded an incorrect choice (per Gunter, 2018; for justification, see

Udell et al., 2010). The session ended when this occurred four times or when the dog completed 30 minutes of testing.

## Analysis

All data were analyzed in R version 3.4.1. Generalized linear mixed models were constructed and tested using the package “lme4” (Bates et al., 2015b).

### Trial outcome.

To test whether age predicted performance, we conducted a binomial regression analysis of trial outcome (correct or incorrect) using a generalized linear mixed model with a binomial error distribution (logit-link function). Recent studies indicate that cognitive performance in pet dogs may peak in middle age rather than decline linearly from adolescence (e.g., Watowich et al., 2020). Thus, to test both the linear and the nonlinear (quadratic) relationships between age and trial outcome we included fixed effects for age (in months) and $age^2$. To control for subject covariates, we included fixed effects for weight (kg), height (cm) and sex.

To test for evidence of learning, we included a fixed effect for trial number (a cumulative count of trials from the start of testing). If the dog completed reversals by learning to search for treats at the correct box rather than by random chance, we should expect an increase in performance with trial number. However, after each reversal the probability of choosing the correct box should abruptly drop. Thus, we also included a fixed effect for reversal number to control for variation between reversals (i.e., to account for this oscillating relationship between trial number and trial outcome).

If the assessment required inhibitory control, previously learned reward contingencies should interfere with the dog’s ability to learn the current reward contingency. Such interference may be additive and thus the task may become more difficult with each additional reversal. Alternatively, dogs may learn to track reversing reward contingencies more rapidly with each additional completed reversal. The fixed effect for reversal number tested each of these hypotheses.

For random effects, we included only subject intercepts because subject slopes for trial and reversal number could not be reliably estimated (they created singularities and prevented convergence), did not improve model prediction, and did not account for any variance.

To test the significance of each predictor we conducted likelihood ratio tests of the difference in total prediction between the full model and the nested model without the predictor. To test overall model prediction, we compared the full model to an intercept-only model.

**Total reversals.**

To test whether the total number of reversals completed during the test provided a stand-alone score which could be used to determine the dog's level of cognitive function, we constructed a generalized linear model of total reversals as a function of age and age$^2$. Because total reversals could take only a handful of discrete values, including zero, we used a Poisson error distribution. This better fit the data than other error distributions for count variables (e.g., negative-binomial, zero-inflated Poisson, etc.). To test the goodness of fit of the overall model, we compared the residual deviance and residual degrees of freedom to the chi-squared distribution ($p$-values < .05 indicate data do *not* fit the model). To control for minor violations of the Poisson distribution assumptions, we calculated robust standard errors for the predictors. We then calculated *Z* and *p*-values using the robust standard errors to test the significance of the parameters.

**Longest streak of perseverative errors.**

To assess the severity of the dog's worst bout of performance, we measured the longest streak of perseverative errors committed during the test. Given that the dog was not informed when a reversal occurred, only incorrect choices in trials after the first of each reversal were considered perseverative errors. To test the relationship between age and bouts of poor performance, we tested the regression of age and age$^2$ on the longest streak of perseverative errors, which was normalized with a square root-transformation. We then conducted an F-test for the overall prediction of the model and two-tailed t tests for the prediction of the individual parameters. Satisfying the assumption of equal variances required removing a strong outlier. Importantly, this did not change the outcome of the analysis (see Appendix D-3).

**Age-weight interactions.**

Although we are unaware of any evidence that the rate of cognitive aging in dogs varies as a function of body size, lifespan is inversely related to body size in dogs and physiological deterioration may progress more rapidly in large breeds (Kraus et al., 2013). To control for potential effects of body size on lifespan changes in inhibitory control, we tested a large enough sample to ensure that age was not related to weight ($R^2 < 0.0001$). In addition, we repeated each analysis with additional coefficients for the interaction between weight and age, and the interaction between weight and age$^2$. These effects were not significant in any analysis and did not meaningfully change the outcome of any analysis (see Appendix D-4).

## Results

Shaping required an average of 20 trials and all dogs that completed shaping also completed the assessment (n = 62). Sixteen dogs failed to complete shaping. Additionally, one dog would not eat treats and one dog would not approach the left box.

Dogs made a choice in all but 30 of the 2878 trials suggesting that they were highly motivated to participate. Moreover, no-choice trials were approximately uniformly distributed among 16 dogs (mean = 1.88 trials) indicating that low motivation was not a major problem for any dog. Although dogs may become less active with age (Salvin et al., 2011b) and may lose interest in cognitive tests more quickly (Salvin et al., 2011a), age was not correlated with the number of 'no-choice' trials during our test ($R^2 = 0.001$). Thus, these shaping and exclusion criteria may effectively control for motivation.

### Trial Outcome

The model significantly predicted trial outcome, $X^2(8) = 29.70$, $p = .0002$ (Table 3-1). The effects of age, $X^2(1) = 4.88$, $p = .027$, and age$^2$ were significant, $X^2(1) = 6.76$, $p = .009$. Middle-aged dogs chose correctly more often than younger and older dogs. The probability of choosing the correct box increased with trial number, $X^2(1) = 7.66$, $p = .006$, and decreased with reversal number, $X^2(1) = 12.39$, $p = .0004$ (Appendix D-5). Random variation among subjects predicted trial outcome, $X^2(1) = 6.96$, $p = .008$.

The effects of height, weight, and sex were not significant. Therefore, to confirm the significance of the other predictors, we constructed a final model without these covariates. This did not change overall prediction and all remaining predictors were significant (Table 3-1).

The variance inflation factors of the coefficients for trial (7.0) and reversal number (7.1) suggested multicollinearity. Thus, the standard errors of these terms may be overestimated and the strength of their prediction, underestimated.

### Total Reversals

Although the goodness-of-fit test was not significant, $X^2(59) = 74.55$, p = .08, neither age, ß = 4.05e$^{-03}$, *SE* = 7.22e$^{-03}$; *p*(>|*Z*|) = .58, nor age$^2$, ß = -2.97e$^{-05}$, *SE* = 4.52e$^{-05}$; *p*(>|*Z*|) = .51, significantly predicted total reversals.

### Longest Streak of Perseverative Errors

The overall model significantly predicted the longest streak of perseverative errors, *F*(2,58) = 8.30, *p* = .0007; $r^2 = 0.22$. The effects of both age, ß = -0.013, *SE* = .004; *t*(58) = -3.20, *p* = .002, and age$^2$ were significant, ß = 8.99e$^{-05}$, *SE* = 2.41e$^{-05}$; *t*(58) = 3.73, *p* = .0004. Perseverative streaks ($\bar{y} = 3.49$, *SEy* = 0.19) were shorter in middle-aged dogs than in young and old dogs (Fig. 3-1).

## Discussion

In this sample of pet dogs, the ability to correctly respond to serially reversing reward contingencies peaked in middle-age and this finding could not be attributed to subject covariates. After accounting for reversal number, performance improved with trial number indicating that dogs completed reversals by learning the correct location rather than choosing at random. Dogs were less likely to choose the correct box after each reversal suggesting that previously learned reward contingencies increasingly interfered with the dog's ability to learn the current reward contingency. In turn, this indicates that completing reversals required inhibitory control. Together, these findings provide evidence that this test detected age-related deficits in inhibitory control.

Importantly, these deficits were also detected with a stand-alone score (the longest streak of perseverative errors) that can be easily measured and interpreted by clinicians and dog owners. In addition, this test required only two visually separated areas, two identical objects large enough to hide food treats, and 30 minutes for testing. Thus, this assessment may provide a viable clinical or in-home assessment for ARCD.

The positive association between age and cognitive flexibility and the negative association between age and perseveration in dogs between 10 and 72 months of age provided additional support for recent indications that executive functions continue to develop until middle-age in pet dogs. Watowich et al. (2020) observed a quadratic relationship between age and performance in a broad range of cognitive assessments implemented by citizen scientists. Experimental studies on pet dogs also align with these findings. Wallis et al. (2014) found that selective attention and sensorimotor coordination peaked in middle-age. Using the same two-box paradigm as the present study, Van Bourg et al. (in press) found that middle-aged dogs could recall the location of a hidden treat more accurately and after longer retention intervals. In addition, young and old dogs displayed stronger box preferences suggesting that the ability to inhibit incorrect responses to a preferred location does not fully develop until midlife.

Given that age and age$^2$ predicted the dog's longest streak of perseverative errors but not the number of reversals completed during the test, older and younger dogs may be prone to more severe bouts of perseveration but not deficits in average efficiency of serial reversal learning. More generally, ARCD in dogs may be better characterized by severe bouts of cognitive dysfunction rather than temporally consistent cognitive deficits. Worst performance and other measurements related to intra-individual variability are underutilized tools which can complement scores for overall performance and may improve prediction of cognitive decline (Hultsch and MacDonald, 2004). Indeed, studies on humans indicate that an individual's worst performance in a multi-trial psychometric test may serve as a useful indicator of ARCD (Wallert et al., 2017; 2018). To our knowledge, the present study is the first to examine the relationship between age and worst performance in a cognitive assessment for dogs. Thus, future studies are needed to confirm this hypothesis.

The finding that age did not predict total reversals completed suggests that this measurement was not a useful assessment score. Particularly in rapid assessments which must be completed in fewer trials, integer scores may be limited to a small number of potential outcomes. Although such measurements may be used to detect average performance differences between age groups, they provide little resolution and therefore, may be less useful for identifying ARCD in individual dogs. This further highlights the importance of analyzing age as a continuous variable rather than grouping dogs into age categories (see Van Bourg et al., 2020; Watowich et al., 2020).

The apparently counterintuitive decrease in performance across reversals displayed by dogs in the present may stem from methodological constraints. When animals are trained in many sessions on spatial serial reversal learning tasks, performance generally improves across reversals as subjects learn the "principal of reversal" or simple strategies like the Win-Stay and Lose-Shift rules (Shettleworth, 1998; 2010). In the present study, dogs were tested in only one short session. Thus, nearly one third of the dogs completed only one reversal and over half of the dogs were unable to complete a third reversal, which is usually the first reversal that subjects complete faster than the initial side-learning event (e.g., Warren, 1966).

# CHAPTER 4

# COGNITIVE FLEXIBILITY AND AGING IN COYOTES

Joshua Van Bourg[1], Julie K. Young[2], Raneem Alkhalifah[1], Stacey Brummer[2], Erik Johansson[3], Jacqueline Morton[4], Victoria Quintana[1], Clive D. L. Wynne[1]

[1] Department of Psychology, Arizona State University

[2] USDA-National Wildlife Research Center, Predator Research Facility, Millville, UT, USA

[3] Department of Biology, Hamilton College

[4] Department of Animal and Food Sciences, University of Delaware

Joshua Van Bourg https://orcid.org/0000-0003-0599-9602; joshvanbourg@asu.edu

Clive D. L. Wynne https://orcid.org/0000-0001-8408-1466; cwynne1@asu.edu

Julie K. Young https://orcid.org/0000-0003-4522-0157; julie.k.young@usda.gov

This research was supported by the Arizona State University Chapter of Sigma Xi and USDA-National Wildlife Research Center. We thank Riley Etcitty, Sylvain Sonntag, and Shelby Stanley for their assistance with data collection. We also thank the animal care staff at the USDA-National Wildlife Research Center, Predator Research Facility for their insight and assistance.

We have no known conflict of interest to disclose.

The findings and conclusions in this publication are those of the author(s) and should not be construed to represent any official USDA or U.S. Government determination or policy.

Correspondence concerning this article should be addressed to Joshua Van Bourg, Department of Psychology, Arizona State University. PO Box 871104, Tempe, AZ 85287-1104. Email: joshvanbourg@asu.edu

Keywords: canid, dog, reversal learning, plasticity, dementia

# Abstract

Cognitive flexibility evolves in species that live in complex and dynamic social systems and habitats, and may enable species to better cope with anthropogenic habitat modification. Aging may also impact the cognitive abilities of canids. Coyotes (*Canis latrans*) and domestic dogs (*Canis lupus familiaris*) differ markedly in their social and trophic ecology but have both been highly successful in adapting to human-modified ecosystems. Aging dogs develop a form of dementia that mirrors Alzheimer's disease in humans, but it is unknown whether similar cognitive deficits develop with age in coyotes and other wild canids. In this study, we modified a spatial serial reversal learning test that was sensitive to cognitive aging in pet dogs to test cognitive flexibility in captive coyotes. We also performed a second experiment using a color discrimination task to test for flexible rule learning. Nineteen of 20 coyotes demonstrated the ability to track shifts in spatial reward contingencies and learned to rapidly complete reversals by employing a win-stay, lose-shift strategy. In addition, coyotes inhibited prepotent win-stay choices in order to acquire the color discrimination task. These findings suggest that behavioral flexibility may help coyotes to detect and respond appropriately to both rapid fluctuations and gradual changes in ecological conditions. Performance did not differ between coyotes and previously tested dogs, but like dogs, behavioral flexibility declined with age in adult coyotes. Thus, cognitive decline and flexibility may be conserved among canines pending additional studies on other *Canis* species.

Behavioral flexibility allows an individual to respond appropriately to changes in environmental stimuli through the modification, reversal, innovation, or inhibition of behavior (see Bond et al., 2007 for discussion). Behavioral flexibility involves basic learning processes and complex executive functions (Reader and MacDonald, 2003) which are heritable and often selectively advantageous (Hendry, 2016; MacLean et al., 2014). Serial reversal learning tasks have been widely used to assess and compare behavioral flexibility (Shettleworth, 2010) and overall cognitive ability (e.g., Bitterman, 1965) in non-human animals. In these operant procedures, reward contingencies are abruptly reversed each time a subject acquires a given discrimination task, usually a two- or three-choice visual or spatial discrimination (Rayburn-Reeves et al., 2013). Reversal learning involves multiple components of behavioral flexibility including inhibitory control, a tendency to explore or probe alternative options, and probabilistic learning (Izquierdo et al., 2017). Subjects trained on repeated reversals may become able to generalize that reward contingencies can be reversed, which is sometimes referred to as the principle of reversal (Shettleworth, 1998). In deterministic two-choice serial reversal learning tasks, subjects may optimize performance by acquiring the win-stay, lose-shift (WSLS) strategy (e.g., Warren, 1966; for discussion, see Bessemer and Stollnitz, 1971), in which the subject repeats the response from the previous trial if that choice was rewarded (i.e., if *win*, then *stay*) but switches to the alternative response if that choice was not rewarded (i.e., if *lose*, then *shift*).

Reversal learning, inhibitory control, and other executive functions have been tested in domestic dogs (*C. lupus familiaris*) and results suggest they may be strongly affected by age-related neurodegenerative conditions which mirror the early stages of

Alzheimer’s Disease (AD) in humans (Head, 2001; 2013). Age-related cognitive deficits (ARCDs) in dogs and AD in humans are both associated with the development of neuritic beta-amyloid plaques (Head, 1998; Vite and Head, 2014) as well as increased anxiety, nocturnal activity, agitation, and disorientation (Madari et al., 2015; Landsberg et al., 2012; Fast et al., 2013).

ARCDs are not uncommon among non-human animals and have been found in phylogenetically diverse species (Youssef et al., 2016). However, ARCDs in dogs are unique in the degree to which the neuroanatomical progression of affected brain regions, underlying changes in neurophysiology, and relative sensitivity of different cognitive functions overlap with Alzheimer’s Disease in humans (Head, 2013). Thus, testing for ARCDs in other canine species may help to identify unique genetic correlates of ARCDs in dogs and, in turn, Alzheimer’s Disease in humans.

Among animal models of Alzheimer’s Disease, pet dogs are also unique in the extent to which they share their environments with humans. Dogs and humans may overlap heavily in their activity patterns, circadian rhythms, and exposure to water, air, noise, and light pollution, food preservatives, pathogens, and other anthropogenic materials (see Kaeberlein et al., 2016 for discussion). To a lesser degree, coyotes (*Canis latrans*) living in urban, suburban, and rural habitats are also exposed to the same environments as humans. However, coyotes often shift their activity and space use patterns to avoid overlapping with peak activity and space use patterns of humans (e.g., Gehrt et al., 2011; Smith et al., 2018). Thus, comparing age effects among coyote populations living in anthropogenic and natural environments (in which dogs cannot survive: Boitani et al., 2007) may provide insight into the role of the local environment in the development of ARCDs.

Coyotes typically survive for a maximum of 10 years in the wild although there have been reports of coyotes surviving for up to 18 years in the wild and 21 years in captivity (for discussion see Way and Strauss, 2004). In dogs, lifespan is inversely related to body size and senescence is generally more rapid in larger breeds (Kraus et al., 2013). Median life-expectancies in smaller breeds may exceed 13 years, while many larger breeds are expected to live half as long (see Watowich et al., 2020). However, age-related cognitive deficits and behavioral changes in dogs do not appear to depend upon size-predicted lifespan (Salvin et al., 2010; 2012; Watowich et al., 2020).

Wild species that exhibit behavioral flexibility should be better equipped to acquire patchy and unpredictable food resources, more likely to exploit new dietary resources, and more likely to innovate new feeding strategies (Bond et al., 2007; MacLean et al., 2014; Zuberbühler and Janmaat, 2010). Importantly, flexible species may be more successful at invading new habitats and exploiting new ecological niches; thus, behavioral flexibility may enhance the ability of a species to cope with anthropogenic habitat modification (Snell-Rood, 2013; Wright et al., 2010). Behavioral flexibility is also thought to evolve in social species, which must navigate diverse and dynamic relationships (Easton, 2005). In addition, flexibility may be particularly advantageous in species that cooperate in pair-bonding, rearing young, hunting, vigilance, or defense (for discussion and review, see Marshall-Pescini et al., 2015). However, selection does not always favor ecological generalization and flexibility. Specialization can be favored in relatively constant environment or arise through interspecific competition (Futuyma and

Moreno, 1988). Insights from foraging theory also suggest specialization on high value food or other resources can lead to broader ecological specialization (Futuyma and Moreno, 1988).

Coyotes and dogs are closely related canines that have been highly successful in adapting to anthropogenic habitats (i.e., rural, urban, and suburban) but differ substantially in their behavior and ecology (Jensen, 2007; Moehlman, 1989). Understanding behavioral flexibility among these and other canid species may provide insight into the relationships between social ecology, diet, cognitive flexibility, and species resilience.

Like free-ranging dogs, which still compose 70-80% of the global dog population (Lord et al., 2013), coyotes scavenge for human waste in anthropogenic habitats (coyotes: Morey et al., 2007; dogs: Sen Majumder et al., 2014). However, coyotes also forage for fruit, scavenge carrion, and effectively hunt small mammals, birds, reptiles, insects, cervids, domestic pets, and livestock (Hernández et al., 1994; Morey et al., 2007; Sacks and Neale, 2002). The composition of this diverse diet may further vary with seasonal changes in resource availability (Andelt et al., 1987; Lingle, 2000), pack size (Bowen, 1981), breeding status (Morey et al., 2007), age (Gese et al., 1996), region (reviewed in Morey et al., 2007), and habitat characteristics (Ward et al., 2018), including the degree of urbanization (Santana and Armstrong, 2017). Coyotes have greatly expanded their range since the late 1800s, from their native distribution in the arid regions of western North America to the entire continental US and much of Alaska, eastern Canada, and Central America (Hody et al., 2018). In doing so, they have invaded many novel habitats, some of which are characterized by daily or seasonal fluctuations in environmental conditions (e.g., alpine and boreal forests).

Coyotes typically form stable breeding pairs (Hennessy et al., 2012) which cooperate in rearing young and defending territory (for review, see Lord et al., 2013). Dyads and packs occasionally collaborate in hunting large prey (Bowen, 1981) and individuals have been observed hunting rodents in synchrony with badgers (Rathbun et al., 1980; Thornton et al., 2018). Social structure varies with food availability (Bowen, 1981; Macdonald and Carr, 1995), reproductive status (Kleiman and Brady, 1978; Sen Majumder et al., 2014), and season (Bowen, 1981; Sen Majumder et al., 2014). However, the social hierarchies of coyote family groups are stable and less complex than the hierarchies of multi-male, multi-female dog packs, which may be linear or nonlinear depending on sex, age, and type of dominance-controlled resource (Cafazzo et al., 2010). After dispersing, coyotes may become solitary residents or transients (i.e., may not form pair-bonds), particularly when small rodents are the primary food resource (Bekoff and Wells, 1980).

Based on diversity of foraging strategies and inhabited habitats, as well as relative tendencies to form pair bonds, hunt in groups, and collaborate in rearing young, coyotes should display more behavioral flexibility than domestic dogs. However, based on relative social complexity and stability, dogs should be more behavioral flexible than coyotes. Furthermore, selection for tractability and selection against fear and aggression in the domestication of dogs may have created direct and indirect selection for inhibitory control, although comparisons between dogs and wolves have not supported this hypothesis (Brucks et al., 2019; Marshall-Pescini et al., 2015).

Here, we designed a serial reversal learning test to assess cognitive aging and flexibility in captive coyotes of varying ages. We modified a test used on pet dogs and applied the same serial reversal learning criterion to a similar two-choice spatial paradigm so that we could directly compare results from coyotes to domestic dogs (Van Bourg et al., 2021). Although coyotes have demonstrated proficiency in other operant tests for flexible learning (e.g., Gilbert-Norton et al., 2009), to the best of our knowledge, serial reversal learning has only been successfully assessed in one coyote (Stanton et al., 2020). In a second experiment, we presented ten of the coyotes with a simple color discrimination task using the same two-choice paradigm. In this assessment, coyotes were required to inhibit prepotent serial reversal learning responses while learning to selectively attend to the different clothing colors of two experimenters. Thus, the acquisition of this task measured flexible rule-learning.

## General Method

This research was approved by Institutional Animal Care and Use Committees at Arizona State University (Protocol Number: 18-1650R) and USDA's National Wildlife Research Center (QA-2967).

### Subjects

We tested 21 coyotes (11 female, 10 male) ranging from 25 - 110 months of age (mean = 70.1, SD = 25.9) at the USDA's National Wildlife Research Center-Predator Research Facility in Millville, Utah. Two additional coyotes were fearful of the experimenters and did not pass the shaping criteria for inclusion in this study. Coyotes at the facility receive a daily 650-g food ration that is composed of diverse high fat and protein food items known locally as "mink food" because it is typically purchased by area mink breeders (Fur Breeders Agricultural Cooperative, Logan, UT). Coyotes are typically housed as mated male-female pairs in outdoor enclosures ranging from 0.1-1.0 hectares in size. For these experiments, coyotes were housed and tested individually in identical 0.1-hectare, tear-drop-shaped enclosures (Fig. 4-1). The first nine coyotes (Group 1) participated in a pilot study for Experiment II after completing Experiment I. Ten of the remaining twelve coyotes (Group 2) participated in Experiment II after completing Experiment I. One coyote from Group 2 stopped participating after completing Experiment I and the last coyote to begin Experiment I did not meet the requirements for Experiment II before the end of the field season (see Table 4-1). Information on the 62 domestic dogs ranging from 10 to 173 months of age and their testing conditions are in Van Bourg et al. (2021). The testing procedure described here for Experiment I is similar to that carried out on pet dogs, with modifications noted here that were used to accommodate coyote welfare and human health and safety considerations while working with coyotes.

### Materials

In each trial, a 75g semi-frozen cube of mink food (a “treat”) was hidden behind one of two identical 0.3 x 0.2 x 0.2m landscaping bricks set 7m to either side of the front entrance to the enclosure (Fig. 4-1). Metal tongs were used to insert these treats into the enclosure through chain-link fencing. Tall grass surrounding each treat blind was cut such that the treat was only visible at coyote height from within 1m of the brick. A bucket of treats hidden outside the enclosure next to each of the blinds remained in place for the duration of each session to prevent the coyote from using its sense of smell to determine whether a blind contained a treat.

## Acclimation and Shaping

Before testing, all coyotes completed a 5-10-day response-dependent acclimation and shaping procedure which involved successively introducing novel testing materials (e.g., the blinds, vehicle, and experimenters) and procedures (e.g., remaining near the enclosure after feeding the coyote, inserting food through the fence, and two people approaching the enclosure simultaneously).

## Testing Procedure

In each trial, two experimenters walked in synchrony to the separate blinds from an observation vehicle which was either parked equidistant to the blinds or at least 75m away from the enclosure. Before returning to the vehicle, both experimenters went through the motions of placing a treat behind their blind but only one did so. The coyote was then allowed to search for the treat behind one of the blinds. When the coyote chose the side with the treat, the experimenters allowed the coyote to finish eating before standing to begin the next trial. When the coyote chose the side without the treat, the experimenters waited for the coyote to move away from this incorrect blind before standing, which was sufficient to prompt most coyotes to move away from the blinds towards the center of the back of the pen. To achieve this reaction from bolder coyotes, the vehicle was parked closer to the pen. To accommodate the boldest coyotes, a recorder who otherwise remained in the observation vehicle stood at the central front-entrance of the enclosure. To avoid cueing, the recorder and experimenters wore dark sunglasses.

Video recordings of all sessions were captured by a camera mounted to the observation vehicle. The recorder's determination of the coyote's choice was confirmed by each experimenter in real time and later verified from the video. For both experiments, each testing session ended when the coyote retrieved its daily ration of mink food (mean = 8 cubes) or stopped participating (whichever came first). If the coyote stopped participating before retrieving at least 75% of its daily ration, additional mink food was left in the enclosure.

### *Analysis - GLMMs*

Analyses of task acquisition (test completion) speed are provided in Appendix E-1. For all other analyses, we examined trial-by-trial performance using generalized linear mixed models (GLMMs) of trial outcome (correct or incorrect choice) with binomial error distributions (logit-link functions).

For each analysis, we first fit a full model with fixed effects for the predictor of interest and all potential covariates, random intercepts for subjects, and all possible random subject slopes (see Barr et al., 2013; Schielzeth and Forstmeier, 2009). To address potential overfitting and underpowering of these maximal models (see Bates et al., 2015a; Matuschek et al., 2017), we then removed non-significant random subject slopes and non-significant fixed effects for covariates beginning with the effect which accounted for the least variation in trial outcome. We continued this stepwise elimination process until all remaining covariates added significant prediction to the final (reduced) model. Because random slopes did not add prediction to any model in any analysis, the random-effects structure of each reduced model included only subject intercepts.

To test the significance of individual effects in models without interactions, we conducted Likelihood Ratio Tests of nested models (i.e., compared models with and without each predictor). In models with interactions, we calculated profiled confidence intervals of fixed-effect estimates and profiled confidence intervals of random-effect variances using parametric bootstraps.

To test for homogeneity of variance and linearity, we inspected boxplots and scatterplots of fitted values and residuals as a function of each predictor. To test for normality, we examined density and quantile plots of fitted values and each random effect. To test for overdispersion, we compared the sum of the squared Pearson residuals to the degrees of freedom of the residuals. All data were analyzed in R version 3.4.1. GLMMs were constructed and tested using the package "lme4" (Bates et al., 2015b).

## Experiment I

### Method

In the serial reversal learning test, the experimenters wore matching camouflage clothing and hats so that the coyote could not easily discriminate between experimenters using visual cues. The same side (blind) was baited with treats until the coyote chose this side in three consecutive trials. In subsequent trials, the other blind was baited with treats until the coyote again chose the correct side in three consecutive trials. These side reversals of S+ and S- were repeated each time this criterion was met. The pairing of the experimenter, side, and treat was counterbalanced and pseudorandomized such that the coyote did not retrieve the treat from the same experimenter in more than three consecutive trials.

#### Scoring Choices.

To consistently retrieve treats, the coyote was required to attend to feedback from recent choices (i.e., to recall whether the previously visited side contained the treat). Thus, in the first trial following a reversal, we scored a correct choice when the coyote visited the side that did *not* contain the treat because this was the correct win-stay response (i.e., the coyote retrieved a treat at this blind in this previous trial). In all other trials, we scored a correct choice when the coyote visited the side that *did* contain the treat because this was either the correct win-stay response (i.e., the coyote retrieved a treat at this blind in the previous trial) or the correct lose-shift response (i.e., the coyote did not retrieve a treat at the other blind in the previous trial).

### Test Completion.

In Experiment I, the point at which the coyote completed the test depended on its performance. Five criteria were used to determine when the coyote’s test ended. These criteria aimed to standardize experience and proficiency in preparation for Experiment II. At a minimum, we attempted to test each coyote until it completed at least 100 trials and ten total reversals. In addition, we attempted to test each coyote until it chose the correct side nine times in ten consecutive trials ($p = .02$). We called this third criterion the WSLS criterion. Furthermore, we attempted to test each coyote until it completed multiple reversals in a single session faster than was likely to occur by random chance. The coyote could accomplish this by completing two reversals within the first nine trials of a session ($p = .05$) or three reversals within the first 17 trials of a session ($p = .01$). We called this fourth criterion the Session Reversals criterion.

Each coyote that participated in Experiment II was also required to complete the Session Reversals criterion in its final session, regardless of whether it fulfilled the Session Reversals criterion in a previous session. We called this fifth criterion the Final Session criterion.

## Results

### Acquisition of the serial reversal learning task.

All coyotes completed ten reversals. Nineteen coyotes satisfied the WSLS and Session Reversals criteria, one coyote failed to meet these criteria within 20 sessions, and one coyote's test was terminated prematurely due to a scoring error which was later identified from video playback (Table 4-1). Age significantly predicted the number of trials to complete ten reversals, trials to meet the WSLS criterion, and sessions to meet the session reversals criterion (see Appendix E-1).

### Performance in the serial reversal learning task.

To assess whether age affected the coyote's ability to track the serial reversal reward contingency, we tested the regression of age in months on trial outcome in Experiment one using a binomial GLMM (see Table 4-2a). Because the coyote did not know which side was baited in the first trial of a session, we omitted these trials from the analysis. To test for sex differences and learning effects, we included fixed effects for sex and test trial number (cumulative across all sessions). To assess whether reversing the rewarded side affected performance, we included a fixed-effect dummy code which indicated whether a reversal had occurred in the current session. In the random effects structure of this model, we included subject slopes for this dummy code, subject slopes for test trials, and subject intercepts.

The probability of choosing correctly decreased with age, $X^2(1) = 9.02$, $p = .003$, and increased with test trial, $X^2(1) = 12.80$, $p < .001$ (Fig. 4-2a). Coyotes were less likely to choose the correct location after the first reversal of the session, $X^2(1) = 17.55$, $p < .0001$. No other effects were significant.

After stepwise elimination of non-significant predictors, the final (reduced) model included fixed effects for age, test trial, and whether a reversal had occurred in the current session. Random subject intercepts were retained to account for data dependency. The significance and approximate estimates of all predictors in this reduced model were nearly unchanged relative to the full model (Table 4-2a).

To test whether age significantly predicted both behavioral inhibition and initial learning speed, we analyzed trials before and after the first reversal of the session separately using identical GLMMs of the trial outcome. In each model, we included random subject intercepts and fixed effects for age (in months) and test trial number. Both models are summarized in Table 4-2b.

The probability of choosing the correct side tended to decrease with age before the first reversal of the session but this effect was not significant, $X^2(1) = 3.57$, $p = .06$. However, performance significantly decreased with age after the first reversal of the session, $X^2(1) = 6.78$, $p = .003$ (Fig. 4-2b). Performance improved with test trial both before, $X^2(1) = 7.96$, $p = .005$, and after, $X^2(1) = 5.60$, $p = .02$, the first reversal of the session (Table 4-2b). Coyotes did not significantly differ in their initial propensities to choose correctly either before or after the first reversal of the session (i.e., random subject intercepts were not significant).

### Comparison to Dogs.

In our previous study (Van Bourg et al., 2021), pet dogs completed as few as 30 trials in a 30-minute, single-session test. Thus, to compare behavioral flexibility between coyotes and dogs, we conducted a binomial regression analysis of trial outcomes in the first 30 trials as a function of species (see Table 4-3). To further control for procedural differences, we included fixed effects for session number, trial number within the session, and reversal number (cumulative across all sessions). Further, we omitted the first trial of each session and excluded dogs that did not fall within the age range of the coyotes (30 – 115 months; *n* = 37). To test for learning, we included a fixed effect for the test trial number (cumulative across all sessions). To control for subject covariates, we included fixed effects for sex and age (in months). For random effects, we included subject slopes for the session, session trial, test trial, and reversal number as well as subject intercepts.

The probability of choosing correctly did not differ between coyotes ($\bar{x}$ = 0.58, *SE* = 0.02) and dogs ($\bar{x}$ = 0.58, *SE* = 0.01) but decreased with age, $X^2(1) = 6.43$, $p = .01$, and reversal number, $X^2(1) = 10.07$, $p = .002$ (Table 4-3). No other effects were significant.

After eliminating non-significant covariates (predictors other than species), the reduced model included fixed effects for species, age, test trial, and reversal number (Table 4-3). As in the full model, the effect of species was not significant. Again, the probability of choosing correctly decreased with age, $X^2(1) = 5.97$, $p = .01$, and reversal number, $X^2(1) = 7.84$, $p = .005$. However, the probability of choosing correctly increased with the test trial number, $X^2(1) = 4.05$, $p = .04$.

## Experiment II

### Method

In the second experiment, one experimenter wore beige clothing (cap and coveralls), one experimenter wore black clothing, and these clothing colors served as the S+ and S- for a visuospatial discrimination task. The pairing of S+ and experimenter did not change during an individual's test, but different coyotes were assigned different pairings. For example, Experimenter 1 always wore black and always provided a treat for some coyotes, and Experimenter 1 always wore beige and never provided a treat for other coyotes. These pairings were counterbalanced among coyotes and pseudorandomly assigned. This reduced the likelihood that coyotes could use the odors of the experimenters as cues because experimenters frequently swapped clothing between tests and visited both locations many times in each short testing session. Furthermore, we did not expect coyotes to attend to the odors of the experimenters given the salience of the smell controls and the fact that coyotes are primarily visual, rather than olfactory hunters (Wells and Lehner, 1978).

When the coyote incorrectly chose the side (blind) without the treat, the sides visited by the experimenters were repeated in the next trial. When the coyote correctly chose the side with the treat, the sides visited by the experimenters in the next trial followed a pseudorandomized (counterbalanced) order which prevented the coyote from retrieving treats from the same side in more than three consecutive trials.

#### Scoring Choices.

To consistently retrieve treats in the second experiment, the coyote was required to attend to the experimenters' clothing and the locations visited by the experimenters in the current trial. Searching the side baited by the S+ experimenter was always scored as a correct choice and searching the side visited by the S- experimenter was always scored as an incorrect choice.

#### Test Completion.

We aimed to test each coyote until it chose correctly in 17 of the last 20 trials (binomial test: $p = .002$). Because of the rule that treats could not be obtained from the same side in more than three consecutive trials, the coyote could only satisfy this criterion by switching sides after at least 7 of these 17 correct choices. Thus, the coyote could not complete this color discrimination task by using the same win-stay, lose-shift strategy it used in the serial reversal-learning task (Fisher's exact test of the difference between 10 of 17 and 17 of 17 win-stay choices: *p* = .007).

## Results

### Acquisition of the Color Discrimination Task.

Nine coyotes satisfied the test completion criterion of choosing the correct side in 85% of the last 20 trials, and 48% of these choices did not follow the WSLS strategy learned in Experiment I. Trials to criterion (mean = 138) decreased with age among these coyotes (see Appendix E-1). The tenth coyote only completed 107 trials before the end of the field season. This coyote chose correctly in 21 of 25 trials (Binomial Test; *p* = .0009) and switched sides after 45% of these correct choices during this period (Fisher exact test; *p* = .002).

### Performance in the Color Discrimination Task.

To test whether age affected the coyote's ability to learn the color discrimination task, we conducted a binomial regression analysis of the trial outcome in Experiment II. In this GLMM we included fixed effects for age (in months), sex, test trial number, and the interaction between age and test trial. We also included random subject intercepts and random subject slopes for the test trial (see Table 4-4).

The probability of correctly choosing the blind visited by the S+ experimenter decreased significantly with age, ß = -0.16, 95% CI: [-0.29, -0.03], and increased significantly with test trial number, ß = 0.27, 95% CI: [-0.12, 0.42] (Fig. 4-3). The rate of improvement in performance across test trials (i.e., learning speed) tended to decrease with age, ß = -0.14, 95% CI: [-0.29, 0.01], but this effect was not significant. No other effects were significant.

Removing the non-significant fixed effect for sex and the random subject slopes for test trial did not change the estimates of the remaining effects (Table 4-4). However, the interaction between age and test trial was significant in this reduced model, ß = -0.14, 95% CI: [-0.30, -0.01] indicating that older coyotes acquired the color discrimination task more slowly (Table 4-4).

