ANALOGICAL REASONING BY PINYON JAYS USING A NEW EXPERIMENTAL PARADIGM

by

Breanna Cheri

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

MASTER OF ARTS

Department of Psychology

University of Manitoba

Winnipeg

Copyright © 2021 by Breanna Cheri

Abstract

An analogy is the comparison of two pairs of ideas or items based upon the relationship of the features or functions that they share (e.g., a puppy is to a dog as a kitten is to a cat). Analogies are useful for gathering new information based on comparisons between familiar and novel items. I adapted a new paradigm, originally designed for humans, to assess analogical reasoning by a highly intelligent species of corvid, pinyon jays (*Gymnorhinus cyanocephalus*). My objective was to determine whether pinyon jays could use a same/different relational rule to progressively solve more difficult analogical reasoning problems. During an operant task, the birds were first presented with two items that could be defined by their similarity or difference along a common property (e.g., shape: two circles). The birds were required to learn the commonality between the first two items (e.g., same shape). They were then presented with a single third item (e.g., a square), followed by two choice items. The birds were required to accurately match the third item with the correct choice (e.g., a square), to create an analogical match (e.g., a same-shape relationship: a circle is to a circle as a square is to a square). Despite using a similar training procedure to that used with humans, the pinyon jays showed difficulty learning the task requirements. This resulted in an inability to test for concept learning. Described herein are aspects of the task, including stimulus presentation, early training procedures, and factors controlling the progression of a session that might have prevented learning. My thesis highlights specific training procedures that should be avoided when attempting to train pinyon jays on relational similarity. As there is currently a need for new methods of assessing analogical reasoning by non-human animals, the failure of the current study procedure provides insight into task requirements that may prove necessary for future studies. Specifically, when employing paradigms which attempt to train subjects on problems that require them to complete a presented analogy.

Acknowledgements

I would like to thank my supervisor, Debbie Kelly, for accepting me as a student and helping me to grow as a researcher. Her support has brought me many steps closer to my future goals and I cannot thank her enough. I thank my committee members, Randy Jamieson and Jim Young for their helpful comments and advice. Their support has been invaluable to my success as a Masters student and I am greatly thankful to them for challenging me. I would like to acknowledge the University of Manitoba's Department of Psychology for granting me the Psychology Graduate Fellowship and Research Manitoba for granting me the Master's Studentship award. Thank you Chris Vattheuer for assisting me with the Python code for my thesis project. I want to give a huge thanks to me fellow lab members who have offered me so much guidance and support! Especially, Alizée for reading countless drafts of my proposal and thesis. Thank you for being an amazing mentor and friend. Also, Kevin for helping me get my project off the ground and answering my endless questions! Thank you to my parents, Kevin and Laura, my siblings, Nicole, Cameron, and Sabrina, and of course all of my family and friends for your love and prayers. They have gotten me this far and continue to comfort me when I need it most. Also thank you Kayla for making sure I ate, exercised, and took care of myself through my program. You have brightened each of my days and I thank you so much for always being there to care for me.

Table of Contents

Abstract ii

Acknowledgments iii

List of Tables vii

List of Figures viii

Chapter I: Introduction 1

Chapter II: Literature Review 5

Chapter III: Methods 17

Subjects 17

Apparatus 17

Stimuli 18

General Procedures 19

Shaping Procedure 19

Analogical Reasoning Task 21

Training and Test Procedures 22

Arrangement of Training/Test Phases 24

Structure of a Trial 29

Dimension Exemplars 30

Chapter IV: Statistical Analyses 35

Dependent Variables 35

Percent Correct 35

Learning Rate 35

Independent Variables 36

Phase 36

Trial Type 36

Statistical Analyses 36

Test Performance 36
Training Performance 36
Chapter V: Training Outcome 37
Dimension 1 Training Modifications 37
Intertrial Interval and Timeout Duration 38
Correction Procedure Removal 38
Correction Procedure Addition 39
Choice Fixed Ratio Increase 40
Warning Stimulus 40
Individuals Restarting Dimension 1 Training 41
Statistical Analyses for Dimension 1 Training Data 42
Dependent Variable 42
Percent Correct 42
Independent Variable 43
Dimension 43
Statistical Analyses 43
Dimensional Performance 43
Results 43
Dimensional Performance 47
Chapter VI: Discussion 48
Thesis Overview 48
Results Summary 48
Objectives and Advantages of the Current Paradigm 49
The Current Analogical Reasoning Paradigm in Practice 50
Mistakes of the Current Study 51
Future Procedural Considerations 52
Stimuli Arrangement and Presentation 53
Initial Training Procedure 55

Training Procedure Elements....58
Correction Procedures....58
Fixed Ratio....59
Warning Stimulus....60
Learning Points and Contributions of the Current Study....61
Conclusions....63
Future Directions....64
Literature Cited....65

List of Tables

1. An Example of all training and test phases.....22
2. An assortment of stimuli arranged by the possible exemplars for each dimension.....31
3. A comparative illustration of the three combinations of dimensions that were same or different between a pair of stimuli.....32

List of Figures

1. Example of a match-to-sample problem that can be solved using conceptual or perceptual processing 9
2. Examples of relational match-to-sample trials for three dimensions a) size, b) colour, and c) shape, adapted from Smirnova et al., 2015.....14
3. Example of three displays for trials comprised of one dimension, the example used here is for shape 25
4. Example of a display for a trial comprised of a) two dimensions (i.e., shape and colour) and b) three dimensions (i.e., shape, colour, and pattern) 27
5. Example of the pseudo-counterbalancing for the order by which the birds will learn the three dimensions (i.e., shape, colour, and pattern).....34
6. Example display of the warning stimulus.....41
7. Percentage of correct choices per session.....44
8. Percentages of correct choices per session for birds that restarted Stage 1 training.....46
9. Examples of identity-match-to-sample (IMTS) and match-to-sample (MTS) trials 58

CHAPTER I

Introduction

Imagine attending a university lecture surrounded by hundreds of students, some of which are having ongoing conversations, whereas you are attempting to pay close attention to the speaker. There are several processes that you engage while listening, including perceiving visual input from presentation slides, processing new information, thinking about and committing this information to memory, and inhibiting noisy distractions. During this scenario you are using several types of cognitive processes—the mental actions by humans and non-human animals that involve attention, comprehension, memory, and the formation of thoughts and ideas. These cognitive processes are active every day and are fundamental components of cognition. Generally, cognition refers to how information, registered through the senses, is processed, retained and acted on (Shettleworth, 2001). However, even recently, controversy over what constitutes cognition and the extent to which cognition manifests in and outside of the body (i.e., how cognitive processes are integrated with or supported by components of the external environment) has caused disagreement across disciplines in how the term should be defined (Akagi, 2018). Even so, investigation of the components of cognition, namely how cognitive processes influence the mental capabilities of different species, remains an important focus of cognitive science (Stout, 2010; Thornton & Lukas, 2012). For instance, cognitive processes are shown to support early childhood development by contributing to academic performance, goal-directed behaviour, self-control, and emotional regulation (for a review see Raver and Blair, 2016). Cognitive processes are also linked to the adaptive behaviours that non-human animals exhibit in response to the demands of their environments, including navigating, foraging, communicating, learning from conspecifics (i.e., individuals of the same species), and problem solving (for reviews see Healy, 2019; Shettleworth, 2001). The interest in investigating the varying extent to which cognitive processing has evolved and is present across species gave rise to the field of comparative cognition.

Comparative cognition is a specific branch of cognitive psychology that is broadly focused on investigating the cognitive processing abilities of non-human animals, the neurological mechanisms that control those processes (Beran et al., 2014), and whether those processes and their resulting behaviours differ from those exhibited by humans (Zentall, 2018).

The origins of comparative cognition date back to the psychologists of the mid-1900s who were pioneers in the field of animal learning (MacDonald & Ritvo, 2016), such as B. F. Skinner. Skinner's interest in behavioural contingencies (i.e., the relationship observed between an animal's response to an outside stimulus and the reinforcing or punishing consequence; Catania, 1984) became a popular topic in the discipline of psychology and resulted from his invention of the operant chamber. Operant chambers are enclosed environments that a non-human animal is placed into and presented a stimulus, such as an image/a shape. Its response to this stimulus, such as time to respond or choice made, is measured. These behavioural responses indicate how the subject perceives the stimulus. For instance, if conditioned to find the stimulus reinforcing, the subject will respond with behaviour that leads to the reinforcing consequence (e.g., receiving a food reward), such as responding to the stimulus when presented. If given a task with multiple stimuli display at the same time, the subject's behaviour would indicate whether it has learned which stimulus response leads to reinforcement. The operant chamber contributed to the field of comparative cognition by advancing the study of conditioned behaviour. Moreover, behavioural (or operant) conditioning is a technique that can be used to understand how non-human animals' think and reason.

The idea that non-human animals are capable of having thoughts or beliefs has warranted much speculation. For instance, Clayton et al. (2006) showed that western scrub jays (*Aphelocoma californica*) possess beliefs about the state of decay and the security of the food they have stored over time in specific locations. The jays changed their food recovery strategies following changes to the conditions of their food stores (i.e., following longer intervals between storing and recovery, or whether they stored food in the presence of a potential thieving conspecific). This suggested that scrub jay behaviour is influenced by the beliefs that their food may perish or be stolen. This example and others showing that non-human animals adjust present behaviour in anticipation of future needs to point to adaptive thinking and planning, possibly even conceptual thought (Beck, 2012), which is defined as ideas about mental constructs that are not physically observable (e.g., time). For instance, some non-human animals have been shown to understand the concept of similarity. Cook (2002) discusses a number of operant chamber experiments in which pigeons (*Columba livia*) were shown *same* and *different* depictions of various classes of stimuli (e.g., rectangular matrices of elements, varying in shape or colour). Across all experiments, pigeons were successful in discriminating between displays of stimuli

representing either same or different relationships. These findings further support non-human animals' abilities to think about conceptual information. Yet, Beck (2012) explains that conceptual thinking by non-human animals is not entirely accepted across disciplines; rather some philosophers suggest that 'thinking' is not the term used for what non-human animals experience and that 'concepts' are beyond their comprehension. However, Beck explains that exhibiting the ability to think, which involves some ability to reason, would advance the idea that conceptual thinking *is* possible among non-human animals, but can only be revealed if the subject is capable of making logical inferences (e.g., *if* the weather is getting colder, *then* winter is drawing near). This is due to conceptual thoughts maintaining a compositional structure (e.g., possessing discrete concepts that can be combined, like the words of a sentence, to reveal information) that coincides with the logical sentence-like structure that is followed when forming a logical inference (Beck, 2012).

One means of investigating this ability to think conceptually is by determining if non-human animals are capable of reasoning about relationships. Specifically, can animals identify a relationship and then infer that the same relationship can apply to something completely novel, thus making an *analogical* inference? Analogy defines how a familiar context shares relational similarity with an unfamiliar context, thus aiding an individual's understanding of the unfamiliar context (Gentner & Maravilla, 2018). A familiar context may possess elements that are shared with the unfamiliar context. Recognizing the relational similarities between the shared elements results in candidate inferences that are applied to the unfamiliar context. An individual must determine whether the candidate inferences are true or false. If true, an analogy is established (Gentner & Maravilla, 2018). This is the basis for analogical reasoning.

Exhibiting analogical reasoning abilities would support the argument that a non-human animal is capable of conceptual thinking. Yet, there is limited research addressing analogical reasoning by non-human animals. Fortunately, analogical reasoning abilities can be investigated using operant conditioning. Analogical problems can be applied to this behavioural shaping technique, through which subjects can learn to recognize conceptual relationships and reveal their abilities to transfer this knowledge to novel stimuli. For this reason, in my thesis, I have created an operant chamber task to answer the question of whether the pinyon jay, a highly intelligent species of corvid which has never been examined for this ability, is capable of

reasoning analogically. This study can provide further evidence for how non-human animals think and reason. Additionally, this can demonstrate how the use of a new analogical reasoning paradigm is able to reveal analogical reasoning abilities which to date have been attributed to few species, mainly primates.

In the following chapter, I provided further explanation of analogical reasoning and reviewed the literature on analogical reasoning and same/different discrimination by non-human animals, some of the concerns regarding these investigations, and a description of my research objective. My review covers some of the speculations concerning whether subjects in earlier studies were reasoning by analogy, or whether other factors, such as language training or recognizing featural similarities, could account for their successful performance. I then propose the use of an analogical reasoning paradigm that addresses these earlier concerns and can provide an improved means of investigating this ability in a non-human animal. In Chapter III, I explained the paradigm and the testing procedures. In Chapter IV, I explain the experimental variables and how they were measured. In Chapter V, I review what problems I encountered and the results of the test subjects' performance. Finally, in Chapter VI, I provide a discussion of the factors future studies should consider to correct these problems and where this leaves the investigation of analogical reasoning by pinyon jays.

CHAPTER II

Literature Review

Impressive cognitive abilities for memory and foresight, communication, and problem solving have evolved in humans and many non-human animals. Some examples of such abilities include spatial and episodic (or episodic-like) memory (Clayton & Dickinson, 1998; Burgess et al., 2002; Dere et al., 2006; Clayton & Russell, 2009; Fagan et al., 2013), future planning (Raby et al., 2007), linguistic capabilities (Snowdon, 1990), and behavioural innovation (e.g., feeding and foraging techniques, and tool use; Tebbich et al., 2016). Recently, one area of cognition that has generated considerable interest and debate is the capacity for non-human animals to understand relational concepts—such as two objects having matching appearances or qualities—and by extension analogies.

Analogy is a form of relational reasoning. It requires the comparison of features or functions between two pairs of items/ideas. Analogies are often established when applying familiar concepts to novel situations, thus revealing informative relationships between items (Gentner & Holyoak, 1997). For instance, an analogy will be used to compare information from one set of items (i.e., set 1: both items are blue) to another set of items (i.e., set 2: both items are red) in order to draw a conclusion about their relatedness (i.e., set 1 and set 2 are alike in that the items in each set match in colour). Initially, pairs of items can appear largely similar or different based on item attributes such as appearance, but the strength of an analogy relies upon the conceptual relatedness between pairs of items (i.e., comparisons are abstract: 4 is twice the value of 2, similarly to how 10 is twice the value of 5) (Gentner, 1983).

An analogy can be broken down into its individual parts, which helps with inferring how those parts relate to one another (Gentner, 1983). An analogy can be formatted as *item A* is to *item A'* as *item B* is to *item B'*. For example, a puppy is to a dog, as a kitten is to a cat. This structure shows how the relationship between the first two items (i.e., puppy and dog) is the same as the relationship between the last two items (i.e., kitten and cat). To understand the analogy, the relationship between the first two items must first be understood; this has been referred to as a *first-order relationship* (Vonk, 2003). In the example above, the first-order relationship is of two animals of the same species, but differing in their developmental stage. Comparing the relationship between two items (determining the first-order relationship) and

constructing the relationship between the second two items is known as a s*econd-order relationship* (Vonk, 2003). Knowing that a puppy is the juvenile life stage of a dog provides the information necessary to conclude that a kitten is the juvenile life stage of a cat. Thus, the second-order relationship forms an analogy because it matches the relation (i.e., developmental stage) between relations (i.e., puppy is to dog, as kitten is to cat) (Gillan et al., 1981).

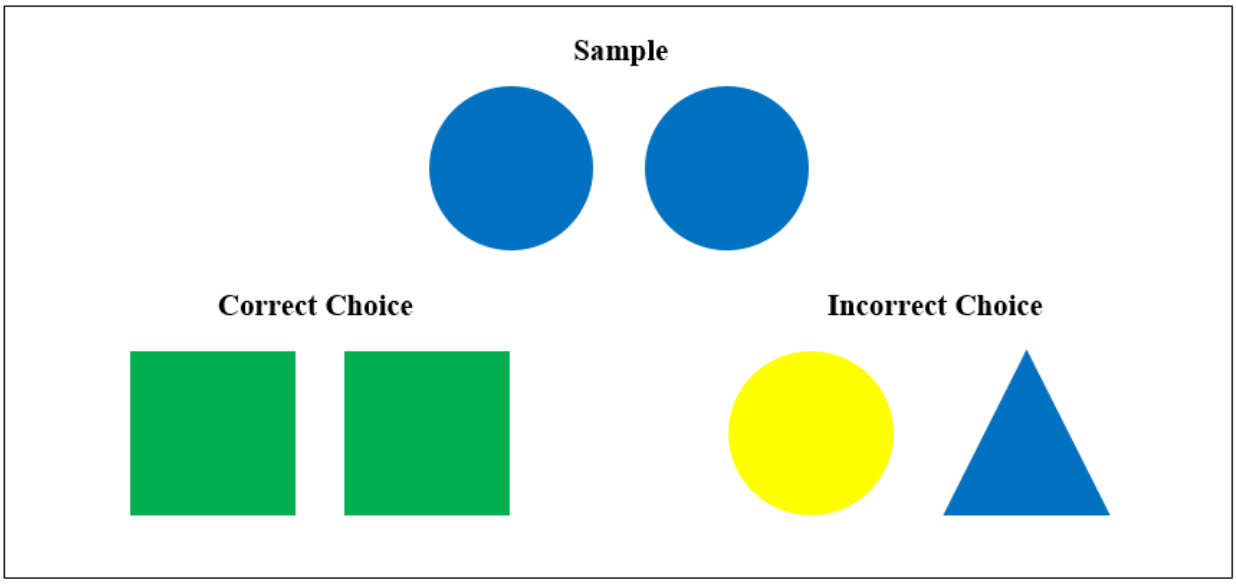
Much of the research into the reasoning abilities of non-human animals has been largely focused on various species of primates due to their evolutionarily relatedness to humans (Barrett et al., 2003; Thompson & Oden, 2000; Vonk & MacDonald, 2002; Vonk & MacDonald, 2004). Gillan et al. (1981) showed that a chimpanzee (*Pan troglodytes*), named Sarah, solved analogical reasoning problems that required an understanding of equivalence. Sarah was presented with the problem A:A′ same B:?, in which the letters represented cardboard cutout shapes and at the center of the display was a token that Sarah was previously trained to recognize as the concept *same*. Sarah was required to pick the correct matching stimulus, B′ that would complete the analogy from an alternative choice, C, which contained some, but not all features of the correct choice. Binomial testing revealed that Sarah chose the correct choice, B′, more often than the alternatives ($p < .001$). Although it was reasoned that choice accuracy could not be due simply to identifying physical similarities among stimuli, a subsequent test was given that presented the same problem displays, only this time carefully controlling for the influence that physical features among the stimuli may have on matching. Sarah again chose the correct alternative(s) significantly more than expected by chance ($p < .01$). Sarah's performance could be explained by her use of a relational rule, which allowed her to accurately match the relations between items. However, critics have argued that having been previously language trained (i.e., given extensive lexical training, which included an understanding of relational concepts through experience with symbolic tokens) may have also factored into Sarah's performance with analogical reasoning (Thompson et al., 1997). To answer the question of whether language training is needed as a prerequisite for conceptual matching, Thompson et al. (1997) used three problem sets to test language-naïve chimpanzees' abilities to solve analogies by having them match a set of stimuli to a corresponding set of stimuli based on physical or conceptual relatedness. Although, some subjects had past experience with symbolic tokens, they had no training on lexical or sentence structure, which was thought to be necessary for animals to exhibit an understanding of relations (Premack, 1988). Problem set 1 tested the chimps' matching abilities despite subjects having no

training on conceptual matching. Subjects were required to reach a performance criterion (i.e., 75% correct choices over two consecutive sessions) before advancing to subsequent problem sets. Problem set 2 tested conceptual matching ability with novel stimuli to ensure that performance with the first set was not due to 'an artifact' (pg. 34) of the previous stimuli. Problem set 3 tested performance after recombining familiar stimuli from the prior sets to determine if an associative strategy was being used for matching. An associative strategy would involve making choices based on a stimulus-specific association between the physical dimensions of a sample stimulus and a choice stimulus. For instance, "if sample AA appears, pick choice BB," therefore "whenever a sample has A in it, pick the choice that contains B." This strategy is not based on an understanding of conceptual relatedness, as this association would led to incorrect choices, "sample is AA, select choice BH (incorrect), rather than CC (correct)". Additionally, differential reinforcement was provided for physical matching, whereas non-differential reinforcement (i.e., all choices, regardless of correctness, were reinforced) was used for conceptual matching in order to examine spontaneous judgments (i.e., made without prior training, nor experience with reinforced learning). One potential concern with using this reinforcement method is the uncertainty in how the subjects could form accurate judgments of conceptual relatedness when given no corrective feedback for their mistakes. However, despite this and their lack of language training, three of the chimps' performed at above chance levels for all problem sets. Also, their performance was statistically identical between physical and conceptual matching and they showed no evidence of using an associative strategy, thus supporting that some species of primate can show concept learning without relying on linguistic knowledge. Instead, the authors suggest that the chimps' prior training in an unrelated task, which used a token that the chimps associated with the concept 'same', provided them with a basis whereby they could develop the comprehensive abilities necessary to detect conceptual relatedness in the current task.

Vonk (2003) tested the understanding of first- and second-order relationships in a gorilla (*Gorilla gorilla gorilla*) and four orangutans (*Pongo abelli*), none of which were language nor token trained. They were given two delayed-match-to-sample tasks that required them to remember a sample image with a particular colour and shape (e.g., a blue square). After a short delay, individuals were shown two comparison images. During the first experiment, the subjects had to choose between the comparison images based on the dimensions common between the

sample and two comparison images (i.e., first-order relationship). During the second experiment, each sample and comparison image depicted two coloured shapes (e.g., a blue square beside a red square) and choices were to be made based on the conceptual relatedness shared between the first pair of images (the sample) and the two pairs of corresponding images (the comparisons) (i.e., second-order relationship). Individuals from both species, who did not have prior testing on relational concepts, nor were familiar with the images, were capable of correctly matching the images in both cases. Vonk claimed that to succeed at the task, subjects had to attend only to relevant dimensions, which tested their ability to identify the correct dimension or concept (i.e., "same" [or different], "colour," and "shape") shared by the images. Additionally, Vonk claimed that having an understanding of second-order relationships might suggest that gorillas and orangutans are capable of more advanced reasoning—interpreting the relation-between-relations (i.e., reasoning by analogy). However, Flemming et al. (2008) argue that the apes in the Vonk (2003) study may have used a perceptual matching strategy to solve the task. They argue that it would have appeared that a subject had solved the task by choosing a comparison image that was perceptually less similar to the sample image, however the subject could still have chosen that image based on a perceptual rule, such as 'sample contains one colour, therefore the correct choice only has one colour'. For instance, Figure 1 shows a sample image of two circles displaying a *single* colour (blue). The incorrect choice displays an image composed of two colours, a yellow circle and a blue triangle. The incorrect choice shares one colour and one shape with the sample image, thus demonstrating perceptual similarities with the sample image that would need to be ignored. The correct choice displays an image of two squares with a *single unique* colour (green). Although the correct choice did not share a colour or shape with the sample, it could have been chosen because together the sample and correct choice exemplify a within-pair same colour *conceptual* relationship. Yet, the correct choice could also have been chosen because, like the sample, it possessed a single colour, thus showing how processing perceptual similarities, even when features are not exactly the same, can account for correct matching. This criticism raised concern that the presence of perceptual similarities between sample and choice item pairs, which could result in the development of perceptual matching strategies, need to be carefully controlled during studies of concept learning.

Figure 1. Example of a match-to-sample problem that can be solved using conceptual or perceptual processing. A conceptual match can be made by understanding that there is a within-pair same colour relationship between the sample and the correct choice. A perceptual match can be made by identifying that the correct choice displays a single colour, similarly to the sample. The perceptual match does not demonstrate an understanding of between-pair relationships that are based on abstract concepts (e.g., having *same* versus *different* dimensional relationships).



Although questions remain regarding whether non-human animals understand problems of conceptual relatedness, various species have shown a comprehension for the abstract (i.e., non-physical) concepts *same* and *different*—an important pre-requisite to analogical reasoning. Bovet and Vauclair (2001) examined same/different concept learning by baboons (*Papio anubis*) by having the primates distinguish whether various food and nonfood objects (e.g., an apple versus a coffee cup) belonged to the same functional category. For instance, classifying an apple and a banana as being the *same* because they both belong to the food category, even though they are non-identical food objects. Prior to experimentation, all objects were placed in the baboons' home enclosures for a day in order for the subjects to have an opportunity to manipulate and familiarize themselves with the various objects. Importantly, during the experiment there were no reliable perceptual features that could have been used to categorize the objects aside from the function they serve. During a trial, two objects were placed on a board in front of a subject's enclosure. A screen was raised to permit the subjects to view the objects and make a choice (pulling one of two ropes) to indicate whether the objects were the same or different. Experiment

1 served to test the baboons' abilities to classify physically identical and non-identical objects, in which they were successful. Experiment 2 tested the baboons' comprehension of functional relatedness between non-identical objects. The baboons were successful at categorizing the objects according to their functional (conceptual) relatedness as well. This study provides an example of non-human animals succeeding in a task that required an understanding of abstract concepts, without the ability to rely on a perceptual processing strategy. However, this study still only examined the understanding of first-order relationships between objects. Moreover, primates are not the only species shown to understand the concepts *same* and *different*. This ability extends to various species of birds.

Studies of Alex, a language trained grey parrot (*Psittacus erithacus*), examined his understanding of first-order relationships by his ability to label a pair of items by category (i.e., according to the singular dimension [colour, shape, or matter] in which items were similar or different; Pepperberg, 1987, 2013). For example, labelling a red wooden rectangular block and a red leather sphere as 'colour', because colour is the dimension along which they match. Alex could not rely of physical similarity (i.e., identity/non-identity) to solve the task, but rather had to process information of multiple dimensions to correctly respond to the question "what's same?" or "what's different?". Alex's correct response scores exceeded chance levels when tested with both familiar (i.e., similar, but not the same as training objects) and completely novel objects. Alex showed a comprehension for same/different concepts and, rather than fixating on and responding to questions with specific physical labels, like 'red', he showed an understanding of another form of abstract concepts (categorical classes). Cook et al. (1997) provide evidence that pigeons are also able to learn an abstract same/different rules, when required to differentiate uniformity in various visual arrays (i.e., the birds had to differentiate arrays that contained identical elements (*same* array), such as repeating circular dots, from arrays with a portion of non-identical elements (*different* array), such as a patch of square dots interspersed among circular dots). The arrays were categorized by 'type' based on their configurations, and included a wide variety of stimuli chosen to prevent the pigeons from relying on perceptual features to solve the task. Pigeons were trained and tested in a chamber equipped with a touch-sensitive colour monitor. During a trial, an array was presented on the screen (counterbalanced by same/different display and by type) and the subject would need to complete a five peck fixed ratio at the array before two choice locations (i.e., food hoppers) were presented. The pigeon was

then required to incline its head into one of the two hoppers to indicate whether it perceived the array to be *same* or *different*. The pigeons learned how to discriminate the *same* and *different* arrays, even as various types of arrays were concurrently presented across trials. Furthermore, the pigeons' rate of learning (i.e., the number of sessions completed before learning to discriminate same/different arrays) was not affected by the type of array. Indeed, the pigeons' responses showed that rather than learning separate discriminatory rules for each of the types of arrays (each containing its own unique stimulus features), they were able to apply a single relational rule across all types, even when tested with completely novel arrays. However, Katz and Wright (2006) suggest that the results obtained from the previous study cannot confirm that pigeons had learned a same/different relational rule. During a subsequent test in which the pigeons were presented with smaller arrays containing fewer (3-4) elements, their performance declined when shown *different* arrays (Cook et al., 1997)—thus supporting the conclusion that they were unable to apply a general rule across all types of arrays. This could suggest that the pigeons were using a perceptual judgment when making a choice, such as determining how *relatively* dissimilar elements of a single array were, and if not perceived as absolutely dissimilar, a bird might incorrectly indicate that it is viewing a *same* array. To address the inconclusive results of pigeon same/different discrimination tasks, Katz and Wright (2006) utilized a previously developed experimental paradigm based on set size expansion where the same/different concept (i.e., comparing a sample and a comparison image, and determining if they are the same or different images) could be learned through a process of systematically increasing the number of images with additional image sets. Hence, the paradigm has been referred to as *set size expansion*. Generally, subjects were trained to make same/different discriminations with a set of 8 images (travel pictures). Once reaching the performance criterion, subjects were tested with trials containing novel images. This test was followed by a training set size expansion, in which the number of items to discriminate were doubled (i.e., a new set of 8 images were added to the original set of 8 imagines). The successful acquisition of abstract concept learning using this paradigm was first observed in rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*) (Katz et al., 2002; Wright et al., 2003), before being examined with birds. Katz and Wright (2006) later found that their design facilitated learning of a relational rule by pigeons—including application of this rule to novel images (i.e., generalizing what they have learned, by processing that a new problem, like a familiar one, can be solved

using the same rule). Research into the cognitive reasoning abilities of the Corvidae, a family of bird species (e.g., crows, magpies, jays), has also yielded exceptional evidence for the presence of same/different abstract concept learning. For instance, Clark's nutcrackers (*Nucifraga columbiana)* and black-billed magpies (*Pica hudsonia*) were able to learn how to classify images correctly as same or as different with fewer training exemplars than any other non-human species using the same task (Magnotti et al., 2015; Magnotti et al., 2017; Wright et al., 2016; Wright et al., 2017). Although these various studies only test first-order same/different concept learning, evidence supporting an understanding of abstract concepts provides support that solving second-order problems by analogical reasoning might be within the abilities of avian species.

Corvids are ideal subjects for studying analogical reasoning as they possess many complex cognitive abilities such as causal reasoning, planning for the future, mirror recognition, and predicting the behaviour of others (Clary & Kelly, 2011; 2016a; 2016b; Clayton & Emery, 2007; Dally et al., 2005; Emery & Clayton, 2001, 2004). In particular, pinyon jays (*Gymnorhinus cyanocephalus*), are suitable for cognitive research as they have demonstrated many impressive cognitive abilities such as extensive spatial memories for storing and retrieving food (Balda & Kamil, 1989; Bednekoff & Balda, 1996) in addition to transitive interference and predicting the intentions of other individuals (Paz-y-Miño et al., 2004; Vernouillet et al., in prep-a). Also, when administered the same/different concept learning task that Katz and Wright (2006) gave to Clark's nutcrackers and black-billed magpies, pinyon jays were able to learn the concept at a fast rate (Vernouillet et al., in press). This provides evidence that pinyon jays possess an understanding of same/different concepts that could be applied to analogical problems.