## Discussion

Coyotes demonstrated behavioral inhibition and flexibility by rapidly tracking changes in spatial rules and shifting to a new foraging strategy when these rules became obsolete. These cognitive abilities may help coyotes to invade novel habitats (Hody et al., 2018), thrive in dynamic environments with patchy resources (see discussion in Gilbert-Norton et al., 2009), and facultatively live in cohesive family groups (Moehlman, 1989). In the first 30 trials of the serial reversal learning assessment, average performance and age-related deficits did not differ between coyotes and dogs. Thus, cognitive decline and flexibility may be conserved among canines.

### Cognitive Flexibility & Discrimination Learning in Coyotes

#### Experiment I.

Of the 20 coyotes that were given sufficient time to complete the first experiment, all but the second oldest learned to track shifts in spatial reward contingencies flexibly and efficiently. Coyotes learned to search for food at a productive location, to inhibit this response when the location became unproductive, and to instead search for food at a different location. By the end of the first assessment, coyotes accurately recruited the

WSLS strategy and required few (1-4) unreinforced choices to learn to inhibit a prepotent response.

This flexibility may help coyotes to efficiently track changes in local resource availability when foraging or scavenging (see Gilbert-Norton et al., 2009) and to rapidly switch between targets when hunting. It has also been suggested that reversal learning assessments approximate the demands of effectively navigating dynamic interactions between group members, which requires rapidly switching between responses to a fixed set of stimuli (Bond et al., 2007). Ultimately, the cognitive flexibility demonstrated by coyotes in this serial reversal learning test may be a product of their flexible dietary and social ecology.

The finding that coyotes, like dogs in a prior study (Van Bourg et al., 2021), were more likely to choose the correct side as trial number increased indicates that they learned and became more proficient at following the WSLS rules. Given that WSLS performance began to improve within just the first 30 trials in coyotes and dogs, both species may be predisposed to detect and respond to fluctuating conditions. Indeed, all coyotes completed reversals and displayed tendencies to probe for alternative feeding locations (i.e., committed win-shift errors) in the earliest sessions of the test. Thus, coyotes displayed baseline flexibility and sensitivity to serially reversing spatial reward contingencies as well as an ability to learn and effectively employ a flexible foraging strategy.

The finding that coyotes were more likely to choose the incorrect side after reward contingencies were reversed supports the hypothesis that this serial reversal learning task assessed behavioral inhibition. Moreover, the coyote’s ability to search for food at the new location was constrained by the ability to inhibit the prepotent response of searching for food at the previously rewarded location. Thus, the finding that performance improved with trial number in the model of trials after the first reversal of the session indicates that coyotes learned to inhibit these prepotent responses. Likewise, the finding that performance improved with trial number in the model of trials before the first reversal of the session indicates that coyotes became more proficient at the underlying spatial discrimination (associative learning) task as they gained more experience.

The high test-completion rate achieved in the present study (21 of 23 coyotes) suggests that this paradigm may be particularly useful for testing neophobic species in captivity. Previous studies have found that coyotes are fearful of large, novel objects (e.g., Mettler and Shivik, 2007; Windberg, 2008). In line with these findings, Stanton et al. (2020) found that coyotes (unlike skunks and raccoons) were generally unwilling to participate in a serial reversal learning test which used an automated device resembling an operant conditioning chamber. Despite strong incentives, at least 40 days to habituate to the apparatus, and at least 25 attempted testing sessions, only one of six coyotes participated in those trials. The apparatus was used as a steppingstone for studies of wild mesocarnivores but the authors concluded it may not be appropriate for testing wild coyotes. Thus, paradigms which use naturalistic materials rather than a mechanical apparatus may help to facilitate testing of neophobic species.

These findings add to a large body of research on serial reversal learning in diverse species ranging from honeybees (e.g., Strang and Sherry, 2014) and octopi (e.g., Mackintosh and Mackintosh, 1964) to birds and fish (e.g., Gonzalez et al., 1967;

Mackintosh, 1969). Early studies which used serial reversal learning tests to compare "intelligence" between species suggested that improvement in performance across reversals depended on attentional processes and the ability to overcome proactive interference (e.g., Mackintosh, 1968; 1974; Behrend et al., 1970; but see Squier, 1969). Our study findings align with more recent studies which suggest that performance improvement reflects the acquisition of abstract learning strategies (i.e., the ability to 'learn to learn'), including the WSLS strategy (Rumbaugh et al., 1996; Shettleworth, 1998; 2010).

### Experiment II.

Coyotes also displayed flexibility by learning the color discrimination task after the serial reversal learning task. At the beginning of Experiment II, the experimenters' new clothing provided the only potential cue that reinforcement contingencies might have changed. Thus, the significant improvement in performance with trial number and the high level of accuracy achieved in the second experiment indicate that coyotes learned to inhibit prepotent win-stay choices and serially reversing response patterns.

The ability to shift from focusing on the productivity of different feeding locations to focusing on visual differences between the experimenters may stem from the flexible feeding strategies and diverse diets of coyotes. Whereas the first experiment required that coyotes attend to a single type of information, the second experiment required that coyotes flexibly attend to and integrate information from multiple sources. Depending on habitat characteristics (Ward et al., 2018), seasonal resource availability (Andelt et al., 1987; Lingle, 2000), pack size (Bowen, 1981), breeding status (Morey et al., 2007), age (Gese et al., 1996), and region (Morey et al., 2007), coyotes switch between foraging for fruit, scavenging carrion, hunting alone, and hunting cooperatively (Hernández et al., 1994; Morey et al., 2007; Sacks and Neale, 2002). Thus, attentional flexibility may help coyotes to effectively switch between feeding modalities.

We primarily retained the lose-shift reward contingency in Experiment II to combat potential frustration effects and non-participation which may result from overtraining the serial reversal learning task (see Amsel, 1962). However, this may have also increased the difficulty of inhibiting win-stay responses if the coyote formed an association between the win-stay rule and the lose-shift rule during Experiment I (i.e., learned a compound WSLS rule). Ultimately, coyotes learned the color discrimination even though the previously learned WSLS strategy could be used to acquire the treat more readily than by choosing at random. This finding aligns with previous indications that coyotes detect and respond appropriately to subtle changes in reinforcement ratios. Specifically, Gilbert-Norton et al. (2009) found that coyotes rapidly matched their relative rate of foraging time to the relative rate of reinforcement in an operant two-choice test with concurrent variable-interval schedules of reinforcement.

The continuity between experiments and the above-chance rate of reinforcement predicted by the continuation of the WSLS strategy in Experiment II may emulate gradual changes in ecological conditions during which an existing suboptimal strategy for acquiring resources continues to be marginally productive. For example, a coyote hunting rodents in summer may alternate between entrances to a burrow in search of signs that a rodent may soon exit. If each sign (e.g., sound of nearby underground movement) is considered a reinforcing event, an optimal hunting strategy may boil down to a WSLS strategy. However, in late fall, only recently or heavily used burrows may remain clear of snowpack while sounds of movement may become increasingly difficult to detect. Thus, the WSLS strategy may continue to be somewhat productive but learning to discriminate signs of heavy use (and to focus hunting efforts at heavily used burrow entrances) may become the optimal rodent hunting strategy.

The second assessment also tested the coyote's ability to form an association between treats and a particular clothing color. However, unlike the first experiment, here we could not assess the independent contributions of associative learning and inhibitory control to performance. To this end, future studies may present this color discrimination task to coyotes that are naïve to this paradigm.

## Cognitive Aging

Both experiments indicated that, as in dogs, coyotes develop age-related deficits in cognitive flexibility. Older coyotes required more trials to complete ten reversals, learn the WSLS rules, and acquire the color discrimination task, as well as more sessions to complete a reversal faster than could be attributed to chance (Appendix E-1). In fact, only the second oldest coyote failed to satisfy the WSLS and Session Reversals criteria. Most importantly, the probability of choosing the correct side decreased with age in both tests and these findings could not be attributed to the subject or test covariates.

These findings suggest that heritable traits which predispose aging individuals to cognitive deficits may be conserved among canines. However, ARCDs may result from several underlying neurological changes (Youssef et al., 2016), which may differ between coyotes and dogs. As with Alzheimer's Disease in humans, ARCDs in dogs are associated with the accumulation of β amyloid proteins in diffuse plaques and cerebral vessels (cerebral amyloid angiopathy – CAA), as well as mitochondrial dysfunction, oxidative damage, and reduced neurogenesis (for review, see Head, 2013). Colony beagles with age-related reversal learning deficits show significantly higher amounts of β amyloid plaque deposition in the prefrontal cortex (Cummings et al., 1996; Head et al., 1998). Similarly, humans with AD-associated orbitofrontal degeneration display deficits in visuospatial reversal learning (Freedman and Oscar-Berman, 1989). The distribution of CAA is also similar in aging dogs and humans (Head et al., 2013).

Assays for these specific neurological changes in coyotes and other closely related species are needed to assess whether the pathology of ARCDs in dogs is unique among canines. More broadly, interspecific comparisons of ARCDs and associated neuropathology in canines may help to distinguish between the environmental and genetic factors which lead to ARCDs in dogs and humans.

The finding that older coyotes were less likely to inhibit prepotent responses after reward contingencies were reversed in the serial reversal learning test supports the hypothesis that aging coyotes develop deficits in inhibitory control. However, age did not significantly predict the trial outcome before the first reversal of the session. Thus, the ability to form simple spatial associations may not decline with age in coyotes. In line with these findings, aging colony beagles display performance deficits in reversal learning tests (e.g., Tapp et al., 2003a) but not simple visual discrimination tasks (e.g., Milgram et al., 1994).

Recent studies on pet dogs suggest that cognitive performance in canids may peak in middle-age (e.g., Van Bourg et al., 2020, Watowich et al., 2020). However, many of the dogs tested in these studies were substantially older and younger than the coyotes tested in the present study. Thus, the linear relationship between age and performance found in the present study does not preclude the possibility of an underlying curvilinear relationship.

In both experiments, the most severe cognitive deficits were displayed by coyotes older than eight years of age (Figs. 4-2b, 4-3). This may help to explain why coyotes rarely survive for more than nine years in the wild (e.g., Nelson and Lloyd, 2005; Young et al., 2006). More broadly, these findings may support previous suggestions that cognitive decline contributes to mortality in non-human animals (Ward et al., 2016). Although many intrinsic senescence processes and extrinsic factors contribute to age-related mortality in wild animals (Hämäläinen et al., 2014), intact behavioral inhibition and flexibility may be particularly important to the survival of coyotes given their dynamic social and dietary ecology. Additional cognitive flexibility assessments for both captive and wild coyotes, particularly individuals older than nine years of age, may help to clarify the role of cognitive decline on coyote mortality.

Given that vision tends to decline with age in diverse species, it is possible that visual impairments in older coyotes may have hindered their performance in the color discrimination task. However, we suggest this is unlikely given that an inability to discriminate between light, tan-colored clothing and dark, black-colored clothing would equate to near blindness. None of the coyotes displayed behavior or showed physical symptoms consistent with severe visual impairments. Furthermore, age-related visual impairments cannot explain the findings of the first experiment given that it was an egocentric spatial task.

## Coyote-Dog Comparisons

The finding that performance in the first 30 trials of this serial reversal learning task did not differ between coyotes and dogs suggests that the ability to learn rapidly alternating spatial reward contingencies may not differ between these closely related canids. These findings failed to support dietary, environmental, and cooperation hypotheses which predict that coyotes should be more behaviorally flexible than dogs, as well as social complexity and domestication hypotheses which predict that dogs should be more flexible than coyotes. However, the lack of support for these hypotheses should not be interpreted as direct counter evidence. For example, similar serial reversal learning performance in dogs and coyotes may be attributed to convergent evolution if the more complex and dynamic social ecology of dogs and the more flexible dietary ecology of coyotes were equally effective in driving evolution towards greater behavioral flexibility. Alternatively, genes that allow for greater flexibility that evolved in a common canine ancestor may be conserved in both coyotes and dogs, regardless of whether these genes became established via genetic drift or indirect selection for social or dietary traits. We suggest that the lack of differences in the performance of these species should be interpreted as evidence that cognitive decline and flexibility are conserved among canines.

Testing additional canine species with this procedure may help to clarify these findings. For example, wolves live in relatively stable packs (Mech and Boitani, 2010) which can be larger and more complex than coyote family groups but not as large or complex as packs of free-roaming dogs (based on Bowen, 1981; Lehman et al., 1992). In addition, wolves cooperate and collaborate more (Marshall-Pescini et al., 2017b), but display less dietary flexibility (Gable et al., 2018) than either coyotes or dogs. Comparing flexibility among these and other canids may provide greater resolution (and therefore, more power to detect differences) by allowing for analyses of social and dietary complexity as continuous rather than binary traits (see MacLean et al., 2014).

Differences in the experimental procedures used to test coyotes and dogs may have obscured underlying species differences in behavioral flexibility (see Appendix E-3). For example, some coyotes required multiple sessions to complete 30 trials, but all dogs were tested in a single session. However, Izquierdo et al. (2017) argued that discrepancies among reversal learning procedures are largely mitigated by the comparability and convergence of different methods. Nonetheless, testing dogs with the same multi-session procedure used to test coyotes may yield new insights. Indeed, coyotes and dogs may perform similarly in the first 30 trials of the serial reversal learning test yet differ in the speed with which they acquire this task across many sessions.

Presenting the color discrimination task to dogs after this extended serial reversal learning test may also help to control for potential task-dependent findings. To this point, serial reversal learning was not correlated with flexible color discrimination learning in coyotes, although caution should be taken in interpreting this finding given the small sample size of this analysis ($n = 9$). In line with this result, Brucks et al. (2017) found that measurements for inhibitory control in pet dogs which were acquired using four different tasks were not correlated. Thus, inhibitory control and flexibility may be multi-faceted cognitive functions which cannot be captured by any single paradigm or assessment. Ultimately, additional tests for cognitive aging and flexibility are needed to better understand how the phenomena differ among canines.

## Conclusion

Coyotes demonstrated the ability to track shifts in spatial reward contingencies and learned to rapidly complete reversals by employing a win-stay, lose-shift strategy. Coyotes also demonstrated flexible rule learning by inhibiting prepotent win-stay choices and thus succeeded at a color discrimination task. These findings suggest that behavioral flexibility may help coyotes to detect and respond appropriately to both rapid fluctuations and gradual changes in ecological conditions.

As in dogs, behavioral flexibility declined with age in adult coyotes. Thus, future studies which examine the pathology and genetics underlying these and other age-related cognitive deficits in coyotes may provide insight into the role of the shared dog-human environment in the development of ARCDs in pet dogs and humans.

Performance in the early stage of the first experiment did not differ between coyotes and dogs (Table 4-3). This finding failed to support, but did not conflict with, hypotheses predicting that various forms of social and dietary complexity drive the evolution of behavioral flexibility and cognitive complexity. Deploying identical procedures to assess behavioral flexibility in additional canid species may help to clarify these findings.

# CHAPTER 5

# COGNITIVE FLEXIBILITY AND AGING IN GRAY WOLVES, RED FOXES, AND GRAY FOXES

## Abstract

Cognitive aging in domestic dogs (*Canis lupus familiaris*) mirrors the early stages of Alzheimer's Disease in humans but it is unclear whether other canids develop a similar form of dementia. In recent studies, we demonstrated that dogs and captive coyotes (*Canis latrans*) develop similar age-related deficits in cognitive flexibility. In the present study, we used the same serial reversal learning assessment to test for such deficits in captive wolves (*Canis lupus lupus*). In addition, we piloted the use of this assessment for red (*Vulpes vulpes*) and gray foxes (*Urocyon cinereoargenteus*). Furthermore, we compared cognitive flexibility among canines to test hypotheses about which ecological traits drive the evolution of behavioral flexibility. Wolves demonstrated less flexibility than did coyotes and dogs, supporting the hypothesis that behavioral flexibility helps species to adapt to human-modified habitats and the hypothesis that reliance on patchy food resources selects for greater behavioral flexibility. Relative to young and old wolves, middle-aged wolves were better able to inhibit unproductive behaviors and to flexibly switch between responses. Thus, age-related deficits in cognitive flexibility may be a conserved trait among canines. Additionally, executive functions like reversal learning and inhibitory control may continue to develop after adolescence in wolves as in dogs. Finally, this assessment required little modification for successful deployment with smaller-bodied red and gray foxes. Thus, this test may be useful for assessing cognitive deficits in a variety of species.

Dogs (*Canis lupus familiaris*) develop behavioral changes and age-related cognitive deficits (ARCDs) which mirror the early stages of Alzheimer's Disease (AD) in humans (behavioral CDS cite, Head, 2001). Although ARCDs have been found in diverse species (Youssef et al., 2016), ARCDs in dogs are unique in the degree to which the neuroanatomical progression of affected brain regions, underlying changes in neurophysiology, and relative sensitivity of different cognitive functions overlap with AD in humans (Head, 2013). Thus, testing for ARCDs in wolves, the closest relative of dogs, may help to identify unique genetic correlates of ARCDs in dogs and in turn, AD in humans.

Both in humans and in dogs, reversal learning, short-term memory, inhibitory control, and other executive functions are more susceptible to age-related decline than other cognitive processes (Head, 2013). Behavioral flexibility — the modification, reversal, innovation, or inhibition of behavior (see Bond et al., 2007 for discussion) — involves many of these age-sensitive cognitive functions. Thus, tests for behavioral flexibility may be particularly sensitive to ARCDs.

Life expectancies in dog breeds range from as short as six years in larger breeds to more than 13 years in smaller dogs (Kraus et al., 2013; Watowich et al., 2020). However, age-related cognitive deficits and behavioral changes in dogs do not appear to depend upon size (Salvin et al., 2012; Watowich et al., 2020). Gray wolves (*C. lupus*), the ancestral species of all dogs (Vilà and Jennifer, 2012), do not typically survive for more than nine years in the wild but some individuals have reached 13-15 years of age (Holyan et al., 2005; Mech, 1988; Theberge and Theberge, 1988). In captivity, wolves commonly reach nine years of age and some individuals have lived over 17 years (Holyan et al., 2005). Coyotes (*C. latrans*), the next closest living relative to wolves and dogs (Vilà and Jennifer, 2012), rarely survive over ten years in the wild although there have been reports of coyotes surviving for up to 18 years in the wild and 21 years in captivity (for discussion see Way and Strauss, 2004).

Serial reversal learning tasks have been widely used to assess and compare behavioral flexibility (Shettleworth, 2009) and intelligence (e.g., Bitterman, 1965) in non-human animals. In these two- or three-choice visual or spatial discrimination tasks, reward contingencies are abruptly switched each time the subject learns to respond to the rewarded stimulus (Rayburn-Reeves et al., 2013). Reversal learning involves multiple aspects of behavioral flexibility including inhibitory control, exploration, and probabilistic learning (Izquierdo et al., 2017). In two-choice serial reversal learning tasks, subjects can maximize rewards by repeating a response that was rewarded in the previous trial (i.e., if *win*, then *stay*), but switching to the other response if their previous choice was not rewarded (i.e., if *lose*, then *shift*: for discussion, see Bessemer & Stollnitz, 1971). These are known as the win-stay and lose-shift (WSLS) rules (Warren, 1966).

Behavioral flexibility allows animals to respond appropriately to changes in environmental stimuli and can therefore be selectively advantageous under a variety of ecological conditions (Hendry, 2016; MacLean et al., 2014). Species which forage for patchy and fluctuating ecological resources (Day et al., 1999) and live in dynamic environments (Davey, 1989) should be more flexible than species which live in stable environments with stable resources (Jones, 2005). In addition, social species should exhibit greater behavioral flexibility than solitary species (Easton, 2005). Moreover, balancing individual and group needs through negotiations between individual actors or groups of actors requires constant modification to, shifts in, and inhibition of behavior (Byrne and Whiten, 1988; Easton, 2005). Similarly, cooperating in rearing young, hunting, vigilance, and defense require rapid and reversible behavioral adjustments to account for the actions of others (Marshall-Pescini et al., 2015). Thus, behavioral flexibility should evolve in species characterized by large social groups (Dunbar, 1998) with complex hierarchical structure (Amici et al., 2009), fission–fusion group dynamics (Amici et al., 2008), and extensive cooperation (Hager, 2005; Marshall-Pescini et al., 2015).

Flexible species and individuals should be more likely to innovate new feeding strategies, exploit new dietary resources, and invade new habitats and ecological niches (check Bond et al., 2007; MacLean et al., 2014; Zuberbühler and Janmaat, 2010). As a result, behavioral flexibility may enhance the ability of a species to cope with anthropogenic habitat modification (Snell-Rood, 2013; Wright et al., 2010). Thus, testing these hypotheses by examining the relationships between various ecological traits and behavioral flexibility may ultimately help to identify populations at risk of decline and extirpation, and species at risk of extinction. Furthermore, comparing behavioral flexibility among closely related species may help to control for evolutionary divergence in cognition which occurred in ancestral species rather than in response to ecological differences among extant focal species.

Canid species differ in their social, dietary, and habitat ecology (Jensen, 2007; Moehlman, 1989) and have had markedly different levels of success in adapting to human-modified ecosystems. Thus, canids are an ideal clade for testing competing hypotheses about the evolutionary drivers of behavioral flexibility.

Coyotes and dogs have been highly successful in adapting to all types of anthropogenic habitats (Hill et al., 1987; Kuijper et al., 2019). Gray wolves were once the most widely distributed terrestrial mammal but were eradicated from much of their range by the early 20th century (Young and Goldman, 1944). However, over the last 50 years their ranges in North America and Europe have been expanding (Mech, 2017) and they are recolonizing some agricultural landscapes (Chapron et al., 2014; Kuijper et al., 2016). Nonetheless, wolves tend to avoid areas with greater human activity (Massolo and Meriggi, 1998) and public intolerance of wolves in suburban and urban areas will likely prevent their recolonization of these habitats (Bruskotter and Wilson, 2014).

Red foxes (*Vulpes vulpes*) have been highly successful in adapting to human habitats around the world and are one of the most widely distributed mammals (Schipper et al., 2008). Gray foxes (*Urocyon cinereoargenteus*) are widely distributed across central America and much of the US. Studies on associations between gray foxes and anthropogenic habitats have yielded mixed results (for review, see Allen et al., 2021). Some studies indicate that gray foxes use an array of urban and agricultural habitats (e.g., Lombardi et al., 2017; Temple et al., 2010) while others report negative associations between gray foxes and human development (e.g., Markovchick-Nicholls et al., 2008; Cooper et al., 2012). Thus, if behavioral flexibility allows species to cope with anthropogenic habitat modification, dogs, coyotes, and red foxes should be more flexible than gray foxes and wolves should the least flexible of these canids.

Although most canids are omnivorous, wolves (Mech, 1970) and red foxes (Jędrzejewski and Jędrzejewski, 1992) rely more heavily on hunting and less on foraging and scavenging than do coyotes (Andelt et al., 1987), dogs (Coppinger and Coppinger, 2002), and gray foxes (Hockman and Chapman, 1983). Thus, if searching for patchy food resources requires more flexibility than hunting, wolves should display less flexibility than these other canids. However, unlike coyotes and gray foxes which are adept hunters, free-ranging dogs, which still compose 70-80% of the global dog population and approximate the ancestral niche of all domestic dogs (Lord et al., 2013), have evolved to be entirely dependent on scavenging and begging for food (Coppinger and Coppinger, 2002; Marshall-Pescini et al., 2017a). Thus, if behavioral flexibility is inversely related to reliance on any single feeding strategy, coyotes and gray foxes should still display greater flexibility than red foxes and wolves, but dogs should be less flexible than any of these other canids.

If behavioral flexibility is positively correlated with the diversity of habitats in which a species can survive, flexibility should decrease from coyotes to red foxes and wolves, to gray foxes, and finally to dogs. Alternatively, if seasonal habitat variability selects for behavioral flexibility, wolves, coyotes, and red foxes should be more flexible than gray foxes while dogs should be less flexible than all other canids.

Coyotes were historically found in arid and semi-arid open habitats but now inhabit almost any habitat in North and Central America and are expanding into the tropics (Hody and Kays, 2018). Wolves can also inhabit a variety of temperate, subarctic, and subtropical habitats (Mech and Boitani, 2007) but generally require forest cover, abundant prey, and low human-impact (Massolo and Meriggi, 1998). Red foxes inhabit diverse temperate and subarctic habitats but require cooler climates than do coyotes or wolves (Lloyd, 1980). Gray foxes primarily inhabit brushy woodlands near water in temperate and subtropical regions of North and Central America (Allen et al., 2021).

Finally, dogs can only survive in or around human habitats (Paul et al., 2016; Sen Majumder et al., 2016), which undergo far less seasonal variation than do natural habitats (Lowry et al., 2013).

If behavioral flexibility is positively correlated with social group complexity, structural variability and fission-fusion group dynamics, flexibility should decrease from dogs to wolves and red foxes, to coyotes, to gray foxes. Alternatively, if behavioral flexibility is negatively associated with solitary activity, flexibility should decrease from wolves to dogs and coyotes, to red foxes, to gray foxes.

Dogs are facultatively social and can form loosely associated groups of 2-30 individuals (Bonanni and Cafazzo, 2014; Marshall-Pescini et al., 2017a). These fluid packs can be characterized by both linear and nonlinear social hierarchies among numerous related and unrelated individuals, including multiple breeding adults (Cafazzo et al., 2010). In contrast, wolf packs are usually composed of a single breeding pair, their adult and sub-adult offspring, and the most recent litter (Packard, 2003). Coyote packs are simpler still, consisting of a breeding pair, their most recent litter, and in some cases yearlings from the previous litter (Bowen, 1981; Hennessy et al., 2012). Coyotes tend to be more solitary where small rodents and vegetation comprise most of their diet (Bekoff and Wells, 1980) whereas in higher latitudes and where larger prey are available, delayed dispersal is more common (Moehlman, 1989). More broadly, the social structures of dogs, wolves, and coyotes all vary with food availability (Bowen, 1981; Macdonald and Carr, 1995, Mech, 1970), reproductive status (Kleiman and Brady, 1978; Sen Majumder et al., 2014; Harrington, 1982), and season (Bowen, 1981; Sen Majumder et al., 2014; Metz et al., 2011).

Red foxes have historically been considered primitively social, although recent studies are challenging this view (for discussion see Dorning and Harris, 2019; Soulsbury, 2008). Red foxes typically form monogamous breeding pairs (Iossa et al. 2008) and adults spend most of their time alone (White and Harris, 1994). However, depending on ecological conditions, a monogamous dominant breeding pair may associate with several adult subordinates or dominant males may engage in polygyny (Moehlman, 1989). The few existing studies on the social ecology of gray foxes indicate that they are solitary but seasonally form small family units of monogamous breeding pairs and their young (Nicholson et al., 1985; Tucker et al., 1993; Chamberlain and Leopold, 2002).

If behavioral flexibility is positively correlated with cooperation, flexibility should decrease from wolves to coyotes, to dogs, to red foxes, to gray foxes. Wolves of the same pack cooperate extensively in hunting, territorial defense, and rearing young (MacNulty et al., 2014; Mech et al., 2021). Pair-bonded coyotes and sometimes their yearlings cooperate in rearing young and defending territory (for review, see Lord et al., 2013). Coyote dyads and packs occasionally collaborate in hunting large prey (Bowen, 1981) and individuals have been observed hunting rodents in synchrony with badgers (Rathbun et al., 1980; Thornton et al., 2018). Free-ranging dogs usually scavenge alone (Sen Majumder et al., 2014) and rarely display allomaternal care (Pal, 2005) but packs sometimes cooperate in territorial defense (Pal, 2015). Owned dogs (particularly working breeds) cooperate with humans, but this may require training (Range and Virányi, 2015). Red and gray foxes are solitary hunters and foragers but on rare occasions, red fox females may rear litters communally (for review, see Moehlman, 1986; 1989).

In a previous study, we designed a serial reversal learning test which rapidly identified cognitive flexibility deficits in elderly pet dogs (Van Bourg et al., 2021). In a follow-up study, we applied the same reversal learning criteria to a similar two-choice spatial paradigm to assess cognitive aging and flexibility in captive coyotes (Van Bourg et al., 2022). As in dogs, behavioral flexibility declined with age in coyotes, and performance did not differ between coyotes and dogs. In the present study, we used a similar test to assess whether gray wolves also develop ARCDs as well as to compare cognitive flexibility among dogs, coyotes, wolves, red foxes, and gray foxes.

## Method

This research was approved by Institutional Animal Care and Use Committees at Arizona State University (18-1650R) and USDA’s National Wildlife Research Center (QA-2967).

### Subjects

Subject details are provided in Appendix F-1. We worked with 12 wolves and completed testing with nine individuals (4 female, 5 male) which ranged in age from 36 - 125 months (mean = 67, SD = 34 months). We also tested two red foxes (an 86-month-old female, and a 110-month-old male) and two gray foxes (both 26-month-old males). Six of the wolves that completed testing, three of the wolves that did not, and all foxes were tested at Wolf Park (WP) in Battleground, IN. The other three wolves were tested at the Wolf Conservation Center (WCC) in South Salem, NY.

There were two trips to WP. Three of nine wolves completed testing trials during the first trip in 2020. The three oldest wolves died before we returned in 2021. Five of the six surviving wolves and all foxes completed testing trials during this second trip. Thus, two wolves at WP were tested twice. All three wolves at the WCC completed testing during a single trip in 2021.

All canids were reared using the same socialization procedures and all served as ambassador animals. Thus, they experienced similar levels of through-the-fence exposure to humans. Only experienced handlers entered the enclosures of the wolves at the WCC whereas unfamiliar humans were allowed to enter the pens of canids at WP for handler-supervised encounters. However, in the present study humans never entered enclosures. Wolves at WP are fed large sections of deer three times per week and receive small pieces of assorted meats as treats during training and enrichment. Wolves at the WCC are fed daily rations of deer and assorted meats.

All canids were tested in their home enclosures. Two of the six wolves tested at WP were housed individually and the other four lived in a pack. The three wolves tested at the WCC lived in a pack.

The testing procedures for wolves and foxes used in the present study were adapted from previous studies on pet dogs (Van Bourg et al., 2021), and captive coyotes (Van Bourg et al., 2022).

### Materials and Layout

Experimenters fed each canid through the fence of its enclosure at three locations which were spaced three meters apart for wolves and two meters apart for foxes. Experimenters wore matching clothing, hats, and dark sunglasses so that the canid could not easily discriminate between experimenters or use the experimenter’s eyes as cues. Treats were an assortment of 1.5cm$^3$ pieces of raw deer, cooked meats (chicken, pork, steak, sausage), and cold cuts.

### Acclimation and Shaping

Before testing, all canids completed two to eight days of response-dependent acclimation and shaping procedures which involved successively introducing novel aspects of the testing procedure (e.g., dropping treats into the enclosure through the fence, one experimenter feeding the wolf at different locations, multiple experimenters walking to different feeding locations at the same time).

### Testing Procedure

At the beginning of each trial both experimenters dropped a treat into the enclosure at the central feeding (start) location. This lure ensured that the canid began equidistant to the two lateral (test) locations. When the canid retrieved this treat from the center, the experimenters walked in synchrony to opposite test locations. There, the experimenters stood facing the enclosure, gazing into the distance until the canid chose (approached within one meter of) the left or right test location.

When the canid chose the correct side, the experimenter at that location provided a treat. Conversely, no treat was provided if the canid chose the incorrect side. In either case, the experimenters then walked in synchrony back to the start location to begin the next trial.

Video recordings of all sessions were captured by a camera set approximately 4 meters away from the start location. The determination of the coyote's choice was confirmed by each experimenter in real time and later verified from the video.

Each testing session ended when the canid stopped participating, was interrupted by another canid in the enclosure, or the maximum time allotted for the canid expired (whichever came first). This allotment of time was determined by the daily schedule of activities at the facility. In general, this allowed for 15 minutes of training or testing per canid, per session.

The amount of food consumed by each animal in each session was reported to animal care staff to ensure that canids received additional food as needed.

### Testing Rules

The same side (testing location) was treated as the correct side until the canid chose this side in three consecutive trials whereupon reward availability was switched to the other location, until the canid again chose this side in three consecutive trials. These side reversals of S+ and S- were repeated each time this criterion was met. The pairing of the experimenter, side, and treat was counterbalanced and pseudorandomized such that the canid did not retrieve the treat from the same experimenter in more than three consecutive trials.

### Scoring Choices

To consistently retrieve treats, the canid was required to attend to feedback from recent choices (i.e., to recall whether it was rewarded for approaching the previously visited side). Thus, in the first trial following a reversal, we scored a correct choice when the canid visited the side that did *not* contain the treat because this was the correct win-stay response (i.e., the canid retrieved a treat at this location in the previous trial). In all other trials, we scored a correct choice when the canid visited the side that *did* contain the treat because this was either the correct win-stay response (i.e., the canid retrieved a treat at this location in the previous trial) or the correct lose-shift response (i.e., the canid did not retrieve a treat at the other side in the previous trial).

### Distracting Packmates

Separating pack-housed animals into different enclosures was not possible. Thus, to allow the experimenters to work with pack-housed wolves individually, assistants lured non-focal packmates away from the testing area and continued to distract these individuals by regularly delivering small treats or by providing foods which could not be eaten quickly (e.g., meat frozen in 25x10x10cm ice blocks).

## Analysis and Results

Further details about models and tests of their assumptions are provided in Appendix F-2.

### Proportion of Correct WSLS Choices

Because participation in this test was voluntary, the number of trials completed by different wolves varied greatly. To control for this variation while examining the effects of age on performance we first tested the regression of age on the proportion of correct choices (WSLS scoring) in the first 10 trials because all wolves completed at least 10 trials. We then repeated this analysis using the first 20 trials because only one wolf did not complete 20 trials, and this provided twice as much data for the remaining wolves. To adjust alpha, we used a Bonferroni correction ($\alpha = .025$). However, we could not repeat this analysis with still larger subsets of trials as the consequent sample sizes could not accommodate alpha corrections for numerous comparisons (e.g., only 6 wolves completed 50 trials and only 4 wolves completed 100 trials).

In our previous study on pet dogs (Van Bourg et al., 2020), we found a curvilinear relationship between age and performance in this test. Thus, we tested both the linear and quadratic effects of age on performance in each analysis.

We did not test enough red or gray foxes to include either species in these analyses. Thus, we have included foxes in the graphs of these variables for visual assessment of potential congruence with age effects observed in wolves.

In the first ten trials, the overall model did not significantly predict the proportion of correct WSLS choices, $F(2,8) = 1.24$, p = .34, $r^2 = 0.24$ (Fig. 5-1). The effects of age, $ß = 0.02$, SE = 0.01; $t(8) = 1.35$, $p = .22$, and age$^2$ were not significant, $ß = -1e^{-04}$, SE = $7e^{-05}$; $t(8) = -1.45$, $p = .19$.

In the first 20 trials, the overall model accounted for substantial variation in the proportion of correct WSLS choices ($r^2 = 0.60$) but was only marginally significant after adjusting for multiple comparisons, $F(2,7) = 5.35$, $p = .04$; $r^2 = 0.60$ (Fig. 5-2). However, the effects of both age, $ß = 0.02$, SE = 0.01; $t(7) = 3.27$, $p = .01$, and age$^2$ were significant, $ß = -1e^{-04}$, SE = $3e^{-05}$; $t(7) = -3.25$, $p = .01$. Compared to younger and older wolves, middle-aged wolves made significantly more correct choices.

**Longest Streak of Perseverative Errors**

To examine whether age affected perseveration in wolves, we tested the regression of age and age$^2$ on the longest streak of perseverative errors made by each canid. Again, we repeated this analysis using the first 10 and 20 trials of the test and included foxes in graphs but not analyses.

In the first 10 trials, the overall model did not significantly predict the longest streak of perseverative errors, $F(2,8) = 2.59$, $p = .14$, $r^2 = 0.39$ (Fig. 5-3). The effects of age, $ß = -0.09$, SE = 0.08; $t(8) = -1.13$, $p = .29$, and age$^2$ were not significant, $ß = 7e^{-04}$, SE = $5e^{-04}$; $t(8) = 1.40$, $p = .20$.

In the first 20 trials, the overall model significantly predicted the longest streak of perseverative errors, $F(2,8) = 8.20$, $p = .01$; $r^2 = 0.67$ (Fig. 5-4). Compared to middle-aged wolves, younger and older wolves displayed significantly longer streaks of perseverative errors. However, the linear, $ß = -0.09$, SE = 0.04; $t(8) = -2.15$, $p = .06$, and curvilinear effects of age$^2$, $ß = 7e^{-04}$, SE = $3e^{-04}$; $t(8) = 2.61$, $p = .03$, were each marginally significant after adjusting for multiple comparisons.