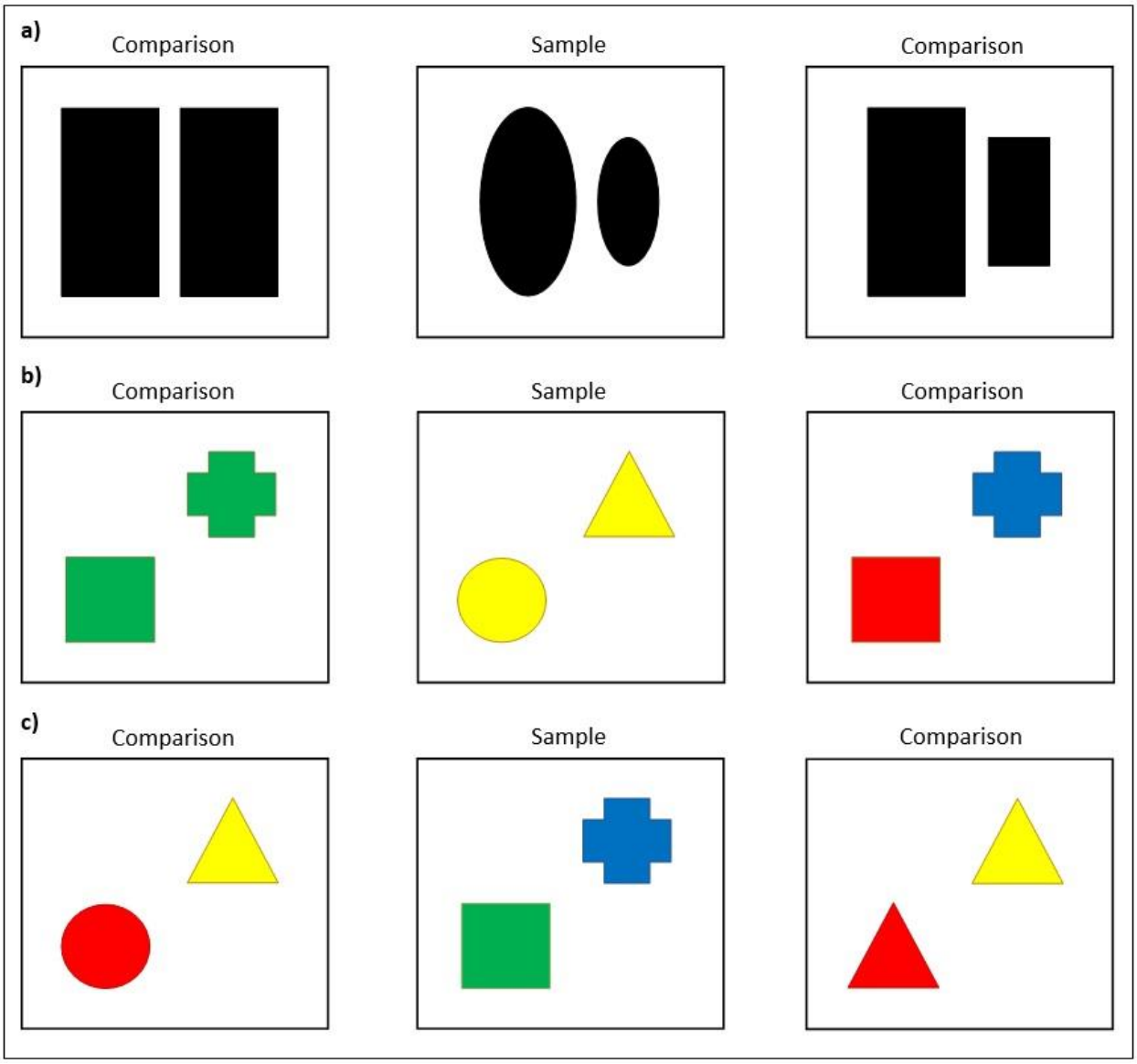
During previous abstract concept learning studies, animals were required to make a decision of which comparison stimulus matched a sample stimulus according to a shared relationship. However, it is often undeterminable, using previous task procedures, which properties of the comparison stimuli (e.g., colour, shape) were driving individual's decisions. To investigate whether subjects can solve an analogical reasoning task, the properties of the stimuli defining a relationship must be isolated. This would allow for examining which stimulus properties a subject relies on when making a choice, but also which properties cause birds to make incorrect choices. Likewise, the concern that perceptual matching strategies could be used during such a task has raised the question of whether or not subjects truly understand analogous

relationships. To address this concern, analogical problems would need to be presented so that only the properties being examined could be used to accurately make conceptual matches. This was recently done in a study testing analogical reasoning by crows.

To better understand which properties of the stimuli are being used by birds during an analogical reasoning task Smirnova et al. (2015) controlled which properties (i.e., dimensions) of the stimuli differed [similar to Vonk (2003)]. Two hooded crows (*Corvus corone*) were used as subjects and were tested individually in an experimental cage. During a trial, a tray was placed outside of the cage with three white square cards arranged horizontally in front of the bird. Each card had pairs of two stimuli illustrated on its face. The sample card was in the center and on either side were two comparison cards each placed overtop a plastic cup. The depicted relationship between the two stimuli on the sample card had to be matched to the relationship depicted by one of the two comparison cards (i.e., according to the dimension being assessed: colour, shape, or size). Following the presentation of the cards, the tray was slid into the cage via an opening and the crow was able to make a choice by knocking over a comparison card to reveal a food reward (when correct). The crows were able to spontaneously answer the relational match-to-sample problems (i.e., matching comparison and sample stimuli based on the relationship between their dimensions) after being trained solely on *identity* match-to-sample problems (i.e., matching a comparison stimulus that is identical to the sample stimulus). Importantly, during relational match-to-sample trials, the specific dimensions that comprised the comparison stimuli were never physically identical to the sample stimuli (e.g., if the sample contained yellow and red colours, the comparisons only contained blue and/or green colours), which was meant to place focus on conceptual relationships (e.g., having different colours) rather than perceptual information. Smirnova et al. (2015) claimed that the crows' concept learning was likely enabled by their prior identity match-to-sample training and that their high accuracy scores in the relational match-to-sample task demonstrated their ability to reason analogically. This success was due to the crows showing an understanding of the concept of "sameness" between sample and test stimuli in the relational task, as they would not have been able to rely on physical identity between sample and test stimuli since physical matches were not made possible in the experimental design. However, Vonk (2015) argued that Smirnova's experimental design tested perceptual similarity rather than relationships between stimuli during relational matching (i.e., the crows matched stimuli according to shared features or patterns associated with their

presentation; see Figure 2). Moreover, some of the inherent issues with Smirnova's experimental design, as pointed out by Vonk, were that: (1) it cannot be distinguished whether the birds perceived the two stimuli on a card as separate or if due to being enclosed by the borders of the card, the birds' perceived the stimuli as a single entity, (2) the stimuli presented side-by-side on the cards were not properly counterbalanced for left-right orientation, and (3) the birds could have been identifying patterns between cards whose stimuli share certain physical characteristics ('perceptual processing account') and then choosing a comparison card that shares the same pattern (e.g., if two colours are present among the stimuli on the sample card, then a subject might choose a comparison card because it also has two colours present, not because the subject understands the same-colour relationship, see Figure 1). Together, these concerns raised the idea that the birds might be viewing each card as a single item that should be compared to one another, which is more indicative of first-order reasoning and/or they may not be attending to relationships between the stimuli on the cards, which is required for second-order reasoning (i.e., analogical reasoning). For future studies of analogical reasoning, Vonk suggested cautious consideration be given to stimulus configurations, in addition to controlling the quantity of shared features depicted between the stimuli. This would allow for a more accurate assessment of analogical reasoning and of the specific features being considered when attempting to solve a relational problem.

Figure 2. **Examples of relational match-to-sample trials for three dimensions a) size, b) colour, and c) shape, adapted from Smirnova et al., 2015.** A subject chooses between two comparison stimuli based on which one is a relational match to the sample along the target dimension. For instance, a) depicts a size relation in which the sample card displays stimuli of different sizes and the correct comparison is the right card which shares the same relation between its stimuli.



My research objective was to evaluate whether pinyon jays are capable of analogical reasoning. To show this ability, pinyon jays would have to be taught to recognize shared properties between a pair of stimuli and then use relational equivalence to correctly form an analogous pair. I adapted a newly developed experimental procedure for the study of analogical reasoning by humans (Leonard et al., in prep) for testing pinyon jays. Rather than determining solely whether subjects could exhibit spontaneous analogical reasoning, as was conducted during other studies, this procedure trained subjects using analogy problems (providing feedback on these problems until concept acquisition), and then presented tests using novel analogy problems. To succeed, the birds would have to recognize the relational rule that is generalizable to all

analogical problems regardless of which stimulus properties are presented. By using this procedure, I would be able to examine the extent to which pinyon jays can use this rule to solve analogical reasoning problems that present relationships with progressively greater variation. The procedure was as follows: pinyon jays participated in an operant experiment during which they were presented with virtual stimuli displayed on LCD monitors (an infrared touch frame was placed around the monitor's screen, such that pecking at the screen would cause a beam break from which the x and y coordinated could be registered). Their task was to complete a visual second-order problem that required the use of analogical reasoning and followed the same structure as the example mentioned earlier – "a kitten is to a cat as a puppy is to a _____". The relationships were composed of *same* or *different* combinations of stimulus dimensions (i.e., shape, colour, and pattern). To accurately complete the analogical relationship that exists between the stimuli (*item A* is to *item A'* as *item B* is to *item B'*), the dimensions displayed among alternative choice stimuli (*item B'* and *item C*) would need to be attended or inhibited, based on which dimensions accurately illustrate the shared relationship. I predicted that pinyon jays would be capable of solving the analogical problems presented during the proposed study. Successful concept learning (i.e., understanding the relational rule) would show for the first time that an avian species was able to solve an analogical task that does not require prior language training and is specifically created to eliminate the possibility for reasoning using perceptual similarity.

My study's design addressed several concerns from previous studies by using the newly adapted analogical reasoning task mentioned above. In particular: 1) each stimulus was presented individually, 2) the position of the stimuli was counterbalanced across trials, 3) perceptual qualities of the stimuli were controlled, and 4) tests were designed to be conducted with a single as well as multiple stimulus dimensions. Additionally, my findings were proposed to provide evidence for whether this procedure is suitable for making direct comparisons of multiple species' performance and would assist with validating a new analogical reasoning paradigm that can be used for the study of complex comparative cognition.

CHAPTER III

Methods

Subjects

The subjects were adult pinyon jays (n = 9, 5 females), wild-caught from populations around Flagstaff, Arizona, and housed in captivity between seven and twenty-three years. They have participated in unrelated experiments involving object and environment exploration (Vernouillet & Kelly, 2020), food-caching (Vernouillet et al., in prep-a; Vernouillet et al., 2020; Vernouillet et al., in prep-b), and same/different abstract concept learning (Vernouillet et al., in press).

Pinyon jays were individually housed in large cages (51 cm wide x 51 cm deep x 72 cm high) within the Psychology Animal Holding facilities at the University of Manitoba. Various wood perches were affixed inside their cages. The colony room is maintained at 22°C with a 12:12 day-night cycle, which serves to replicate the late-summer and early-fall conditions of their natural habitat. Light onset in the colony room is at 0700 (Central time). The birds were provided with water and grit *ad libitum*. During the experiment, they were kept at approximately 90% of their free-feeding weight. This procedure allowed them to maintain a healthy weight while encouraging their participation during the experiment using food rewards. Their usual diet was composed of a mixture of parrot pellets, stripped sunflower seeds and chicken starter, as well as, peanuts, pine nuts, powder of oyster shells, and a Vitaprime® supplement. Freeze-dried mealworms were used as food rewards during the experiment.

Approval for all experimental procedures was granted by the University of Manitoba's Animal Care Committee (protocol #F2014-037) and were in accordance with the Canadian Council on Animal Care ethical standards.

Apparatus

The birds were tested in identical custom-built wooden operant chambers (61 cm wide x 31 cm deep x 56 cm high). A fan mounted on the back wall provided ventilation and diffuse noise. A house light (24V, 0.04A Eiko bulb), positioned on the center of the chamber's ceiling, was illuminated when the chamber was in use. The chamber was equipped with a LCD monitor

(ThinkVision®; resolution: 800 x 600 pixels; dimensions: approximately 33 cm wide x 26 high cm; diagonal length: approximately 42 cm). The monitor, viewable through a cut-out in a clear acrylic sheet (33 cm wide x 26 high cm) at the front wall of the chamber, was connected to a Dell® computer (i5-3570 CPU with 60 Hz refresh rate) that ran on Windows XP. The birds viewed the screen at a visual angle of approximately 172° (horizontal) x 142° (vertical) from a perch positioned 14.5 cm from the screen and 15.9 cm from the floor. This experiment was viewed on the monitor screens. The monitors were a minimum of eight years old. Colour accuracy of monitor displays change variably over time; thus the colour and contrast (i.e., changes in brightness) displayed by the chamber monitors were calibrated for consistent colour exposure. An X-Rite® ColorMunki Display calibrator was used to generate a colour profile, which was set to the Windows Default colour spectrum and adjusted to the ambient light conditions inside the enclosed chambers. This profile was assigned to all monitors.

This experiment was administered using a custom software designed with Python 3.7. The software also controlled a rotating wheel (read below) and chamber lights via a USB digital I/O panel (Phidget). The chamber contained an Elo touch frame mounted to the monitor. The touch frame was lined with LEDs. These LEDs emitted infrared beams that passed across the monitors. When a peck was made to the monitor, it would break the infrared beam. The x and y coordinates where a beam was broken was used to determine the location of the peck. For each correct response, birds received a food reward from a rotating wheel containing 100 slots, positioned under a wooden shelf with a small cut-out (approximately 7.5 cm) below the monitor. The birds were only able to access a single slot (one reward) from the wheel at a time.

Stimuli

Stimuli used in the current study's Analogical Reasoning Task were created using Sumo Paint and GIMP image editing software. The stimuli varied along three dimensions (i.e., shape, colour, and pattern), each comprised of 17 variations (i.e., exemplars). The exemplars within the shape dimension included: circle, square, triangle, star, arc, pentagon, plus, heart, dress, boot, leaf, cloud, cactus, moth, puzzle piece, gem, and blob, and the surface area of each shape was between 7600 – 6200 pixels. The colour dimension included: orange, yellow, green, cyan, blue, violet, pink, grey, red, light red, lime green, light green, teal, light purple, light pink, fuchsia, and white, which are composed of wavelengths between 400 – 700 nm. Colours were chosen to

provide the greatest possible colour range along the continuum while accounting for their discriminability on the monitor's screen [(pinyon jays have violet sensitive vision (Håstad et al., 2005; Ödeen & Håstad, 2003) and the wavelengths used for the stimuli in this experiment were chosen to be within their visible light spectrum (Håstad & Ödeen, 2014)]. The pattern dimension included: vertical stripes, horizontal stripes, diamonds, dots, swirls, open-circles, X's, Z's, whirl, starburst, waves, flowers, slant, paint, glass, zebra, and plain (solid-fill).

These dimensions were chosen based on their suitability for the analogical problems of this task and for their use in prior avian visual discrimination tasks. Shape, colour, and pattern exemplars (e.g., English letters, triangle, square, circle, vertical lines, horizontal lines, flowers, blue, green, cyan, red, purple, orange, and yellow) have been used in previous tasks, such as visual discrimination tasks presenting geometric shapes, objects or textured arrays to pigeons (Cook et al., 1997). Additionally, hundreds of unique travel images consisting of multidimensional features consisting of these dimensions are discriminable by corvids (e.g., pinyon jays, Clark's nutcrackers, and magpies; Wright et al., 2017; Vernouillet et al., in press). Item exemplars of colour (e.g., blue, red, green, yellow, or purple) and shape (e.g., circle, square, plus, and triangle) have been used for investigations of analogical reasoning by parrots and crows (Obozoba et al., 2015; Smirnova et al., 2015) and during two-item same/different discrimination by pigeons in which stimuli were presented on computer screens (Blaisdell & Cook, 2005). Vonk (2015) suggested that avian relational discrimination tasks could incorporate a condition that uses shapes with multi-colored patterns or larger shapes containing patterns of smaller shapes, the latter of which was used here with several of the pattern exemplars.

Prior to beginning the Analogical Reasoning Task, the birds completed initial shaping procedures (see below) which incorporated two types of stimuli. During Stage 1 of the shaping procedure a *blank* stimulus was used (i.e., a plain white blob), and during Stage 2 the stimuli appeared as asymmetrical shapes, randomly generated during trials using Python 3.7, and overlaying multi-coloured stock images that contained no identifiable pattern. Thus, they had no characteristics (i.e., shape, colour, or pattern) that matched those of the stimuli in the Analogical Reasoning Task.

General Procedures

The birds received one session per day (maximum: 4 hours), five days a week between 0900-1600 hours (each individual's daily start time was consistent). To begin a session, a bird was transported from the colony room to the experimental room via an opaque transportation box, and at the conclusion of the session, was returned in a similar fashion.

Shaping Procedure

The current experiment required the pinyon jays to complete an initial shaping procedure, which was divided into two stages. Stage 1 was designed to train birds to peck at a visual stimulus presented on the monitor. Stage 1 was composed of shaping trials, each of which presented a single stimulus (i.e., the blank stimulus). Across shaping trials, the blank stimulus appeared in pseudo-randomly chosen positions, such that it would appear in any location to be used during analogical reasoning trials. This procedure was used to familiarize the birds with pecking all areas of the screen. During any particular shaping trial, once the stimulus was pecked a single time, the shaping trial ended and the bird received a food reward. Following an intertrial interval (ITI) of 15 s (as used in previous same/different experiments with corvids, Magnotti et al., 2017; Wright et al., 2016; Wright et al., 2017) another trial began and the stimulus appeared in a different position. A single shaping procedure session consisted of 40 trials. Criterion to complete Stage 1 of the shaping procedure was the completion of an entire session in a single day.

Stage 2 of the shaping procedure served to familiarize each bird with the sequence in which stimuli would be presented during an experimental trial, as the order in which the stimuli had to be pecked would be important. Over the course of the shaping procedure birds learned to peck the stimuli in the appropriate order. During Stage 2 of the shaping procedure, and all following phases of the experiment, a white outline framed the area of the screen that would contain the first three stimuli to appear on the screen during the analogical reasoning trials. This framed area was displayed approximately 5 cm below the upper left-hand corner of the computer screen—this 5 cm height adjustment was made in accordance with the visual angle in which the birds viewed the screen and thus allowed stimuli to be in reach. Additionally, the frame was designed to assist the bird in distinguishing between stimuli based on the order in which they appear, and whether they appeared inside or outside of the white frame. A single session

included 20 shaping trials. Initial shaping trials began with the appearance of a random stimulus—herein referred to as a *primary stimulus*—in either the left or right position of the upper rectangular area of the white outline. When the bird pecked this first primary stimulus it received a reward, thus marking the end of a trial. The fixed ratio of one peck (FR1) for the first primary stimulus progressively increased across trials within a session. By the end of the season, five successive pecks to the primary stimulus was required before receiving a reward. After completing two consecutive sessions, the following session included the addition of a second primary stimulus in the sequence. From this point forward, all trials began with an FR5 requirement for the first primary stimulus. Following this requirement, the second primary stimulus was presented and was pecked once, ending the trial with a reward. The fixed ratio (FR) for the second primary stimulus increased across trials in the same manner as the first primary stimulus. Upon completion of two consecutive sessions, the bird advanced. Subsequent sessions included the addition of a third *bridge stimulus* to the sequence. Throughout these sessions the bird was trained to peck the stimuli in each position within the frame (left or right) to prevent the development of a side-bias. The final addition in the sequence was as a single *choice stimulus* outside the frame—its location on the screen was pseudorandomized, as in Stage 1. By the completion of Stage 2 shaping, the bird was required to complete an FR5 for each stimulus. A reward always followed the completion of the FR requirement for the choice stimulus. Once the bird achieved the criterion for completing Stage 2 (i.e., completing two consecutive sessions following the addition of each stimulus to the sequence) it progressed to the Analogical Reasoning Task.

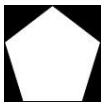
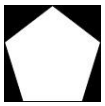
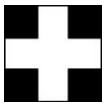
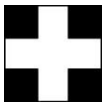
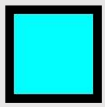
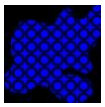
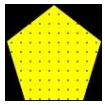
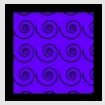
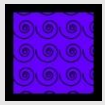
Analogical Reasoning Task

This task was designed to evaluate analogical reasoning abilities. The task was composed of trials that progressed in the same manner as Stage 2 shaping. First, the pair of primary stimuli were used to determine a within-pair relationship. This within-pair relationship was defined according to one or more stimulus dimensions (i.e., shape, colour, or pattern). Second, a corresponding pair of stimuli had to be matched to create an identical within-pair relationship. This corresponding pair was established by matching the bridge stimulus with the correct choice stimulus (along the same dimension(s)). Thus, the matching within-pair relationship represented either the concept *same* or *different* across each specific dimension. Finally, an analogy (i.e., the

relation between relations) was created. For example, if the primary stimuli are two circles and the bridge-choice stimuli are two triangles, then these pairs both demonstrate a within-pair relationship in which their stimuli share the same shape. Comparing these within-pair relationships illustrates a matching "same shape" between-pair relationship (i.e., the primary pair compared to the bridge-choice pair).

Training and Test Procedures. The Analogical Reasoning Task was comprised of training phases followed by test phases. For instance, during the first training phase of the Analogical Reasoning Task (i.e., *Dimension 1 Training*), each bird was trained to relate stimuli that varied along the first dimension it was assigned. Immediately following this training, the *Dimension 1 Test* was administered, which was designed to test analogical reasoning abilities using novel first dimension exemplars. As a bird advanced through testing and training phases, the within-pair relationships that defined the stimuli varied progressively along two and three dimensions (see Table 1).

Table 1. An example of all training and test phases. Each phase below includes an example of stimuli that could be presented during a trial, starting with the shape dimension and concluding with the pattern dimension. The relationship between the primary stimuli must match the relationship between the bridge and choice stimuli, along the dimension(s) presented. The relationships presented during any phase could be either *same* or *different*. 'Dimensions Represented' indicates the number of dimensions presented during a particular phase, and 'Dimensions per Trial' indicates how many of those dimensions vary among the stimuli during each individual trial. For instance, during the *Dimension 3 Test* and *Dimension 3 Training*, all three dimensions are included throughout trials, but on any individual trial only two dimensions vary (these cases are indicated by the numbers in bold font).

Phase	Dimensions Represented	Dimensions per Trial	First Primary Stimulus	Second Primary Stimulus	Bridge Stimulus	Correct Choice
Dimension 1 Training	1	1				
Dimension 1 Test	1	1				
Dimension 2 Test	1	1				
Dimension 2 Training	1	1				
Dimensions 1&2 Combination Test	2	2				
Dimensions 1&2 Combination Training	2	2				
Dimension 3 Test	**3**	**2**				
Dimension 3 Training	**3**	**2**				
All Combination Test	3	3				
All Combination Training	3	3				

I designed this task so that birds would progress through a total of ten phases, alternating between five training phases and five test phases. Each test phase presented a new transfer test to introduce novel exemplars, a novel dimension, or a novel combination of dimensions. The complexity of the task (i.e., the addition of more dimensions) was designed to increase with successive training and test phases. Training and test phases were composed of sessions containing a minimum of 60 trials—correction procedures (see Structure of a Trial) during training sessions may have required additional trials. Trials were separated by a 15 s ITI. The birds were required to complete a training criterion before proceeding to the following test phase.

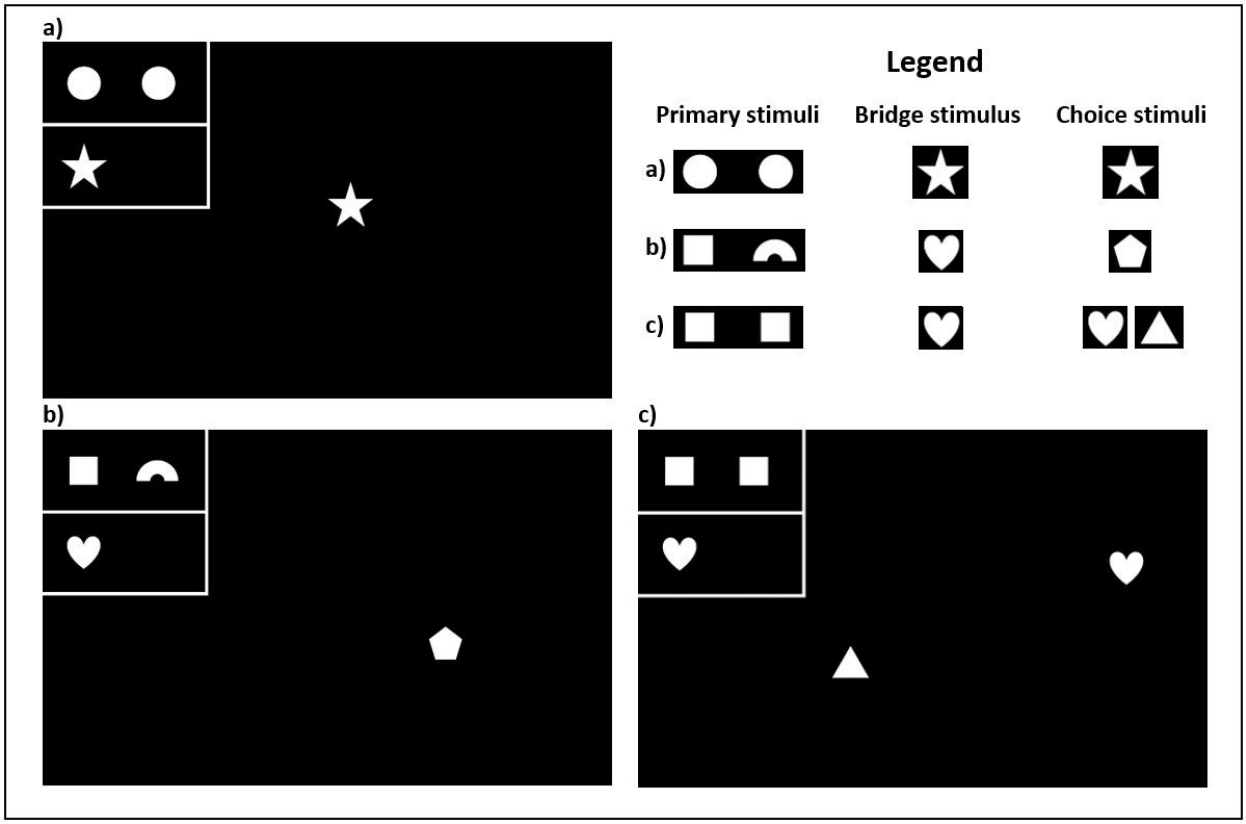
The training criterion required the birds to reach 80% correct choices or better for two consecutive training sessions with correction procedures in place. The birds were then required to repeat this criterion without correction procedures to show that they were learning without each trial ending in reinforcement. Thus, following a non-reinforced peck to an incorrect stimulus, a bird would begin a new trial after the ITI, rather than repeating the incorrect trial until reinforced for pecking the correct choice.

Test phases were composed of five sessions. Each test session consisted of 48 training trials (composed of trials identical to those from the prior training phase) and 12 test trials (any trial testing novel exemplars, a novel dimension, or a novel combination of dimensions), with one test trial following every four to eight training trials. During test trials, a bird received a reward regardless of a correct or incorrect choice—to avoid extinguishing response behaviour due to novel stimuli being associated with no reward. Performance during test trials was designed to be compared to training trial performance.

Arrangement of Training/Test Phases. Training on the first dimension (*Dimension 1 Training*) began with a session that was divided into two stages—Stage 1 included one choice stimulus in each trial (i.e., see Figure 3a for a display depicting a *same* relationship and Figure 3b for a display depicting a *different* relationship), whereas Stage 2 included two choice stimuli in each trial (see Figure 3c). Stage 1 comprised the first 30 trials of the first session. After becoming familiar with pecking at a choice stimulus (i.e., completing the first 30 trials, each ending with a peck to the choice stimulus followed by a reward), the session proceeded to 30 trials containing two choice stimuli to choose between (Stage 2). All following sessions of *Dimension 1 Training* contained only trials with two choices. Table 1 provides an example of all training and test phases in order. The *Dimension 1 Test* presented a novel set of stimulus exemplars from the first dimension a bird was trained on. The *Dimension 2 Test* presented stimuli that varied along the second dimension the bird was assigned, followed by training on this dimension during *Dimension 2 Training*. Most test phases were designed to be given before a subsequent training phase (e.g., the *Dimension 2 Test* followed by *Dimension 2 Training*). This is because test phases each represented a *transfer test* which was meant to determine if a bird was able to generalize (or transfer) a relational rule, acquired during training (e.g., *Dimension 1 Training*), in order to solve a problem with either a novel exemplar, dimension, or combination

of dimensions. In this case, training on these novel problems came after testing, to provide the bird with further experience solving these more challenging problems. Completion of this training advanced the bird to the next test phase.

Figure 3. **Example of three displays for trials comprised of one dimension, the example used here is for shape.** Each trial represents *Dimension 1 Training* in which: a) presents a *same* relationship (Stage 1), b) presents a *different* relationship (Stage1), and c) contains two choice stimuli (Stage 2). For c) the *primary stimuli* (i.e., squares presented within the top white-outlined frame) are the same shape. The *bridge stimulus* (i.e., heart) is presented directly below the *primary stimuli* in its own frame. In this example, the correct *choice stimulus* among the two options in the open area to the right side of the screen, is the heart, because the relationship between this *bridge stimulus* and the *choice stimulus* matches the relationship presented by the *primary stimuli*. A legend defining the stimulus classifications is shown in the upper right.