**Canine Comparisons**

In our previous study on pet dogs (Van Bourg et al., 2021), all subjects completed at least 30 trials in a 30-minute, single-session test. Thus, to compare behavioral flexibility between coyotes and dogs (Van Bourg et al., 2022), we conducted a binomial regression analysis of trial outcomes in the first 30 trials as a function of species (see Table 5-1). In the present study, we added wolves to this analysis. To control for procedural differences, we included fixed effects for session number, trial number within the session, and reversal number (cumulative across all sessions). To test whether reversing the rewarded side affected performance, we included a dummy code (fixed effect) for whether a reversal had occurred in the session (Session Reversal). To test for learning, we included a fixed effect for the test trial number (cumulative across all sessions). To control for subject covariates, we included fixed effects for sex and age (in months). For random effects, we included subject intercepts and subject slopes for variables which were not constant for individuals.

The effect of species was not significant but the probability of choosing correctly decreased with age, $X^2(1) = 4.21$, $p = .04$ and increased with test trial number, $X^2(1) = 3.91$, $p = .048$ (Table 5-1). Canids were more likely to choose the correct side before completing a reversal during a session, $X^2(1) = 12.07$, $p = .0005$. No other effects were significant.

After eliminating covariates (predictors other than species and age) that were not significant, the reduced model included fixed effects for species, age, test trial, and session reversal (Table 5-1). Younger canids tended to choose the correct side more frequently, but this effect was not significant, $X^2(1) = 2.90$, $p = .09$. However, the effect of species was significant, $X^2(1) = 6.44$, $p = .04$. Based on Tukey post-hoc pairwise comparisons, wolves ($\bar{x} = 0.51$, SE = 0.03) chose incorrectly more often than coyotes ($\bar{x} = 0.58$, SE = 0.02; $z = 2.51$, $p = .02$) or dogs ($\bar{x} = 0.58$, SE = 0.01; $z = 2.27$, $p = .04$) but dogs and coyotes did not differ significantly. Again, the probability of choosing correctly increased with test trial number, $X^2(1) = 5.78$, $p = .02$ and canines were more likely to choose the correct side before completing a reversal during a session, $X^2(1) = 21.23$, $p < .0001$. Random subject intercepts were not significant.

## Canid Comparisons

Given that wolves were the focus of this study and that we tested only two individuals of each fox species, we did not include foxes in our primary analysis of trial-by-trial performance. Thus, to explore whether red and gray foxes differed from the other canids when controlling for age differences, we repeated the previously described final model of trial outcome with foxes included (Table 5-2).

Younger canids tended to choose the correct side more frequently, but this effect was not significant, $X^2(1) = 2.96$, $p = .09$. The effect of species was not significant, $X^2(1) = 8.70$, $p = .07$. Again, the probability of choosing correctly tended to increase with test trial number but this effect was not significant, $X^2(1) = 2.99$, $p = .09$. However, canids were more likely to choose the correct side before completing a reversal during a session, $X^2(1) = 17.73$, $p < .0001$. Random subject intercepts were not significant.

### Ecological Comparisons

To examine whether species ecology may explain why coyotes and dogs outperformed wolves in this cognitive flexibility assessment, we tested eight hypotheses related to feeding, habitat, and social ecology (Table 5-3). We could not include foxes in this qualitative analysis because age would be an uncontrolled variable in these comparisons. Of these hypotheses, two aligned with our findings. The hypothesis that behavioral flexibility allows species to adapt to human-modified habitats and the hypothesis that foraging for spatially distributed food resources requires more behavioral flexibility than hunting were supported by the finding that coyotes and dogs displayed greater flexibility than wolves.

## Discussion

### Assessment Validity

The fact that session reversal was the strongest predictor of trial outcome in all GLMMs indicates that the canid’s ability to choose the correct side was strongly constrained by its ability to inhibit the prepotent response of searching for food at the previously rewarded location when a reversal occurred. In turn, this finding provides strong evidence that this serial reversal learning task assessed behavioral inhibition as intended.

The finding that canines were more likely to choose correctly as trial number increased indicates that they became more proficient at following the Win-Stay and Lose-Shift rules as the test progressed. The fact that this occurred within just the first 30 trials suggests that canines may be predisposed to quickly detect and respond to fluctuating environmental conditions.

## Cognitive Aging

The significant quadratic relationship between age and the proportion of correct choices in the first 20 trials suggests that relative to young and old wolves, middle-aged wolves were more sensitive to the Win-Stay rule, the Lose-Shift rule, or both contingencies. In other words, middle-aged wolves were better able to identify the more productive of the two possible responses, either by repeating behaviors that were rewarded or by inhibiting behaviors that were not rewarded. Taken alone, this measurement cannot separate between the associative learning component of the Win-Stay rule and the cognitive flexibility component of the Lose-shift rule. However, the significant quadratic relationship between age and the longest streak of perseverative errors committed during the first 20 trials supports the hypothesis that middle-aged wolves were better able to inhibit unproductive behaviors and flexibly switch between responses.

These findings agree with recent studies on pet dogs which suggest that executive functions like reversal learning, inhibitory control, and working memory may continue to develop beyond adolescence and then decline in old age. Van Bourg et al. (2021) found that young and old dogs displayed longer streaks of perseverative errors and were more likely to choose the incorrect side in this serial reversal learning test. Additionally, Van Bourg et al. (2020) found that young and old dogs displayed stronger side biases in a short-term memory assessment which used the same two-choice, spatial paradigm. Furthermore, Watowich et al. (2020) found that old and young dogs displayed deficits in social communication, inhibitory control, selective attention, reasoning, and working memory.

The fact that age effects were not observed in the analyses of performance in the first ten trials, which included all subjects, may indicate that all wolves required more than ten trials to begin responding appropriately to the reward contingencies. Alternatively, these analyses of only the first ten trials may have been underpowered due to insufficient individual data. In line with this hypothesis, both trends were strong ($r^2 = .24$ and $r^2 = .43$) and both were similar in shape to the significant regressions of age on performance in the first 20 trials.

Ultimately the present study indicates that, as in dogs and coyotes, wolves develop age-related deficits in cognitive flexibility. This suggests that ARCDs in dogs may not be a unique product of the shared dog-human environment but rather a conserved trait which evolved in an ancestor common to dogs, wolves, and coyotes. However, additional studies are needed to determine whether the physiological changes that are associated with ARCDs in dogs and Alzheimer's Disease in humans are also associated with ARCDs in wolves and coyotes. In addition, future studies on ARCDs in wild canines should aim to differentiate between (1) minor deficits which gradually develop in old age and are typical of normal (healthy) aging, and (2) AD-like deficits which appear earlier, progress more rapidly, and become more severe. Finally, the successful piloting of this test for red and gray foxes suggests that this assessment can be used to test for cognitive deficits in a variety of canid species.

## Ecological Predictors of Behavioral Flexibility

### Supported hypothesis.

Coyotes (Andelt et al., 1987) and dogs (Coppinger and Coppinger, 2002) rely more heavily on patchy resources than do wolves (Mech, 1970). Thus, the finding that coyotes and dogs displayed greater flexibility than wolves supports the hypothesis that reliance on patchy food resources selects for greater behavioral flexibility.

Studies on the evolution of behavioral flexibility have primarily focused on primates (MacLean et al., 2014). A prominent hypothesis from these studies is that reliance on fruit contributed to the evolution of complex and flexible cognition in primates because the ability to move flexibly between different fruiting plants as resources are exhausted and replenished should increase fitness (Zuberbühler and Janmaat, 2010). Although to a lesser degree than primates, coyotes can also rely heavily on fruit depending on location and season. For example, in a coniferous forest in Oregon, fruits were found in 83% of coyote scats collected during summer months and fruits were the most common dietary component when percent frequency was averaged across seasons (Toweill and Anthony, 1987).

Similar selection pressures may drive behavioral flexibility in free-ranging dogs, which are obligate scavengers that actively search for human waste and handouts (Coppinger and Coppinger, 2002). Aside from large municipal waste facilities (dumps), these anthropogenic food resources are continuously depleted and replenished. Thus, dogs should benefit from flexibly rotating between different trash cans, dumpsters, storefronts, and other sources of handouts within their territories in responses to local changes in food quality and quantity.

Although wolves will opportunistically browse for fruit, this typically comprises a small (< 10% frequency) component of their diet (for discussion see Gable et al., 2017). Instead, wolves rely primarily on hunting and opportunistically scavenging large ungulates (Mech, 1970). Thus, heavy reliance by coyotes and dogs (but not wolves) on resources which are depleted and replenished over short time periods may explain the behavioral flexibility deficits displayed by wolves in the present study.

**Alternative interpretations.**

In interpreting these findings, it is important to note that the various hypotheses about ecological drivers of behavioral flexibility tested in the present study were not mutually exclusive. Equally important, failure to support a hypothesis should not be treated as evidence counter to that hypothesis. Thus, social, dietary and habitat complexity, diversity and instability may all drive the evolution of behavioral flexibility, and these effects may be additive or interactive. Some of these selective pressures may simply be stronger than others. For example, our findings do not preclude the possibility that cooperative tendencies promoted the evolution of behavioral flexibility in wolves. Moreover, this effect may have been overshadowed by the importance of foraging for patchy resources if the latter selected for even greater flexibility in dogs and coyotes.

It is also possible that this test was more likely to detect effects of foraging ecology on behavioral flexibility because the paradigm it used modeled a foraging situation. In another paradigm assessing behavioral flexibility in social situations, species differences in gregariousness or other aspects of social ecology may better predicts differences in performance. Finally, our findings may not generalize to other species. For example, it is possible that social complexity is a better determinant of behavioral flexibility in primates given that most primate species rely heavily on fruit.

#### Limitations of comparisons.

Differences in the experimental procedures used for different canids may have obscured underlying species differences. Experimenters walked to and from coyote enclosures but remained close to wolves, dogs, and foxes throughout testing. Thus, experimenters may have been more salient and distracting for wolves, dogs, and foxes whereas coyotes may have been able to better focus on which side was baited in the previous trial. However, this would not explain why both coyotes and dogs outperformed wolves. Given that the food blinds were roughly 15m apart for coyotes but only 3m apart for wolves and 2m apart for dogs and foxes, coyotes may have devoted less attention to the incorrect blind. Thus, inhibiting incorrect choices may have been easier for coyotes. Incorrect choices may also have been more costly for coyotes because traversing greater distances required more time and energy. This may have provided a stronger incentive for coyotes to choose correctly. However, these procedural differences still do not explain why both coyotes and dogs outperformed wolves.

#### Future directions.

Testing identically reared and housed canids, such as pack-housed dogs and wolves, should help to address procedural inconsistencies. However, as this study has demonstrated, disentangling hypotheses about the numerous potential ecological drivers of cognitive evolution will likely require testing more than two species. Thus, future studies which test larger samples of red and gray foxes, as well as other canids are needed to clarify the findings of the present study.

## Conclusions

The present study showed that canines may be predisposed to quickly detect and respond to fluctuating environmental conditions and that age-related deficits in cognitive flexibility may be a conserved trait among canines. In addition, our findings indicate that executive functions like reversal learning and inhibitory control continue to develop after adolescence in wolves as in dogs. Together, these findings suggest that ARCDs in dogs may not be a product of the shared dog-human environment, however additional studies are needed to assess the utility of canines other than dogs as animal models of Alzheimer’s Disease in humans. Finally, the successful deployment of this test with both captive and pet, large- and small-bodied canids suggests that this assessment may be used to test for cognitive deficits in a variety of species.

The present study also indicated that coyotes and dogs are more behaviorally flexible than wolves. This finding supports only two of many hypotheses about potential ecological drivers of cognitive flexibility — that reliance on patchy food resources selects for greater behavioral flexibility and that behavioral flexibility allows species to adapt to human-modified habitats. However, additional studies are needed to rule out alternative explanations. Importantly, our findings indicated that controlling for age in these future studies will be essential.

# CHAPTER 6

# GENERAL DISCUSSION

This dissertation describes a series of studies on comparative cognition in dogs and their wild relatives. The first two studies aimed to design cognitive assessments for dementia in pet dogs which can be implemented easily and rapidly. These tests may aid in the diagnosis of Canine Cognitive Dysfunction syndrome and thereby improve care for elderly pet dogs while also facilitating the use of pet dogs in translational research on Alzheimer's Disease in humans. The third and fourth studies tested for similar cognitive changes in wild canids and examined which ecological traits drive the evolution of behavioral flexibility and species resilience.

## Cognitive Aging in Dogs and Humans

In the first experiment, I designed and deployed an assessment for spatial working memory (SWM) in which 26 pet dogs were required to recall the location of a treat hidden behind one of two identical boxes following a delay. To efficiently measure the duration and accuracy of the dog's short-term memory, I used a response-dependent testing procedure in which the dog was given trials with longer delays following correct choices and trials with shorter delays following incorrect choices. This test was sensitive to age effects but was hindered by a high rate of attrition. Thus, I modified the shaping and testing procedures and retested the dogs. In this second experiment, I acquired reliable evidence that dogs learned the rules of the assessment, that the test measured SWM as intended, and that performance declined with age. This study demonstrated that up-down psychometric staircase methods and adaptive testing procedures more broadly, may be used to efficiently assess SWM and identify cognitive decline in pet dogs.

Although designing a single-session assessment was an important step towards a widely deployable test, the SWM task was complicated and lengthy (approximately one hour). As a result, in the next study I applied a faster and less complicated serial reversal learning (SRL) procedure to the same two-box paradigm to assess cognitive flexibility and behavioral inhibition in 80 pet dogs. In this assessment, an experimenter hid treats in the same box until the dog learned to search for treats only in this box. Each time this occurred, the experimenter began hiding treats in the other box. All dogs that completed training also completed the assessment. Age did not predict how frequently the dog learned the serially reversing reward contingency, but older dogs chose the incorrect box more often and displayed longer streaks of perseverative errors. Importantly, this age-related cognitive decline was detectable with a stand-alone score that could be easily measured and interpreted. And as with the SWM test, this SRL assessment required only two visually separated areas and two identical objects large enough to hide food treats. Ultimately, this widely deployable test may aid in the diagnosis of CDS in dogs and as a result, facilitate research on ARCDs in dogs and the use of dogs as an animal model of AD in humans.

## Cognitive Aging in Other Canids

Comparing cognitive, behavioral, and neurodegenerative changes between canid species to determine whether ARCDs are a conserved or unique trait in dogs may be an effective means of disentangling the genetic and environmental components of dementia in dogs. Moreover, if wild canids and pet dogs are equally prone to cognitive, behavioral, and neurodegenerative senescence, future studies on ARCDs in dogs should focus on the genetics underlying these changes rather than the dog-human environment. To this end, I conducted a series of cognitive tests on captive North American canids.

First, I assessed cognitive flexibility in 20 coyotes using a modified version of the SRL test. After the coyotes completed this assessment, I immediately tested their ability to learn a color discrimination task which required that they stop following the obsolete rules of the SRL test. In both tasks, multiple facets of performance declined with age suggesting that coyotes, like dogs, may develop ARCDs.

In the next study, I slightly modified the SRL assessment to test for age-related cognitive flexibility deficits in gray wolves. In addition, I successfully piloted the use of this test for red and gray foxes so that future studies may test for ARCDs in these and other small-bodied wild canids. As in dogs and coyotes, performance in wolves declined in old age suggesting that ARCDs may be a conserved trait among canines.

Studies on the neuropathology and epidemiology of ARCDs in wild canids are needed to confirm this hypothesis. Specifically, future studies must differentiate between (1) moderate cognitive decline which occurs late in life in many species and (2) more rapid and severe changes which begin earlier in life as a result of neuritic plaque formation and consequent cortical neurotrophy. This will likely require repeated, lifetime testing of the same individuals and postmortem neurological assays. Testing for other cognitive and behavioral changes in these individuals will also be valuable.

Future studies which compare populations of wild canids living in human-modified and natural habitats may help to identify environmental pollutants that contribute to cognitive aging. Coyotes and red foxes in particular may be ideal focal species given their success in highly developed and densely populated human habitats.

For example, it may be the case that populations of canids in these habitats are more prone to AD-like senescence than are populations of canids in remote, undisturbed habitats. Follow-up studies could then attempt to identify the specific anthropogenic pollutants contributing to this phenomenon. Ultimately, the identification of ARCDs in coyotes and wolves has paved the way for multiple promising avenues for research on dementia.

## Cognitive Aging in Other Species

Animal models that closely mimic the age-related cognitive decline that occurs in humans are essential for understanding and designing more effective treatments for AD in humans (Bizon and Woods, 2009). AD-like amyloid plaque and neurofibrillary tangle accumulation has been observed in several diverse animals including Bactrian camels, polar bears, degus, and woodpeckers (Youssef et al., 2016). However, the effects of these neurotrophic changes on the cognitive function of these species are poorly documented (Youssef et al., 2016). New cognitive tests which are known to detect age-related deficits in diverse species are needed to remedy this limitation. Thus, the ability of the serial reversal learning test developed in this dissertation to detect age-related cognitive flexibility deficits in both domesticated and wild small- to large-bodied animals varying greatly in neophobia and boldness suggests that this test may be valuable for assessing and comparing cognitive aging in diverse species.

## Cognitive Development in Dogs

Little is known about the development of non-social cognition in dogs (Bray et al., 2021b). Thus, in addition to investigations into senescence, I also tested for cognitive changes early in adulthood. To this end, I examined both the linear and curvilinear relationships between age and performance.

In both the spatial working memory and serial reversal learning assessments, young dogs displayed performance deficits relative to middle-aged dogs, indicating that executive functions like reversal learning, inhibitory control, and working memory may develop beyond adolescence and then decline in old age. These insights are in line with similar findings in other recent studies on cognitive aging in pet dogs (e.g., Watowich et al., 2020).

Understanding such developmental changes may help owners and handlers to better care for and train both pet and working dogs. Indeed, several aspects of trainability including selective attention (Wallis et al., 2014), short-term memory (Chapter 2), and behavioral inhibition (Chapter 3) appear to improve through much of adult life. Working dogs are often expected to complete training and begin working in the field around two years of age (e.g., Duffy and Serpell, 2010). The findings of this dissertation and other recent studies suggest that these young dogs may be best suited for placement with experienced handlers who may not require that their dogs have achieved peak performance. In addition, working dogs are often released from training programs between one and two years of age, at which point programs have already invested heavily in their training (Bray et al, 2021a). If these decisions are based on behaviors which involve slow-developing cognitive processes like attentiveness and flexibility, releasing these young individuals may be premature. Moreover, future studies should assess whether underperforming young dogs can ultimately achieve the necessary level of proficiency if allowed more time to develop. Conversely, working dogs such as police and military dogs generally retire around eight years of age, usually due to physical limitations resulting from musculoskeletal degeneration (Worth et al., 2013). However, the findings of this dissertation suggest that middle-aged dogs may be in the prime of their cognitive ability. Thus, transitioning middle-aged working dogs to less physically demanding, but still valuable, tasks rather than fully retiring these individuals may maximize their utility.

Clarifying when dogs reach peak cognitive performance should improve the sensitivity of cognitive tests for ARCDs. For example, considering the longest streak of perseverative errors made at any point during the SRL test, dogs under two years of age committed as many or more  consecutive errors than dogs over ten years of age. Thus, had I used these young adult dogs as a baseline, I would have failed to detect the substantial decline that was found when comparing middle-aged and elderly dogs. Most studies on colony-living beagles have included only young dogs in baseline comparison groups, have not tested middle-aged dogs, or have not tested for curvilinear age effects. For example, Head et al. (1995) found a linear decline in SWM with age but tested no dogs between three and seven years of age. These studies may have failed to detect important cognitive changes or underestimated the magnitude of cognitive decline from middle to old age. Ultimately, this updated view on lifespan changes in cognition may improve our understanding of, and ability to detect ARCDs.

## Cognitive Development in Wild Canids

Interestingly, middle-aged wolves were also better able to inhibit unproductive behaviors and to flexibly switch between responses compared to young adult individuals. However, performance declined linearly with age in coyotes, both in the serial reversal learning and discrimination learning tests. Initially, I attributed this to the restricted age range of the coyote sample relative to the dog samples — many of the dogs I tested were substantially older and younger than the coyotes I tested. Thus, I hypothesized that a curvilinear relationship between age and performance in coyotes may be detected if older and younger coyotes were tested. However, the age ranges of the wolf (three to ten years) and coyote (two to nine years) samples were not substantially different. Thus, cognitive flexibility may develop more slowly in dogs and wolves than coyotes. This may indicate that flexibility yields fitness benefits earlier in development of coyotes, but additional studies are needed to test this hypothesis.

In line with the findings of this dissertation, Wobber et al. (2010) found that the development of social inhibition occurred more slowly in bonobos relative to chimpanzees. Interestingly, social tolerance also decreased with age in chimpanzees but not bonobos. Thus, aspects of cognition in adult bonobos may represent developmentally delayed forms of traits in chimpanzees. Given that coyotes disperse within their first or second year while adult wolves often remain with their natal pack for many years, social tolerance may also decrease more rapidly with age in coyotes than wolves. Like bonobos, dogs are also characterized by multiple paedomorphic traits (Frank and Frank, 1982). Thus, differences in behavioral inhibition between these closely related canines may mirror differences between closely related apes. Future studies which examine the development of social inhibition, rather than non-social inhibition are needed to test this hypothesis.

## Evolution of Cognitive Flexibility

Behavioral plasticity, which can arise through a variety of processes including ontogenetic expression, individual variation, and cognitive flexibility, allows individuals and populations to adapt to environmental changes. Thus, cognitive flexibility may enable species to invade human habitats and to cope with anthropogenic habitat modification. In turn, identifying the evolutionary drivers and ecological predictors of cognitive flexibility may help practitioners and policy makers to identify and protect species at risk of population decline. Canid species differ in their social, dietary, and habitat ecology and have had markedly different levels of success in adapting to human-modified ecosystems. Thus, canids are an ideal clade for testing competing hypotheses about the evolutionary drivers of behavioral flexibility.

Coyotes and dogs thrive in almost every type of human habitat (Hill et al., 1987; Kuijper et al., 2019) while stable populations of gray wolves are generally restricted to areas with low densities of humans (Massolo and Meriggi, 1998). Wolves also rely less on foraging and scavenging than do coyotes and dogs. In the SRL tests, wolves demonstrated less flexibility than did coyotes and dogs, but performance did not differ between coyotes and dogs. Thus, these findings support the hypothesis that behavioral flexibility helps species to adapt to human-modified habitats as well as the hypothesis that reliance on patchy food resources selects for greater behavioral flexibility. It follows that species which do not rely on patchy resources may be ill-equipped to adapt to anthropogenic environmental change. Although this hypothesis requires additional testing, if supported this knowledge could help practitioners to identify species in need of protection and help conservationist to garner support for the protection of these species.

The ability of an animal to adjust to novel situations through win-stay lose-shift rules should reflect the past selection pressures in which this species evolved (Sih, 2013). However, the ability to make such adjustments may not generalize to all aspects of the animal's behavior. Moreover, species should display greater flexibility in the specific aspects of their ecology which have been characterized by greater complexity and unpredictability (Bond et al., 2007). For example, animals that have evolved in dynamic or complex environments should exhibit greater perceptual and attentional flexibility. In contrast, animals that have evolved to rely on fluctuating resources should exhibit greater flexibility in allocating time and energy on acquiring spatially distributed resources. These insights suggest that the findings of this dissertation may not hold when using a different cognitive flexibility assessment. For example, coyotes may display less flexibility than dogs and wolves in a paradigm which examines responses to different social partners. In line with this prediction, previous studies have found that wolves demonstrate greater flexibility than dogs in a cooperative context by appropriately adjusting their behaviors to coordinate their actions with a partner (e.g., Marshall-Pescini et al., 2017b). In addition, Marshall-Pescini et al. (2015) found that dogs displayed greater behavioral inhibition in a cylinder task while the opposite was true in a different detour task.

Additional studies are needed to assess the extent to which the findings of this dissertation generalize to other species. For example, it is possible that social complexity is a better determinant of flexibility in primates, even in SRL tasks, given that most primate species rely heavily on fruit. Few studies have used SRL assessments to compare cognitive flexibility in closely related, non-primate species. Bond et al. (2007) compared performance in three SRL tasks between pinyon jays (*Gymnorhinus cyanocephalus*), which are highly social, Clark's nutcrackers (*Nucifraga columbiana*), which are mostly solitary but have superior spatial memory, and western scrub jays (*Aphelocoma californica*), which are ecological generalists. Pinyon jays outperformed nutcrackers and scrub jays suggesting that in North American Corvids, the ability to respond to shifts in reward associations evolved in response to social rather than ecological or spatial complexity.

Importantly, cognitive flexibility is only one of several forms of behavioral plasticity which have likely contributed to the success of both dogs and coyotes in adapting to anthropogenic habitat modification. For example, Breck et al. (2019) found that urban coyotes are bolder and more exploratory than rural coyotes, a product of differential selection, lifetime learning, or perhaps both processes. In addition, the domestication of dogs, an example of extensive adaptation to human environments, may be largely attributed to behavioral plasticity. A defining feature of this process was the evolution of a prolonged social critical period which afforded dogs more opportunities to imprint on co-occurring humans, or to learn that humans need not be completely avoided (Coppinger and Coppinger, 2002). In turn, this allowed free-roaming dogs to live in densely populated human habitats. However, dogs are still capable of developing avoidance behaviors consistent with wild canids if imprinting on humans does not occur. Thus, domestication can be thought of as an expression of behavioral plasticity. Ultimately, canids may be an ideal clade for examining multiple forms of behavioral plasticity and their roles in adapting to human habitats.

## Conclusions

The studies described in this dissertation further our understanding of comparative cognition and cognitive aging in canids. In the first two studies, I designed novel cognitive tests for dogs which may aid in the diagnosis of CDS in clinical and in-home setting. In the next two studies, I identified for the first time ARCDs in coyotes and wolves. In addition, I created a cognitive aging assessment which can be used for diverse captive animals. My hope is that future studies use these tests to advance veterinary care for elderly pet dogs and to facilitate translational research on AD in humans.

Unexpectedly, I also detected cognitive deficits in young, adult dogs. I hope that this insight will lead others to question the instantiated narrative that dogs are most trainable and best able to perform in a working capacity as young adults. In turn, I hope future studies on cognitive development in dogs may improve training and care for pets and working dogs. Furthermore, I believe that better describing lifespan cognitive changes will ultimately improve the sensitivity of cognitive aging tests. Interestingly, I also observed delayed development of cognitive flexibility in wolves but not coyotes which may reflect earlier selection for flexibility in coyotes.

Finally, I found that coyotes and dogs displayed greater flexibility than wolves suggesting that species which have not evolved to forage or scavenge for unstable resources may be unable to cope with human habitat modification. Ultimately, I hope that this and similar research on the evolutionary drivers of cognitive flexibility will aid in wildlife conservation.

# REFERENCES

Adams, B., Chan, A., Callahan, H., & Milgram, N. W. (2000). The canine as a model of human cognitive aging: recent developments. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *24*(5), 675-692. https://doi.org/10.1016/S0278-5846(00)00101-9

Adams, B., Chan, A., Callahan, H., Siwak, C., Tapp, D., Ikeda-Douglas, C., ... & Milgram, N. W. (2000). Use of a delayed non-matching to position task to model age-dependent cognitive decline in the dog. *Behavioural brain research*, *108*(1), 47-56. https://doi.org/10.1016/S0166-4328(99)00132-1

Aiello, L. C., & Wheeler, P. (1995). The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. *Current anthropology*, *36*(2), 199-221.

Allen, M.L, Avrin, A. C., Farmer, M. J., Whipple, L. S., Alexander, E. P., Cervantes, A. M., & Bauder, J. M. (2021). Limitations of current knowledge about the ecology of Grey Foxes hamper conservation efforts. *Journal of Threatened Taxa*, *13*(8), 19079–19092. https://doi.org/10.11609/jott.7102.13.8.19079-19092

American Veterinary Medical Association. (2018). *AVMA pet ownership and demographics sourcebook* (Doctoral dissertation, Colorado State University. Libraries).

Amici, F., Aureli, F., & Call, J. (2008). Fission-fusion dynamics, behavioral flexibility, and inhibitory control in primates. *Current Biology*, 18(18), 1415-1419. https://doi.org/10.1016/j.cub.2008.08.020

Amici, F., Call, J., Aureli, F. (2009). Variation in withholding of information in three monkey species. *Proceedings of the Royal Society B: Biological Sciences*, 276, 3311–3318. https://doi.org/10.1098/rspb.2009.0759

Amsel, A. (1962). Frustrative nonreward in partial reinforcement and discrimination learning: Some recent history and a theoretical extension. *Psychological review*, 69(4), 306. https://doi.org/10.1037/h0046200

Andelt, W. F., Kie, J. G., Knowlton, F. F., & Cardwell, K. (1987). Variation in coyote diets associated with season and successional changes in vegetation. *The Journal of Wildlife Management*, 51(2), 273-277. https://doi.org/10.2307/3801002

Araujo, J. A., Baulk, J., & Rivera, C. D. (2017). The aged dog as a natural model of Alzheimer's disease progression. In: G. Landsberg, A. Madari, & N. Žilka (Eds.), *Canine and Feline Dementia* (pp. 69–94). Springer, Cham.

Araujo, J. A., Studzinski, C. M., & Milgram, N. W. (2005). Further evidence for the cholinergic hypothesis of aging and dementia from the canine model of aging. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *29*(3), 411-422. https://doi.org/10.1016/j.pnpbp.2004.12.008

Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 780.

Azkona, G., García-Belenguer, S., Chacón, G., Rosado, B., León, M., & Palacio, J. (2009). Prevalence and risk factors of behavioural changes associated with age-related cognitive impairment in geriatric dogs. *Journal of Small Animal Practice*, *50*(2), 87-91.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255-278. https://doi.org/10.1016/j.jml.2012.11.001

Barton, R. A. (1996). Neocortex size and behavioural ecology in primates *Proc R Soc B*, 263(1367), 173-177. https://doi.org/10.1098/rspb.1996.0028

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015a). Parsimonious mixed models. *arXiv preprint*, 1506.04967.

Bates, D., Mächler, M., Bolker, B., Walker, S. (2015b). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Behrend, E. R., Powers, A. S., & Bitterman, M. E. (1970). Interference and forgetting in bird and fish. *Science*, 167(3917), 389-390. https://doi.org/10.1126/science.167.3917.389

Bekoff, M., & Wells, M. C. (1980). The social ecology of coyotes. *Scientific American*, *242*(4), 130-151.

Benson-Amram, S., Dantzer, B., Stricker, G., Swanson, E. M., & Holekamp, K. E. (2016). Brain size predicts problem-solving ability in mammalian carnivores. *Proceedings of the National Academy of Sciences*, *113*(9), 2532-2537. https://doi.org/10.1073/pnas.1505913113

Beran, M. J., & Hopkins, W. D. (2018). Self-control in chimpanzees relates to general intelligence. Current Biology, 28(4), 574-579. https://doi.org/10.1016/j.cub.2017.12.043

Bessemer, D. W., & Stollnitz, F. (1971). Retention of discriminations and an analysis of learning set. In A. M. Schrier & F. Stollnitz (Eds.), *Behavior of nonhuman primates* (Vol. 4, pp. 1-58). New York: Academic Press, Inc. https://doi.org/10.1016/B978-0-12-629104-9.50008-X

Bitterman, M. E. (1965). Phyletic differences in learning. *American Psychologist*, 20(6), 396-410. https://doi.org/10.1037/h0022328

Boitani, L., Ciucci, P., & Ortolani, A. (2007). Behaviour and social ecology of free-ranging dogs. In P. Jensen (Eds.), *The behavioural biology of dogs* (pp. 147-165). Oxfordshire: CAB International.

Bonanni, R., & Cafazzo, S. (2014). The social organisation of a population of free-ranging dogs in a suburban area of Rome: a reassessment of the effects of domestication on dogs' behaviour. In J. Kaminski (Eds.), *The social dog* (pp. 65-104). San Diego, CA: Academic Press.

Bond, A. B., Kamil, A. C., & Balda, R. P. (2007). Serial reversal learning and the evolution of behavioral flexibility in three species of North American corvids (*Gymnorhinus cyanocephalus, Nucifraga columbiana, Aphelocoma californica*). *Journal of Comparative Psychology*, *121*(4), 372–379. https://doi.org/10.1037/0735-7036.121.4.372

Bowen, W. D. (1981). Variation in coyote social organization: the influence of prey size. *Canadian Journal of Zoology*, 59(4), 639-652. https://doi.org/10.1139/z81-094

Brambor, T., Clark, W. R., & Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political analysis*, 63-82.

Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., ... & Reed, B. R. (2001). Context processing in older adults: evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*, *130*(4), 746. https://doi.org/10.1037/0096-3445.130.4.746

Bray, E. E., Gruen, M. E., Gnanadesikan, G. E., Horschler, D. J., Levy, K. M., Kennedy, B. S., ... & MacLean, E. L. (2021). Dog cognitive development: a longitudinal study across the first 2 years of life. *Animal Cognition*, *24*(2), 311-328. https://doi.org/10.1007/s10071-020-01443-7

Bray, E. E., MacLean, E. L., & Hare, B. A. (2014). Context specificity of inhibitory control in dogs. *Animal Cognition*, *17*(1), 15-31. https://doi.org/10.1007/s10071-013-0633-z

Brucks, D., Marshall-Pescini, S., & Range, F. (2019). Dogs and wolves do not differ in their inhibitory control abilities in a non-social test battery. Animal cognition, 22(1), 1-15. https://doi.org/10.1007/s10071-018-1216-9

Brucks, D., Marshall-Pescini, S., Wallis, L. J., Huber, L., & Range, F. (2017). Measures of dogs' inhibitory control abilities do not correlate across tasks. *Frontiers in Psychology*, *8*, 849. https://doi.org/10.3389/fpsyg.2017.00849

Bruskotter, J. T., & Wilson, R. S. (2014). Determining where the wild things will be: using psychological theory to find tolerance for large carnivores. *Conservation Letters*, *7*(3), 158-165.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron*, *33*(2), 301-311. https://doi.org/10.1016/S0896-6273(01)00583-9

Byrne R. W., Whiten, A. (1988). Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans. Oxford: Oxford University Press.

Cafazzo, S., Valsecchi, P., Bonanni, R., & Natoli, E. (2010). Dominance in relation to age, sex, and competitive contexts in a group of free-ranging domestic dogs. *Behavioral Ecology*, 21(3), 443-455. https://doi.org/10.1093/beheco/arq001

Chamberlain, M. J., & Leopold, B. D. (2002). Movements and space use of gray foxes (Urocyon cinereoargenteus) following mate loss. *The American midland naturalist*, *147*(2), 409-412.

Chan, A. D., Nippak, P., Murphey, H., Ikeda-Douglas, C. J., Muggenburg, B., Head, E., ... & Milgram, N. W. (2002). Visuospatial impairments in aged canines (*Canis familiaris*): the role of cognitive-behavioral flexibility. *Behavioral neuroscience*, *116*(3), 443. https://doi.org/10.1037/0735-7044.116.3.443

Chapagain, D., Range, F., Huber, L., & Virányi, Z. (2018). Cognitive aging in dogs. *Gerontology*, *64*(2), 165-171. https://doi.org/10.1159/000481621

Chapron, G., Kaczensky, P., Linnell, J. D., Von Arx, M., Huber, D., Andrén, H., ... & Boitani, L. (2014). Recovery of large carnivores in Europe's modern human-dominated landscapes. *Science*, *346*(6216), 1517-1519.

Cooper, S. E., Nielsen, C. K., & McDonald, P. T. (2012). Landscape factors affecting relative abundance of gray foxes Urocyon cinereoargenteus at large scales in Illinois, USA. *Wildlife Biology*, *18*(4), 366-373.

Coppinger, R., & Coppinger, L. (2002). *Dogs: a new understanding of canine origin, behavior and evolution*. Chicago: University of Chicago Press.

Cornsweet, T. N. (1962). The staircase-method in psychophysics. *The American journal of psychology*, *75*(3), 485-491. https://doi.org/10.2307/1419876

Cotman, C. W., & Head, E. (2008). The canine (dog) model of human aging and disease: dietary, environmental and immunotherapy approaches. *Journal of Alzheimer's Disease*, *15*(4), 685-707.

Craik, F. I., & Bialystok, E. (2006). Cognition through the lifespan: mechanisms of change. *Trends in cognitive sciences*, *10*(3), 131-138. https://doi.org/10.1016/j.tics.2006.01.007

Cummings, B. J., Head, E., Afagh, A. J., Milgram, N. W., & Cotman, C. W. (1996). β-amyloid accumulation correlates with cognitive dysfunction in the aged canine. *Neurobiology of learning and memory*, *66*(1), 11-23. https://doi.org/10.1006/nlme.1996.0039

Davey, G. (1989). Ecological learning theory. London: Routledge.

Day, L. B., Crews, D., & Wilczynski, W. (1999). Spatial and reversal learning in congeneric lizards with different foraging strategies. *Animal behaviour*, *57*(2), 393-407.

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental review*, *12*(1), 45-75. https://doi.org/10.1016/0273-2297(92)90003-K

Diamond, A. (1990). Developmental time course in human infants and infant monkeys, and the neural bases of, inhibitory control in reaching a. *Annals of the New York Academy of Sciences*, *608*(1), 637-676. https://doi.org/10.1111/j.1749-6632.1990.tb48913.x

Diamond, A. (2013). Executive functions. *Annual review of psychology*, *64*, 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Dorning, J., & Harris, S. (2019). Understanding the intricacy of canid social systems: Structure and temporal stability of red fox (*Vulpes vulpes*) groups. *PloS one*, *14*(9), e0220792.

Dunbar, R. I. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews*, 6(5), 178-190. https://doi.org/10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8

Easton, A. (2005). Behavioural flexibility, social learning, and the frontal cortex. In A. Easton & N. J. Emery (Eds.), *The cognitive neuroscience of social behaviour* (pp. 59–80). New York: Psychology Press.

Ehrenstein, W. H., & Ehrenstein, A. (1999). Psychophysical methods. In *Modern techniques in neuroscience research* In U. Windhorst & H. Johansson (Eds.), *Modern techniques in neuroscience research* (pp. 1211-1241). Berlin: Springer.

Fast, R., Schütt, T., Toft, N., Møller, A., & Berendt, M. (2013). An observational study with long-term follow-up of canine cognitive dysfunction: Clinical characteristics, survival, and risk factors. *Journal of Veterinary Internal Medicine, 27*(4), 822-829. https://doi.org/10.1111/jvim.12109

Fish, J. L., & Lockwood, C. A. (2003). Dietary constraints on encephalization in primates. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, *120*(2), 171-181.

Fiset, S., Beaulieu, C., & Landry, F. (2003). Duration of dogs' (*Canis familiaris*) working memory in search for disappearing objects. *Animal Cognition*, *6*(1), 1-10. https://doi.org/10.1007/s10071-002-0157-4

Fiset, S., Gagnon, S., & Beaulieu, C. (2000). Spatial encoding of hidden objects in dogs (*Canis familiaris*). *Journal of Comparative Psychology*, *114*(4), 315. https://doi.org/10.1037/0735-7036.114.4.315

Fox, M. W. (1971). Integrative development of brain and behavior in the dog. Chicago: University of Chicago Press.

Freedman, M., & Oscar-Berman, M. (1989). Spatial and visual learning deficits in Alzheimer's and Parkinson's disease. *Brain and cognition*, *11*(1), 114-126. https://doi.org/10.1016/0278-2626(89)90009-2

Gable, T. D., Windels, S. K., Bruggink, J. G., & Barber-Meyer, S. M. (2018). Weekly summer diet of gray wolves (*Canis lupus*) in northeastern Minnesota. *The American Midland Naturalist*, *179*(1), 15-27. https://doi.org/10.1674/0003-0031-179.1.15

Gagnon, S., & Doré, F. Y. (1994). Cross-sectional study of object permanence in domestic puppies (Canis familiaris). *Journal of Comparative Psychology*, *108*(3), 220. https://doi.org/10.1037/0735-7036.108.3.220

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental psychology*, *40*(2), 177.

Gese, E. M., Ruff, R. L., & Crabtree, R. L. (1996). Foraging ecology of coyotes (*Canis latrans*): the influence of extrinsic factors and a dominance hierarchy. Canadian Journal of Zoology, 74(5), 769-783. https://doi.org/10.1139/z96-089

Gilbert-Norton, L. B., Shahan, T. A., & Shivik, J. A. (2009). Coyotes (*Canis latrans*) and the matching law. *Behavioural processes*, 82(2), 178-183. https://doi.org/10.1016/j.beproc.2009.06.005

Gilmore, K. M., & Greer, K. A. (2015). Why is the dog an ideal model for aging research?. *Experimental gerontology*, *71*, 14-20. https://doi.org/10.1016/j.exger.2015.08.008

Gonzalez, R. C., Behrend, E. R., & Bitterman, M. E. (1967). Reversal learning and forgetting in bird and fish. *Science*, *158*(3800), 519-521. https://doi.org/10.1126/science.158.3800.519

González-Martínez, Á., Rosado, B., Pesini, P., García-Belenguer, S., Palacio, J., Villegas, A., ... & Sarasa, M. (2013). Effect of age and severity of cognitive dysfunction on two simple tasks in pet dogs. *The Veterinary Journal*, *198*(1), 176-181. https://doi.org/10.1016/j.tvjl.2013.07.004

Greer, K. A., Canterberry, S. C., & Murphy, K. E. (2007). Statistical analysis regarding the effects of height and weight on life span of the domestic dog. *Research in veterinary science*, *82*(2), 208-214. https://doi.org/10.1016/j.rvsc.2006.06.005

Gunter, L. (2018). *Understanding the impacts of breed identity, post-adoption and fostering interventions, & behavioral welfare of shelter dogs*. Arizona State University.

Hager, R. (2005). Foreword. In C. B. Jones (Eds.), *Behavioral flexibility in primates: Causes and consequences* (pp. XI). New York: Springer-Verlag.

Hämäläinen, A., Dammhahn, M., Aujard, F., Eberle, M., Hardy, I., Kappeler, P. M., ... & Kraus, C. (2014). Senescence or selective disappearance? Age trajectories of body mass in wild and captive populations of a small-bodied primate. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1791), 20140830. https://doi.org/10.1098/rspb.2014.0830

Harada, C. N., Love, M. C. N., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in geriatric medicine*, *29*(4), 737-752. https ://doi.org/10.1016/j.cger.2013.07.002

Harrington, F. H., Paquet, P. C., Ryon, J., & Fentress, J. C. (1982). Monogamy in wolves: a review of the evidence. In F. H. Harrington & P. C. Paquet (Eds.), *Wolves of the World* (pp. 209-222). Park Ridge, N.J: Noyes Publications.

Head, E. (2001). Brain aging in dogs: parallels with human brain aging and Alzheimer's disease. *Veterinary Therapeutics: Research in Applied Veterinary Medicine*, 2(3), 247-260.

Head, E. (2013). A canine model of human aging and Alzheimer's disease. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease,* 1832(9), 1384-1389. https://doi.org/10.1016/j.bbadis.2013.03.016

Head, E., Callahan, H., Muggenburg, B. A., Cotman, C. W., & Milgram, N. W. (1998). Visual-discrimination learning ability and β-amyloid accumulation in the dog. *Neurobiology of aging*, *19*(5), 415-425. https://doi.org/10.1016/S0197-4580(98)00084-0

Head, E., Mehta, R., Hartley, J., Kameka, M., Cummings, B. J., Cotman, C. W., ... & Milgram, N. W. (1995). Spatial learning and memory as a function of age in the dog. *Behavioral neuroscience*, *109*(5), 851.

Hendry, A. P. (2016). Key questions on the role of phenotypic plasticity in eco-evolutionary dynamics. *Journal of Heredity*, *107*(1), 25-41. https://doi.org/10.1093/jhered/esv060

Hennessy, C. A., Dubach, J., & Gehrt, S. D. (2012). Long-term pair bonding and genetic evidence for monogamy among urban coyotes (*Canis latrans*). *Journal of Mammalogy*, *93*(3), 732-742. https://doi.org/10.1644/11-MAMM-A-184.1

Hernández, L., Delibes, M., & Hiraldo, F. (1994). Role of reptiles and arthropods in the diet of coyotes in extreme desert areas of northern Mexico. *Journal of Arid Environments*, *26*(2), 165-170. https://doi.org/10.1006/jare.1994.1020

Hill, E. P., Sumner, P. W., & Wooding, J. B. (1987). Human influences on range expansion of coyotes in the southeast. *Wildlife Society Bulletin (1973-2006)*, *15*(4), 521-524.

Hockman, J. G., & Chapman, J. A. (1983). Comparative feeding habits of red foxes (*Vulpes vulpes*) and gray foxes (*Urocyon cinereoargenteus*) in Maryland. *American Midland Naturalist*, 276-285.

Hody, J. W., & Kays, R. (2018). Mapping the expansion of coyotes (*Canis latrans*) across North and Central America. *ZooKeys*, (759), 81-97. https://doi.org/10.3897/zookeys.759.15149

Holyan, J., Boyd, D. K., Mack, C. M., & Pletscher, D. H. (2005). Longevity and productivity of three wolves, *Canis lupus*, in the wild. *The Canadian field-naturalist*, *119*(3), 447-448.

Hultsch, D. F., & MacDonald, S. W. (2004). Intraindividual variability in performance as a theoretical window onto cognitive aging. *New frontiers in cognitive aging*, 65-88. In R. A. Dixon, L. Bäckman, & L. G. Nilsson (Eds.), *New frontiers in cognitive aging* (pp. 65–88). New York, NY: Oxford University Press.

Hunter, W. S. (1913). The delayed reaction in animals and children. *Behavior Monograhs*, *2*, 1-85. https://doi.org/10.1037/0735-7044.109.5.851

Inoue, M., Kwan, N. C., & Sugiura, K. (2018). Estimating the life expectancy of companion dogs in Japan using pet cemetery data. *Journal of Veterinary Medical Science*, 17-0384.

Iossa, G., Soulsbury, C. D., Baker, P. J., & Harris, S. (2008). Body mass, territory size, and life-history tactics in a socially monogamous canid, the red fox (*Vulpes vulpes*). *Journal of Mammalogy*, *89*(6), 1481-1490.

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: an updated perspective. *Neuroscience*, *345*, 12-26. https://doi.org/10.1016/j.neuroscience.2016.03.021

Jensen, P. (Ed.). (2007). The behavioural biology of dogs. Oxfordshire: CAB International.

Jędrzejewski, W., & Jędrzejewska, B. (1992). Foraging and diet of the red fox (*Vulpes vulpes*) in relation to variable food resources in Biatowieza National Park, Poland. *Ecography*, *15*(2), 212-220.

Jones, C. B. (2005). Behavioral flexibility: interpretations and prospects. In *Behavioral Flexibility in Primates: Causes and Consequences* (pp. 123-138). Springer, Boston, MA.

Kaeberlein, M., Creevy, K. E., & Promislow, D. E. (2016). The dog aging project: translational geroscience in companion animals. *Mammalian genome*, 27(7-8), 279-288. https://doi.org/10.1007/s00335-016-9638-7

Kensinger, E. A., Shearer, D. K., Locascio, J. J., Growdon, J. H., & Corkin, S. (2003). Working memory in mild Alzheimer's disease and early Parkinson's disease. *Neuropsychology*, *17*(2), 230.

Kleiman, D.G., & Brady, C.A. (1978). Coyote behavior in the context of recent canid research: problems and perspectives. In M. Bekoff (Eds.), *Coyotes: biology, behavior, and management* (pp. 163-188). New York, NY: Academic Press Inc.

Kraus, C., Pavard, S., & Promislow, D. E. (2013). The size–life span trade-off decomposed: why large dogs die young. *The American Naturalist*, *181*(4), 492-505. https://doi.org/10.1086/669665

Kuijper, D. P. J., Churski, M., Trouwborst, A., Heurich, M., Smit, C., Kerley, G. I. H., & Cromsigt, J. P. G. M. (2019). Keep the wolf from the door: How to conserve wolves in Europe's human-dominated landscapes? *Biological Conservation*, *235*, 102-111.

Lai, Z. C., Moss, M. B., Killiany, R. J., Rosene, D. L., & Herndon, J. G. (1995). Executive system dysfunction in the aged monkey: spatial and object reversal learning. *Neurobiology of aging*, *16*(6), 947-954. https://doi.org/10.1016/0197-4580(95)02014-4

Landsberg, G. M., DePorter, T., & Araujo, J. A. (2011). Clinical signs and management of anxiety, sleeplessness, and cognitive dysfunction in the senior pet. *Veterinary Clinics: Small Animal Practice*, *41*(3), 565-590.

Landsberg, G. M., Nichol, J., & Araujo, J. A. (2012). Cognitive dysfunction syndrome: a disease of canine and feline brain aging. *Veterinary Clinics: Small Animal Practice*, *42*(4), 749-768. https ://doi.org/10.1016/j.cvsm.2012.04.003

Lehman, N., Clarkson, P., Mech, L. D., Meier, T. J., & Wayne, R. K. (1992). A study of the genetic relationships within and among wolf packs using DNA fingerprinting and mitochondrial DNA. *Behavioral Ecology and Sociobiology*, *30*(2), 83-94. https://doi.org/10.1007/BF00173944

Lingle, S. (2000). Seasonal variation in coyote feeding behaviour and mortality of white-tailed deer and mule deer. *Canadian Journal of Zoology*, *78*(1), 85-99. https://doi.org/10.1139/z99-171

Lloyd, H.G. (1980). Habitat Requirements of the Red Fox. In E. Zimen (Eds), *The Red Fox*. Dordrecht: Springer. https://doi.org/10.1007/978-94-017-5592-4_2

Lombardi, J. V., Comer, C. E., Scognamillo, D. G., & Conway, W. C. (2017). Coyote, fox, and bobcat response to anthropogenic and natural landscape features in a small urban area. *Urban Ecosystems*, *20*(6), 1239-1248.

Lord, K., Feinstein, M., Smith, B., & Coppinger, R. (2013). Variation in reproductive traits of members of the genus Canis with special attention to the domestic dog (*Canis familiaris*). *Behavioural processes*, 92, 131-142. https://doi.org/10.1016/j.beproc.2012.10.009

Lowry, H., Lill, A., & Wong, B. B. (2013). Behavioural responses of wildlife to urban environments. *Biological reviews*, *88*(3), 537-549.

Macdonald, D. W., & Carr, G. M. (1995). Variation in dog society: between resource dispersion and social flux. In J. Serpell (Ed.), *The domestic dog: its evolution, behaviour and interactions with people* (pp. 199-216). Cambridge: Cambridge University Press.

Mackintosh, N. J. (1969). Habit-reversal and probability learning: Rats, birds and fish. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning* (pp. 175–184). New York: Academic Press Inc.

Mackintosh, N. J. (1974). The psychology of animal learning. New York: Academic Press Inc.

Mackintosh, N. J., & Mackintosh N. J., (1964). The Effect of Overtraining on a Nonreversal Shift in Octopus. *The Journal of Genetic Psychology*, 106, 373-377. https://doi.org/10.1080/00221325.1964.10533071

Mackintosh, N. J., Mcgonigle, B., & Holgate, V. (1968). Factors underlying improvement in serial reversal learning. *Canadian Journal of Psychology/Revue canadienne de psychologie, 22*(2), 85–95. https://doi.org/10.1037/h0082753

MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., ... & Zhao, Y. (2014). The evolution of self-control. *Proceedings of the National Academy of Sciences*, *111*(20), E2140-E2148. https://doi.org/10.1073/pnas.1323533111

MacNulty, D. R., Tallian, A., Stahler, D. R., & Smith, D. W. (2014). Influence of group size on the success of wolves hunting bison. *PloS one*, *9*(11), e112884.

Madari, A., Farbakova, J., Katina, S., Smolek, T., Novak, P., Weissova, T., ... & Zilka, N. (2015). Assessment of severity and progression of canine cognitive dysfunction syndrome using the CAnine DEmentia Scale (CADES). *Applied Animal Behaviour Science*, *171*, 138-145. https://doi.org/10.1016/j.appla nim.2015.08.034

Manrique, H. M., & Call, J. (2015). Age-dependent cognitive inflexibility in great apes. *Animal Behaviour*, *102*, 1-6. https://doi.org/10.1016/j.anbehav.2015.01.002

Marinissen, E. J., Singh, A., Glotter, D., Esposito, M., Carulli, J. M., Nahar, A., ... & Portelli, C. (2010, March). Adapting to adaptive testing. In *2010 Design, Automation & Test in Europe Conference & Exhibition (DATE 2010)* (pp. 556-561). IEEE.

Markovchick-Nicholls, L. I. S. A., Regan, H. M., Deutschman, D. H., Widyanata, A., Martin, B., Noreke, L., & Ann Hunt, T. I. M. O. T. H. Y. (2008). Relationships between human disturbance and wildlife land use in urban habitat fragments. *Conservation Biology*, *22*(1), 99-109.

Marshall-Pescini, S., Cafazzo, S., Viranyi, Z., & Range, F. (2017). Integrating social ecology in explanations of wolf–dog behavioral differences. *Current Opinion in Behavioral Sciences*, *16*, 80-86.

Marshall-Pescini, S., Schwarz, J. F., Kostelnik, I., Virányi, Z., & Range, F. (2017). Importance of a species' socioecology: Wolves outperform dogs in a conspecific cooperation task. *Proceedings of the National Academy of Sciences*, *114*(44), 11793-11798. https://doi.org/10.1073/pnas.1709027114

Marshall-Pescini, S., Virányi, Z., & Range, F. (2015). The effect of domestication on inhibitory control: wolves and dogs compared. *PloS one*, 10(2), e0118469. https://doi.org/10.1371/journal.pone.0118469

Massolo, A., & Meriggi, A. (1998). Factors affecting habitat occupancy by wolves in northern Apennines (northern Italy): a model of habitat suitability. *Ecography*, *21(2)*, 97-107.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305-315. https://doi.org/10.1016/j.jml.2017.01.001

McDowd, J. M., & Oseas-Kreger, D. M. (1991). Aging, inhibitory processes, and negative priming. *Journal of Gerontology*, *46*(6), P340-P345. https://doi.org/10.1093/geronj/46.6.P340

Mech, L. D. (1970). The Wolf: Ecology and Social Behavior of an Endangered Species. New York, NY: Natural History Press.

Mech, L. D. (1988). Longevity in wild wolves. *Journal of Mammalogy*, *69*(1), 197-198.

Mech, L. D. (2017). Where can wolves live and how can we live with them?. *Biological conservation*, *210*, 310-317.

Mech, L. D., & Boitani, L. (Eds.). (2007). *Wolves: behavior, ecology, and conservation*. Chicago: University of Chicago Press.

Mech, L. D., & Boitani, L. (Eds.). (2010). *Wolves: behavior, ecology, and conservation*. Chicago: University of Chicago Press.

Mech, L. D., Smith, D. W., & MacNulty, D. R. (2021). *Wolves on the hunt*. Chicago: University of Chicago Press.

Mettler, A. E., & Shivik, J. A. (2007). Dominance and neophobia in coyote (*Canis latrans*) breeding pairs. *Applied Animal Behaviour Science*, *102*(1-2), 85-94. https://doi.org/10.1016/j.applanim.2006.03.012

Metz, M. C., Vucetich, J. A., Smith, D. W., Stahler, D. R., & Peterson, R. O. (2011). Effect of sociality and season on gray wolf (*Canis lupus*) foraging behavior: implications for estimating summer kill rate. *PloS one*, *6*(3), e17332.

Mikhalevich, I., Powell, R., & Logan, C. (2017). Is behavioural flexibility evidence of cognitive complexity? How evolution can inform comparative cognition. *Interface focus*, 7(3), 20160121. https://doi.org/10.1098/rsfs.2016.0121

Milgram, N. W., Adams, B., Callahan, H., Head, E., Mackay, B., Thirlwell, C., & Cotman, C. W. (1999). Landmark discrimination learning in the dog. *Learning & Memory*, *6*(1), 54-61. https://doi.org/10.1101/lm.6.1.54

Milgram, N. W., Head, E., Weiner, E., & Thomas, E. (1994). Cognitive functions and aging in the dog: Acquisition of nonspatial visual tasks. *Behavioral Neuroscience*, *108*(1), 57–68. https://doi.org/10.1037/0735-7044.108.1.57

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

Moehlman, P. D. (1986). Ecology of cooperation in canids. In D. I. Rubenstein & R. W. Wrangham (Eds.), *Ecological aspects of social evolution, birds and mammals* (pp. 64-86). Princeton, NJ: Princeton University Press.

Moehlman, P. D. (1989). Intraspecific variation in canid social systems. In J. L. Gittleman (Ed.), *Carnivore behavior, ecology, and evolution* (pp. 143-163). Boston: Springer. https://doi.org/10.1007/978-1-4757-4716-4

Mongillo, P., Araujo, J. A., Pitteri, E., Carnier, P., Adamelli, S., Regolin, L., & Marinelli, L. (2013). Spatial reversal learning is impaired by age in pet dogs. *Age*, *35*(6), 2273-2282. https://doi.org/10.1007/s11357-013-9524-0

Mongillo, P., Scandurra, A., D'Aniello, B., & Marinelli, L. (2017). Effect of sex and gonadectomy on dogs' spatial performance. *Applied Animal Behaviour Science*, *191*, 84-89. https://doi.org/10.1016/j.appla nim.2017.01.017

MoMozawa, Y. (2019). The potential of translational research in dogs in human medicine. *Translational and Regulatory Sciences*, *1*(1), 31-36.

Morey, P. S., Gese, E. M., & Gehrt, S. (2007). Spatial and temporal variation in the diet of coyotes in the Chicago metropolitan area. *The American Midland Naturalist*, 158(1), 147-161. https://doi.org/10.1674/0003-0031(2007)158[147:SATVIT]2.0.CO;2

Nelson, T. A., & Lloyd, D. M. (2005). Demographics and condition of coyotes in Illinois. *The American midland naturalist*, *153*(2), 418-427. https://doi.org/10.1674/0003-0031(2005)153[0418:DACOCI]2.0.CO;2

Nicholson, W. S., Hill, E. P., & Briggs, D. (1985). Denning, pup-rearing, and dispersal in the gray fox in east-central Alabama. *The Journal of wildlife management*, *49*(1), 33-37.

Osella, M. C., Re, G., Odore, R., Girardi, C., Badino, P., Barbero, R., & Bergamasco, L. (2007). Canine cognitive dysfunction syndrome: prevalence, clinical signs and treatment with a neuroprotective nutraceutical. *Applied Animal Behaviour Science*, *105*(4), 297-310.

Packard, J. M. (2003). Wolf behaviour: reproductive, social and intelligent. In L. D. Mech (Ed.), *Wolves: behavior, ecology, and conservation* (pp. 35-65). Chicago: University of Chicago Press.

Pal, S. K. (2005). Parental care in free-ranging dogs, *Canis familiaris*. *Applied Animal Behaviour Science*, *90*(1), 31-47.

Pal, S. K. (2015). Factors influencing intergroup agonistic behaviour in free-ranging domestic dogs (*Canis familiaris*). *acta ethologica*, *18*(2), 209-220.

Paul, M., Sen Majumder, S., Sau, S., Nandi, A. K., & Bhadra, A. (2016). High early life mortality in free-ranging dogs is largely influenced by humans. *Scientific reports*, *6*(1), 1-8.

Pickering, S. J. (2001). The development of visuo-spatial working memory. *Memory*, *9*(4-6), 423-432.

Piotti, P., Szabó, D., Wallis, L., Bognár, Z., Stiegmann, B. S., Egerer, A., ... & Kubinyi, E. (2017). The effect of age on visuo-spatial short-term memory in family dogs. *Pet Behaviour Science*, (4), 17-19. https://doi.org/10.21071/pbs.v0i4.10130

Piotti, P., Szabó, D., Bognár, Z., Egerer, A., Hulsbosch, P., Carson, R. S., & Kubinyi, E. (2018). Effect of age on discrimination learning, reversal learning, and cognitive bias in family dogs. *Learning & behavior*, *46*(4), 537-553. https://doi.org/10.3758/s13420-018-0357-7

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, *32*(1), 3-25.

Mühlhoff, N., Stevens, J. R., & Reader, S. M. (2011). Spatial discounting of food and social rewards in guppies (*Poecilia reticulata*). *Frontiers in psychology*, *2*, 68.

Range, F., & Virányi, Z. (2015). Tracking the evolutionary origins of dog-human cooperation: the “Canine Cooperation Hypothesis”. *Frontiers in psychology*, *5*, 1582.

Rathbun, A. P., Wells, M. C., & Bekoff, M. (1980). Cooperative predation by coyotes on badgers. *Journal of Mammalogy*, *61*(2), 375-376. https://doi.org/10.2307/1380074

Rayburn-Reeves, R. M., Laude, J. R., & Zentall, T. R. (2013). Pigeons show near-optimal win-stay/lose-shift performance on a simultaneous-discrimination, midsession reversal task with short intertrial intervals. *Behavioural Processes*, 92, 65-70. https://doi.org/10.1016/j.beproc.2012.10.011

Reader, S. M., & MacDonald, K. (2003). *Environmental variability and primate behavioural flexibility*. In S. M. Reader & K. N. Laland (Eds.), *Animal innovation* (pp. 83–116). Oxford: Oxford University Press.

Reisberg, B., Borenstein, J., Salob, S. P., Ferris, S. H., et al. (1987). Behavioral symptoms in Alzheimer's disease: Phenomenology and treatment. *The Journal of Clinical Psychiatry, 48*(5, Suppl), 9–15.

Ridderinkhof, K. R., Van Den Wildenberg, W. P., Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and cognition*, *56*(2), 129-140. https://doi.org/10.1016/j.bandc.2004.09.016

Rofina, J. E., Van Ederen, A. M., Toussaint, M. J. M., Secreve, M., Van Der Spek, A., Van Der Meer, I., ... & Gruys, E. (2006). Cognitive disturbances in old dogs suffering from the canine counterpart of Alzheimer's disease. *Brain research*, *1069*(1), 216-226.

Rowan, A., & Kartal, T. (2018). Dog population & dog sheltering trends in the United States of America. *Animals*, *8*(5), 68.

Rumbaugh, D. M., Savage-Rumbaugh, E. S., & Washburn, D. A. (1996). Toward a new outlook on primate learning and behavior: complex learning and emergent processes in comparative perspective 1. *Japanese Psychological Research*, *38*(3), 113-125. https://doi.org/10.1111/j.1468-5884.1996.tb00016.x

Sacks, B. N., & Neale, J. C. (2002). Foraging strategy of a generalist predator toward a special prey: coyote predation on sheep. *Ecological Applications*, *12*(1), 299-306. https://doi.org/10.1890/1051-0761(2002)012[0299:FSOAGP]2.0.CO;2

Salvin, H. E., McGreevy, P. D., Sachdev, P. S., & Valenzuela, M. J. (2010). Under diagnosis of canine cognitive dysfunction: a cross-sectional survey of older companion dogs. *The veterinary journal*, *184*(3), 277-281. https ://doi.org/10.1016/j.tvjl.2009.11.007

Salvin, H. E., McGreevy, P. D., Sachdev, P. S., & Valenzuela, M. J. (2012). The effect of breed on age-related changes in behavior and disease prevalence in cognitively normal older community dogs, *Canis lupus familiaris*. *Journal of Veterinary Behavior*, *7*(2), 61-69. https ://doi.org/10.1016/j.jveb.2011.06.002

Santana, E., & Armstrong, J. (2017). Food habits and anthropogenic supplementation in coyote diets along an urban-rural gradient. *Human–Wildlife Interactions*, *11*(2), 6. https://doi.org/10.26077/vhdx-1033

Sayol, F., Sol, D., & Pigot, A. L. (2020). Brain size and life history interact to predict urban tolerance in birds. *Frontiers in Ecology and Evolution*, 8, 58. https://doi.org/10.3389/fevo.2020.00058

Schipper, J., Chanson, J., Chiozza, F., Cox, N., Hoffmann, M., Katariya, V., ... & Smith, A. T. (2008). The biogeography of diversity, threat, and knowledge in the world's terrestrial and aquatic mammals. *Science*, *322*(225-230).

Schütt, T., Toft, N., & Berendt, M. (2015). A comparison of 2 screening questionnaires for clinical assessment of canine cognitive dysfunction. *Journal of Veterinary Behavior*, *10*(6), 452-458.

Sen Majumder, S., Bhadra, A., Ghosh, A., Mitra, S., Bhattacharjee, D., Chatterjee, J., ... & Bhadra, A. (2014). To be or not to be social: foraging associations of free-ranging dogs in an urban ecosystem. *acta ethologica*, *17*(1), 1-8. https://doi.org/10.1007/s10211-013-0158-0

Sen Majumder, S., Paul, M., Sau, S., & Bhadra, A. (2016). Denning habits of free-ranging dogs reveal preference for human proximity. *Scientific reports*, *6*(1), 1-8.

Schielzeth, H., & Forstmeier, W. (2009). Conclusions beyond support: overconfident estimates in mixed models. *Behavioral Ecology*, *20*(2), 416-420. https://doi.org/10.1093/beheco/arn145

Shettleworth, S. J. (1998). *Cognition, evolution, and behavior*. Oxford: Oxford University Press.

Shettleworth, S. J. (2009). *Cognition, evolution, and behavior*. Oxford: Oxford University Press.

Shettleworth, S. J. (2010). *Cognition, evolution and behavior* (2nd ed.). New York, NY: Oxford University Press.

Simone, P. M., & Baylis, G. C. (1997). Selective attention in a reaching task: effect of normal aging and Alzheimer's disease. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(3), 595.

Siwak-Tapp, C. T., Head, E., Muggenburg, B. A., Milgram, N. W., & Cotman, C. W. (2007). Neurogenesis decreases with age in the canine hippocampus and correlates with cognitive function. *Neurobiology of learning and memory*, *88*(2), 249-259. https://doi.org/10.1016/j.nlm.2007.05.001

Smyth, M. M., & Scholey, K. A. (1994). Interference in immediate spatial memory. *Memory & Cognition*, *22*(1), 1-13. https://doi.org/10.3758/BF03202756

Snell-Rood, E. C. (2013). An overview of the evolutionary causes and consequences of behavioural plasticity. *Animal Behaviour*, 85(5), 1004-1011. https://doi.org/10.1016/j.anbehav.2012.12.031

Sol, D., Duncan, R. P., Blackburn, T. M., Cassey, P., & Lefebvre, L. (2005). Big brains, enhanced cognition, and response of birds to novel environments. *Proceedings of the National Academy of Sciences*, 102(15), 5460-5465. https://doi.org/10.1073pnas.0408145102

Solt, F., & Hu, Y. (2019). interplot: Plot the effects of variables in interaction terms (R package version 0.2. 2). Retrieved from https://CRAN.R-project.org/package=interplot

Soulsbury, C. D., Baker, P. J., Iossa, G., & Harris, S. (2008). Fitness costs of dispersal in red foxes (*Vulpes vulpes*). *Behavioral Ecology and Sociobiology*, *62*(8), 1289-1298.

Stanton, L. A., Bridge, E. S., Huizinga, J., Johnson, S. R., Young, J. K., & Benson-Amram, S. (2020). Variation in reversal learning by three generalist mesocarnivores. *Animal Cognition*, 1-14. https://doi.org/10.1007/s10071-020-01438-4

Statham, M. J., Sacks, B. N., Aubry, K. B., Perrine, J. D., & Wisely, S. M. (2012). The origin of recently established red fox populations in the United States: translocations or natural range expansions?. *Journal of Mammalogy*, *93*(1), 52-65.

Strang, C. G., & Sherry, D. F. (2014). Serial reversal learning in bumblebees (*Bombus impatiens*). *Animal Cognition*, 17(3), 723-734. https://doi.org/10.1007/s10071-013-0704-1

Studzinski, C. M., Christie, L. A., Araujo, J. A., Burnham, W. M., Head, E., Cotman, C. W., & Milgram, N. W. (2006). Visuospatial function in the beagle dog: an early marker of cognitive decline in a model of human aging and dementia. *Neurobiology of learning and memory*, *86*(2), 197-204. https://doi.org/10.1016/j.nlm.2006.02.005

Szabó, D., Gee, N. R., & Miklósi, Á. (2016). Natural or pathologic? Discrepancies in the study of behavioral and cognitive signs in aging family dogs. *Journal of Veterinary Behavior*, *11*, 86-98.

Szabó, D., Miklósi, Á., & Kubinyi, E. (2018). Owner reported sensory impairments affect behavioural signs associated with cognitive decline in dogs. *Behavioural processes*, *157*, 354-360. https://doi.org/10.1016/j.beproc.2018.07.013

Tapp, P. D., Siwak, C. T., Estrada, J., Head, E., Muggenburg, B. A., Cotman, C. W., & Milgram, N. W. (2003). Size and reversal learning in the beagle dog as a measure of executive function and inhibitory control in aging. *Learning & Memory*, *10*(1), 64-73. https://doi.org/10.1101/lm.54403

Tapp, P. D., Siwak, C. T., Estrada, J., Holowachuk, D., & Milgram, N. W. (2003). Effects of age on measures of complex working memory span in the beagle dog (*Canis familiaris*) using two versions of a spatial list learning paradigm. *Learning & Memory*, *10*(2), 148-160. https://doi.org/10.1101/lm.56503

Tapp, P. D., Siwak, C. T., Gao, F. Q., Chiou, J. Y., Black, S. E., Head, E., ... & Su, M. Y. (2004). Frontal lobe volume, function, and β-amyloid pathology in a canine model of aging. *Journal of Neuroscience*, *24*(38), 8205-8213.

Temple, D. L., Chamberlain, M. J., & Conner, L. M. (2010). Spatial ecology, survival and cause-specific mortality of gray foxes (*Urocyon cinereoargenteus*) in a longleaf pine ecosystem. *The American Midland Naturalist*, *163*(2), 413-422.

Theberge, J., & Theberge, M. (2013). *Wolf country: eleven years tracking the Algonquin wolves*. Toronto: McClelland & Stewart.

Thornton, D., Scully, A., King, T., Fisher, S., Fitkin, S., & Rohrer, J. (2018). Hunting associations of American badgers (*Taxidea taxus*) and coyotes (*Canis latrans*) revealed by camera trapping. *Canadian Journal of Zoology*, *96*(7), 769-773. https://doi.org/10.1139/cjz-2017-0234

Toweill, D. E., & Anthony, R. G. (1988). Coyote foods in a coniferous forest in Oregon. *The Journal of Wildlife Management*, 507-512.

Tucker, R. L., Jacobson, H. A., & Spencer, M. R. (1993). Territoriality and pairbonding of gray foxes in Mississippi. In *Proceeding of the Annual Conference of Southeastern Association of Fish and Wildlife Agencies* (Vol. 47, pp. 90-98).

Udell, M. A., & Wynne, C. D. (2010). Ontogeny and phylogeny: both are essential to human-sensitive behaviour in the genus Canis. *Animal Behaviour*, *79*(2), e9-e14.

Van Bourg, J., Gilchrist, R., & Wynne, C. D. L. (2020). Adaptive spatial working memory assessments for aging pet dogs. *Animal Cognition*, 1-21. https://doi.org/10.1007/s10071-020-01447-3

Van Bourg, J., Gunter, L. M., & Wynne, C. D. (2021). A Rapid Serial Reversal Learning Assessment for Age-related Cognitive Deficits in Pet Dogs. *Behavioural Processes*, 104375. https://doi.org/10.1016/j.beproc.2021.104375

Van Bourg, J., Young, J. K., Alkhalifah, R., Brummer, S., Johansson, E., Morton, J., Quintana, V., & Wynne, C. D. L. (2022). Cognitive flexibility and aging in coyotes (*Canis latrans*). *Journal of Comparative Psychology*, *136*(1), 54–67. https://doi.org/10.1037/com0000307

Vilà, C., & Jennifer, A. L. (2012). Canid phylogeny and origin of the domestic dog. In E. A. Ostrander & A. Ruvinsky (Eds.), *The genetics of the dog* (pp. 1-9). Oxfordshire: CABI.

Vite, C. H., & Head, E. (2014). Aging in the canine and feline brain. *Veterinary Clinics: Small Animal Practice*, *44*(6), 1113-1129. https ://doi.org/10.1016/j.cvsm.2014.07.008

Wainer, H., & Lewis, C. (1990). Toward a psychometrics for testlets. *Journal of Educational Measurement*, *27*(1), 1-14. https://doi.org/10.1111/j.1745-3984.1990.tb00730.x

Wallert, J., Ekman, U., Westman, E., & Madison, G. (2017). The worst performance rule with elderly in abnormal cognitive decline. *Intelligence*, *64*, 9-17. https://doi.org/10.1016/j.intell.2017.06.003

Wallert, J., Westman, E., Ulinder, J., Annerstedt, M., Terzis, B., & Ekman, U. (2018). Differentiating patients at the memory clinic with simple reaction time variables: a predictive modeling approach using support vector machines and bayesian optimization. *Frontiers in aging neuroscience*, *10*, 144. https://doi.org/10.3389/fnagi.2018.00144

Wallis, L. J., Range, F., Müller, C. A., Serisier, S., Huber, L., & Virányi, Z. (2014). Lifespan development of attentiveness in domestic dogs: drawing parallels with humans. *Frontiers in Psychology*, *5*, 71. https://doi.org/10.3389/fpsyg.2014.00071

Wallis, L. J., Virányi, Z., Müller, C. A., Serisier, S., Huber, L., & Range, F. (2016). Aging effects on discrimination learning, logical reasoning and memory in pet dogs. *Age*, *38*(1), 1-18. https://doi.org/10.1007/s11357-015-9866-x

Walton, A. C. (1915). The influence of diverting stimuli during delayed reaction in dogs. *Journal of Animal Behavior*, *5*(4), 259. https://doi.org/10.1037/h0074449

Ward, J. N., Hinton, J. W., Johannsen, K. L., Karlin, M. L., Miller, K. V., & Chamberlain, M. J. (2018). Home range size, vegetation density, and season influences prey use by coyotes (*Canis latrans*). *PloS one*, *13*(10), e0203703. https://doi.org/10.1371/journal.pone.0203703

Ward, J. M., Youssef, S. A., & Treuting, P. M. (2016). Why animals die: an introduction to the pathology of aging. *Veterinary pathology*, *53*(2), 229-232. https://doi.org/10.1177/0300985815612151

Warren, J. M. (1966). Reversal learning and the formation of learning sets by cats and rhesus monkeys. *Journal of Comparative and Physiological Psychology*, *61*(3), 421–428. https://doi.org/10.1037/h0023269

Watowich, M. M., MacLean, E. L., Hare, B., Call, J., Kaminski, J., Miklósi, Á., & Snyder-Mackler, N. (2020). Age influences domestic dog cognitive performance independent of average breed lifespan. *Animal cognition*. https://doi.org/10.1007/s10071-020-01385-0

Watson, A. B., & Fitzhugh, A. (1990). The method of constant stimuli is inefficient. *Perception & psychophysics*, *47*(1), 87-91. https://doi.org/10.3758/BF03208169

Way, J. G., & Strauss, E. G. (2004). Old-aged coyote in an urbanised landscape. *Canid News*, *7*, 1-3.