The complexity of the task was designed to increase during the *Dimensions 1&2 Combination Test* as test trials combined the first two dimensions learned. From this point forward, the choice pool contained four choice stimuli. These four choices represented each

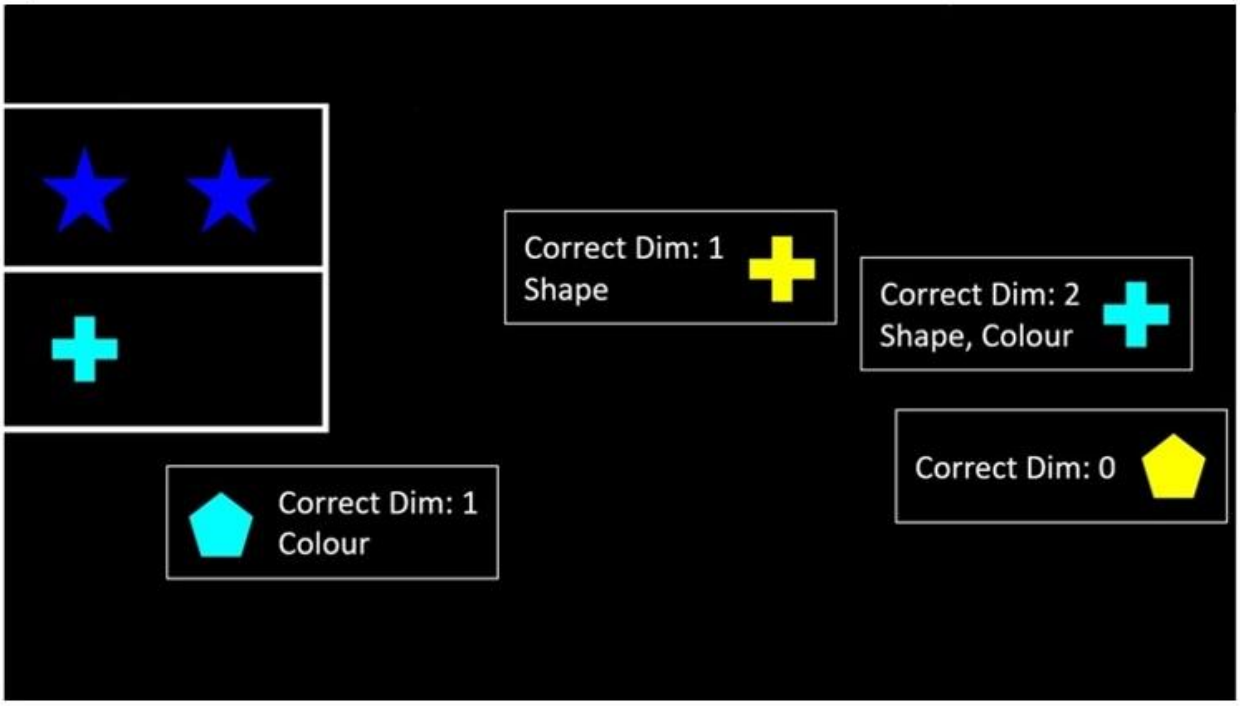
combination of correct or incorrect dimensions that a choice stimulus can vary along. For instance, one choice stimulus was correct along both dimensions, two choice stimuli were correct along one of either dimensions, and one choice stimulus was incorrect along both dimensions. There was also consistency among the varying dimension exemplars of the choice stimuli. For instance, if during a trial, the bridge stimulus was a cyan plus, then the choice stimuli were composed of shapes that are either a plus or one incorrect exemplar of shape, and colour options that were cyan or one incorrect exemplar of colour (see Figure 4a). This consistency in the choice stimuli was designed to prevent the bird from attending to a certain dimensional exemplar, such as a colour that appears less frequently than others, which would enable individuals to use a process of elimination to solve the task—especially if one stimulus contained a certain exemplar that was not shared by any other choice stimulus. Rather, this balanced design was meant to promote learning of the relational rule (i.e., understanding the relation between relations) to solve the task.

Following the *Dimensions 1&2 Combination Test*, the *Dimensions 1&2 Combination Training* trained birds on the combination of the first two dimensions learned. The *Dimension 3 Test* included test trials with all three dimensions, but only two of the three dimensions varied during any given trial, whereas the third was held constant. *Dimension 3 Training* presented training trials identical to the previous phase's test trials. The *All Combination Test* presented test trials in which all three dimensions were combined. The choice stimuli were balanced to ensure that there was equal representation of the correct and incorrect exemplars for each dimension. For instance, one of the four presented choice stimuli was correct along all three dimensions, one was correct along two dimensions, one was correct along the third dimension, and one was incorrect along all dimensions. To illustrate this, Figure 4b presents a trial with a relationship defined as *different shape, different colour*, and *same pattern*. The bridge stimulus is shown as a yellow heart with a diamond pattern. The correct choice stimulus (correct along three dimensions) is a green star with a diamond pattern, a second choice stimulus (correct along two dimensions: different colour, same pattern) is a green heart with a diamond pattern, a third choice stimulus (correct along one dimension: different shape) is a yellow star with a swirl pattern, and a fourth choice stimulus is a yellow heart with a swirl pattern (incorrect along all dimensions). Each dimensional exemplar (green/yellow colour, star/heart shape, and diamond/swirl pattern) was represented an equal number of times (twice) among the choice

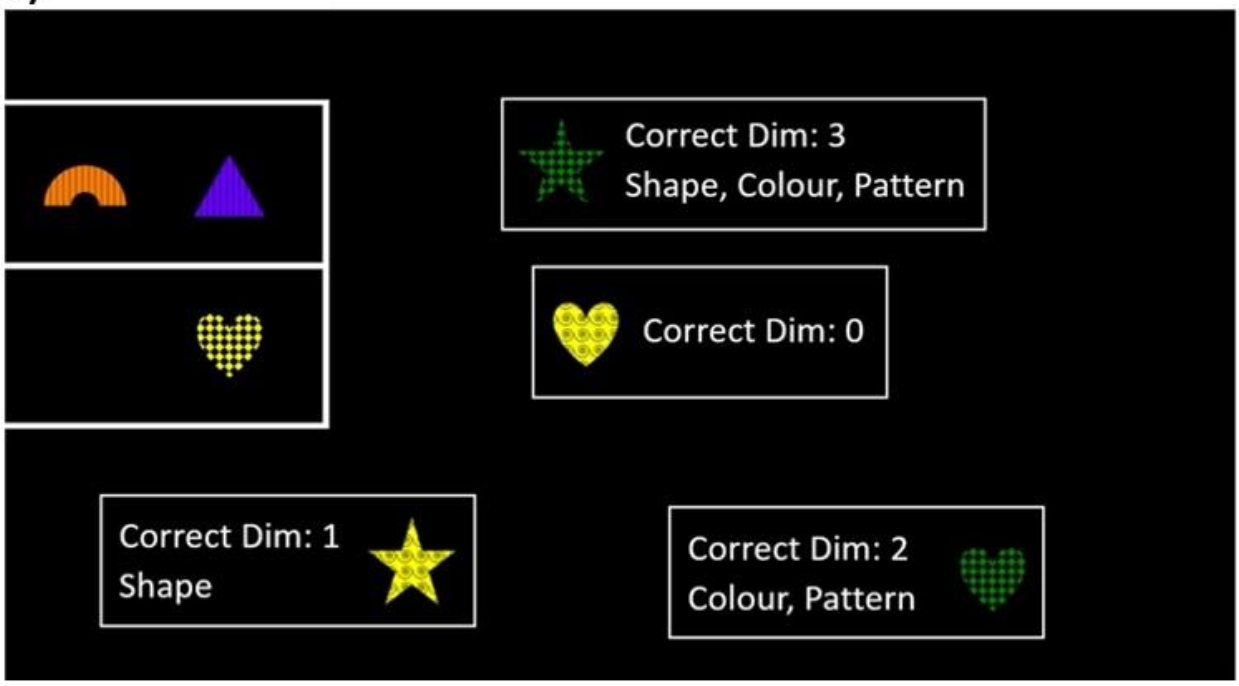
stimuli. During *All Combination Training*, birds were trained on all three dimensions combined. Participation in the experiment was designed to end when the training criterion was reached.

Figure 4. Example of a display for a trial comprised of a) two dimensions (i.e., shape and colour) and b) three dimensions (i.e., shape, colour, and pattern). For display a), the *primary stimuli* (i.e., blue stars) have a *same* shape and *same* colour relationship. The correct *choice stimulus* is the cyan plus because it has the same shape and the same colour as the *bridge stimulus* (i.e., cyan plus). The remaining *choice stimuli* are correct along fewer dimensions than necessary to complete the analogical relationship. The correct dimensional exemplars (i.e., cyan/yellow colours and plus/pentagon shapes) represented by the *choice stimuli* are counterbalanced so that each exemplar is equally represented among the choices. For display b), the *primary stimuli* (i.e., orange arc and violet triangle, both with a vertical pattern) have a *different* shape, *different* colour and *same* pattern relationship. The correct *choice stimulus* is the green star with the diamond pattern because it has a different shape, is a different colour, and has the same pattern as the bridge stimulus (i.e., yellow heart with diamond pattern). The dimensional exemplars (green/yellow colour, star/heart shape and diamond/swirl pattern) represented by the *choice stimuli* are also balanced. The boxes around the *choice stimuli* and the descriptions are for illustrative purposes only.

a)


Correct Dim: 1
Shape
Correct Dim: 2
Shape, Colour
Correct Dim: 0
Correct Dim: 1
Colour

b)


Correct Dim: 3
Shape, Colour, Pattern
Correct Dim: 0
Correct Dim: 1
Shape
Correct Dim: 2
Colour, Pattern

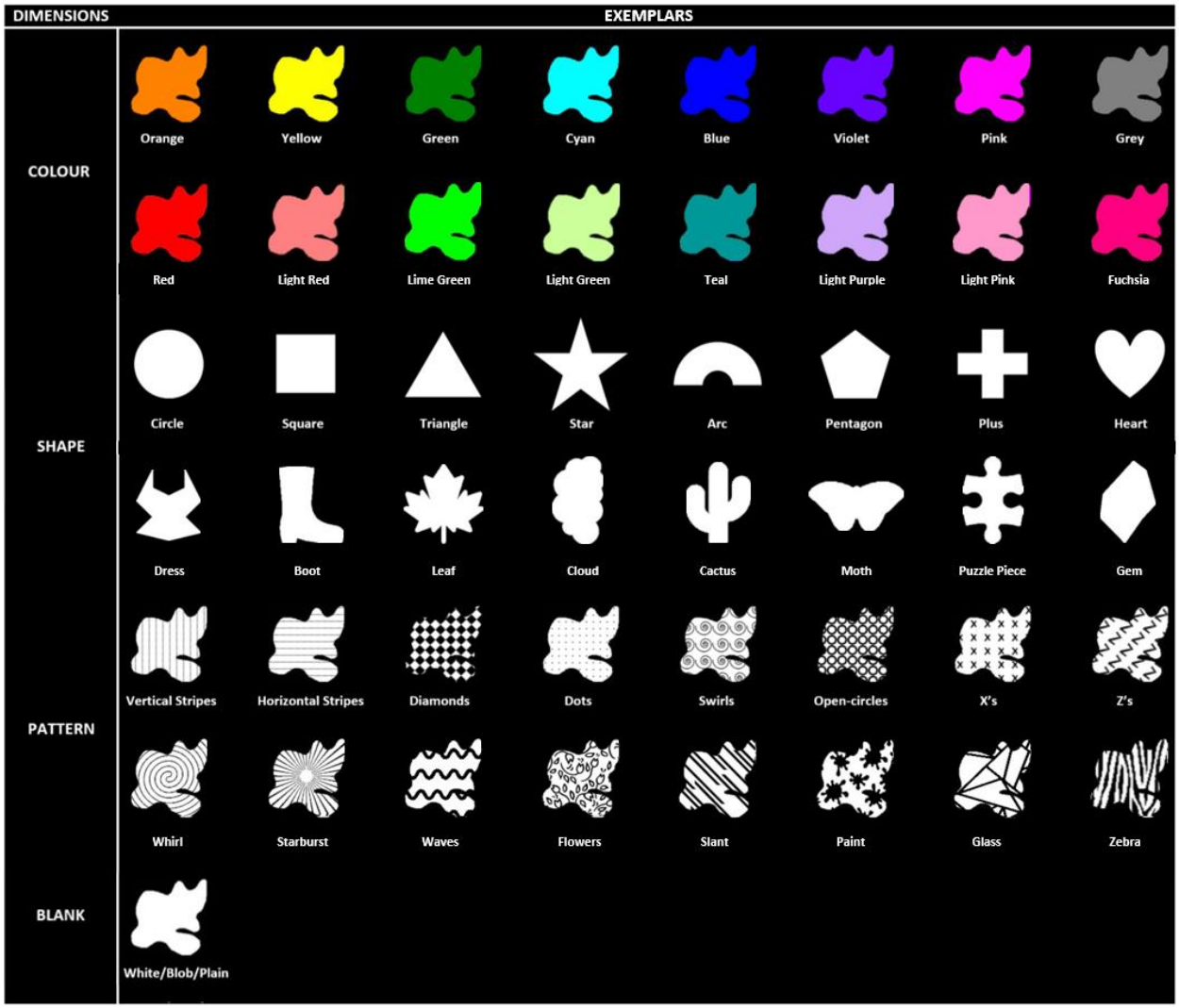
Structure of a Trial. During a *Dimension 1 Training* trial, a bird first viewed the pair of primary stimuli, which were defined by a relationship along a single dimension (e.g., both were red). The primary stimuli were presented inside the top frame of the white outline. The primary stimuli were successively presented side by side, counterbalanced for location order. Similar to the Stage 2 shaping procedure, the birds were required to make five pecks to the first primary stimulus before the second primary stimulus was displayed. Likewise, the second primary stimulus was pecked five times before the bridge stimulus was presented (with the location counterbalanced in either the left or right position directly below the primary stimuli, inside the bottom frame of the white outline). Once the bridge stimulus was pecked five times, a pool of choice stimuli (ranging from one to two) appeared in pseudorandomized positions on the screen. These choice stimuli provided different matching options by presenting different exemplars of the relevant dimension(s). As a pairing, the correct choice stimulus and the bridge stimulus formed a relationship that matched the relationship represented by the primary stimuli. For example, a pair of primary stimuli having a "same shape" relationship was presented, followed by a bridge stimulus and the option of picking between two choice stimuli. One choice stimulus, i.e., the *correct choice stimulus*, matched the relationship between the primary stimuli (i.e., if the two primary stimuli had the same shape, this choice stimulus would also have the same shape as the bridge stimulus, thus matching the presented "same shape" relationship), whereas the second stimulus, i.e., the *incorrect choice stimulus*, did not match (i.e., the second choice stimulus differed in shape to the bridge stimulus, thus not matching the presented "same shape" relationship). The birds were required to peck the *single correct* choice stimulus that, when paired with the bridge stimulus, completed the analogy depicted by the primary stimuli. If a bird pecked on the correct choice stimulus, it received one food reward (i.e., mealworm). Throughout *Dimension 1 Training* and later phases, if the bird pecked on a choice stimulus that was correct along fewer dimensions than required to complete the relationship, or a choice stimulus that was incorrect along all dimensions, it would not receive a reward. Instead, a correction procedure was instituted—the bird would receive a 15 s timeout, followed by an additional 10 s ITI, and then the same trial would repeat, with the same choice stimuli presented in the same positions (i.e., correction trial). Chamber lights were extinguished and the screen was darkened during the timeout and subsequent ITI. The lights were once again illuminated when the correction trial

started. Correction procedures were only instituted during training phases, or during training trials within a test phase. Results from correction procedures were not included in analyses.

Dimension Exemplars. The stimuli presented across trials of the Analogical Reasoning Task were generated using random combinations of the 17 exemplars of each dimension (see Table 2). The stimuli varied along the number of dimensions appropriate for each phase of training or test trials. For instance, stimuli used during *Dimension 1 Training* varied along one dimension. The bridge and choice stimuli varied along exemplars that were not present in the primary stimuli (see Figure 3 for an example). During the *Dimensions 1&2 Combination Test* and subsequent phases of the experiment, a stimulus was comprised of a combination of two or three exemplars, each from a separate dimension (e.g., a pink triangle with vertical stripes).

When a bird began *Dimension 1*- and *Dimension 2 Training*, it was presented with stimuli that varied along the first or second dimension learned (i.e., shape, colour, or pattern). During these phases, the stimuli appeared as they do in the rows of Table 2, where only exemplars of the relevant dimension were visible (e.g., only colour varies in the Colour rows). There were *blank* exemplars that were used in place of the two remaining dimensions (*white* for the colour dimension, an asymmetrical *blob* for the shape dimension (7600 pixels), and *plain* for the pattern dimension). For instance, when shape was the first dimension learned during *Dimension 1 Training*, the blank exemplars for colour and pattern were used in order to prevent exposure to exemplars of these two dimensions before each was learned during *Dimension 2*- and *Dimension 3 Training*, respectively. This was done to allow the colour and pattern exemplars to be used as novel dimensions during test phases. During *Dimension 3 Training*, when a bird was first presented with a third dimension, exemplars of this dimension appeared in conjunction with one of the other two dimensions already learned. Thus, the blank exemplar was used for the dimension held constant during any individual trial (see *Dimension 3 Test/Training* in Table 1 for examples).

Table 2. An assortment of stimuli arranged by the possible exemplars for each dimension. Each of three dimensions (colour, shape, and pattern) had 17 exemplars. The 'blank dimension' presents a single stimulus that represents the *blank* exemplars of each of the three dimensions (colour: white, shape: blob, pattern: plain). These blank exemplars were used such that a bird experienced exemplars of the first dimension learned, such as shape, without exposure to exemplars of colour or pattern (i.e., the shape rows present stimuli that vary in shape, but are consistently white and plain).



Following *Dimension 2 Training*, test and training phases contained stimuli comprised of two or three relevant dimensions. The primary stimuli could have been defined by various combinations of same/different relationships that existed among those dimensions. Table 3 provides an illustration of the various combinations of dimensions that can be considered *same* or *different* between two stimuli. There are eight possible same/different relationships when a trial includes three relevant dimensions. When a trial includes two relevant dimensions, with the third held constant (same across all stimuli), there are four possible same/different relationships—as illustrated by the relationships above the bold line in Table 3, where the relationships between colour and shape vary, but pattern is held constant. The order in which the birds learned the three dimensions was pseudo-counterbalanced. Each dimension was assigned as either the first, second, or third dimension three times across all nine birds (see Figure 5). Furthermore, 16 exemplars for each dimension, not including the 17th or blank exemplar, were divided into two blocks. One block consisted of the eight exemplars that were presented throughout most of training and testing, and the second block consisted of the remaining eight exemplars that were only used to generate the novel first dimension stimuli presented during *Dimension 1 Test* trials.

Table 3. A comparative illustration of the three combinations of dimensions that were same or different between a pair of stimuli. A pair of primary stimuli contain two items, the first and second stimulus. These stimuli are composed of dimensions (pattern, colour, and shape). The sameness or difference of dimensions between both stimuli represents a relationship. The stimuli can be defined by a relationship comprising three dimensions that are entirely *same* or *different*, or are a combination of both. The relationships above the bold line represent when one of three dimensions (e.g., pattern) is held constant.

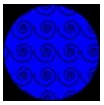
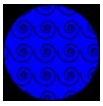
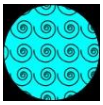
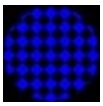
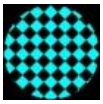
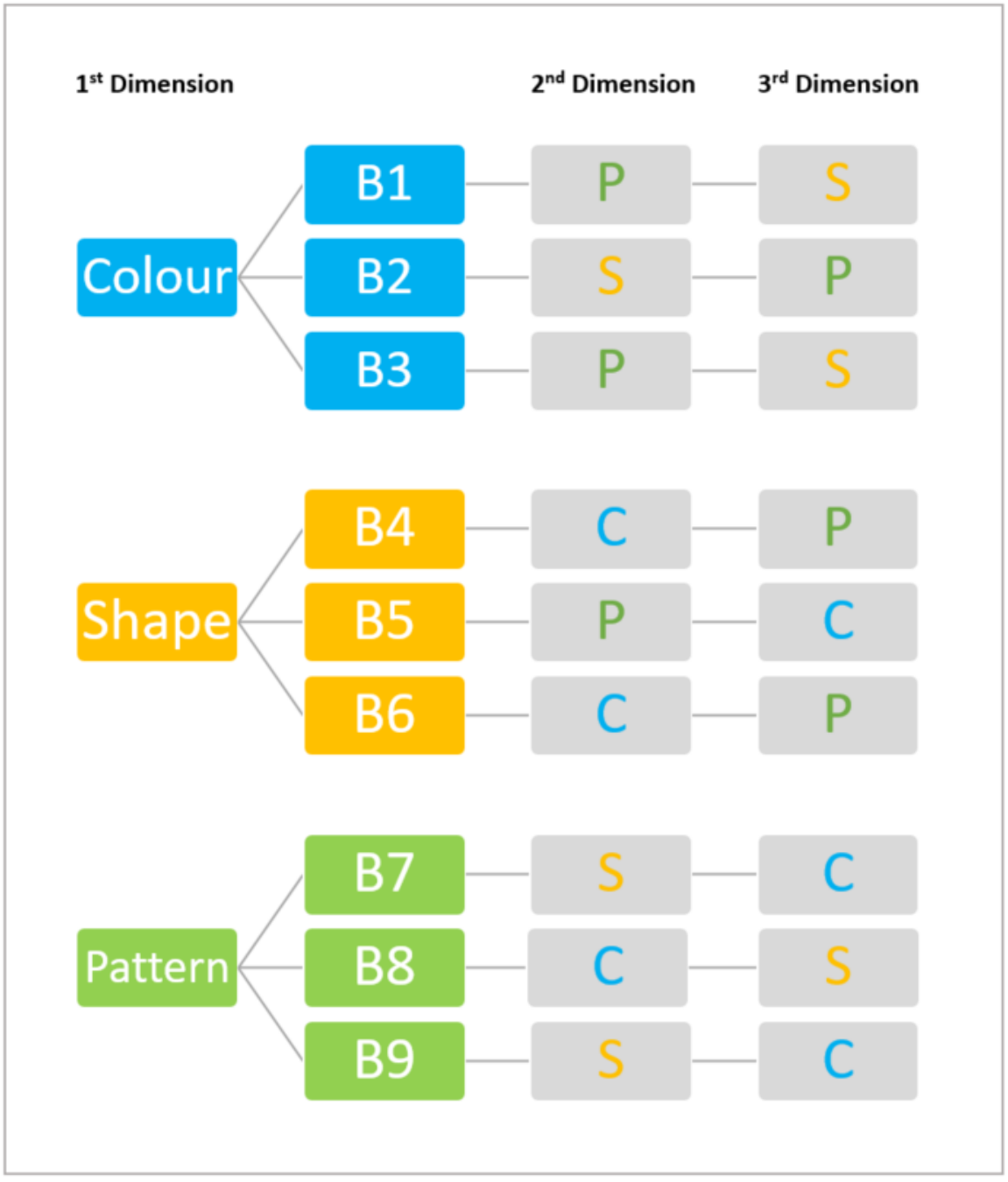
First Stimulus	***Second Stimulus***	***Same*** *Pattern*	*Colour*	*Shape*
		✓	✓	✓
		✓	✓	✗
		✓	✗	✓
		✓	✗	✗
		✗	✓	✓
		✗	✓	✗
		✗	✗	✓
		✗	✗	✗

Figure 5. Example of the pseudo-counterbalancing for the order by which the birds would learn the three dimensions (i.e., shape, colour, and pattern). The nine birds are denoted as "B" followed by a number. The dimensions are represented as S (shape), C (colour) and P (pattern). Each dimension is represented three times as the first, second, or third dimension learned.



CHAPTER IV

Statistical Analyses

Dependent Variables

Percent Correct

I designed the Analogical Reasoning Task so that the percentage of correct trials could be used to compare the birds' performance during test phases. Percent correct was designed to be evaluated against chance (i.e., 25% or 50% when there were four or two choice stimuli in the pool, respectively) and then compared between test trials and training trials for each test phase. When analyzing test performance, three potential outcomes may arise: 1) if the percentage of correct trials was not significantly different from chance, the birds did not learn to transfer knowledge of a relational rule when encountering novel stimuli, novel dimensions, or novel dimension combinations, 2) if the percentage of correct trials was statistically greater than chance, but statistically less than that of training, this would be considered partial transfer, or 3) if the percentage of correct trials was statistically greater than chance, and not significantly different from training, this would be considered full transfer (Wright et al., 2017). Full transfer indicates learning of the relational rule to solve analogies (i.e., recognizing same/different within-pair relationships and then solving the task by identifying the shared relationship between-pairs), partial transfer indicates partial learning or possibly the use of a different strategy to solve analogies, and no transfer indicates the inability to use the learned analogy during training with novel exemplars.

Learning Rate

My design allowed for assessing the average learning rate across birds by evaluating the number of training trials needed to reach criterion during each phase of training. An increase or no change in the learning rate needed to reach the training criterion during a certain phase may indicate a difficulty with learning the task requirements. For instance, how to apply the relational rule to the problems and/or difficulties disambiguating particular dimensions or relationships presented during a phase.

Independent Variables

Phase

My design allowed for evaluating whether the percentage of correct trials and the learning rate across birds significantly changed among test and training phases, respectively.

Trial Type

Trial type was used to compare test and training trials to determine if the birds met the criterion for full or partial transfer.

Statistical Analyses

Test Performance

My design allowed for comparing percent correct test trial performance to chance using one-sample t-tests and percent correct performance between test and training trials using paired sample t-tests. These analyses were included to determine if the birds learned to transfer a relational rule to novel problems.

My design allowed for examining the complexity of the test phases using a 2x5 factorial ANOVA for trial type (training or test) and test phase (*Dimension 1 Test*, *Dimension 2 Test*, *Dimensions 1&2 Combination Test*, *Dimension 3 Test*, and the *All Combination Test*) with the dependent variable of percent correct trials. Complexity comprises the specific number of dimensions present in each phase, as well as, how many of those dimensions are present per trial (e.g., the *Dimension 3 Test* contains three dimensions, but presents variations of only two out of the three per trial). This analysis was included to determine how well the birds performed during each test phase as complexity increased. Main effects and interactions would be examined using a Tukey HSD test.

Training Performance

Each consecutive training phase varied by the substitution or addition of a single dimension compared to the previous phase. My design allowed for examining how increasing training phase complexity effects learning rate using a repeated measures ANOVA with the independent variable being training phase (*Dimension 1 Training*, *Dimension 2 Training*, *Dimensions 1&2 Combination Training*, *Dimension 3 Training*, and *All Combination Training*)

and the dependent variable being learning rate. Main effects would be examined using a Tukey HSD test.

CHAPTER V

Training Outcome

Seven birds participated in *Dimension 1 Training*. Two additional birds were scheduled to participate, however my study was suspended due to the Covid-19 pandemic. During *Dimension 1 Training*, no bird met the criterion of 80% or more correct trials over two consecutive sessions. I was unable to train the birds to use the relational rule required to solve analogies using my procedures. Thus, as a result of difficulties attributed to the training design and failed attempts to execute suitable solutions, the Analogical Reasoning Task training failed. Although data were collected for parts of the training procedures, no testing data were collected. Therefore, this section is organized to first describe the modifications I included in the *Dimension 1 Training* procedure to help facilitate learning, followed by new statistical analyses and the results of the training data.

Dimension 1 Training Modifications

Within the first weeks of *Dimension 1 Training*, my procedures did not appear to promote learning. I was responsible for determining how to make my task successful. I conducted individual observations of the birds during training, inspected the training data, and consulted the literature for similar training methods to identify possible problems with my design. As the birds proceeded through training, I determined potential problems (see below) that might have impeded their learning. Subsequently, modifications were applied to the training procedure in order to promote learning. These modifications were applied consecutively and remained in place throughout the remainder of training. Each bird progressed through training at its own pace, therefore modifications were made at relatively consistent time periods to ensure each bird experienced a similar training progression. These modifications are described below.

Intertrial Interval (ITI) and Timeout Duration

The function of an ITI is to denote the separation between trials. The function of a correction procedure (CP) is to present the same trial repeatedly until a correct choice is made. Immediately preceding the CP was a timeout period that was meant to build an aversion to making incorrect choices. The ITI and timeout duration for the current study was based on previous same/different concept learning operant studies, in which correct choices were followed by an ITI of 15 s with the lights on and incorrect choices were followed by a timeout of 15 s with the lights off, preceding a CP and subsequent ITI of 15 s both with the lights on again (pigeons [Katz & Wright 2006]; nutcrackers [Wright et al., 2016]; magpies [Magnotti et al., 2017]). The current study followed the same procedure with the exception that the ITI following a CP was shortened to 10 s. However, over the first few weeks of training many of the birds were not consistently completing all trials within a daily session, or birds were unable to make a correct choice, resulting in extinction of choice behaviour. I hypothesized that the ITI was too long following correct choices, causing the birds to extinguish their choice behaviour. Short ITI durations have been shown to have effects on choice behaviour, due to the fast rate of reward between trials resulting in sustained attention. During a visual discrimination task, pigeons were reported to make fewer errors during continued training with short ITIs (i.e., 4 s, 15 s), compared to longer ITIs, which yielded no improvements in learning (Williams, 1971). In addition to ITIs, I hypothesized that in the current study *incorrect choices* needed to be followed by a longer timeout, as the current timeout duration was not improving the birds' accuracy. Extending the duration would result in a more aversive timeout, which was intended to prompt careful choice-making, while also minimizing guessing by making it less motivating to risk a longer wait before the next opportunity for reward. Therefore, in the current study, the ITI for correct choices was lowered to 10 s and the timeout was raised to 30 s (followed by an ITI of 5 s).