Wells, M. C., & Lehner, P. N. (1978). The relative importance of the distance senses in coyote predatory behaviour. *Animal Behaviour*, *26*, 251-258. https://doi.org/10.1016/0003-3472(78)90025-8

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological bulletin*, *120*(2), 272. https://doi.org/10.1111/jvim.12109

White, P. C., & Harris, S. (1994). Encounters between red foxes (*Vulpes vulpes*): implications for territory maintenance, social cohesion and dispersal. *Journal of Animal Ecology*, 315-327.

Windberg, L. A. (1996). Coyote responses to visual and olfactory stimuli related to familiarity with an area. *Canadian Journal of Zoology*, *74*(12), 2248-2253. https://doi.org/10.1139/z96-255

Wright, T. F., Eberhard, J. R., Hobson, E. A., Avery, M. L., & Russello, M. A. (2010). Behavioral flexibility and species invasions: the adaptive flexibility hypothesis. *Ethology Ecology & Evolution*, *22*(4), 393-404. https://doi.org/10.1080/03949370.2010.505580

Young, J. K., Andelt, W. F., Terletzky, P. A., & Shivik, J. A. (2006). A comparison of coyote ecology after 25 years: 1978 versus 2003. *Canadian Journal of Zoology*, *84*(4), 573-582. https://doi.org/10.1139/z06-030

Young, S. P., & Goldman, E. A. (1944). *The wolves of North America: Part I. Their history, life habits, economic status, and control*. American Wildlife Institute.

Youssef, S. A., Capucchio, M. T., Rofina, J. E., Chambers, J. K., Uchida, K., Nakayama, H., & Head, E. (2016). Pathology of the aging brain in domestic and laboratory animals, and animal models of human neurodegenerative diseases. *Veterinary pathology*, *53*(2), 327-348. https://doi.org/10.1177/0300985815623997

Zuberbühler, K., & Janmaat, K. (2010). Foraging cognition in non-human primates. In M. L. Platt & A. A. Ghazanfar (Eds.), *Primate neuroethology* (pp. 64-83). Oxford: Oxford University Press.

# APPENDIX A

# TABLES

Table 2-1

*Example of Training Progression in Experiment I*

| Trial | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Level | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 |
| Choice | **L** | **R** | **L** | **L** | **L** | **R** | **L** | **L** | **L** | **R** | **L** | **R** |
| PCC | 1.0 | 1.0 | 1.0 | .50 | .67 | .75 | 0.0 | 0.0 | 0.33 | .50 | .60 | .67 |

PCC = proportion correct at the level attempted. L = left box. R = right boxes. Green letters indicate correct choices, red and underlined letters indicate incorrect choices

Table 2-2

*Example of Training Progression in Experiment II*

| Trial | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Level | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 2 | 2 | 3 | 3 | 3 | 3 |
| Choice | **L** | **R** | **L** | **R** | **r** | **L** | **R** | **r** | **r** | **l** | **R** | **L** | **R** | **R** | **L** |
| PCC | 1.0 | 1.0 | 1.0 | .50 | .67 | .75 | 0.0 | 0.0 | 0.0 | .80 | .83 | .25 | .40 | .50 | .57 |

PCC = proportion correct at the level attempted. L = left box. R = right boxes. Green letters indicate correct choices, red and underlined letters indicate incorrect choices. Lowercase letters indicate correction trials

Table 2-3

*Experiment II Analyses*

| Purpose | Analysis | Criterion | Predictors | *n* | Tab./Fig. |
|---|---|---|---|---|---|
| Task Acquisition & Retention | Binomial Tests | PCC in Zero-delay Trials [a] (one test per dog) | | 6-33 [b] | Fig. 2-6<br>Appendix C-4 |
| Task Acquisition & Retention | Binomial Tests | PCC in Delays Trials [a] (one test per dog) | | 8-29 [b] | Fig. 2-6<br>Appendix C-4 |
| Age-Dependent Performance | Quadratic Regression | CMS | AGE + AGE $^2$ | 23 | Fig. 2-7 |
| Delay-Dependent Performance [a] | Binomial GLMM | Trial Outcome | **AGE** + **AGE$^2$** + **Delay** + **Trial** + AGE*Delay + AGE$^2$*Delay + AGE*Trial + AGE$^2$*Trial + (1 + Delay + Trial \| Dog) | 696 | Tab. 2-4 |
| Head & Body Orientations | Binomial GLMM | Delay trial Outcome | H.C.Dl + B.C.Dl + H.F.Dl +B.F.Dl + H.F.Dm + B.F.Dm + (1 \| Dog) | 426 | Tab. 2-5 |
| | Binomial GLMM | Delay trial Outcome | AGE + AGE $^2$ + Delay + AGE*B.F.Dl + AGE$^2$*B.F.Dl + RSS B.F.Dl + (1 + B.F.Dl \| Dog) | 426 | Tab. 2-6 |
| Perseveration | Quadratic Regression | \|Side Bias\| | AGE + AGE $^2$ | 23 | Fig. 2-8 |
| Delay Optimization | 5 x Quadratic Regressions | PCC in the 15s, 30s, 45s, 60s, & 90s Delays | AGE + AGE $^2$ | 20-22 [c] | Tab. 2-7<br>Fig. 2-9 |

GLMM = generalized linear mixed model (logit error distribution). PCC = proportion or correct choices. CMS = cumulative memory score. Predictors in the final GLMM of trial outcome are in bold text. GLMM equation notation follows from the package lme4 in R. For head and body orientations, H = head, B = body, C = oriented towards the correct box, F = oriented forwards (towards the testing room), Dl = during the entire delay, Dm = during the entire demonstration

[a] binomial tests compared observed to expected number of correct choices in a given number of attempts

[b] number of trials differed among dogs

[c] number of dogs differed among delays

Table 2-4

*Delay Sensitivity – GLMM of Test-trial Outcome*

| **Model** | | Estimates | | | | | Random Effects | | | Model Fit | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Predictor | Type | ß | *SE* | ß *CI* | $X^2$ | *p* | $s^2$ *CI* | $X^2$ | *p* | *AIC* | *BIC* | $\frac{r^2}{r(df)}$ |
| **Full** | | | | | | | | | | 614 | 682 | .96 |
| AGE | F | 2.46 | 0.13 | [1.6,3.3] | -- | -- | -- | -- | -- | | | |
| AGE$^2$ | F | -2.25 | 0.51 | [-3.1,-1.4] | -- | -- | -- | -- | -- | | | |
| Delay | F,RS | -0.61 | 0.14 | [-.9,-.3] | -- | -- | [.01,.4] | 0.17 | .98 | | | |
| Delay* AGE | F | 0.81 | 0.54 | [-.02,1.8] | 2.2 | .69 | -- | -- | -- | | | |
| Delay *AGE$^2$ | F | -0.82 | 0.54 | [-1.8,.01] | 2.3 | .68 | -- | -- | -- | | | |
| Trial | F,RS | -0.10 | 0.13 | [-.35,.2] | -- | -- | [.01,.3] | 0.99 | .09 | | | |
| Trial *AGE | F | -1.57 | 0.47 | [-2.1,-.8] | 9.3 | .05 | -- | -- | -- | | | |
| Trial *AGE$^2$ | F | 1.55 | 0.47 | [0.8,2.6] | 8.9 | .06 | -- | -- | -- | | | |
| Intercept | I | 1.61 | 0.51 | [1.4,1.9] | -- | -- | -- | -- | -- | | | |
| Subject | RI | -- | -- | -- | -- | -- | [.01,.4] | 0.09 | .99 | | | |
| **Final** | | | | | | | | | | 608 | 635 | .92 |
| AGE | F | 2.78 | 0.87 | [1.5,4.4] | 11.7 | <.01 | -- | -- | -- | | | |
| AGE$^2$ | F | -2.58 | 0.86 | [-4.2,-1.3] | 10.5 | <.01 | -- | -- | -- | | | |
| Delay | F | -0.66 | 0.16 | [-.9,-.4] | 20.9 | <.01 | -- | -- | -- | | | |
| Trial | F | 0.20 | 0.17 | [-.04,.5] | 1.3 | .26 | -- | -- | -- | | | |
| Intercept | F | 1.86 | 0.20 | [1.5,2.3] | -- | -- | -- | -- | -- | | | |
| Subject | RI | -- | -- | -- | -- | -- | [0,.8] | 0.48 | .50 | | | |

F = fixed effect. RS = random slope. RI = random intercept. ß = estimate (predictor coefficient). $s^2$ 95% *CI* = confidence interval of the random effect variance (values near zero are not significant). *AIC* & *BIC* = Akaike & Bayesian Information Criterion (smaller values indicate better fit). $r^2/r(df)$ = sum of squared Pearson residuals divided by residuals degrees of freedom (deviations from 1.0 larger than |0.2| may indicate over/ underdispersion). Chi-square and *p* values were obtained from Likelihood Ratio Tests of the difference in total prediction between the full model and the nested model without the predictor. Profiled confidence intervals for ß were calculated using parametric bootstraps

Table 2-5

*Predictive Value of Orientation Behaviors – GLMM of Delay Trial Outcome*

| Dimension | Predictor | Levels | Group 1 (no) | | Group 2 (yes) | | LRT | |
|---|---|---|---|---|---|---|---|---|
| | | | *n* | PCC (*M* ± *SE*) | *n* | PCC (*M* ± *SE*) | $X^2$ | *p* |
| Rehearsal | Head correct (delay) | Correct Orientation (yes/no) | 454 | 0.78 ± 0.02 | 20 | 0.70 ± 0.11 | 1.11 | 0.29 |
| | Body correct (delay) | | 362 | 0.75 ± 0.02 | 112 | 0.85 ± 0.03 | 2.45 | 0.12 |
| Engagement | Head forward (delay) | Forward Orientation (yes/no) | 380 | 0.77 ± 0.02 | 94 | 0.77 ± 0.04 | 0.13 | 0.72 |
| | Body forward (delay) | | 113 | 0.65 ± 0.04 | 361 | 0.81 ± 0.02 | 8.46 | 0.004** |
| | Head forward (demonstration) | | 137 | 0.79 ± 0.04 | 337 | 0.77 ± 0.02 | 1.23 | 0.27 |
| | Body forward (demonstration) | | 74 | 0.74 ± 0.05 | 400 | 0.78 ± 0.02 | 0.29 | 0.59 |
| Individual variation | Subject | Random Intercepts | - | - | - | - | 1.34 | 0.25 |

Group 1: trials in which the dog did not maintain the corresponding orientation. Group 2: trials in which the dog did maintain the corresponding orientation. LRT = Likelihood Ratio Test of nested models. The full model included all orientation behaviors and the random intercept. n: number of trials in which the dog did or did not maintain the orientation. See Table 2-3 for remaining abbreviations

Table 2-6

*Engagement, Delay, and Age – GLMM of Delay-trial Outcome*

| **Model** | | Estimates | | | | | Random Effects | | | Model Fit | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Predictor | Type | ß | *SE* | ß *CI* | $X^2$ | *p* | $s^2$ *CI* | $X^2$ | *p* | *AIC* | *BIC* | $\underline{r^2}$<br>$r(df)$ |
| **Full** | | | | | | | | | | 502 | 543 | 0.95 |
| Age | F | 1.86 | 0.78 | [0.5,3.4] | -- | -- | -- | -- | -- | | | |
| Age$^2$ | F | -1.72 | 0.72 | [-3.1,-0.5] | -- | -- | -- | -- | -- | | | |
| B.F.Dl | F,RS | 0.76 | 0.28 | [0.1,1.3] | -- | -- | [0,0.8] | 0.92 | 0.63 | | | |
| B.F.Dl *AGE | F | -0.08 | 0.96 | [-2.1,1.6] | 0.01 | 0.94 | -- | -- | -- | | | |
| B.F.Dl *AGE$^2$ | F | -0.14 | 0.92 | [-1.8,1.7] | 0.02 | 0.88 | -- | -- | -- | | | |
| Delay | F | -0.26 | 0.14 | [-0.5,0] | 3.49 | 0.06 | -- | -- | -- | | | |
| Intercept | F | 0.78 | 0.22 | [0.4,1.3] | -- | -- | -- | | | | | |
| Subject | RI | -- | -- | -- | -- | -- | [0,0.7] | 0.19 | 0.91 | | | |

See Tables 2-3 and 2-5 for abbreviations

Table 2-7

*Regression of Age on PCC in Each Delay Attempted by at Least 20 Dogs*

| | | | Model Prediction | | AGE.CENT | | AGE.CENT$^2$ | |
|---|---|---|---|---|---|---|---|---|
| Delay (s) | *n* | $r^2$ | *F* | *p* | *t* | *p* | *t* | *p* |
| 15 | 22 | 0.20 | 2.35 | 0.12 | 2.12 | 0.05 | -2.16 | 0.04 |
| 30 | 22 | 0.15 | 1.63 | 0.22 | 1.80 | 0.09 | -1.77 | 0.09 |
| 45 | 22 | 0.33 | 4.67 | 0.02 | 0.33 | 0.75 | -3.06 | 0.01 |
| 60 | 22 | 0.35 | 5.11 | 0.02 | -1.31 | 0.21 | -2.71 | 0.01 |
| 90 | 20 | 0.28 | 3.38 | 0.06 | 0.90 | 0.38 | -2.57 | 0.02 |

Table 3-1

*Analysis of Trial Outcome*

| | **Coefficients** | | **LRT (Nested)** | | **Model Fit** | | | **LRT (vs. Null)** | | **LRT (vs. Full)** | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **ß** | ***SE*** | **$X^2$** | ***p*** | ***AIC*** | ***BIC*** | **$\frac{r^2}{r(df)}$** | **$X^2$** | ***p*** | **$X^2$** | ***p*** |
| **Null Model** | | | | | 3912 | 3918 | 1.004 | | | | |
| Intercept | 0.11 | 0.04 | -- | -- | | | | | | | |
| **Full Model** | | | | | 3898 | 3952 | 0.978 | 29.70 | <.01 | | |
| Age | 0.55 | 0.25 | 4.88 | <.05 | | | | | | | |
| $Age^2$ | -0.65 | 0.24 | 6.76 | <.01 | | | | | | | |
| Trial | 0.29 | 0.11 | 7.66 | <.01 | | | | | | | |
| REV | -0.45 | 0.13 | 12.39 | <.01 | | | | | | | |
| Height | 0.08 | 0.15 | 0.27 | .60 | | | | | | | |
| Weight | -0.19 | 0.15 | 1.55 | .21 | | | | | | | |
| Sex | -0.09 | 0.12 | 0.63 | .43 | | | | | | | |
| Intercept | 0.15 | 0.08 | -- | -- | | | | | | | |
| Rnd.Sbj.Int | -- | 0.07 | 6.96 | <.01 | | | | | | | |
| **Final Model** | | | | | 3898 | 3933 | 0.975 | 24.38 | <.01 | 5.33 | 0.15 |
| Age | 0.50 | 0.25 | 3.95 | <.05 | | | | | | | |
| $Age^2$ | -0.59 | 0.25 | 5.63 | <.05 | | | | | | | |
| Trial | 0.30 | 0.10 | 8.31 | <.01 | | | | | | | |
| REV | -0.45 | 0.13 | 13.04 | <.01 | | | | | | | |
| Intercept | 0.10 | 0.06 | -- | -- | | | | | | | |
| Rnd.Sbj.Int | -- | 0.05 | 9.19 | <.01 | | | | | | | |

The three sections of the table identify the three generalized linear mixed models of trial outcome (correct or incorrect choice) with binomial error distributions. The first row of each section, in which the model is named, provides information about the fit and significance of the overall model. Each subsequent row provides information about a predictor in the model. The full model was constructed based on a priori predictions and included fixed effects for test and subject variables, a fixed intercept, and random subject intercepts (Rnd.Sbj.Int). Likelihood Ratio Tests (LRTs) of nested models were used to assess the significance of each predictor. The final model with only the significant predictors from the full model was constructed to address potential overfitting. To determine whether these models accounted for significant variation in trial outcome, each was compared to the null model using an LRT. A third LRT was used to compare overall prediction between the full and final models. Akaike & Bayesian Information Criterion (AIC and BIC) were used to compare model fit. The ratio of the sum of the squared Pearson residuals to the residuals' degrees of freedom ($r^2/r(df)$) was used to assess dispersion. REV: reversal number.

Table 4-1

*Individual Data for Coyotes: Covariates, Test Duration, and Acquisition Speed*

| | | | **Experiment I** | | | | | | **Experiment II** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | **Test Duration** | | **Acquisition Criteria** | | | | **Test Duration** | | **Final 20 Trials** |
| **Age (m)** | **Sex** | **Grp** | **Trials** | **Sess** | **10 Revs** | **WSLS** | **Sess Revs** | **Final Sess** | **Trials** | **Sess** | **WSLS Prop. Crct** |
| 25 | M | 1 | 107 | 7 | 93 | --[a] | --[a] | 2:17 | -- | -- | -- |
| 38 | F | 1 | 136 | 12 | 73 | 36 | 1 | 3:14 | -- | -- | -- |
| 39 | F | 1 | 100 | 10 | 69 | 15 | 1 | 2:9 | -- | -- | -- |
| 39.5 | F | 2 | 157 | 12 | 85 | 32 | 5 | 2:8 | 64 | 10 | 0.63 |
| 39.5 | M | 2 | 177 | 13 | 124 | 91 | 6 | 2:7 | 77 | 6 | 0.58 |
| 40 | M | 2 | 75[b] | 10 | 73 | 38 | 1 | 1:5 | -- | -- | -- |
| 53 | M | 2 | 124 | 10 | 83 | 64 | 10 | 3:12 | 136 | 6 | 0.50 |
| 53 | F | 2 | 171 | 11 | 107 | 30 | 5 | 2:13 | -- | -- | -- |
| 62 | F | 1 | 120 | 7 | 87 | 61 | 3 | 2:9 | -- | -- | -- |
| 76 | F | 2 | 135 | 12 | 103 | 67 | 5 | 2:7 | 182 | 10 | 0.53 |
| 76 | M | 2 | 150 | 10 | 139 | 86 | 6 | 2:8 | 134 | 9 | 0.42 |
| 86 | M | 1 | 150 | 10 | 128 | 145 | 10 | 2:7 | -- | -- | -- |
| 86 | M | 1 | 243 | 17 | 140 | 83 | 5 | 2:12 | -- | -- | -- |
| 87.5 | F | 2 | 165 | 12 | 103 | 70 | 12 | 2:8 | 136 | 12 | 0.61 |
| 88 | F | 1 | 161 | 13 | 102 | 60 | 13 | 2:9 | -- | -- | -- |
| 88.5 | M | 2 | 132 | 11 | 87 | 94 | 5 | 2:8 | 86 | 10 | 0.37 |
| 88.5 | M | 2 | 100 | 8 | 81 | 75 | 2 | 2:8 | 107[d] | 12 | 0.45 |
| 98 | F | 1 | 206 | 20 | 144 | --[c] | --[c] | 0:10 | -- | | -- |
| 99 | F | 2 | 131 | 12 | 87 | 71 | 8 | 2:6 | 172 | 15 | 0.56 |
| 99.5 | M | 2 | 130 | 9 | 97 | 128 | 9 | 2:7 | 259 | 12 | 0.50 |
| 110 | F | 1 | 147 | 7 | 137 | 138 | 7 | 3:16 | -- | -- | -- |

10 Revs: trials to ten reversals. WSLS: trials to 9 of 10 correct win-stay and lose-shift choices. Sess Revs: session in which the coyote first completed either two reversals in the first nine trials, or three reversals in the first 17 trials. Final Sess: number of reversals (x) completed in the first (y) trials of the final session given as the ratio of x to y. a: Test terminated prematurely due to a scoring error later identified from video playback. b: Stopped participating before completing 100 trials. c: Failed to meet these criteria within 20 sessions. d: Did not satisfy the criterion of 17 correct choices in 20 trials before the end of the study but did chose correctly in 21 of 25 trials (Binomial Test; *p* = .0009).

Table 4-2a

*Models of WSLS Trial Outcome in Experiment I*

| | Fixed Effects | | | | Random Effects | | | Model Fit | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **ß** | ***SE*** | ***$X^2$*** | ***p*** | **Type** | ***$X^2$*** | ***p*** | ***AIC*** | ***BIC*** | **$\frac{r^2}{r(df)}$** | ***n*** |
| **Full Model** | | | | | | | | 3652 | 3717 | 1.004 | 2786 |
| Age | -0.13 | 0.04 | 9.02 | 3e-3 | -- | -- | -- | | | | |
| Sex | 0.02 | 0.08 | 0.05 | 0.82 | | | | | | | |
| Test Trial | 0.15 | 0.04 | 12.80 | 3e-4 | SS | 0 | 1 | | | | |
| Sess.Rev | -0.26 | 0.06 | 17.55 | 3e-5 | SS | 0 | 1 | | | | |
| Intercept | 0.54 | 0.05 | -- | -- | SI | 0 | 1 | | | | |
| **Reduced Model** | | | | | | | | 3640 | 3669 | 1.002 | 2786 |
| Age | -0.13 | 0.04 | 9.48 | 2e-3 | -- | -- | -- | | | | |
| Test Trial | 0.15 | 0.04 | 13.20 | 2e-4 | -- | -- | -- | | | | |
| Sess.Rev | -0.26 | 0.06 | 21.66 | 3e-6 | -- | -- | -- | | | | |
| Intercept | 0.55 | 0.04 | -- | -- | SI | 0 | 1 | | | | |

Rows are separated into two sections corresponding to two generalized linear mixed models of trial outcome (correct or incorrect choice based on the win-stay, lose shift rule) with binomial error distributions. The first row of each section provides information about the overall model. Each subsequent row provides information about a predictor. The Full model was based on a priori predictions. The Reduced model omitted non-significant predictors from the full model to address potential overfitting. Chi-square and *p* values were obtained from Likelihood Ratio Tests of nested models. ß: estimate (coefficient). *AIC* & *BIC*: Akaike & Bayesian Information Criterion. $r^2/r(df)$: sum of squared Pearson residuals divided by residuals degrees of freedom. SS: Subject slope. SI: Subject intercept. Sess.Rev: session reversal number. Age was measured in months. Sex: binary dummy code (1 = male).

Table 4-2b

*Models of WSLS Trial Outcome Before and After the First Reversal of a Session*

| | **Fixed Effects** | | | | **Random Effects** | | | **Model Fit** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **ß** | ***SE*** | **$X^2$** | ***p*** | **Type** | **$X^2$** | ***p*** | ***AIC*** | ***BIC*** | **$\frac{r^2 .}{r(df)}$** | ***n*** |
| **Before Reversal** | | | | | | | | 1693 | 1714 | 1.003 | 1337 |
| Age | -0.11 | 0.06 | 3.57 | 0.06 | -- | -- | -- | | | | |
| Test Trial | 0.17 | 0.06 | 7.96 | 5e-3 | -- | -- | -- | | | | |
| Intercept | 0.72 | 0.06 | -- | -- | SI | 0 | 1 | | | | |
| **After Reversal** | | | | | | | | 1957 | 1973 | 1.003 | 1449 |
| Age | -0.15 | 0.05 | 6.78 | 9e-3 | -- | -- | -- | | | | |
| Test Trial | 0.13 | 0.05 | 5.60 | 0.02 | -- | -- | -- | | | | |
| Intercept | 0.39 | 0.05 | -- | -- | SI | 0 | 1 | | | | |

Trials in which the coyote had yet to complete a reversal in the current session and trials in which the coyote had already completed a reversal in the session were analyzed in separate generalized linear mixed models of trial outcome (correct or incorrect WSLS choice). Rows in this table are separated into two sections corresponding to these two models. The first row of each section provides information about the overall model. Each subsequent row provides information about a predictor. Chi-square and *p* values for predictors were obtained from Likelihood Ratio Tests of nested models. See Table 4-2a for additional abbreviations.

Table 4-3

*Coyote-Dog Comparison: Models of WSLS Trial Outcome*

| | **Fixed Effects** | | | | **Random Effects** | | | **Model Fit** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **ß** | ***SE*** | $X^2$ | ***p*** | **Type** | $X^2$ | ***p*** | ***AIC*** | ***BIC*** | $\frac{r^2}{r(df)}$ | ***n*** |
| **Full Model** | | | | | | | | 2392 | 2517 | 1.00 | 1735 |
| Species | -0.03 | 0.16 | 0.03 | 0.87 | -- | -- | -- | | | | |
| Sex | -0.06 | 0.11 | 0.28 | 0.60 | -- | -- | -- | | | | |
| Age | -0.15 | 0.06 | 6.43 | 0.01 | -- | -- | -- | | | | |
| Reversal | -0.33 | 0.14 | 10.07 | 2e-3 | SS | 0.98 | 0.96 | | | | |
| Test Trial | 0.30 | 0.19 | 2.61 | 0.11 | SS | 0.61 | 0.99 | | | | |
| Sess.Trial | -0.12 | 0.17 | 0.50 | 0.48 | SS | 0.50 | 0.99 | | | | |
| Session | 0.01 | 0.13 | 0.01 | 0.92 | SS | 0.13 | 0.99 | | | | |
| Intercept | 0.35 | 0.13 | -- | -- | SI | 1.11 | 0.95 | | | | |
| **Reduced Model** | | | | | | | | 2361 | 2394 | 1.00 | 1735 |
| Species | -0.10 | 0.10 | 0.90 | 0.34 | -- | -- | -- | | | | |
| Age | -0.12 | 0.05 | 5.97 | 0.01 | -- | -- | -- | | | | |
| Reversal | -0.20 | 0.07 | 7.84 | 5e-3 | -- | -- | -- | | | | |
| Test Trial | 0.14 | 0.07 | 4.05 | 0.04 | -- | -- | -- | | | | |
| Intercept | 0.39 | 0.08 | -- | -- | SI | 0 | 1 | | | | |

Rows are separated into two sections corresponding to two generalized linear mixed models of trial outcome (correct or incorrect WSLS choice). Both models include only the first 30 trials of each canid's test. The full model was based on a priori predictions. The Reduced model omitted non-significant predictors from the full model. Chi-square and *p* values were obtained from Likelihood Ratio Tests. Species: binary dummy code (1 = coyote, 0 = dog). Reversal: test reversal number (cumulative across sessions). Sess.Trial: Trial number within the session. See Table 4-2a for additional abbreviations.

Table 4-4

*Models of Trial Outcome in Experiment II – Full Model*

| | **Fixed Effects** | | | **Random Effects** | | **Model Fit** | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **ß** | ***SE*** | **ß 95% *CI*** | **Type** | ***s*$^2$ 95% *CI*** | ***AIC*** | ***BIC*** | $\frac{r^2\ .}{r(df)}$ | ***n*** |
| **Full Model** | | | | | | 1732 | 1773 | 1.006 | 1353 |
| Sex | .006 | 0.12 | [-0.27,0.25] | -- | -- | | | | |
| Age | -0.16 | 0.07 | [-0.29,-0.03] | -- | -- | | | | |
| Test Trial | 0.27 | 0.07 | [0.12,0.42] | SS | [0.00,0.19] | | | | |
| Age *Test Trial | -0.14 | 0.08 | [-0.29,0.01] | -- | -- | | | | |
| Intercept | 0.72 | 0.09 | [0.53,0.92] | SI | [0.00,0.16] | | | | |
| **Reduced Model** | | | | | | 1726 | 1752 | 1.004 | 1353 |
| Age | -0.16 | 0.07 | [-0.29,-0.01] | -- | -- | | | | |
| Test Trial | 0.27 | 0.07 | [0.14,0.42] | -- | -- | | | | |
| Age *Test Trial | -0.14 | 0.08 | [-0.30,-0.01] | -- | -- | | | | |
| Intercept | 0.73 | 0.06 | [0.58,0.86] | SI | [0.00,0.16] | | | | |

Rows are separated into two sections corresponding to two generalized linear mixed models of trial outcome (correct or incorrect choice). The full model was based on a priori predictions. The final model omitted non-significant predictors from the full model. Chi-square and *p* values were obtained from Likelihood Ratio Tests. Profiled confidence intervals which include zero are not significant. s2: random effect variance. See Table 4-2a for additional abbreviations.

Table 5-1

*Canine Comparisons: Models of WSLS Trial Outcome*

| | **Fixed Effects** | | | | **Random Effects** | | | **Model Fit** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **ß** | ***SE*** | $X^2$ | ***p*** | **Type** | $X^2$ | ***p*** | ***AIC*** | ***BIC*** | $\frac{r^2}{r(df)}$ | ***n*** |
| **Full Model** | | | | | | | | 2726 | 2900 | 0.99 | 1975 |
| Species | -- | -- | 3.40 | 0.18 | -- | -- | -- | | | | |
| Sex | -0.08 | 0.12 | 0.47 | 0.49 | -- | -- | -- | | | | |
| Age | -0.12 | 0.06 | 4.21 | 0.04 | -- | -- | -- | | | | |
| Test Trial | 0.32 | 0.17 | 3.91 | 0.05 | SS | 1.32 | .97 | | | | |
| Sess.Trial | -0.07 | 0.15 | 0.25 | 0.62 | SS | 1.27 | .97 | | | | |
| Reversal | -0.20 | 0.12 | 3.64 | 0.06 | SS | 0.02 | 1 | | | | |
| Sess.Rev | -0.59 | 0.18 | 12.07 | 5e-4 | SS | 0.52 | 1 | | | | |
| Sess | -0.03 | 0.12 | 0.07 | 0.80 | SS | 0.89 | .99 | | | | |
| Intercept | 0.48 | 0.25 | -- | -- | SI | 0 | 1 | | | | |
| **Reduced Model** | | | | | | | | 2683 | 2722 | 1.00 | 1975 |
| Species | -- | -- | 6.44 | 0.04 | -- | -- | -- | | | | |
| Age | -0.08 | 0.05 | 2.90 | 0.09 | -- | -- | -- | | | | |
| Test Trial | 0.13 | 0.05 | 5.78 | 0.02 | -- | -- | -- | | | | |
| Sess.Rev | -0.55 | 0.12 | 21.23 | 4e-6 | -- | -- | -- | | | | |
| Intercept | 0.39 | 0.15 | -- | -- | SI | 0 | 1 | | | | |

Rows are separated into two sections corresponding to two generalized linear mixed models of trial outcome (correct or incorrect choice based on the win-stay, lose shift rule) with binomial error distributions. The first row of each section provides information about the overall model. Each subsequent row provides information about a predictor. The Full model was based on a priori predictions. The Reduced model omitted non-significant predictors from the full model to address potential overfitting. Both models include only the first 30 trials of each canid's test. Chi-square and *p* values were obtained from Likelihood Ratio Tests.

ß: estimate (coefficient). *AIC* & *BIC*: Akaike & Bayesian Information Criterion. $r^2/r(df)$: sum of squared Pearson residuals divided by residuals degrees of freedom. SS: Subject slope. SI: Subject intercept. Species: categorical predictor with three levels (wolf, coyote, dog) and thus, ß cannot be described as a linear slope. Sex: dummy code (Female = 0, Male =1). Age: in months. Test Trial: total trial number (cumulative across sessions). Sess.Trial: trial within the session. Reversal: test reversal number (cumulative across sessions). Sess.Rev: dummy code (0 = before first reversal of session, 1 = after first reversal of session). Sess: session.

Table 5-2

*Canid Comparisons: Models of WSLS Trial Outcome*

| | **Fixed Effects** | | | | **Random Effects** | | | **Model Fit** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **ß** | ***SE*** | $X^2$ | ***p*** | **Type** | $X^2$ | ***p*** | ***AIC*** | ***BIC*** | $\frac{r^2}{r(df)}$ | ***n*** |
| **Reduced Model** | | | | | | | | 2856 | 2907 | 1.00 | 2095 |
| Species | -- | -- | 8.70 | 0.07 | -- | -- | -- | | | | |
| Age | -0.08 | 0.05 | 2.96 | 0.09 | -- | -- | -- | | | | |
| Test Trial | 0.09 | 0.05 | 2.99 | 0.08 | -- | -- | -- | | | | |
| Sess.Rev | -0.47 | 0.11 | 17.73 | 3e-5 | -- | -- | -- | | | | |
| Intercept | 0.39 | 0.15 | -- | -- | SI | 0 | 1 | | | | |

Abbreviations and additional information about the model structure can be found in Table 5-1. The predictors included here were taken from the reduced model in the canine analysis (with foxes excluded).

Table 5-3

*Ecological Predictions about Behavioral Flexibility in Canines*

| Ecological Hypothesis | Flexibility Predictions |
|---|---|
| Behavioral flexibility allows species to cope with anthropogenic habitat modification | Coyotes & Dogs > Wolves * |
| Foraging for spatially distributed food resources require more flexibility than hunting | Coyotes & Dogs > Wolves * |
| Behavioral flexibility is positively correlated with diversity of feeding strategies (e.g., foraging for seeds and berries, solitary hunting for small vertebrates, cooperative hunting for large ungulates, scavenging carrion or human waste, begging for scraps) | Coyotes > Wolves > Dogs |
| Behavioral flexibility is positively correlated with the diversity of habitats in which a species can survive (e.g., tundra, boreal forest, deciduous forest, desert, montane, cities, rural) | Coyotes > Wolves > Dogs |
| Seasonally variable habitats should select for greater flexibility than should stable habitats | Wolves & Coyotes > Dogs |
| Behavioral flexibility is positively correlated with social group complexity, structural variability, and fission-fusion group dynamics | Dogs > Wolves > Coyotes |
| Behavioral flexibility is negatively associated with solitary activity | Wolves > Dogs & Coyotes |
| Species typical cooperation in activities such as hunting, vigilance, defense, rearing young is positively correlated with behavioral flexibility | Wolves > Coyotes > Dogs |

* Supported by findings

# APPENDIX B

# FIGURES

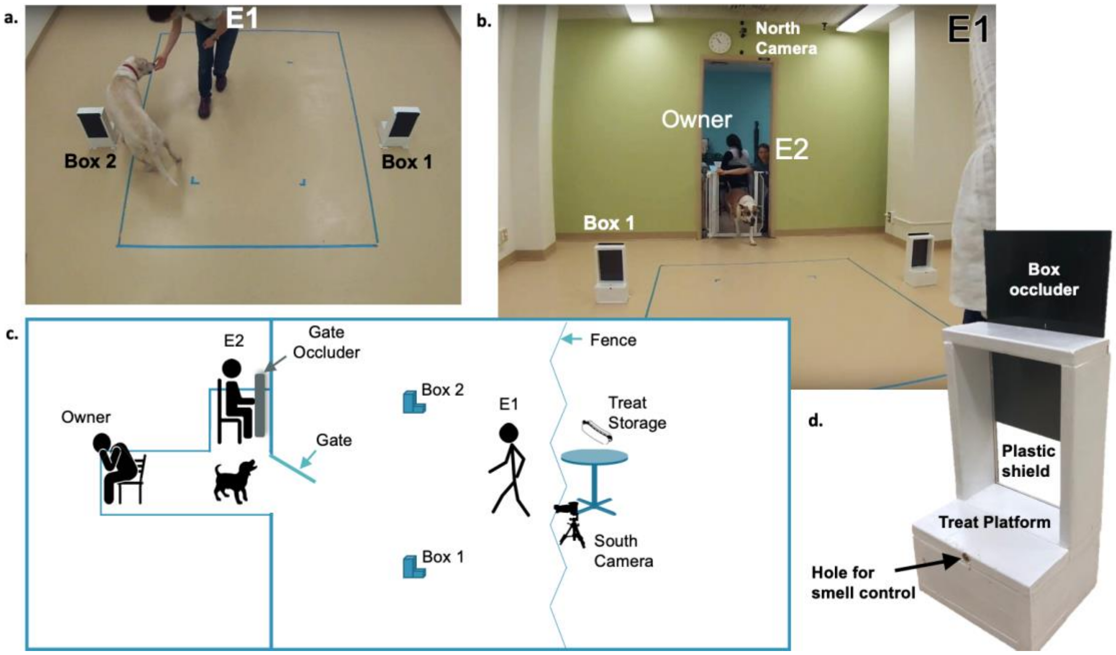

*Figure 2-1.* Schematic of the experimental layout. a View from north camera. b View from south camera. c Layout of the rooms in which the experiments were conducted. d Picture of a treat box.

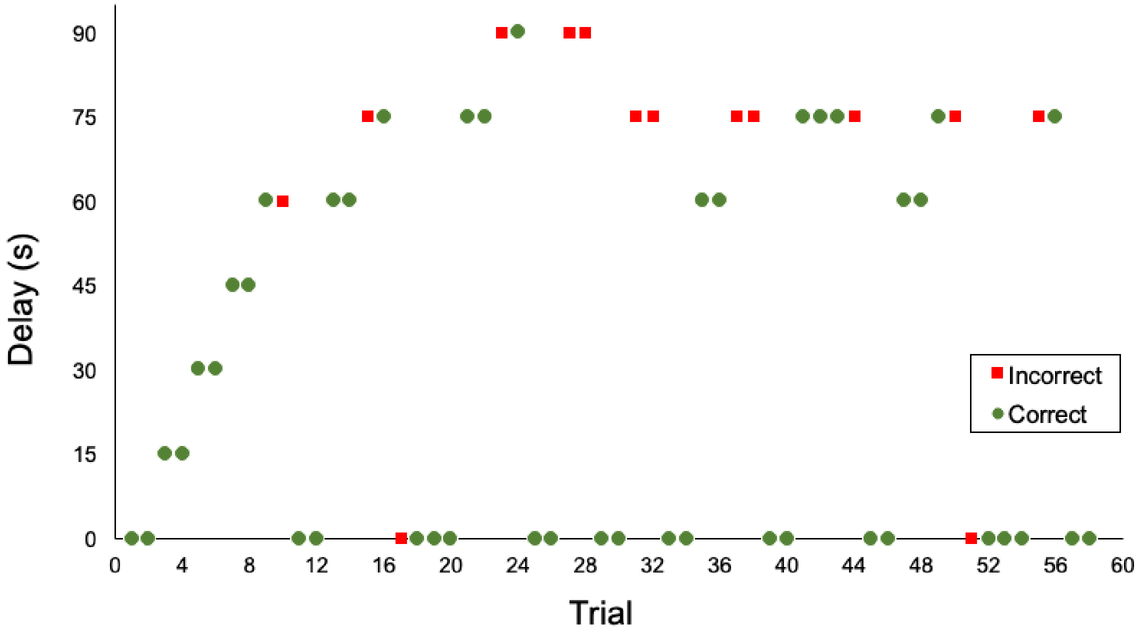

*Figure 2-2.* Example of test progression in Experiment I. Green circles indicate correct choices. Red squares indicate incorrect choices.

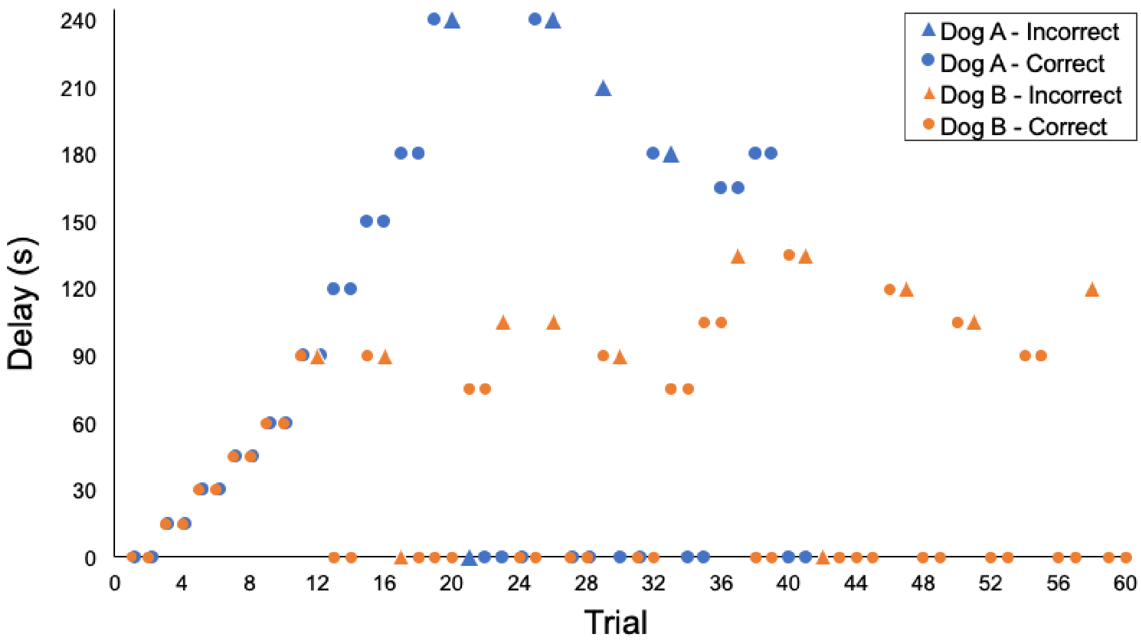

*Figure 2-3.* Examples of test progression in Experiment II. Blue and orange points represent example dogs A and B, respectively. Dots indicate correct choices and triangles indicate incorrect choices. Note that dog A attempted longer delays and therefore, fewer trials within the hour-long test.

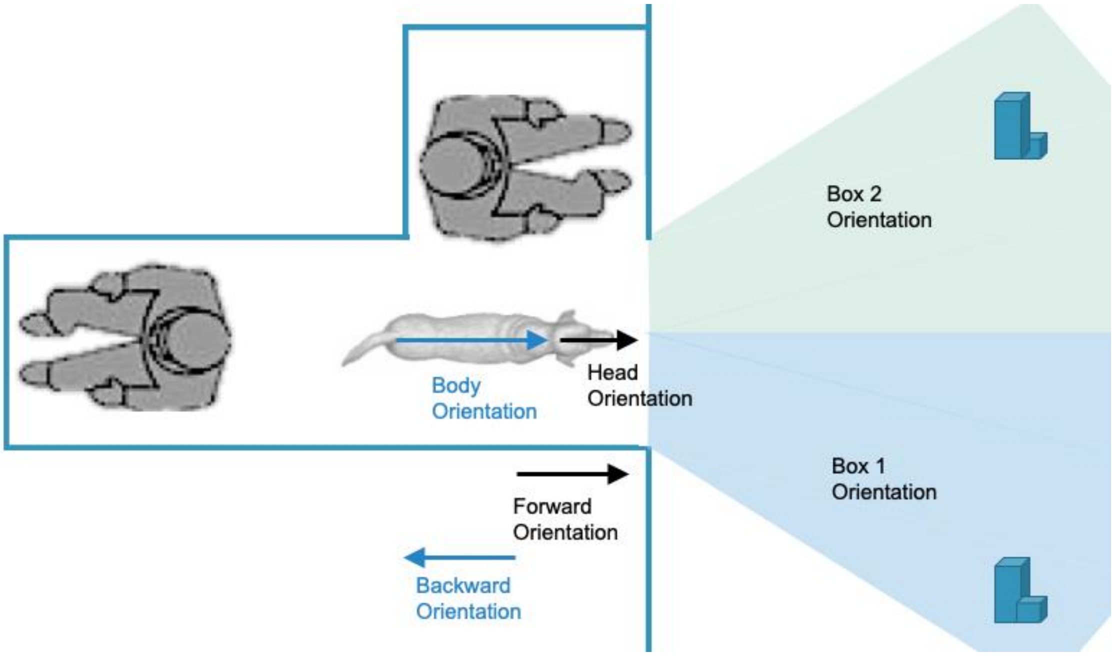

*Figure 2-4.* Diagram of the orientation coding criteria.

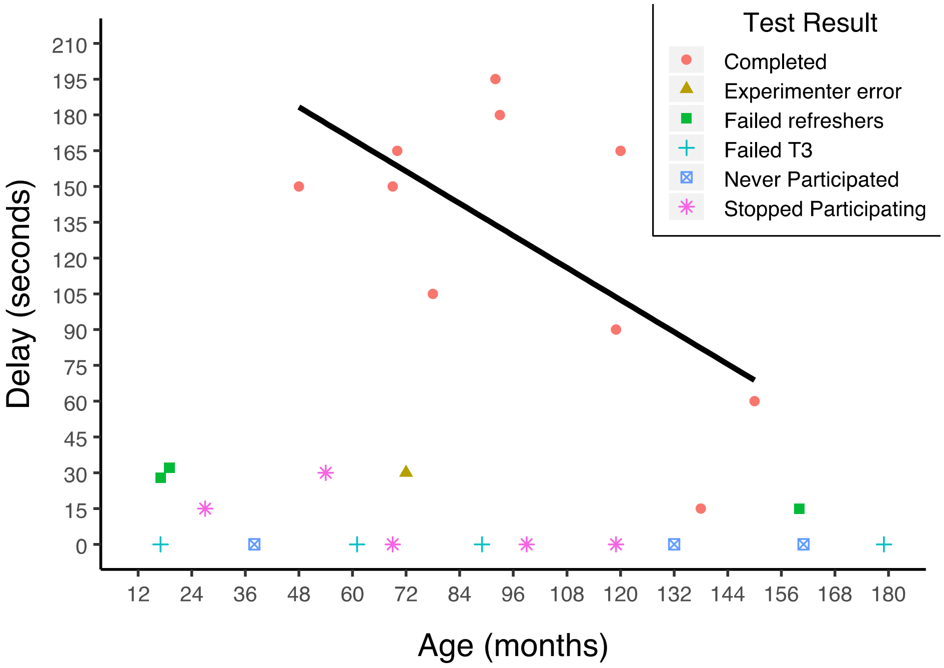

*Figure 2-5.* Longest delay attempted in Experiment I by age. The regression line is fitted only to data of dogs that completed the test (red dots). Blue squares indicate dogs that never participated. Pink strikes indicate dogs that stopped participating. Blue crosses indicate dogs that participated fully but failed to pass training. Green squares indicate dogs that passed training but failed two consecutive blocks of refresher trials. One test was ended prematurely due to an experimenter error (green triangle).

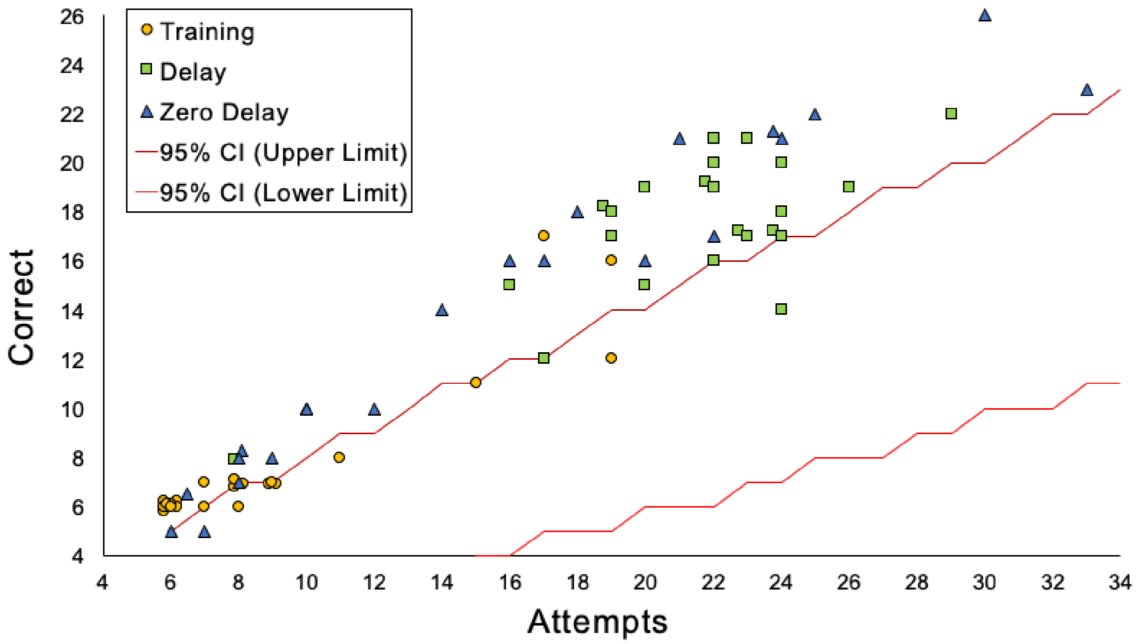

*Figure 2-6.* Correct choices by trials attempted in Experiment II. Each dog is represented by three points. The red lines indicate the maximum and minimum number of correct choices within the 95% confidence interval for random chance. Points within these lines did not significantly differ from chance. Data for dogs that did not differ from chance in either zero-delay trials (green squares) or delay trials (blue triangles) are provided in Appendix C-4. Dogs were not required to exceed chance in training trials (yellow dots) because T3 and refresher trials were identical assessments for task comprehension.

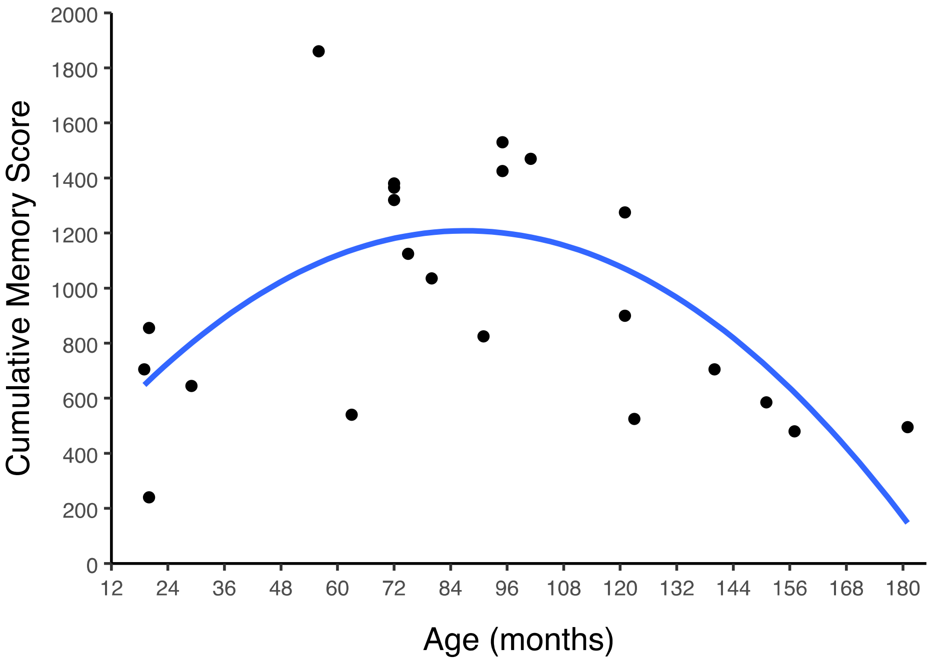

*Figure 2-7.* Cumulative memory score (CMS) in Experiment II by age. Each black dot corresponds to one dog's observed CMS – the total amount of time in the first 18 delay trials that the dog remembered the correct location of the hidden treat. The equation for the blue OLS regression line is, CMS = 292.31 + 21.02*AGE - 0.12*AGE$^2$, where AGE = age in months.

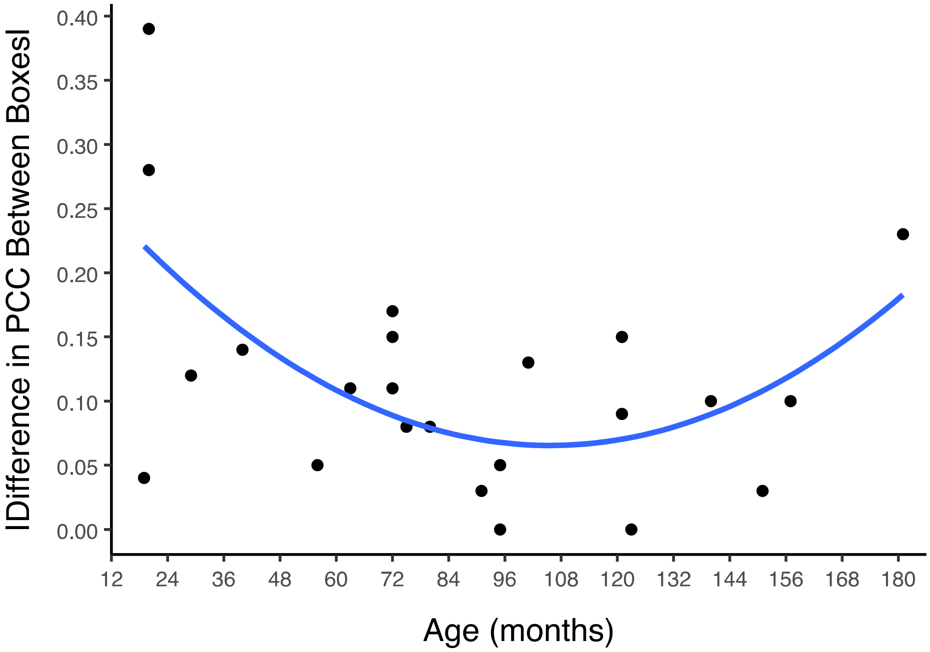

*Figure 2-8.* Magnitude of box preference in Experiment II by age. Each black dot corresponds to one dog's observed perseverative tendency to choose its preferred box (i.e., side bias), which was calculated as the absolute value of the difference between the proportion of correct choices (PCC) on Box 1 and PCC on Box 2. The equation for the blue OLS regression line is, Box Preference = 0.03 - 0.004*AGE + 0.00002*AGE$^2$, where AGE = age in months.

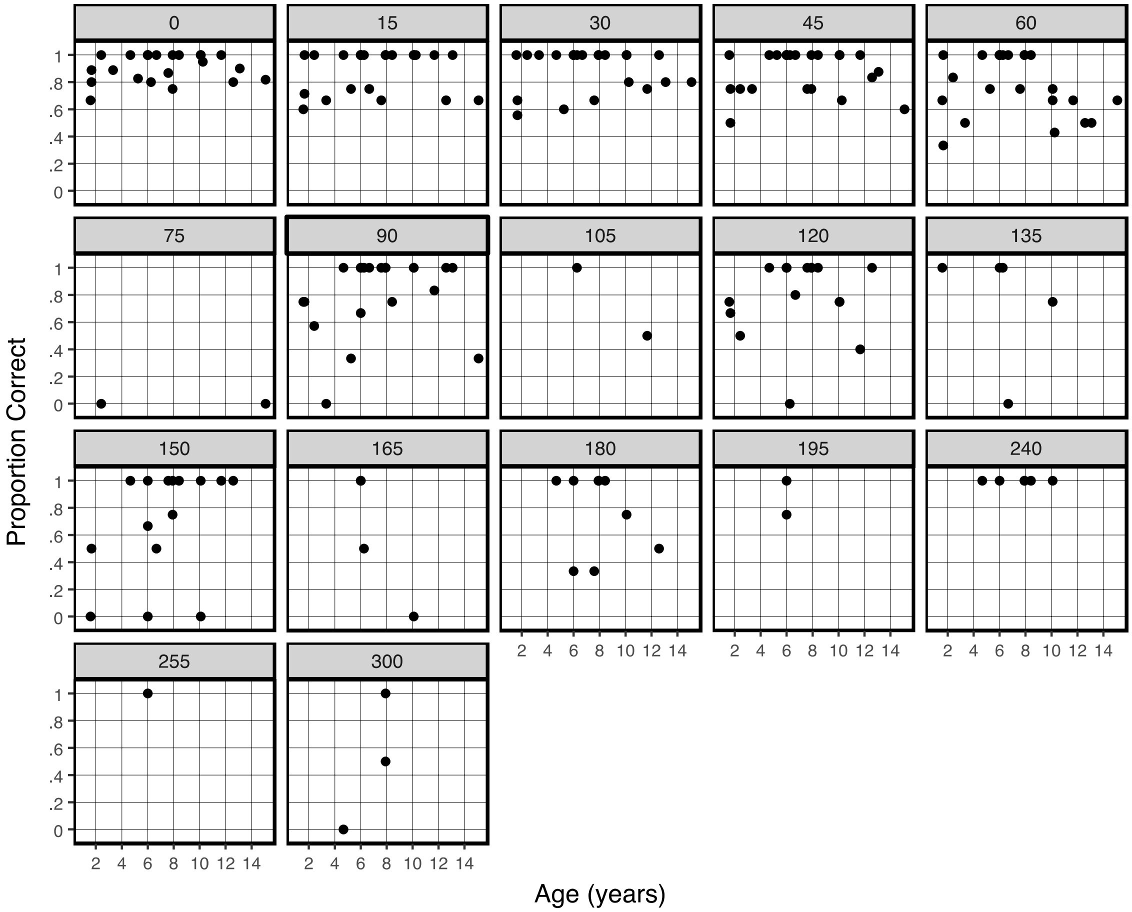

*Figure 2-9*. Multi-plot of the proportion of correct choices (PCC) by age for each delay in Experiment II. Each panel shows PCC in the delay indicated in the gray header

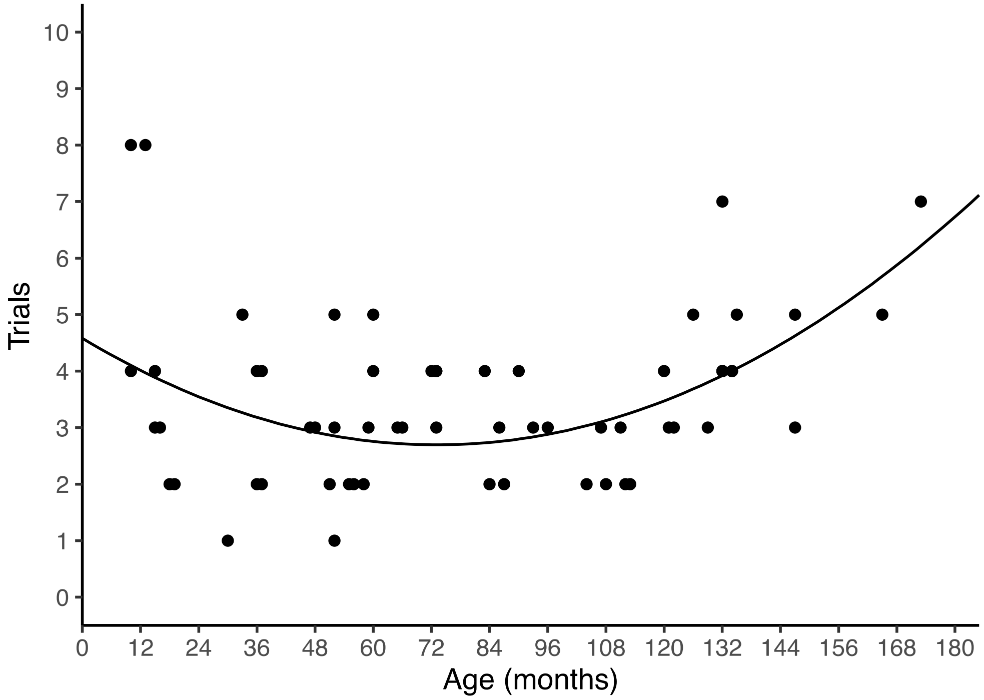

*Figure 3-1*. Longest streak of perseverative errors by age. A square root transformation was used for analysis. The displayed values and regression line are back-transformed.

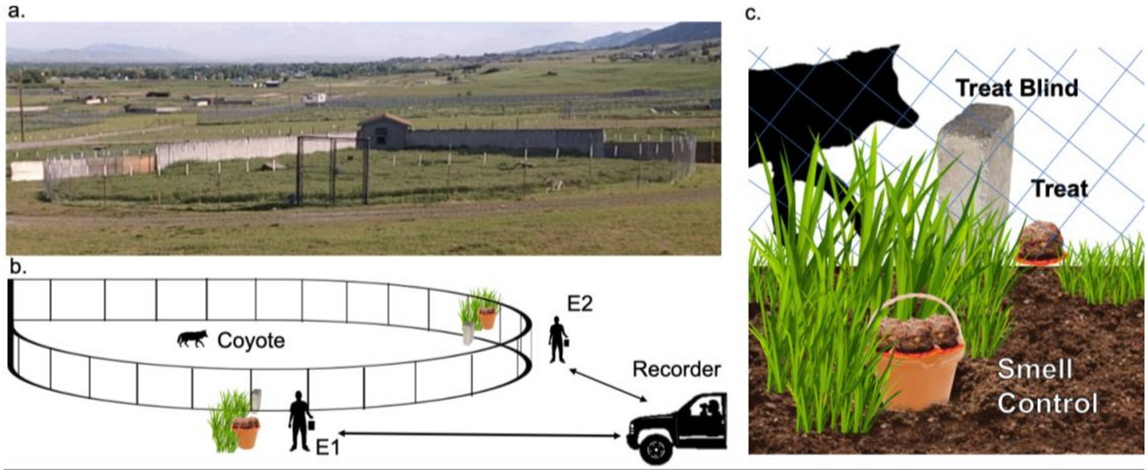

*Figure 4-1.* Materials and layout. a. example of an enclosure in which the experiments were conducted. b. experimental layout. c. baited treat blind and adjacent smell control.

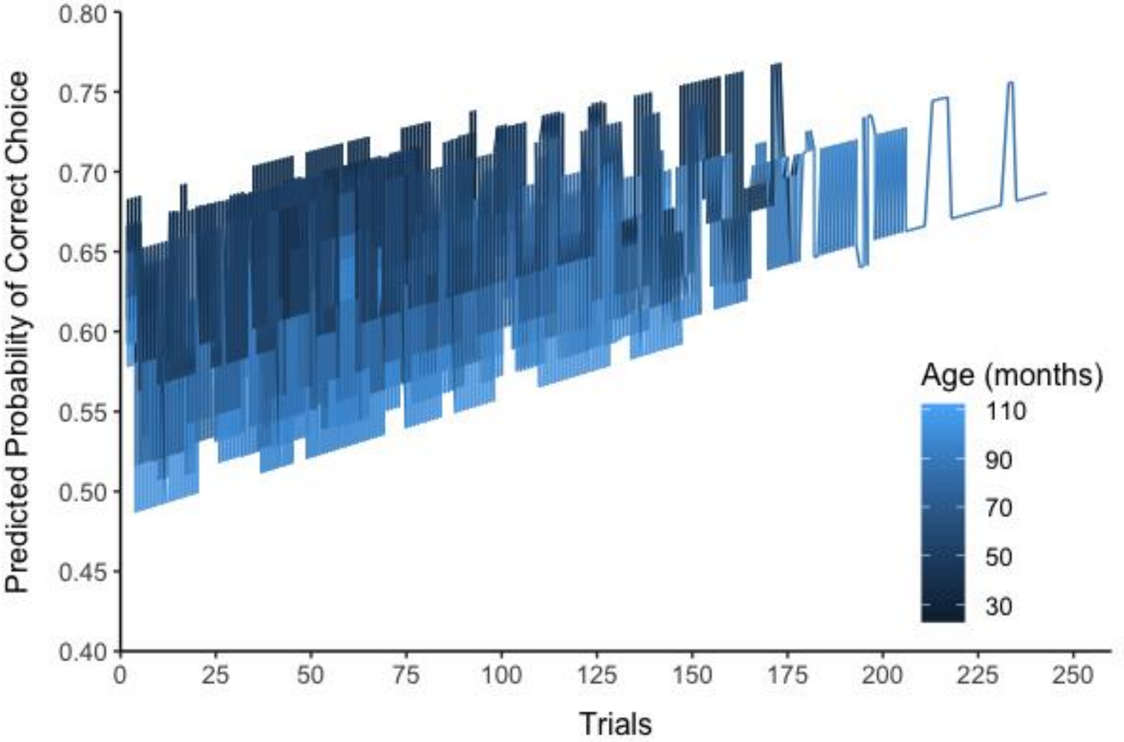

*Figure 4-2a.* Predicted probability of choosing the correct side in Experiment I by test trial number and age in months. Fitted values were calculated using the final model summarized in Table 4-3a. Each coyote is represented by a single line and darker lines indicate younger coyotes. The oscillations of individual lines depict the decline and subsequent increase in performance following each reversal. Peaks & valleys of a similar shade (color) which are close together (within approximately 8 trials or less of each other) correspond to different coyotes of a similar age.

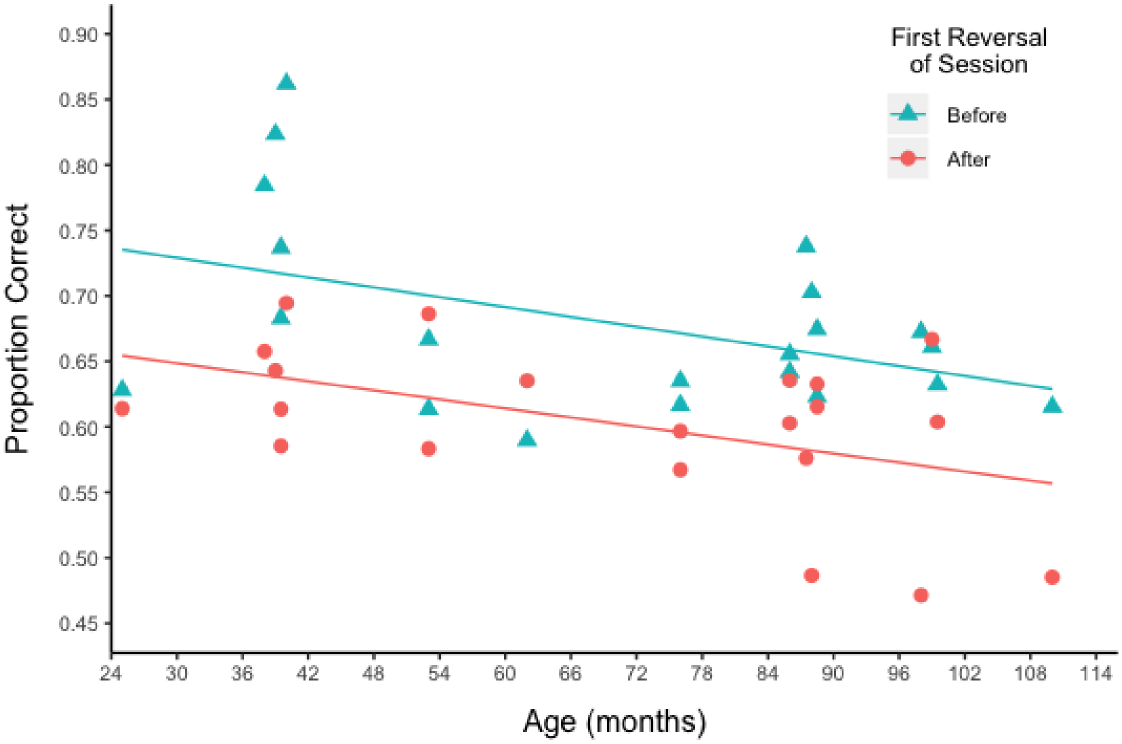

*Figure 4-2b.* Proportion of correct WSLS choices by age before and after the first reversal of the session in Experiment I.  Each coyote is represented by two points. Blue triangles and the top regression line correspond to trials before the first reversal of the session. Red circles and the bottom regression line correspond to trials after the first reversal of the session.

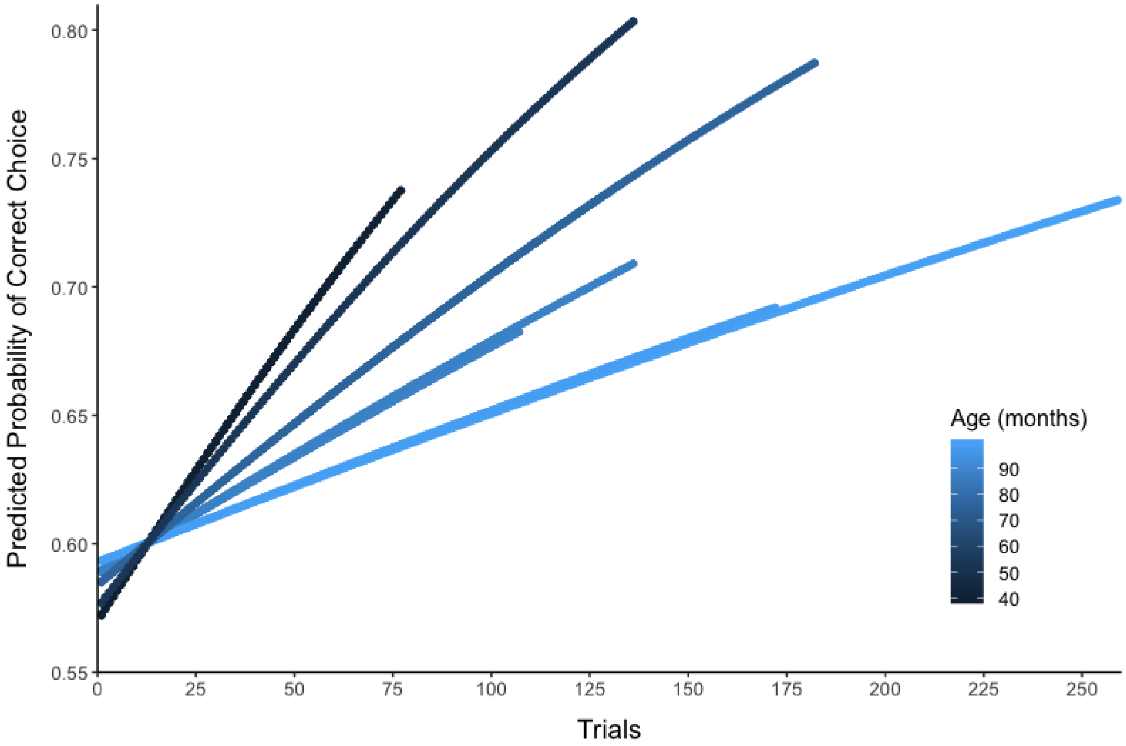

*Figure 4-3.* Predicted probability of choosing the correct side in Experiment II by total trial and age. Fitted values were calculated using the final model summarized in Table 4-4. Darker points indicate younger coyotes. Three pairs of coyotes were the same age (hence seven lines for ten coyotes).

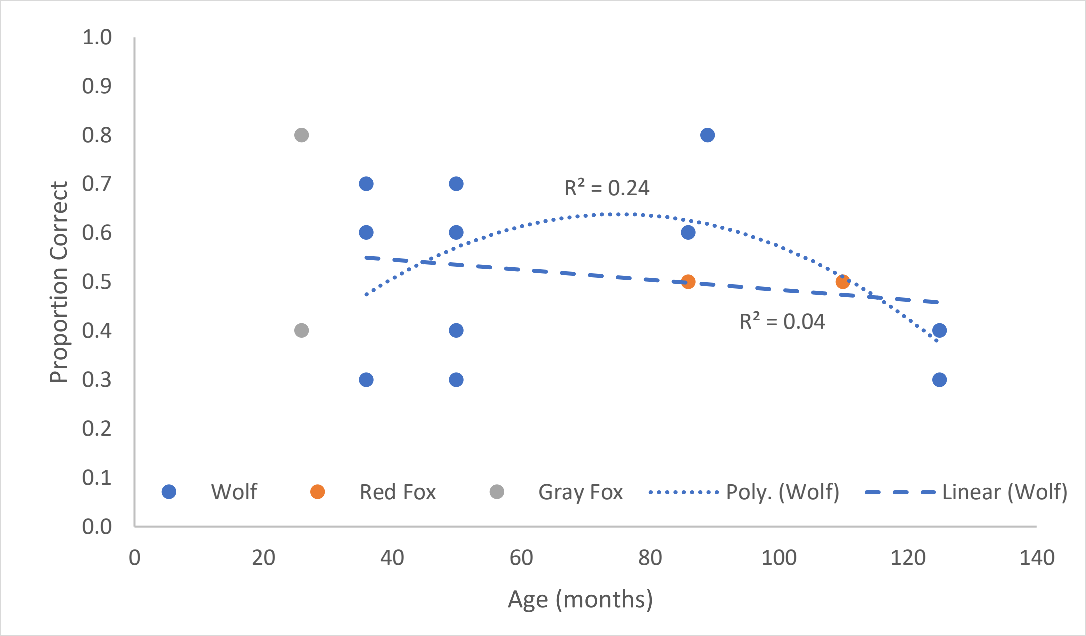

*Figure 5-1*. Proportion of correct WSLS choices in the first 10 trials.