Correction Procedure Removal

Performance did not improve after five sessions (e.g., approximately one to three weeks depending on the subject). Throughout training the birds needed to learn through *trial-and-error*. However, not learning from their errors was perhaps resulting in failure to understand the task. Therefore, the outcome that followed an incorrect choice needed to better facilitate learning. An alternative to the *trial-and-error* training method is *errorless learning* (Terrace, 1963), in which

a subject is gradually taught to discriminate with little to no possibility of making errors. Notably, during training, pigeons given errorless learning procedures are shown to learn discriminations faster and to make significantly fewer errors to incorrect stimuli, compared to controls that were given trial-and-error learning procedures (Arantes & Berg, 2009; Arantes & Machado, 2011). For the current study, I did not adopt an errorless learning procedure, rather I determined that a modification to the current procedure could serve a similar function in reducing errors. I hypothesized that temporarily removing the CP from sessions and ensuring that no trial ended with an error could assist with learning. The modified procedure was as follows: when a bird pecked a correct choice stimulus, it was rewarded and the bird advanced to the next trial, however if the bird pecked the incorrect choice, the trial would not advance. Pecks to the incorrect choice were inconsequential. Therefore, trials could only end with pecks to the correct stimulus. This procedure was intended to increase the rate of reward for the birds over a daily session. Additionally, it was meant to decrease pecks to incorrect stimuli and increase attention to the rewarded stimulus. Directing attention to the correct stimulus should prompt birds to focus on its characteristics, which was intended to aid them in building connections between the correct stimulus and the characteristics of the primary and bridge stimuli.

Correction Procedure Addition

It became evident after 20 sessions of having the CP removed that this procedure was not adequate for improving learning and thus the CP was reinstituted. Prior to the removal of the CP, the birds were completing on average 37 correction trials per session. Upon observation, it was determined that the high number of correction trials were likely due to the birds experiencing difficulty with distinguishing between new and repeated trials. For instance, if it was not understood that a correction trial was a previously incorrect trial repeatedly presented, then the correction trial could be falsely perceived as a new trial. Thus, the incorrect choice on a first attempt may not be recognized as still incorrect during the correction trial. To address this concern, CPs were now modified so that on a new trial all stimuli that were on the screen when a bird made its first choice would now reappear simultaneously on the screen when a correction trial started. The important difference in this CP is that a bird was not required to peck through the stimuli in the manner that it would when given a new trial. Instead, the correction trial would appear on the screen exactly as the screen had looked before the timeout. The only peck required

would be to the correct choice stimulus, which would result in reward, whereas a peck to the incorrect choice stimulus would repeated the trial. Thus, this modification served to improve trial recognition (e.g., recognizing a second attempt at the same trial).

Choice Fixed Ratio Increase

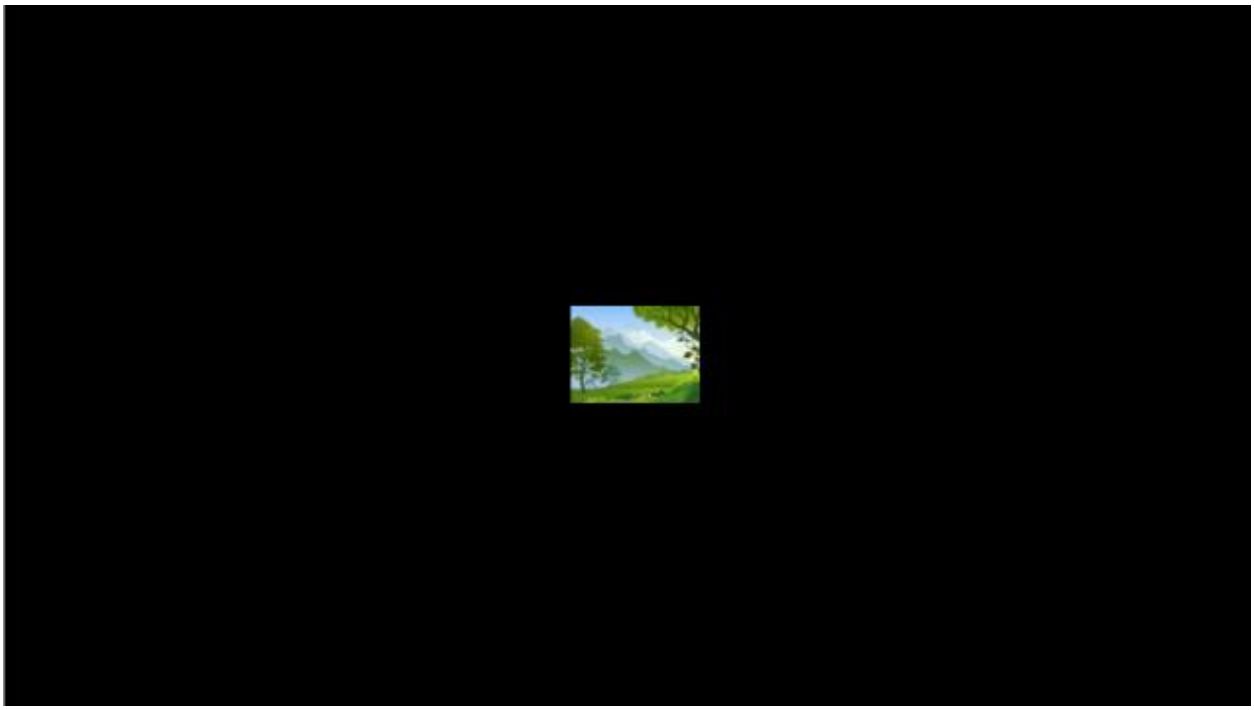
When *Dimension 1 Training* began, one peck (FR1) was required to be made to a choice stimulus. Within the first week of training the FR requirement for choice stimuli was increased to two pecks. Following 57-76 sessions without reaching the training criterion, the FR to choice stimuli was raised further to help improve correct choice making by increasing the effort of selecting a choice. An FR5 was used for each of the primary and bridge stimuli throughout training. Thus, the FR for choice stimuli was gradually raised from two to five with the criterion that a bird had to complete an entire session at each FR before preceding to the next session. Sacks et al. (1972) showed that under operant conditions, pigeons given a two-alternative choice match-to-sample task (i.e., in which they had to choose which choice stimulus matched a sample stimulus) performed with higher rates of acquisition (i.e., concept learning) and transfer to novel stimuli when the FR to the sample stimulus was increased. Similarly, Wright (1997) showed that concept learning by pigeons, in a similar match-to-sample task, increased with a higher sample FR. Whereas, these findings show that learning improves with attention to the sample (in our case the *primary stimuli*) I rationalized that greater attention to choice stimuli may assist my birds with making connections between the relational features of the stimulus pairings. This change was made as opposed to raising the FR to the primary and bridge stimuli further, as a change of this magnitude might cause ratio strain (i.e., a reduction in behaviour resulting from an increase in the ratio of required responses preceding reinforcement), which could result in pauses in progression.

Warning Stimulus

After an additional 8-12 sessions following an increase of five pecks to the choice stimulus, the birds that progressed the furthest in training, Gonzo and Theseus, had remained below the criterion for advancing from training to testing. With further observation and procedural assessment, it became evident that an additional procedural aid—a warning stimulus—would be useful to help the birds understand how trials were presented (i.e., new or repeated), in addition to acting as a ‘ready’ signal. A warning stimulus (WS) is a visual aid that

precedes an event or the presentation of another stimulus, thus creating a preparatory period during which a subject is trained to expect something to occur soon after its presentation. These stimuli have been used in operant studies that investigate a variety of processes (e.g., visual memory capacity [nutcrackers; Qadri et al., 2018]; working memory [crows; Goto & Watanabe, 2009]; metamemory [pigeons; Sutton & Shettleworth, 2008]; form-perception [pigeons; Kelly et al., 2001]; and short-term memory [pigeons; Roitblat, 1980]). A WS was created using a clipart image of a forest landscape (141 x 105 pixels). Prior to the start of a new trial, the WS appeared at the center of the screen (Figure 6). Once a bird made a single peck to the WS it disappeared and a new trial began. However, the WS never appeared before a correction trial. This distinction was instituted to further distinguish a repeated trial from a new trial.

Figure 6. Example display of the warning stimulus. How the warning stimulus appeared on the screen prior to the start of a new trial.



Individuals Restarting Dimension 1 Training

During their first weeks of training, two birds (Kermit and Capulet) responded with difficulty to the *Dimension 1 Training* procedures, as they were frequently inactive or slow to complete sessions. This difficulty sustaining motivation could not be explained by fluctuations in body weight, as there was no clear trend between weight over time and session performance.

Due to their declining participation, both birds were restarted on Stage 1 of *Dimension 1 Training* (i.e., Stage 1 Retraining)**,** in which the first 30 trials of a session presented one choice stimulus and the remaining 30 trials presented two choice stimuli. To simplify the Stage 1 procedure, the ITI was now 10 s and no CP was in place. Whereas correct choices were still followed by reinforcement, the ITI (house light on), and the next trial, incorrect choices were followed immediately by the ITI (house light off) and the next trial. The criterion for advancement to the next stage was to complete 85% of trials for two consecutive days. Additionally, a bird did not resume a session from the previous day. Rather each day began with a new session. This was to ensure that each session began with the simplest type of trials (those with one choice stimulus), which served to increase the daily rate of reward and maintain motivation to participate. Once this criterion was met, the bird would advance to Stage 2 Retraining, which was identical to the previous stage, except that all 60 trials presented two choice stimuli. The criterion for advancement was to complete an entire session for two consecutive days. Once this criterion was met, the bird advanced to regular Stage 2 sessions in which the ITI was 10 s, a 30 s CP was introduced, and incomplete sessions resumed where a bird left off the prior day.

Statistical Analyses for Dimension 1 Training Data

Originally, there were two dependent variables of interest for the analyses: the learning rate during training and the percentage of correct trials during testing. Only *Dimension 1 Training* data was available for analysis. As learning rate could not be evaluated across multiple training phases, the percentage of correct trials was instead examined.

Dependent Variable

Percent Correct

The percentage of correct trials per session was used to evaluate if a bird scored significantly greater than chance (i.e., 50%). For ease of communicating when birds experienced training procedure modifications in the Results section, these modifications were henceforth termed ‘divisions’ of *Dimension 1 Training* (i.e., the span of sessions when each modification was applied to the training procedure). Additionally, I evaluated if there were significant differences in percent correct performance as an effect of dimension trained.

Independent Variable

Dimension

I evaluated whether training dimension (colour, shape, or pattern) had an effect on percentage of correct trials.

Statistical Analyses

I was planning on conducting parametric analyses on my data. However due to having a small sample size for this study, I used non-parametric tests for my analyses, as no more than five birds' data were analyzed per test.

Dimensional Performance

The percentage of correct choices between the three groups of birds trained on the shape, colour, and pattern dimensions were examined using a Kruskal-Wallis test. This analysis was used to examine if the percentage of correct trials was higher when trained on one dimension compared to another, as I expected that training dimension may have affected individual performance. A separate analysis was conducted for each division of training. The percentage of correct trials was analyzed for five of the subjects with dimension as the grouping variable (two trained on shape, two trained on colour, and one trained on pattern). Only sessions in which birds completed 60 trials were included. Kermit and Capulet were not included in this analysis, as their training experience varied from the remaining birds. If significant effects were found, they were further examined using a Dwass-Steel-Critchlow-Fligner multiple comparisons test, which contains protections for family-wise error rate.