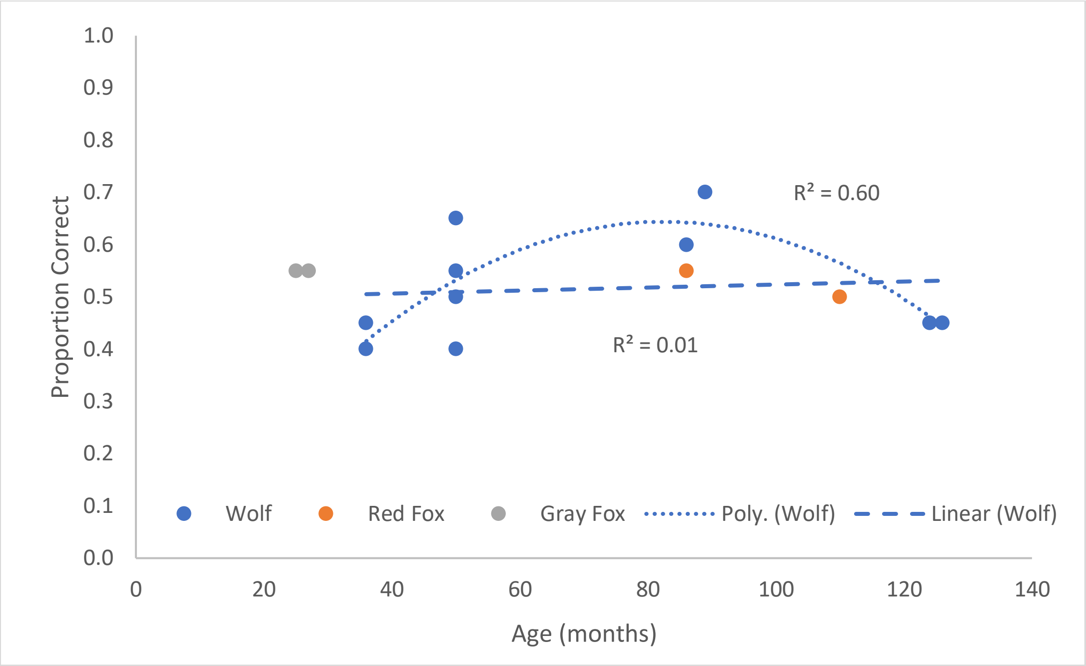

*Figure 5-2*. Proportion of correct WSLS choices in the first 20 trials.

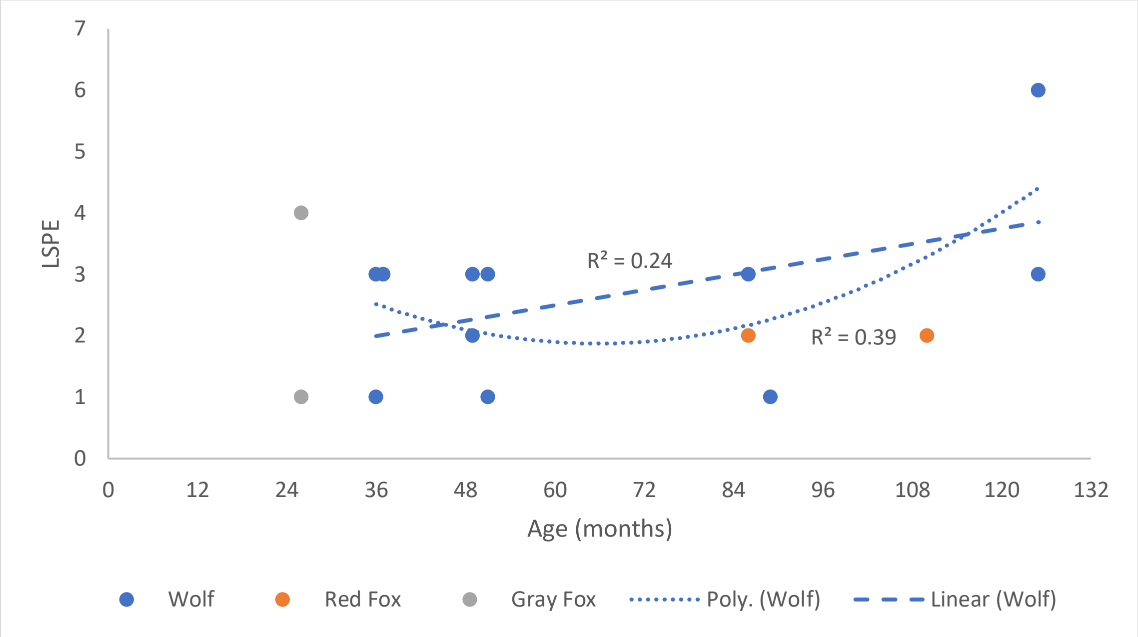

*Figure 5-3*. Longest streak of perseverative errors (LSPE) in the First 10 Trials.

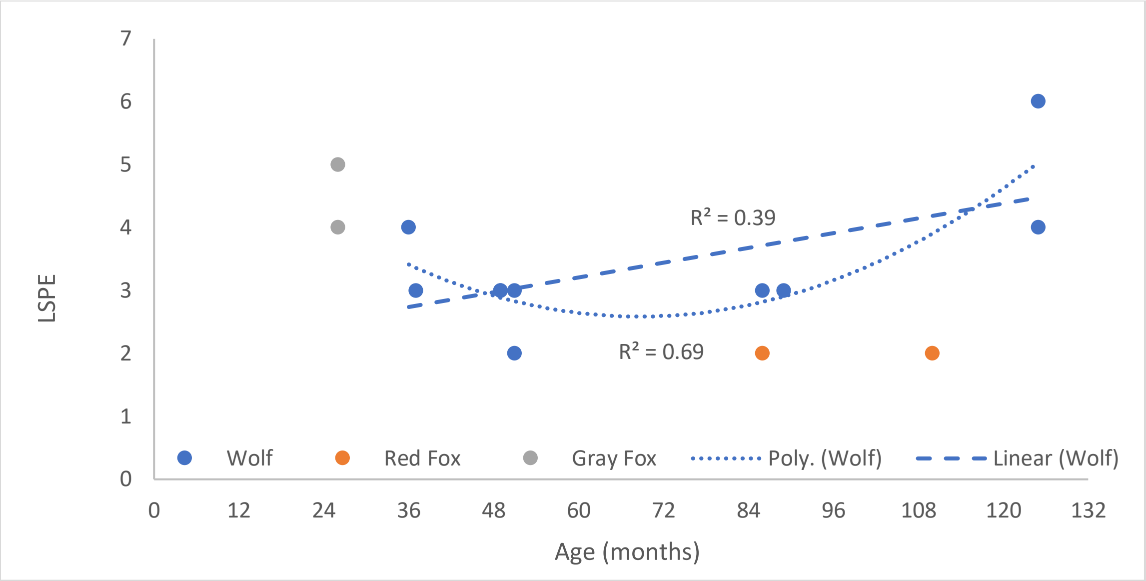

*Figure 5-4*. Longest streak of perseverative errors in the First 20 Trial

# APPENDIX C

# SUPPLEMENTARY MATERIAL FOR CHAPTER 2

Table C-1

*Study Subjects*

| Dog | PSL[a] | Age (yrs)[b] | Predicted Lifespan (yrs)[c] | Height (in)[d] | Weight (lbs)[d] | Sex | Breed |
|---|---|---|---|---|---|---|---|
| Greta | 0.15 | 2 | 13.41 | 12.00 | 19.5 | F | Swedish Vallhund |
| Moses | 0.79 | 8 | 10.06 | 27.50 | 102.0 | M | Akita |
| Miss Piggy | 0.89 | 11 | 12.38 | 23.50 | 53.8 | F | Mix |
| Lucy | 0.89 | 10 | 11.22 | 24.25 | 76.2 | F | Mix |
| Max | 1.14 | 13 | 11.43 | 22.50 | 70.0 | M | Mix |
| Baby Blue | 0.15 | 2 | 13.22 | 23.75 | 38.4 | F | Mix |
| Penny | 0.15 | 2 | 13.15 | 24.00 | 40.0 | F | Mix |
| Brinkley | 1.05 | 14 | 13.39 | 12.00 | 20.0 | M | Havanese |
| Beau | 1.18 | 15 | 12.71 | 23.00 | 47.0 | M | Mix |
| Kukkula | 0.53 | 6 | 11.41 | 28.25 | 78.0 | F | Yellow Lab |
| Basquiat | 1.00 | 12 | 12.00 | 28.25 | 67.0 | M | Mix |
| Winston | 0.52 | 7 | 13.42 | 15.50 | 24.0 | M | Mix |
| Aunika | 0.37 | 5 | 13.41 | 20.00 | 30.0 | F | Mix |
| Auberon | 0.37 | 5 | 13.38 | 31.00 | 45.0 | M | Standard Poodle |
| Ziggy | 0.91 | 10 | 11.03 | 27.00 | 83.4 | M | Chocolate Lab |
| Xuxa | 0.58 | 8 | 13.80 | 9.50 | 9.0 | F | Mix |
| Dodi | 0.22 | 3 | 13.57 | 11.50 | 16.0 | F | Havanese |
| Tibbie | 0.58 | 8 | 13.90 | 13.50 | 12.5 | M | Tibetan Spaniel |
| Daisy | 0.48 | 6 | 12.38 | 14.50 | 42.0 | F | Corgi |
| Kiba | 0.62 | 8 | 12.89 | 11.00 | 28.0 | M | Corgi |
| Kona | 0.94 | 13 | 13.79 | 20.00 | 23.0 | F | Mix |
| Ni | 0.51 | 6 | 11.68 | 23.00 | 66.0 | F | Black Lab |
| Finn | 0.45 | 6 | 13.45 | 13.00 | 20.2 | M | Cockalier |
| Cairo | 0.15 | 2 | 13.18 | 27.50 | 44.0 | F | Mix |
| Dozer | 0.90 | 10 | 11.16 | 22.50 | 75.0 | M | Mix |
| Flynn | 0.38 | 4 | 10.53 | 28.00 | 94.0 | M | German Shepherd |

[a] Proportion of size-predicted lifespan

[b] Rounded to the nearest year to account for any inaccuracy in the dog's initial age estimation

[c] Calculated using the following equation from Greer et al. (2013):

size-predicted lifespan =13.62 + (0.0702*Height) - (0.0538*Weight)

[d] Height and weight were measured in inches and pounds to conform to Greer et al. (2013)

## C-2 Assumption Tests, Case-wise Diagnostics, and Violations

### Quadratic Regressions

To test for homoscedasticity and linearity, Non-constant Variance Score Tests were conducted and residual vs. fitted, residual vs. AGE and residual vs. $\text{AGE}^2$ plots were inspected. To assess whether residuals were normally distributed, standard normal quantile plots and residual density plots were inspected. To test for multicollinearity, predictor correlation plots and matrices, tolerances ($T > 0.1$), variation inflation factors ($\text{VIF} < 5$), eigenvalues ($< 0.5$) and condition indices ($> 10$) were assessed.

To test for outliers, leverage ($h_{ii} > 2p/n$), externally studentized residuals ($\text{SDRESID} > t(n-p-1)$), standardized DFFIT ($\text{DFFITS}_i > 1$), and Standardized DFBETA were examined. For all outliers, the AGE of the dog and the values exceeding preset cutoffs for case-wise diagnostics are reported. When outliers were identified, the analysis was rerun with the corresponding data omitted to confirm that the model was not dependent on individual test subjects. Notation for regression diagnostics follows Cohen et al. (2003).

#### Cumulative Memory Score.

Two potential outliers were identified (AGE = 56: SDRESID = 2.68; AGE = 181: DFFITS = 1.70, $\text{SDBETA}_{AGE2} = 1.27$). With the outliers removed, overall prediction of CMS improved, $F(2,17) = 9.16$, $p = .002$; $r^2 = 0.52$. Again, the effect of AGE.CENT was not significant, $b_1 = -0.10$, $SE_{b1} = 1.56$; $t(17) = -0.06$, $p = .95$, but $\text{AGE.CENT}^2$ significantly predicted CMS, $b_2 = -0.15$, $SE_{b2} = 0.04$; $t(17) = -4.27$, $p = .0005$. Memory scores were higher in middle-aged dogs than in young and old dogs.

### Perseveration.

A Non-Constant Variance Score Test and visual inspection of a residual vs. fitted plot suggested that this model did not meet the assumption of homogeneity of variance, $X^2(1) = 8.89$, $p = .003$. Moreover, variation was greater in young dogs than in middle-aged and old dogs. Thus, the significance of the coefficients may be under-estimated as the variances of the estimates were positively biased.

Two potential outliers were identified (AGE = 19: SDRESID = -3.30, DFFITS = -1.84, $\text{SDBETA}_{AGE} = 1.41$, $\text{SDBETA}_{AGE2} = -1.17$; AGE = 20: SDRESID = 3.05, DFFITS = 1.65, $\text{SDBETA}_{AGE} = -1.27$, $\text{SDBETA}_{AGE2} = 1.02$). With the outliers excluded, the model's overall prediction of side bias slightly increased, $F(2,18) = 6.92$, $p = .006$; $r = 0.43$. The linear effect of AGE.CENT was not significant, $b_1 = 0.0005$, $SE_{b1} = 0.0003$; $t(18) = -1.84$, $p = .08$, but $\text{AGE.CENT}^2$ significantly predicted CMS; $b_2 = 0.00002$, $SE_{b2} = 0.00006$; $t(18) = 2.87$, $p = .002$. Box preferences were stronger in young and old dogs than in middle-aged dogs.

### Delay Optimization.

For each delay, visual inspection of the standard normal quantile plot and residual density plot indicated that residuals were not normally distributed. For samples of moderate size ($n \approx 25$), the effects of nonnormality of error on significance testing are small (Cohen et al. 2003). However, the significance of the coefficients may be slightly under-estimated as the variances of the estimates may be positively biased.

Although potential outliers were identified in each delay, re-running the analyses without these outliers did not change the outcome of the larger delay optimization analysis. Moreover, AGE and AGE$^2$ still accounted for far more variation in PCC at the 45s and 60s delays than in the 15s, 30s, and 90s delays.

### ***15s Delay.***

One potential outlier was identified (AGE = 91: SDRESID = -2.39). With the outlier removed, AGE and AGE$^2$ accounted for a significant proportion of variation in PCC at the 30s delay, $F(2,18) = 4.03$, $p = .04$; $r^2 = 0.30$. Both AGE, $t(18) = 2.78$, $p = .01$, and AGE$^2$, $t(18) = -2.83$, $p = .01$, predicted PCC. Thus, delays shorter than 45s may in fact be used to detect ARCD in future SWM assessments for pet dogs if larger sample sizes, which are more robust to outliers, are used.

### ***30s Delay.***

Two potential outliers were identified (AGE = 63: SDRESID = -2.42; AGE = 20: DFFITS = -1.11, SDBETA$_{AGE}$ = -1.07). Removing the outliers did not change the outcome of the analysis. Moreover, model fit and overall prediction were nearly unchanged ($r^2 = 0.17$, $p = .20$ compared to $r^2 = 0.15$, $p = .22$ with potential outliers).

### *45s Delay*

Three potential outliers were identified (AGE = 20: SDRESID = -2.52, DFFITS = -1.44, SDBETA$_{intercept}$ = -1.39, SDBETA$_{AGE}$ = 1.10; AGE = 19: SDRESID = 2.29, DFFITS = 1.35, SDBETA$_{intercept}$ = 1.30, SDBETA$_{AGE}$ = 1.05; AGE = 123: SDRESID = -2.17). With the outliers removed, overall prediction of PCC improved, $F(2,16) = 8.06$, $p = .004$; $r^2 = 0.50$. Both AGE, $t(16) = 3.57$, $p = .003$, and AGE$^2$, $t(16) = -3.90$, $p = .001$, predicted PCC.

### *60s Delay.*

Three potential outliers were identified (AGE = 20: SDRESID = -3.07, DFFITS = -1.75, $\text{SDBETA}_{intercept}$ = -1.69, $\text{SDBETA}_{AGE}$ = 1.34, $\text{SDBETA}_{AGE2}$ = -1.07; AGE = 181: SDRESID = 2.62, DFFITS = 2.86, $\text{SDBETA}_{AGE}$ = -1.61, $\text{SDBETA}_{AGE2}$ = 2.14; AGE = 123: SDRESID = -2.27). With the outliers removed, overall prediction of PCC improved, $F(2,16) = 17.84$, $p < .0001$; $r^2 = 0.69$. Both AGE, $t(16) = 3.47$, $p = .003$, and $\text{AGE}^2$, $t(16) = -4.58$, $p = .0003$, predicted PCC.

### *90s Delay.*

Two potential outliers were identified (AGE = 63: SDRESID = -3.86, DFFITS = -1.12; AGE = 181: SDRESID = 2.34, DFFITS = -2.61, $\text{SDBETA}_{AGE}$ = 1.40, $\text{SDBETA}_{AGE2}$ = -1.88). With the outliers removed, the curvilinear effect of $\text{AGE}^2$ was not significant, $t(15) = -1.72$, $p = .11$.

## GLMMs

To test for linearity and homogeneity of variance, fitted vs. residual plots, fitted vs. predictor (transformed by the logit link function), and predictor vs. residual plots were inspected. To assess whether random intercepts were normally distributed, normal quantile plots and density plots of fitted intercepts were inspected. To test for overdispersion, the sum of the squared Pearson residuals was compared to the chi square distribution (see $r^2/r(df)$ in Table 2-4 and 2-6), and standardized residual plots were inspected.

### Test-trial Outcome by Delay, Trial, Age, etc.

The full model did not converge but met all other assumptions. The maximal model that converged met all assumptions.

### Delay-trial Outcome by Head and Body Orientations.

Fitted vs. predictor plots indicated that variance in delay-trial outcome was greater when the dog's body did not remain oriented towards the correct box during the delay. Thus, the variance of this estimate (coefficient) may be positively biased. All other assumptions were met.

### Delay-trial Outcome by B.F.Dl, Delay, Age, etc.

Fitted vs. predictor plots indicated that variance in delay-trial outcome was greater when the dog's body did not remain oriented towards the testing room during the delay. Thus, the variance of this estimate (coefficient) may be positively biased. All other assumptions were met.

1. Cohen J, Cohen P, West SG, Aiken L (2003) Applied multiple regression/correlation analysis for the behavioral sciences, 3rd Edn. Lawrence Erlbaum, Mahwah.

## C-3 Sex Effects

Sex did not predict performance in any analysis. Barring two minor exceptions in the delay optimization analysis, sex did not affect the overall model or the other predictors in the model. Given the small sample size of each sex and that we did not strictly age-match male and female dogs, our study was not well equipped to detect sex effects. Thus, future studies are needed to test whether age-related deficits in SWM and other cognitive functions are sex dependent in dogs.

Table C-3a summarizes the ages of the female and male dogs which were tested, and which completed the test. For each analysis, we provide sex-specific summary statistics followed by the equation(s) of the model(s), a brief description of any changes caused by the addition of the coefficient for sex, and then a detailed description or table outlining the outcome of the analysis with sex included.

Table C-3a

*Age by Sex*

| | All Dogs | | Test Completed | |
|---|---|---|---|---|
| Sex | *n* | AGE | *n* | AGE |
| Females | 13 | 69.7 ± 45.6 | 12 | 115.4.3 ± 37.5 |
| Males | 12 | 111.1 ± 36.0 | 11 | 111.1 ± 36.0 |

*n* = number of dogs. AGE = age in months; mean and standard deviation ($\bar{x}$ ± *SD*).

### Cumulative Memory Score

Table C-3b

*Summary statistics for CMS by sex*

| Sex | *n* | AGE | CMS |
|---|---|---|---|
| Females | 11 | 57.5 ± 34.9 | 1315.5 ± 460.8 |
| Males | 11 | 111.1 ± 36.0 | 1565.5 ± 764.1 |

*n* = number of observations (dogs). CMS = Cumulative Memory Score

OLS Quadratic Regression: CMS = Sex + AGE.CENT + AGE.CENT$^2$

Sex did not predict CMS and accounting for sex had no effect on the analysis.

The model significantly predicted CMS, $F(3,18) = 4.57$, $p = .02$; $r^2 = 0.43$. AGE.CENT$^2$ significantly predicted CMS, $b_2 = -0.12$, $SE_{b2} = 0.03$; $t(18) = -3.53$, $p = .002$, but AGE.CENT, $b_1 = -0.37$, $SE_{b1} = 2.0$; $t(18) = -0.19$, $p = .86$, and Sex did not, $b_3 = -4.54$, $SE_{b3} = 179$; $t(18) = -0.03$, $p = .98$.

## Test-trial Outcome

Table C-3c

*Summary statistics for test-trial outcome by sex*

| Sex | *n* | AGE | $y_i$ (Full Model) | $y_i$ (Final Model) |
|---|---|---|---|---|
| Females | 396 | 65.9 ± 43.4 | 0.81 ± 0.11 | 0.81 ± 0.10 |
| Males | 295 | 102.3 ± 30.3 | 0.86 ± 0.08 | 0.86 ± 0.08 |

*n* = number of observations (trials). $y_i$ = predicted probability of choosing the correct box

Full GLMM (logit-link): Test-trial Outcome = Sex + AGE + AGE$^2$ + Delay + Trial + AGE*Delay + AGE$^2$*Delay + AGE*Trial + AGE$^2$*Trial + (1 + Delay + Trial | Dog)

Final GLMM (logit-link): Test-trial Outcome = Sex + AGE + AGE$^2$ + Delay + Trial + (1 | Dog)

Sex did not predict test-trial outcome in either the Full of the Final model. Accounting for sex had no effect on the analysis. The full models failed to converge with, but not without the coefficient for sex included.

Table C-3d

*GLMM of test-trial outcome with a coefficient for sex*

| **Model** | | Estimates | | | | | Random Effects | | | Model Fit | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Predictor | Type | ß | *SE* | ß *CI* | $X^2$ | *p* | $s^2$ *CI* | $X^2$ | *p* | *AIC* | *BIC* | $r^2$<br>*r*(*df*) |
| **Full** | | | | | | | | | | 616 | 688 | .99 |
| AGE | F | 2.37 | 0.58 | [1.5,3.3] | -- | -- | -- | -- | -- | | | |
| AGE$^2$ | F | -2.18 | 0.55 | [-3.1,-1.4] | -- | -- | -- | -- | -- | | | |
| Sex | F | 0.11 | 0.31 | [-.4,.8] | | | -- | -- | -- | | | |
| Delay | F,RS | -0.61 | 0.14 | [-.9,-.4] | -- | -- | [.01,.4] | 0.12 | .99 | | | |
| Delay *AGE | F | 0.82 | 0.54 | [0,1.9] | 2.30 | 0.68 | -- | -- | -- | | | |
| Delay *AGE$^2$ | F | -0.85 | 0.55 | [-1.8,-.03] | 2.42 | 0.66 | -- | -- | -- | | | |
| Trial | F,RS | -0.10 | 0.13 | [-.4,.2] | -- | -- | [.01,.4] | 0.04 | 1 | | | |
| Trial *AGE | F | -1.57 | 0.46 | [-2.3,-.8] | 9.36 | 0.05 | -- | -- | -- | | | |
| Trial *AGE$^2$ | F | 1.57 | 0.47 | [0.70,2.4] | 9.00 | 0.06 | -- | -- | -- | | | |
| Intercept | I | 1.57 | 0.18 | [1.3,12] | -- | -- | -- | -- | -- | | | |
| Subject | RI | -- | -- | -- | -- | -- | [0,0.4] | 0.08 | .99 | | | |
| **Final** | | | | | | | | | | 610 | 642 | .92 |
| AGE | F | 2.72 | 0.92 | [1.0,4.4] | 9.70 | .002 | -- | -- | -- | | | |
| AGE$^2$ | F | -2.54 | 0.88 | [-4.2,-1] | 9.36 | .002 | -- | -- | -- | | | |
| Sex | F | 0.07 | 0.41 | [-.7,.8] | 2.27 | 0.13 | -- | -- | -- | | | |
| Delay | F | -0.66 | 0.16 | [-09,-.4] | 20.9 | <.001 | -- | -- | -- | | | |
| Trial | F | 0.20 | 0.17 | [-.02,0.5] | 1.30 | 0.25 | -- | -- | -- | | | |
| Intercept | F | 1.83 | 0.26 | [1.3,2.4] | -- | -- | -- | -- | -- | | | |
| Subject | RI | -- | -- | -- | -- | -- | [0,.8] | | | | | |

See Table 2-4 of the main article for abbreviations

### Delay-trial Outcome (B.F.Dl Analysis)

Table C-3e

*Summary statistics for delay-trial outcome by sex*

| Sex | n | AGE | $y_i$ |
|---|---|---|---|
| Females | 248 | 63.5 ± 42.4 | 0.76 ± 0.11 |
| Males | 226 | 111.7 ± 34.5 | 0.79 ± 0.10 |

*n* = number of observations (trials). $y_i$ = predicted probability of choosing the correct box

GLMM (logit-link) Equation: Delay-trial Outcome = Sex + AGE + AGE $^2$ + Delay + AGE*B.F.Dl + AGE$^2$ *B.F.Dl + RSS B.F.Dl + (1 + B.F.Dl | Dog)

Sex did not predict delay-trial outcome. Accounting for sex had no effect on the analysis.

The model failed to converge with, but not without the coefficient for sex included.

Table C-3f

*GLMM of delay-trial outcome with a coefficient for sex*

| **Model** | | Estimates | | | | | Random Effects | | | Model Fit | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Predictor | Type | ß | *SE* | ß *CI* | $X^2$ | *p* | $s^2$ *CI* | $X^2$ | *p* | *AIC* | *BIC* | $\frac{r^2}{r(df)}$ |
| **Full** | | | | | | | | | | 504 | 549 | 0.95 |
| Age | F | 1.77 | 0.80 | [0.4,3.2] | -- | -- | -- | -- | -- | | | |
| Age$^2$ | F | -1.66 | 0.72 | [-2.9,-0.5] | -- | -- | -- | -- | -- | | | |
| Sex | F | 0.14 | 0.33 | [-0.6,0.9] | 0.18 | 0.68 | -- | -- | -- | | | |
| B.F.Dl | F,RS | 0.77 | 0.28 | [0.2,1.4] | -- | -- | [0.0,0.7] | 0.98 | 0.61 | | | |
| B.F.Dl *AGE | F | -0.07 | 0.97 | [-1.7,1.6] | 0.01 | 0.95 | -- | -- | -- | | | |
| B.F.Dl *AGE$^2$ | F | -0.15 | 0.92 | [-1.8,1.4] | 0.03 | 0.87 | -- | -- | -- | | | |
| Delay | F | -0.26 | 0.14 | [-0.5,0.0] | 3.43 | 0.06 | -- | -- | -- | | | |
| Intercept | F | 0.71 | 0.28 | [0.2,1.4] | -- | -- | -- | | | | | |
| Subject | RI | -- | -- | -- | -- | -- | [0.0,0.6] | 0.16 | 0.92 | | | |

See Tables 2-3 and 2-4 of the main article for abbreviations

**Perseveration**

Table C-3g

*Summary statistics for side bias by sex*

| Sex | *n* | AGE | \|Side Bias\| |
|---|---|---|---|
| Females | 12 | 64.3 ± 43.3 | 0.14 ± 0.10 |
| Males | 11 | 111.1 ± 36.0 | 0.09 ± 0.07 |

*n* = number of observations (dogs). |Side Bias| = the absolute value of the difference in the proportion of correct choices made at the left and the right boxes

OLS Quadratic Regression: $|\text{Side Bias}| = \text{Sex} + \text{AGE.CENT} + \text{AGE.CENT}^2$

Sex did not predict side bias and accounting for sex had no effect on the analysis.

The overall model, $F(3,19) = 3.40$, $p = .04$; $r = 0.35$, and the coefficient for $\text{AGE.CENT}^2$ significantly predicted side bias, $b_2 =$ 2e-5, $SE_{b2} =$ 8e-6 ; $t(19) = 2.70$, $p = .01$. The coefficients for AGE.CENT, $b_1 =$ -8e-4, $SE_{b1} =$ 5e-4; $t(19) = -1.72$, $p = .10$, and sex, $b_3 = -0.001$, $SE_{b3} = 0.04$; $t(19) = -0.03$, $p = .01$, were not significant.

## Delay Optimization

Table C-3h

*Summary statistics for PCC by sex in each delay attempted by at least 20 dogs*

| | Females | | | Males | | |
|---|---|---|---|---|---|---|
| Delay (s) | *n* | AGE | $PCC_d$ | *n* | AGE | $PCC_d$ |
| 15 | 11 | 66.5 ± 44.7 | 0.91 ± 0.16 | 11 | 111.1 ± 36.0 | 0.89 ± 0.15 |
| 30 | 11 | 66.5 ± 44.7 | 0.88 ± 0.17 | 11 | 111.1 ± 36.0 | 0.90 ± 0.14 |
| 45 | 11 | 66.5 ± 44.7 | 0.88 ± 0.17 | 11 | 111.1 ± 36.0 | 0.90 ± 0.16 |
| 60 | 11 | 66.5 ± 44.7 | 0.80 ± 0.23 | 11 | 111.1 ± 36.0 | 0.79 ± 0.22 |
| 90 | 10 | 71.2 ± 44.2 | 0.87 ± 0.17 | 10 | 109.9 ± 37.8 | 0.83 ± 0.27 |

$PCC_d$ = proportion of correct choices at a given delay

OLS Quadratic Regression (separate model for each delay): $PCC_d = Sex + AGE.CENT + AGE.CENT^2$

Sex did not predict PCC in any delay. The effects of Sex on the significance and predictive strength of the overall model and the other predictors in the model were negligible save for two potential, albeit minor, exceptions. The 45-s delay model was significant without ($p = .02$), but not with ($p = .06$) a coefficient for Sex. In addition, the overall prediction of the 90-s delay model was substantially lower without a coefficient for Sex ($r^2 = 0.28$) but neither model was significant ($p = .06$).

Table C-3i

*Models of PCC by sex and age and in each delay attempted by at least 20 dogs*

| | | Model Prediction | | | AGE.CENT | | AGE.CENT$^2$ | | Sex | |
|---|---|---|---|---|---|---|---|---|---|---|
| Delay (s) | $n$ | $r^2$ | $F$ | $p$ | $t$ | $p$ | $t$ | $p$ | $t$ | $p$ |
| 15 | 22 | 0.23 | 1.80 | 0.18 | 0.83 | 0.42 | -2.30 | 0.03 | -0.88 | 0.39 |
| 30 | 22 | 0.15 | 1.05 | 0.40 | 0.61 | 0.55 | -1.72 | 0.10 | -0.21 | 0.84 |
| 45 | 22 | 0.33 | 3.00 | 0.06 | 0.42 | 0.68 | -2.96 | 0.01 | -0.28 | 0.78 |
| 60 | 22 | 0.35 | 3.24 | 0.05 | -1.15 | 0.27 | -2.53 | 0.02 | 0.12 | 0.91 |
| 90 | 20 | 0.37 | 3.11 | 0.06 | 1.51 | 0.15 | -2.91 | 0.01 | -1.46 | 0.17 |

AGE.CENT = centered age in months

Table C-4

*Dogs below chance on zero-delay or delay trials*

| | Zero-delay Trials | | Delay Trials | |
|---|---|---|---|---|
| Age (months) | Correct / Attempts | Binomial Test: p(Correct \| Attempts) | Correct / Attempts | Binomial Test: p(Correct \| Attempts) |
| 101 | 7/8 | 0.07 | 19/20 | 0.00004 |
| 56 | 5/7 | 0.45 | 18/19 | 0.00007 |
| 95 | 5/6 | 0.22 | 21/22 | 0.00001 |
| 20 | 26/30 | 0.0001 | 14/24 | 0.54 |
| 181 | 22/25 | 0.0002 | 12/17 | 0.14 |
| 19 | 20/21 | 0.00002 [a] | 17/24 | 0.06 |
| 29 | 16/16 | 0.00003 | 17/24 | 0.06 |

[a] This dog made incorrect choices in 10/12 trials before the side bias counter-training procedure. It still exceeded chance when accounting for these trials

# APPENDIX D

# SUPPLEMENTARY MATERIAL FOR CHAPTER 3

Table D-1

*Study Subjects*

| Dog | Age (months) | Height (cm) | Weight (kg) | Sex | Test Completed | Owner-Reported Breed |
|---|---|---|---|---|---|---|
| Olive | 10 | 48.26 | 13.61 | F | Y | Mix |
| Booster | 10 | 66.04 | 34.93 | M | Y | Mix |
| Sterling | 13 | 43.18 | 9.98 | M | Y | Mix |
| Hamilton | 15 | 58.42 | 19.41 | M | Y | Mix |
| Buddy | 15 | 71.12 | 36.29 | M | Y | Yellow Lab |
| Summer | 16 | 50.80 | 19.05 | F | Y | Mix |
| Taryn | 18 | 48.26 | 11.79 | F | Y | Mix |
| Chomsky | 19 | 63.50 | 24.77 | F | Y | Golden Retriever |
| Finn | 30 | 48.26 | 15.88 | M | Y | Mix |
| Finley | 33 | 40.64 | 19.50 | F | Y | Mix |
| Ty | 36 | 33.02 | 10.89 | M | Y | CKC Spaniel |
| Gizmo | 36 | 41.91 | 9.62 | M | Y | Mini Australian Shepherd |
| Quest | 37 | 53.34 | 23.59 | F | Y | Black Lab |
| Oliver | 37 | 24.13 | 3.54 | M | Y | Chihuahua |
| Kristoff | 47 | 58.42 | 29.94 | M | Y | Black Lab |
| Radar | 48 | 62.23 | 42.55 | M | Y | German Shepherd |
| Chai | 51 | 60.96 | 26.13 | F | Y | Mix |
| Bodhi | 52 | 35.56 | 8.62 | F | Y | Mini Autralian Shepherd |
| Lacey | 52 | 53.34 | 20.41 | F | Y | Mix |
| Florence | 52 | 63.50 | 35.83 | F | Y | Golden Retriever |
| Calvin | 55 | 25.40 | 4.17 | M | Y | Mini Poodle |
| Shadow | 56 | 48.26 | 14.88 | M | Y | Mix |
| Sam | 58 | 56.90 | 25.49 | M | Y | Golden Retriever |
| Nikki | 59 | 63.50 | 29.03 | F | Y | Golden Retriever |
| Collins | 60 | 55.88 | 20.41 | F | Y | Mix |
| Misha | 60 | 49.53 | 19.23 | F | Y | Portuguese Water Dog |
| Tara | 65 | 60.96 | 32.84 | F | Y | German Shepherd |
| Winston | 66 | 33.02 | 11.07 | M | Y | Pug |
| Chase | 72 | 53.34 | 22.23 | M | Y | Mix |
| Truffles | 73 | 63.50 | 36.29 | F | Y | Golden Retriever |
| Maggie | 73 | 58.42 | 19.50 | F | Y | Border Collie |
| Rocky | 73 | 53.34 | 22.86 | M | Y | Portuguese Water Dog |
| Lilly | 83 | 29.97 | 4.99 | F | Y | Mix |
| Nubi | 84 | 50.80 | 12.70 | F | Y | Mix |
| Bandit | 86 | 54.61 | 19.23 | M | Y | Border Colllie |
| Hapa | 87 | 34.29 | 9.07 | F | Y | Mix |
| Bixby | 90 | 60.96 | 34.47 | F | Y | Mix |

| Dog | Age (months) | Height (cm) | Weight (kg) | Sex | Test Completed | Owner-Reported Breed |
|---|---|---|---|---|---|---|
| Rinzler | 93 | 55.88 | 22.86 | M | Y | Mix |
| Cleo | 96 | 54.61 | 19.78 | F | Y | Mix |
| Syd | 104 | 76.20 | 51.53 | M | Y | Rottweiler |
| Sadie | 107 | 50.80 | 25.85 | F | Y | Mix |
| Gizmo | 108 | 26.67 | 7.62 | F | Y | Mix |
| Pita | 111 | 60.96 | 27.22 | M | Y | Mix |
| Dex | 112 | 58.42 | 28.21 | M | Y | Mix |
| Nutella | 113 | 22.86 | 4.08 | M | Y | Mix |
| Duke | 120 | 53.34 | 24.49 | M | Y | Mix |
| Rosie | 121 | 27.31 | 4.08 | F | Y | Mix |
| Rosita | 122 | 24.13 | 3.18 | F | Y | Mix |
| Moezie | 126 | 55.25 | 19.87 | F | Y | Mix |
| Marley | 129 | 71.12 | 37.65 | M | Y | Mix |
| Harold | 132 | 35.56 | 8.62 | M | Y | Mini Poodle |
| Obie | 132 | 30.48 | 6.89 | M | Y | Mix |
| Murphy | 132 | 60.96 | 25.04 | M | Y | Wheaton Terrier |
| Daisy | 132 | 53.34 | 25.67 | F | Y | Blue Heeler |
| Chaco | 134 | 66.04 | 34.02 | M | Y | Mix |
| Devyn | 134 | 45.72 | 14.51 | F | Y | Border Collie |
| Barleigh | 135 | 66.04 | 25.40 | F | Y | Mix |
| Venzy | 147 | 58.42 | 25.85 | F | Y | Golden Retriever |
| Odie | 147 | 58.42 | 25.67 | M | Y | Mix |
| Kate | 165 | 38.10 | 9.89 | F | Y | Tibetan Terrier |
| Jin | 170 | 68.58 | 37.42 | F | Y[a] | Akita |
| Klev'r | 173 | 33.02 | 8.26 | M | Y | Border Terrier |
| Reese | 18 | 62.23 | 22.68 | M | N | Standard Poodle |
| Clarabelle | 28 | 53.34 | 19.05 | F | N | Mix |
| Pumpkin | 28 | 53.34 | 21.14 | M | N | Brittany |
| Bear | 35 | 76.20 | 49.90 | M | N | Great Pyrenees |
| Otis | 47 | 71.12 | 36.29 | M | N | Mix |
| Zelda | 50 | 83.82 | 65.77 | F | N | Great Dane |
| Beulah | 51 | 53.34 | 24.95 | F | N | Mix |
| Toshiko | 71 | 41.91 | 9.98 | M | N | Shiba Inu |
| Repede | 72 | 58.42 | 27.94 | M | N | Siberian Husky |
| Miley | 72 | 48.26 | 27.22 | F | N | Mix |
| Tegan | 79 | 53.34 | 19.41 | F | N | Mix |
| Tipper | 90 | 48.26 | 13.61 | F | N | Mix |
| Bo | 96 | 67.31 | 31.30 | M | N | Poodle |
| Soba | 114 | 55.88 | 36.29 | F | N | Mix |
| Cody | 129 | 55.88 | 25.85 | M | N | Australian Shephard |

| **Dog** | **Age (months)** | **Height (cm)** | **Weight (kg)** | **Sex** | **Test Completed** | **Owner-Reported Breed** |
|---|---|---|---|---|---|---|
| Baxter | 135 | 30.48 | 7.26 | M | N | Pug |
| Indie | 144 | 55.88 | 20.41 | F | N | Mix |
| Hershey | 170 | 66.04 | 30.21 | F | N | Standard Poodle |

a. See supplementary material D-3 for analysis of LSPE with Jin included.

## D-2 Materials and Procedures

### D-2.1 Materials and Layout

Throughout the test, a demonstrator (E1) remained in a testing room, an assistant (E2) remained in an adjacent holding room, and the dog moved between these rooms (Fig. D-2.1). The dog’s owner sat at the back of the holding room facing away from and ignoring the dog. One-meter tall collapsible, plastic fencing was used to restrict the dog to a 5 by 5m area in the testing room and a 2 by 1m area in the adjacent holding room. E2 slid a 1.5 x 1.0 x 0.04m piece of foam insulation board into the doorway between rooms to block the dog’s view of, and access to the testing room.

Treats were hidden in one of two identical boxes on the ground of the testing room, set 2m apart and 2.5m from both the doorway and E1. Plastic dog fencing surrounding a tripod and camera created a semi-transparent barrier between the left and right box. Treat boxes consisted of a 0.2 x 0.2 x 0.1m wooden base filled with sand upon which E1 placed treats and a 0.1 x 0.2 x 0.2m shield facing the holding room which blocked the dog's view of, and access to treats from this direction (Fig D-2.3). The shield consisted of a transparent plastic face housed in a wooden frame which contained a slot for an opaque plastic sheet (an occluder). The base contained a hidden compartment for smell controls: treats placed in the boxes before each session to prevent dogs from locating treats by smell.

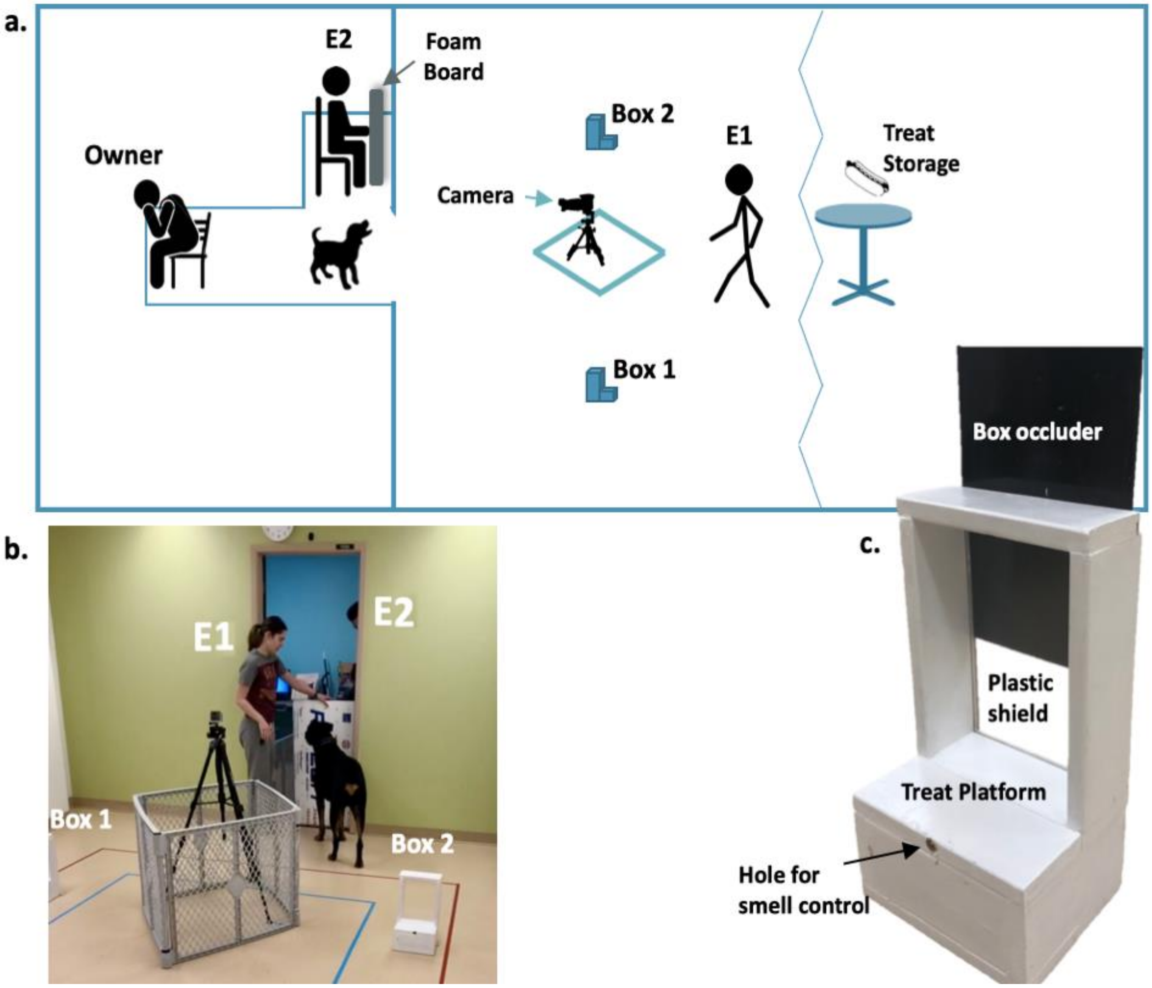

*Figure D-2.1.* Schematic of the experimental layout. **a)** Diagram of the rooms in which the experiment was conducted. **b)** E1 acclimating a dog to the foam board used to separate the rooms. **c)** A treat box with a removable visual occluder partially inserted.

### D-2.2 Acclimation and Shaping Procedures

To prevent satiation, dogs were fasted for six hours before the test. Prior to shaping, each dog was allowed to freely explore the testing and waiting areas. This acclimation period concluded when the dog ceased all exploratory behaviors and at least five minutes had elapsed.

After acclimating, each dog was required to complete four steps of shaping. As in testing, only one location was baited in any given trial. In shaping Steps 1 and 2 the foam board was not used so that the dog could observe where E1 placed treats. In Step 1, the treat was placed on the ground 1 m in front of the holding area. For the remainder of shaping, treats were placed on a box but only one box was set out in each trial. The box that was baited followed a pseudo-randomized order (L, R, L, R, L, L, R, L, R, R; L= Left, R = Right). In Steps 2 and 3 the occluder was not placed in the box shield so that the dog could see the treat through the Plexiglas window. In Step 4, the box occluder was added so that the dog was required to walk past the front of the box in order to see the treat.

To progress to Step 2, the dog was required to immediately retrieve the treat in two successive trials. To progress to Steps 3 and 4, the dog was required to retrieve the treat within two minutes in four consecutive trials. If the dog failed to retrieve the treat, the trial was repeated at the same box.

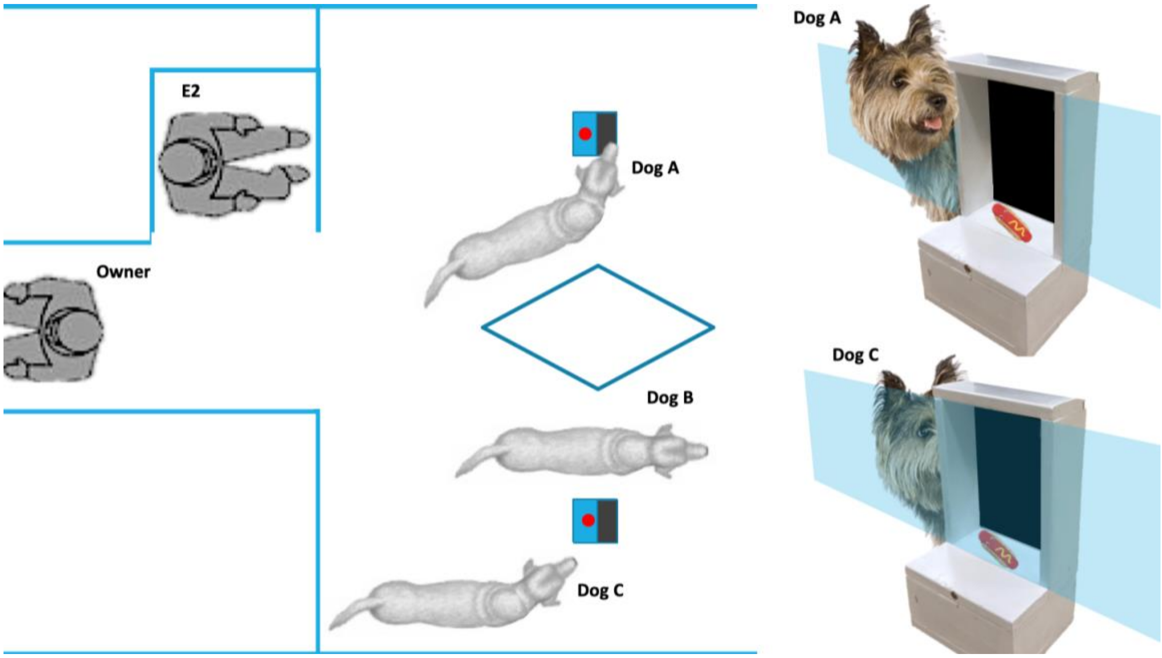

*Figure D-2.3.* Coding choices. The dog's choice was operationally defined as the box the dog first turned its head towards once its eyes were beyond the side of the box shields. **Dog A)** In almost all trials the dog walked directly toward and brought its snout to within a few centimeters of the chosen box. **Dog B)** The dog has not made a choice because it has not turned toward one of the boxes. **Dog C)** The dog has not made a choice because the treat is not yet in view. The distance between boxes was great enough that a treat on the closer box would become visible long before a treat on the box across the room. Thus, the closest box would be scored as the dog's choice if it continued on its current trajectory.

## D-3 Analyses of the Dog's Longest Streak of Perseverative Errors (LSPE)

With all 62 dogs included in the analysis, the model still significantly predicted LSPE, $F(2,59) = 4.57$, $p = .01$; $r^2 = 0.13$, and the effects of both AGE, ß = -9.90, $SE = .004$; $t(59) = -2.37$, $p = .02$, and AGE$^2$ were still significant, ß = 6.58e$^{-05}$, $SE = 2.38\text{e}^{-05}$; $t(59) = 2.76$, $p = .008$. Again, perseverative streaks were longer in middle-aged dogs than in young and old dogs. However, this model failed a Non-Constant Variance Score Test of fitted values, ncvTest: $X^2(1) = 5.25$, $p = .02$. Case-wise diagnostics revealed that this was a result of a single outlier, Jin (SDRESID = -2.80; DFFITS = -1.39, SDBETA$_{AGE2}$ = -1.07; notation follows from Cohen et al. 2003). Jin was the second oldest dog tested (AGE = 170) but never committed more than two consecutive perseverative errors, which was the second lowest LSPE displayed by any dog. With this outlier removed the ncvTest was not significant, $X^2(1) = 0.13$, $p = .72$.

Cohen J, Cohen P, West SG, Aiken L (2003) Applied multiple regression/correlation analysis for the behavioral sciences, 3rd ed. Lawrence Erlbaum, Mahwah.

# D-4 Analyses of Age-Weight Interactions

## D-4.1 Trial Outcome

To test whether the relationship between age and performance varied as a function of body size, we repeated the analysis of trial outcome (main text) but included additional fixed effects for the interaction between weight and age, and the interaction between weight and age$^2$. The model did not converge (due to overfitting) and thus, the estimates may not be reliable.

Likelihood ratio tests of competing models cannot be used test the significance of main effects included in interactions. Therefore, we calculated 95% profiled confidence intervals of each fixed-effect estimate and random-effect variance using parametric bootstraps to assess the significance of each predictor. Coefficients with confidence intervals that did not overlap with zero were considered significant ($\alpha = .05$).

The overall model significantly predicted trial outcome, $X^2(10) = 31.86$, $p = .0004$. Neither the interaction between age and weight nor the interaction between age$^2$ and weight were significant. The linear effect of age was not significant. All other effects were nearly identical to the model without age-weight interactions (see main text). In summary, including age-weight interactions in the model did not change the outcome of the analysis.

Table D-4.1

*Model of Trial Outcome with Age-Weight Interactions*

| | **Coefficients** | | | **Model Fit** | | | **LRT (vs. Null)** | |
|---|---|---|---|---|---|---|---|---|
| **Predictor** | **ß** | ***SE*** | **95% CI** | ***AIC*** | ***BIC*** | $\frac{r^2}{r(df)}$ | $X^2$ | ***p*** |
| | | | | 3900 | 3966 | 0.98 | 31.86 | 0.0004 |
| Age | 0.45 | 0.25 | [-0.01, 0.93] | | | | | |
| Age$^2$ | -0.53 | 0.25 | [-1.00, -0.07] * | | | | | |
| Trial | 0.28 | 0.11 | [0.13, 0.42] * | | | | | |
| REV | -0.43 | 0.13 | [-0.58, -0.26] * | | | | | |
| Height | 0.01 | 0.15 | [-0.28, 0.36] | | | | | |
| Weight | -0.13 | 0.15 | [-0.44, 0.16] | | | | | |
| Sex | -0.12 | 0.12 | [-0.34, 0.11] | | | | | |
| Age*Weight | -0.01 | 0.24 | [-0.50, 0.41] | | | | | |
| Age$^2$*Weight | 0.10 | 0.25 | [-0.36, 0.62] | | | | | |
| Intercept | 0.17 | 0.08 | [0.01, 0.32] | | | | | |
| Rnd.Sbj.Int | -- | 0.10 | [0.12, 0.40] * | | | | | |

LRT: Likelihood Ratio Test. AIC and BIC: Akaike & Bayesian Information Criterion (smaller values indicate better fit). $r^2/r(df)$: sum of squared Pearson residuals divided by residuals degrees of freedom (deviations from 1.0 larger than |0.2| may indicate over/ under-dispersion). REV: cumulative reversal number. Rnd.Sbj.Int: Random intercepts for study subjects (dogs). Asterisks indicate significant predictors (note the fixed effect intercept is not a predictor).

## D-4.2 Total Reversals

To test whether the relationship between age and the total number of reversals completed during the test varied as a function of body size, we repeated the analysis of total reversals (main text: section 2.3.2) but with additional effects for the interaction between weight and age, the interaction between weight and age$^2$, and the main effect of weight. However, we did not have sufficient degrees of freedom to include these additional covariates and thus, tests of all predictors in this model were likely underpowered.

As in the original analysis, the goodness-of-fit test was not significant, $X^2(59) = 70.58$, p = .09. Neither age, ß = 6.73e$^{-03}$, *SE* = 1.59e$^{-04}$; $p(>|Z|) = .67$, nor age$^2$, ß = -6.81e$^{-}$

$^{05}$, *SE* = 1.00e$^{-04}$; *p*(>|*Z*|) = .50, predicted total reversals. The main effect of weight was not significant, ß = -1.91e$^{-02}$; *SE* = 2.37e$^{-02}$; *p*(>|*Z*|) = .42. The interaction between age and weight, ß = -9.98$^{-05}$, *SE* = 6.26e$^{-04}$; *p*(>|*Z*|) = .87, and the interaction between age$^{2}$ and weight, ß = 1.74e$^{-06}$, *SE* = 13.63e$^{-06}$; *p*(>|*Z*|) = .63, did not predict total reversals.

### D-4.3 Longest Streak of Perseverative Errors

To test whether the relationship between age and bouts of poor performance varied as a function of body size, we repeated the analysis of the longest streak of perseverative errors committed by the dog during the test (main text: section 2.3.3) but with additional effects for the interaction between weight and age, the interaction between weight and age$^{2}$, and the main effect of weight. Again, we did not have sufficient degrees of freedom to include these additional covariates and thus, tests of all predictors in this model were underpowered.

The overall model significantly predicted the longest streak of perseverative errors, *F*(5,55) = 3.73, *p* = .005; $r^2$ = 0.25. The linear effect of age was not significant, ß = -0.012, *SE* = .009; *t*(55) = -1.34, *p* = .19. However, the curvilinear of effect of age$^{2}$ was nearly significant, ß = 9.62e$^{-05}$, *SE* = 4.98e$^{-05}$; *t*(55) = 1.93, *p* = .06. Middle-aged dogs still tended to display shorter perseverative streaks than did young and old dogs. The main effect of weight was not significant, ß = 0.01, *SE* = 0.02; *t*(55) = 0.72, *p* = .47. Neither the interaction between weight and age, ß = -3.64e$^{-05}$, *SE* = 4.13e$^{-04}$; *t*(55) = -0.08, *p* = .93, nor the interaction between weight and age$^{2}$, ß = -5.85e$^{-07}$, *SE* = 2.48 e$^{-06}$; *t*(55) = -0.24, *p* = .81, did not predict the dogs longest streak of perseverative errors.

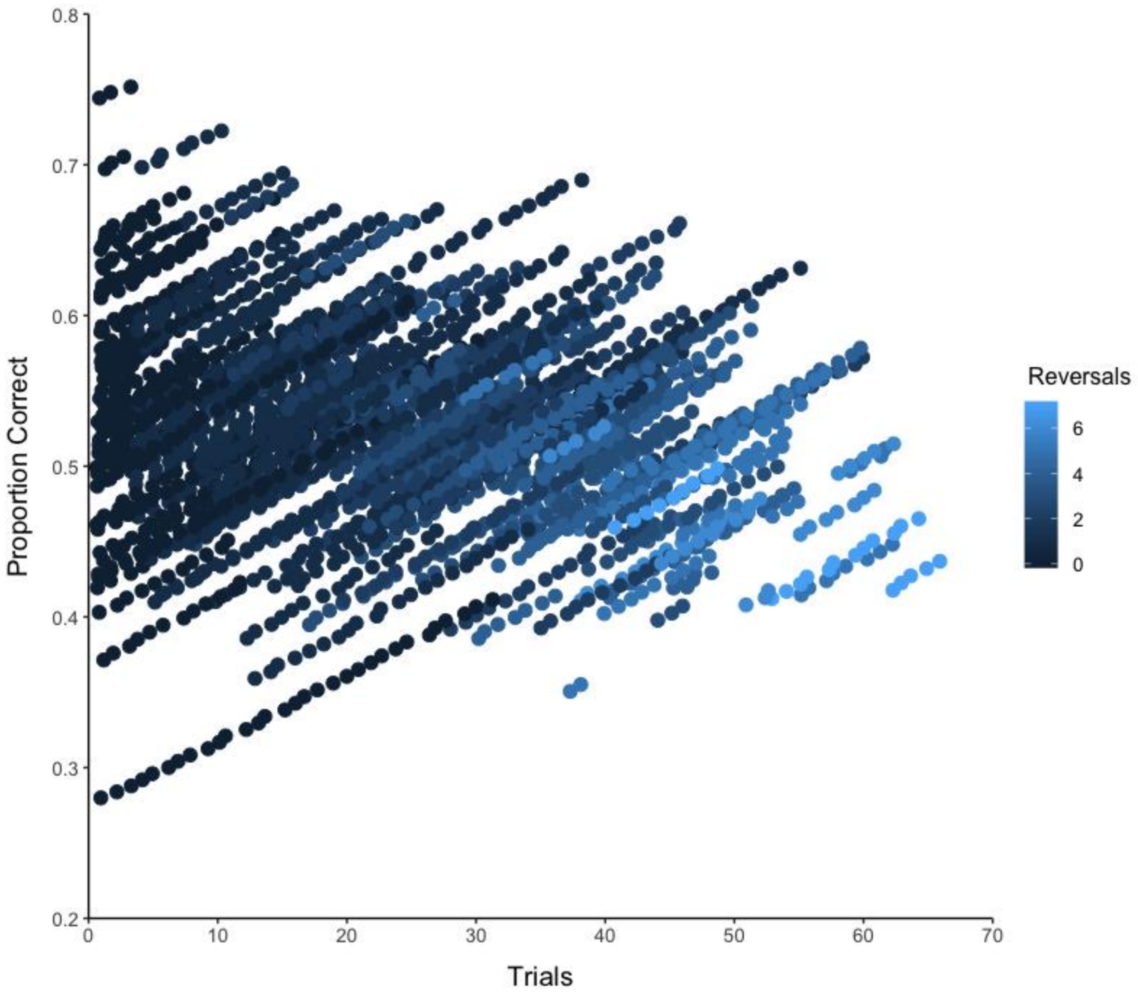

*Figure D-5.* Fitted values of trial outcome as a function of trial number and reversal number. Fitted values for the likelihood of choosing the correct box were estimated from the full model of trial outcome (generalized linear mixed model with a binomial error distribution) described in the main text. Darker points indicate higher reversal numbers (i.e., that the dog had completed more reversals). Different lines of points with the same reversal number (stratified points of the same color) correspond to different individuals. After accounting for reversal number and individual variation, dogs were more likely to choose the correct box with increasing trial number indicating that they did not complete reversals by simply choosing at random. After accounting for trial number and individual variation, dogs were less likely to choose the correct box after each reversal indicating that previous reward contingencies interfered with the current reward contingency and that this interference was additive across repeated reversals.

# APPENDIX E

# SUPPLEMENTARY MATERIAL FOR CHAPTER 4

# E-1 Supplementary Analyses

## Experiment I

To test whether age predicted cumulative serial reversal learning speed, we conducted a linear regression analysis of the number of errors committed before the coyote completed ten reversals. To test whether age predicted speed of acquisition of the win-stay and lose-shift rules, we conducted a linear regression analysis of trials to nine correct WSLS choices in ten consecutive trials (natural log-transformed). To test whether age predicted acquisition of rapid reversal learning within a single session, we constructed a generalized linear mixed model (GLM) of sessions to meet the Session Reversals criterion with a zero-truncated negative binomial error distribution. To confirm that accounting for overdispersion was necessary, we compared the overall prediction of this model to that of GLM with a zero-truncated Poisson error distribution (which assumes dispersion is equal to one). In all models, age in months was the only predictor.

Older coyotes made significantly more errors before completing ten reversals, $F(1,19) = 6.39$, $p = .02$; $r^2 = 0.25$, ß = 0.25 (Fig. E-1), took significantly longer to reach the WSLS threshold, $F(1,18) = 24.1$, $p = .0001$; $r^2 = .57$, ß = 0.02, and took significantly longer to satisfy the Session Reversals criterion, $Z = 3.1$, $p = .002$; ß = 0.02. Overdispersion in the GLM of session to complete the Session Reversals criterion was significantly greater than one, $X^2(1) = 7.89$, $p = .005$.

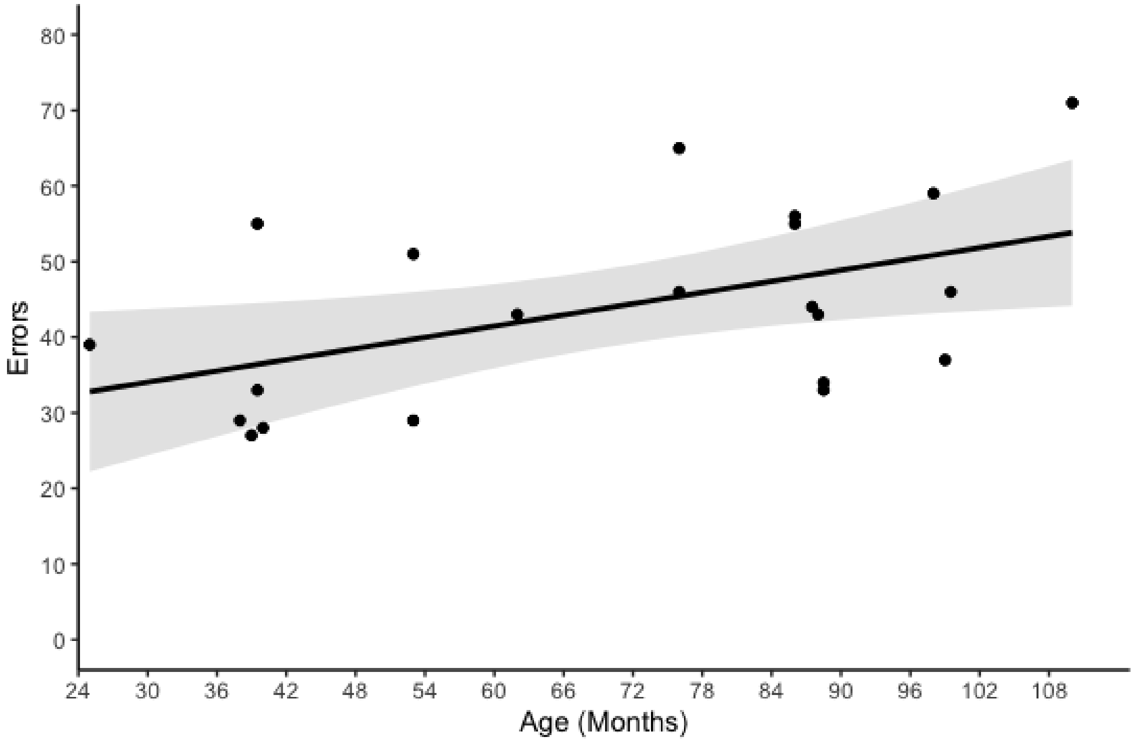

*Figure E-1.* Errors to complete ten reversals by age in months.

## Experiment II

To test whether age predicted speed of task acquisition in Experiment II, we conducted a linear regression analysis of trials to 17 correct choices in 20 trials. Older coyotes took significantly longer to acquire the task, $F(1,7) = 6.23$, $p = .04$; $r^2 = 0.47$, ß = 1.77.

## Assumption Tests, Case-wise Diagnostics, and Violations

To test for homoscedasticity and linearity, Non-constant Variance Score Tests were conducted and residual vs. fitted and residual vs. age plots were inspected. To assess whether residuals were normally distributed, standard normal quantile plots and residual density plots were inspected.

To test for outliers, leverage ($h_{ii} > 2p/n$), externally studentized residuals (SDRESID > t(n-p-1)), standardized DFFIT ($DFFITS_i > 1$), and Standardized DFBETA were examined. When potential outliers were identified, the analysis was rerun with the corresponding data omitted to confirm that the model was not dependent on individual test subjects. Notation for regression diagnostics follows Cohen et al. (2003).

## E-2 Experiment II Method

Although color discrimination tasks have historically used inanimate objects of different colors as the S+ and S-, coyotes are highly neophobic and would have found novel, colored objects/ stimuli aversive. Though it is possible that habituation procedures could have been used to overcome such aversions, these procedures would have been time-consuming (likely would have spanned weeks) and would have interrupted the necessary continuity between Experiments 1 and 2. In contrast, coyotes had already overcome aversions to the experimenters through hundreds of shaping, training, and testing trials for Experiment 1 (at least to the extent necessary for participation).

## E-3 Differences in the Procedures Used to Test Coyotes and Dogs

The procedures used to test dogs and coyotes were designed for different purposes. The aim of Van Bourg et al. (2021) was to design a clinical assessment for age-related deficits in inhibitory control which could be administered rapidly within a single session. As in the present study, dogs were asked to learn which of two locations contained a hidden food treat, and each time the dog chose the correct location in three consecutive trials the procedure was repeated using the other location. The purpose of the present study was not to design a rapid clinical assessment for coyotes. Whereas cognitive deficits in elderly dogs are well documented, it was previously unknown whether coyotes develop similar age-related deficits. Thus, the present study used an extensive multi-session protocol to examine behavioral flexibility and cognitive aging in more depth than was necessary for the purposes of Van Bourg et al. (2020).

Furthermore, the previous serial reversal learning procedure used to test pet dogs, could not be used to test these captive coyotes. Coyotes were neophobic, fearful of humans, and could not be safely handled or herded away from a given location within a small room.

The fact that some coyotes required multiple session to complete 30 trials may have hindered the performance of these individuals if continuity between trials was beneficial for learning the reversing reward contingencies. Alternatively, pauses between testing sessions may have provided the advantages of reducing recent interference from prepotent behaviors and interrupting streaks of perseverative choices. However, the inclusion of session number as a covariate in the species comparison model should have addressed this procedural difference.

The two food blinds were 2m apart for dogs but were roughly 15m apart for coyotes. Thus, the shorter distance between food blinds may have made inhibiting incorrect choices more difficult for dogs. In addition, the cost of an incorrect choice may have been greater for coyotes given that more time and energy were required to search either location for food. This may have provided a stronger incentive for coyotes to choose the correct side. Furthermore, coyotes may have been more motivated to learn the task because their treats were composed of portions of their daily food rations rather than supplementary treats. However, it is also possible that the high-value treats used for dogs (hot dogs) were more rewarding than typical coyote feed.

Incorrect locations may have been more salient for coyotes than for dogs given that coyotes watched an experimenter walk to each of the blinds in every trial while dogs watched only one experimenter walk to only one blind in each trial. Moreover, experimenter-induced stimulus enhancement of the incorrect location may have increased the difficulty of the task for coyotes.

# APPENDIX F

# SUPPLEMENTARY MATERIAL FOR CHAPTER 5

Table F-1

*Study Subjects*

| **EID** | **Name** | **Species** | **Age (months)** | **Sex** | **Facility** | **Test Date** | **Total Trials** | **Total Reversals** |
|---|---|---|---|---|---|---|---|---|
| 1 | Máni | Gray Wolf | 36 | M | WP | 3/20 | 20 | 2 |
| 2 | Khewa | Gray Wolf | 36 | F | WP | 3/20 | 31 | 2 |
| 3 | Sparrow | Gray Wolf | 36 | F | WP | 3/20 | 145 | 8 |
| 4 | Aspen | Gray Wolf | 36 | M | WP | 3/20 | 7 | 0 |
| 5 | Niko | Gray Wolf | 36 | M | WP | 3/20 | 16 | 0 |
| 1 | Máni | Gray Wolf | 50 | M | WP | 6/21 | 12 | 1 |
| 2 | Khewa | Gray Wolf | 50 | F | WP | 6/21 | 78 | 3 |
| 3 | Sparrow | Gray Wolf | 50 | F | WP | 6/21 | 149 | 6 |
| 4 | Aspen | Gray Wolf | 50 | M | WP | 6/21 | 35 | 2 |
| 5 | Niko | Gray Wolf | 50 | M | WP | 6/21 | 37 | 3 |
| 6 | Timber | Gray Wolf | 86 | F | WP | 6/21 | 98 | 9 |
| 7 | Nikai | Gray Wolf | 89 | M | WCC | 10/21 | 222 | 17 |
| 8 | Zephyr | Gray Wolf | 125 | M | WCC | 10/21 | 256 | 16 |
| 9 | Alawa | Gray Wolf | 125 | F | WCC | 10/21 | 40 | 4 |
| 10 | Scarlette | Red Fox | 86 | F | WP | 6/21 | 165 | 7 |
| 11 | Joker | Red Fox | 110 | M | WP | 6/21 | 76 | 5 |
| 12 | Kestrel | Gray Fox | 25 | M | WP | 6/21 | 107 | 7 |
| 13 | Lark | Gray Fox | 25 | M | WP | 6/21 | 56 | 4 |

EID: Experiment identification number. WP: Wolf Park. WCC: Wolf Conservation Center

## F-2 Assumptions Testing

### OLS Regressions Analyses

To test for homoscedasticity and linearity, Non-constant Variance Score Tests were conducted and residual vs. fitted and residual vs. predictor plots were inspected. To assess whether residuals were normally distributed, standard normal quantile plots and residual density plots were inspected.

To test for outliers, leverage (hii > 2p/n), externally studentized residuals (SDRESID > t(n-p-1)), standardized DFFIT (DFFITSi > 1), and Standardized DFBETA were examined. When potential outliers were identified, the analysis was rerun with the corresponding data omitted to confirm that the model was not dependent on individual test subjects. Notation for regression diagnostics follows Cohen et al. (2003).

### Proportion Correct

The model of the proportion of correct WSLS choices in the first 10 trials met all assumptions and no outliers were identified. The model of the proportion of correct WSLS choices in the first 20 trials met all assumptions but several potential outliers were identified. With some of these outliers omitted, the model was no longer significant. However, the multiple R squared remained high with any individual omitted. The model of the proportion of correct WSLS choices in the first 75 trials met all assumptions but three potential outliers were identified. With two of these potential outliers omitted the model was unchanged. With the third removed the model yielded no prediction.

### LSPE

In the model of the longest streak of perseverative errors made in the first 10 trials, the Non-constant Variance Score Test was nearly significant ($p = .05$) suggesting that homoscedasticity may be present. Thus, the significance of this model may be underestimated. Four potential outliers were identified in this model. Removing two of these data points had minimal effects on the model. Removing the third potential outlier substantially improved prediction and removing the fourth potential outlier substantially reduced prediction.

In the model of the longest streak of perseverative errors in the first 20 trials, the Non-constant Variance Score Test was significant ($p$ = .02) suggesting that homoscedasticity may be present. Thus, the significance of this model may be underestimated. Four potential outliers were identified. Four potential outliers were identified. Removing three of these had no effect on the model. Removing the fourth data point reduced prediction and the model was no longer significant. However, the multiple r square remained high ($r^2$ = .45).

The model of the longest streak of perseverative errors in the first 75 trials met all assumptions. Three potential outliers were identified in this model. Removing two of these had no effect on the model. Without the third outlier omitted, the model yielded no prediction.

### GLMMs of Trial Outcome

We examined trial-by-trial performance using generalized linear mixed models (GLMMs) of trial outcome (correct or incorrect choice) with binomial error distributions (logit-link functions).

We first fit a full model with fixed effects for the predictor of interest and all potential covariates, random intercepts for subjects, and all possible random subject slopes (see Barr et al. 2013; Schielzeth and Forstmeier, 2009). To address potential overfitting and underpowering of these maximal models (see Bates et al., 2015a; Matuschek et al., 2017), we then removed non-significant random subject slopes and non-significant fixed effects for covariates beginning with the effect which accounted for the least variation in trial outcome. We continued this stepwise elimination process until all remaining covariates added significant prediction to the final (reduced) model. Because random slopes did not add prediction to any model in any analysis, the random-effects structure of each reduced model included only subject intercepts.

To test the significance of individual effects, we conducted Likelihood Ratio Tests of nested models (i.e., compared models with and without each predictor). To test for homogeneity of variance and linearity, we inspected boxplots and scatterplots of fitted values and residuals as a function of each predictor. To test for normality, we examined density and quantile plots of fitted values and each random effect. To test for overdispersion, we compared the sum of the squared Pearson residuals to the degrees of freedom of the residuals. All data were analyzed in R version 3.4.1. GLMMs were constructed and tested using the package “lme4” (Bates et al. 2015b).

All models met all assumptions.

# APPENDIX G

# AUTHOR PERMISSIONS FOR PUBLISHED WORKS

All authors of all published works (Chapters 2, 3, and 4, as well as their corresponding tables, figures, and appendices) have granted permission for these works to be included in this dissertation.