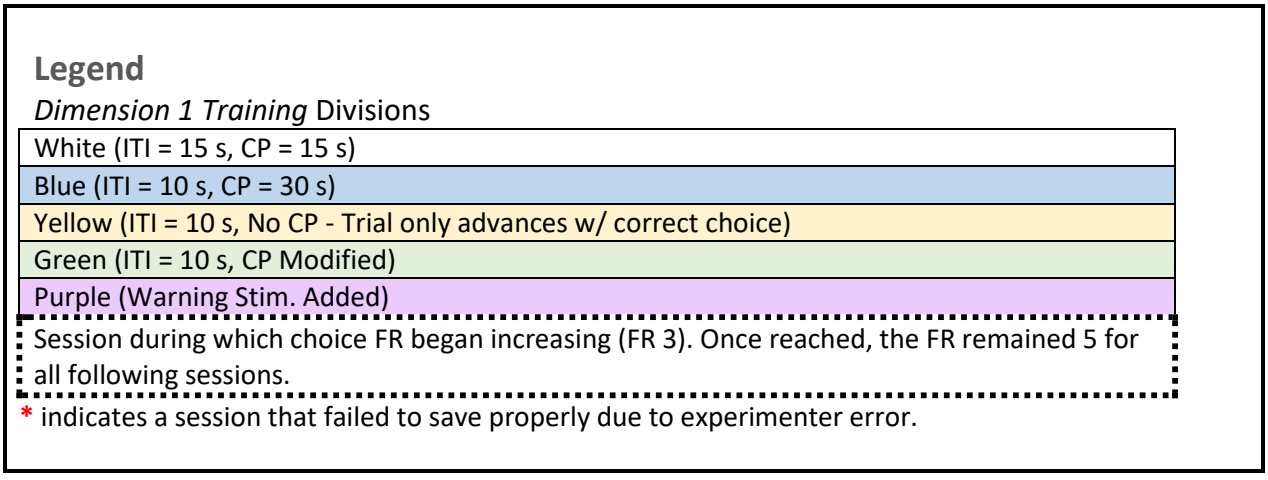
Results

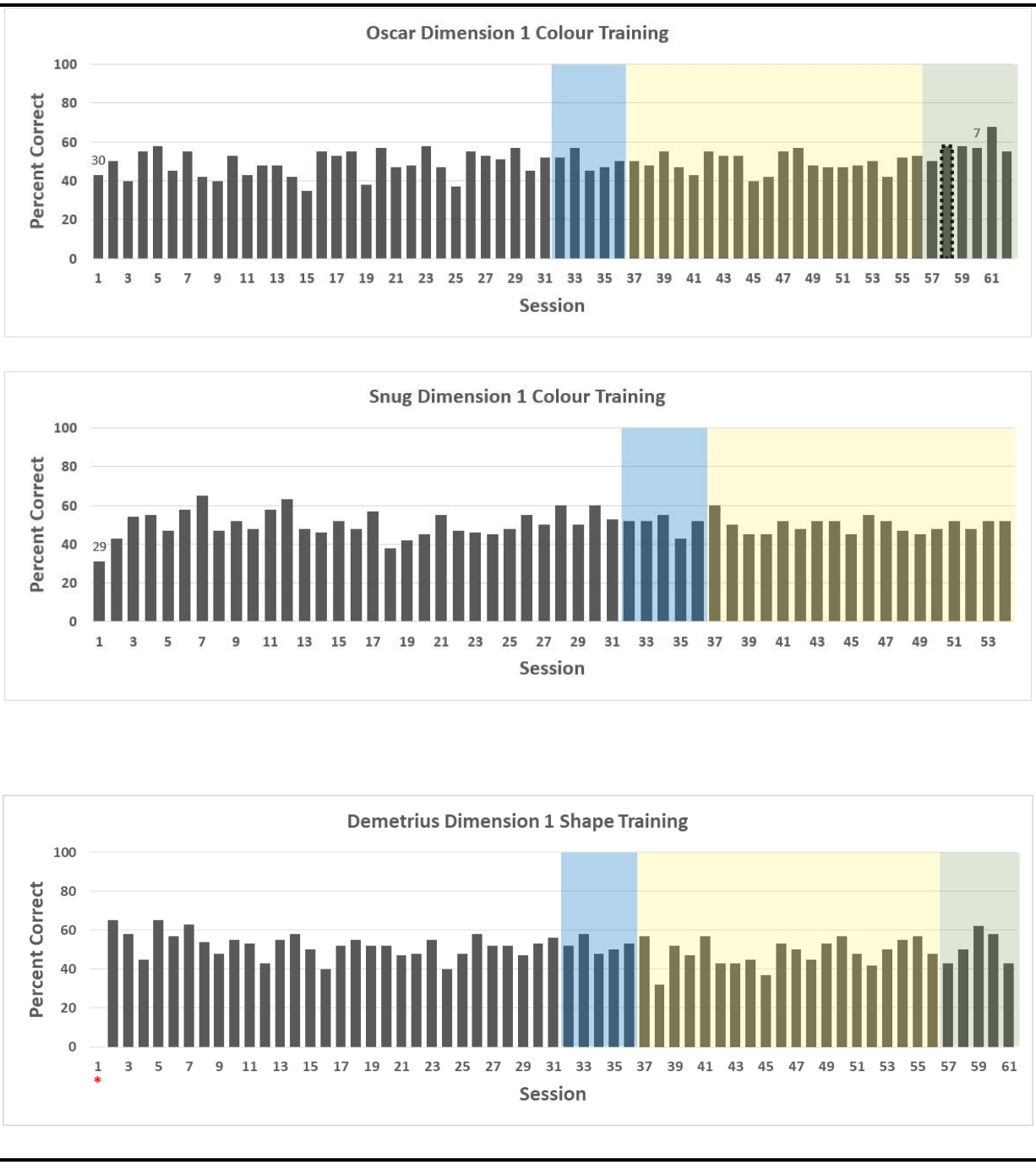
The birds required an average of 2.7 sessions (5 sessions maximum) to complete Stage 1 of the shaping procedures and an average of 16 sessions (23 sessions maximum) to complete Stage 2. During *Dimension 1 Training* individuals completed 54-104 sessions (see Figures 7 and 8). The average percentage of correct trials across all of the birds' sessions (excluding Capulet and Kermit) was 50.3 ± 0.33% (only data from two-choice sessions during which birds completed all 60 trials were included). Notably, during Stage 1 Retraining, an error was detected in the code which resulted in the program not advancing from trial to trial correctly. This error

only occurred following the beginning of two-choice stimulus trials and affected a subset of Kermit's Stage 1 Retraining sessions. These sessions are denoted by a blue asterisk in Figure 8.

The divisions of *Dimension 1 Training* were colour-coded to better illustrate when they occurred during the training time-span (see Figure 7). For instance, sessions preceding any modifications were coded *White* and shown in Figure 7 with a white background, whereas the first modification, made to the ITI and timeout duration, was coded as *Blue* with a blue background. The next modifications resulted in the removal of CPs (Yellow), followed by the addition of a modified CP (Green). The following modification resulted in increasing the FR requirement for choice stimuli. This change coincided with the prior modification and was not colour coded, but rather was indicated in figures using a dashed outline. The final modification was the addition of the warning signal (Purple). Additionally, for Capulet and Kermit's data, Stage 1 Retraining and Stage 2 Retraining were coded as Grey and Red, respectively (see Figure 8).

Figure 7. Percentage of correct choices per session. Percent correct was calculated for all two-choice trials completed for each session. The number of trials completed per session was 60, unless indicated otherwise by the number at the top of the bar. Session 1 was divided between 30 one-choice trials and 30 two-choice trials, in which case percent correct was calculated out of 30 trials (or fewer).




Oscar Dimension 1 Colour Training
Percent Correct
100
80
60
40
20
0
30
7
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61
Session
Snug Dimension 1 Colour Training
Percent Correct
100
80
60
40
20
0
29
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53
Session
Demetrius Dimension 1 Shape Training
Percent Correct
100
80
60
40
20
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61
*
Session

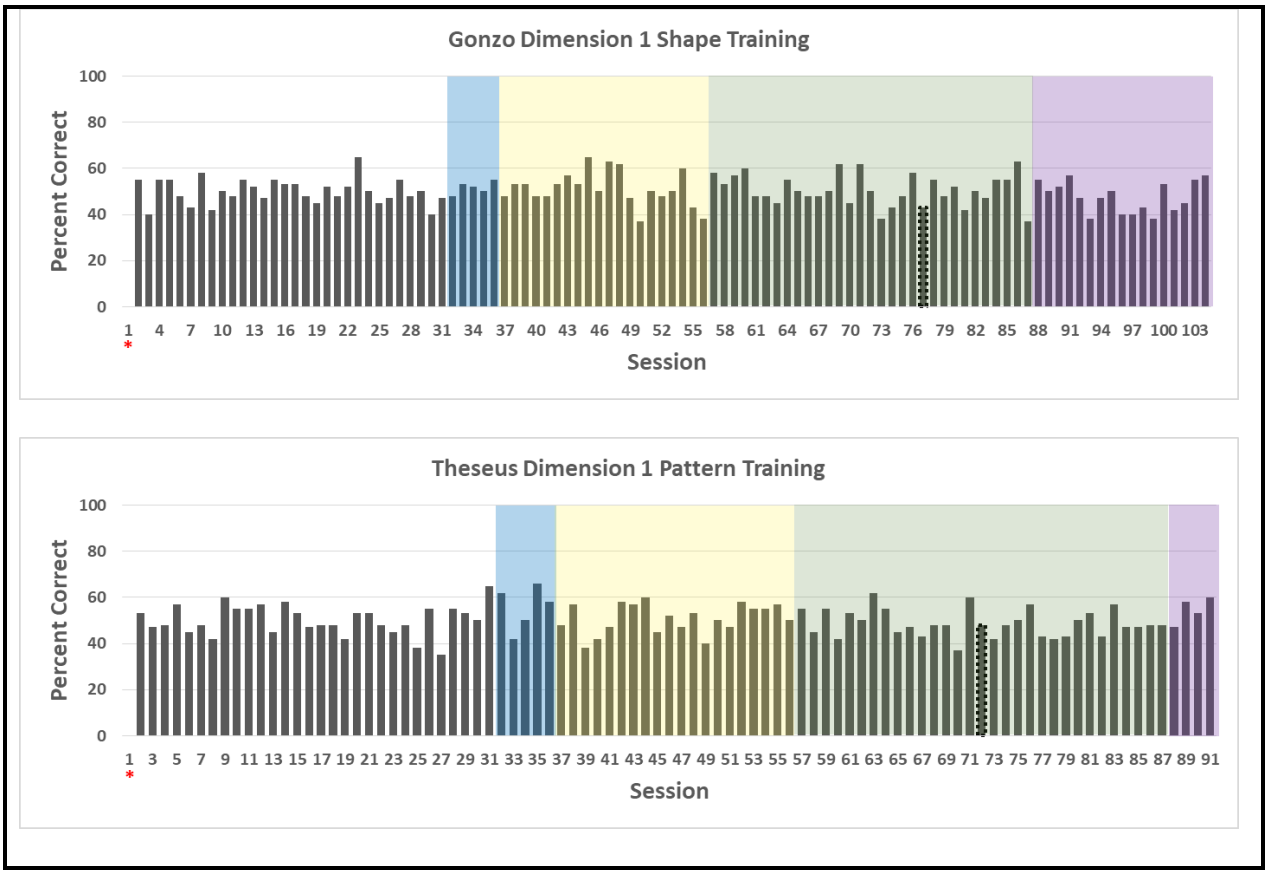


Figure 8. Percentage of correct choices per session for birds that restarted Stage 1 training. Percent correct was calculated for all two-choice trials completed for each session. The number of trials completed per session was 60, unless indicated otherwise by the number at the top of the bar. Session 1 and sessions within the grey area were divided between 30 one-choice trials and 30 two-choice trials, in which case percent correct was calculated out of 30 trials (or fewer).

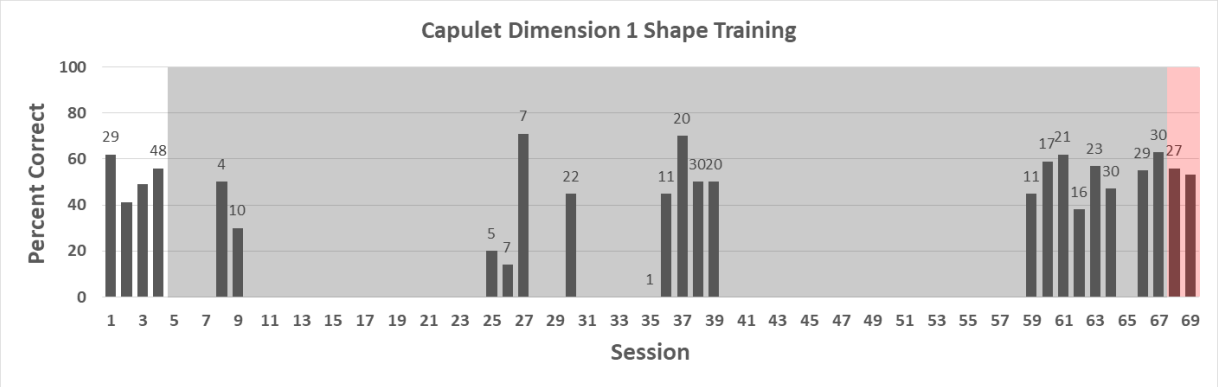
Legend

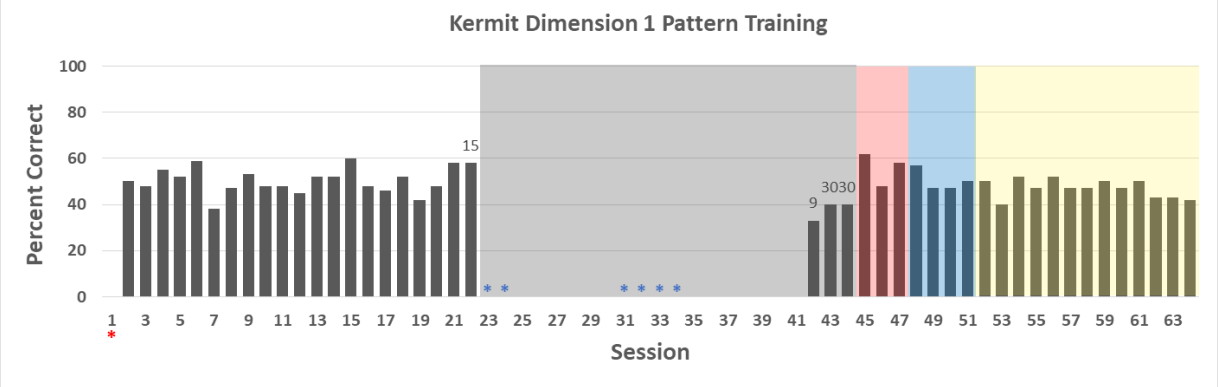
Dimension 1 Training Division

White (ITI = 15 s, CP = 15 s), 60 two-choice trials
Grey - Stage 1 Retraining (No CP), 30 two-choice trials
Red - Stage 2 Retraining (No CP), 60 two-choice trials
Blue (ITI = 10 s, CP = 30 s), 60 two-choice trials
Yellow (ITI = 10 s, No CP - Trial only advances w/ correct choice)

* indicates a session that was excluded due to code error.

* indicates a session that failed to save properly due to experimenter error.





Dimensional Performance

There were no statistically significant differences in the percentage of correct trials among dimensions during any division of training, except the Green division (White: χ^2 (2, N = 3) = 0.94, p = .626, ε^2 = .006; Blue: χ^2 (2, N = 3) = 1.58, p = .454, ε^2 = .066; Yellow: χ^2 (2, N = 3) = 0.85, p = .654, ε^2 = .009; Green: χ^2 (2, N = 3) = 7.58, p = .023, ε^2 = .107; Purple: χ^2 (1, N = 2) = 3.39, p = .065, ε^2 = .170). Post hoc analysis revealed that within the Green division, there

was a difference between the colour and pattern dimensions (p = .022; Ms and SEM: 57.8 ± 2.94, 48.7 ± 1.06, for colour and pattern respectively).

CHAPTER VI

Discussion

Thesis Overview

My research objective was to determine whether pinyon jays are capable of reasoning by analogy. Displaying this ability would indicate that pinyon jays can make logical inferences about the relationships shared between familiar and novel stimuli. Additionally, successfully choosing the correct stimulus to complete the analogy would suggest that the paradigm used was suitable for testing analogical reasoning across multiple species. During my study, the pinyon jays were trained to recognize the relational equivalence among stimuli starting with one of three dimensions (shape, colour, or pattern). The task was designed to present progressively more complex problems by increasing the dimensional variation among the stimuli, both in terms of addition and combination. The pinyon jays progressed through their first phase of training, which consisted of *Dimension 1 Training* sessions. However, the lack of improved correct choice responding indicated that this training phase was not adequate for training the birds. Aspects of the experimental design or procedures may have prevented the pinyon jays from learning the task. In this chapter, I will discuss each of these aspects and suggest further steps to address these problems.

Results Summary

All birds remained on *Dimension 1 Training* for the duration of the experiment (i.e., approximately seven months). Throughout their training, the percentage of correct trials per session did not statistically differ from chance. This result might suggest that there were procedural aspects of the task that unexpectedly prevented the birds from learning how to solve the analogical problems.

I analyzed whether stimulus dimension might have contributed to learning difficulties. I determined that the need for the birds' prolonged training period cannot be explained by the first

dimension learned, as stimulus dimension did not have an effect on task acquisition. Whereas there was a significant difference between the colour and pattern dimensions during the Green division of training, only a single bird trained on colour had reached this division. This colour trained bird completed a low number of sessions during this division (colour = 5, pattern = 31), thus there was not sufficient evidence to determine the cause of this effect (e.g., individual differences or reinstating the modified CP). Moreover, two birds restarted their training, however other birds that were trained on the same dimensions continued training without ceasing to participate (i.e., failing to complete any trials during a session). Therefore, it remains uncertain if the extinction of the two birds' behaviour was due to stimulus dimension or a different cause. The pinyon jays' difficulty advancing past *Dimension 1 Training* was likely due to certain aspects of the training procedure and the procedural modifications failing to promote learning. This reveals that the current paradigm used to evaluate analogical reasoning by humans requires additional adjustments to the training design to successfully aid pinyon jays on learning the task requirements.

Objectives and Advantages of the Current Paradigm

The current paradigm was designed to circumvent common problems associated with the relational match-to-sample (RMTS) and analogical reasoning paradigms from which it was based. The concerns with these paradigms led to uncertainty in whether previous experiments were adequate tests of analogical reasoning ability, given that other strategies (e.g., perceptual matching of surface similarities) could result in success on the task. Additionally, this paradigm was intended to be a true test of analogy by presenting questions in which analogical reasoning was the only reliable strategy that could be used to solve the task. For this reason, the current experimental procedure deviated from the more standard RMTS approach for assessing analogical reasoning. During prior studies subjects were often required to 'match the relations' by presenting a sample pair of stimuli and having them pick between two comparison pairs, based on relational similarity. The current approach required a different form of analogical task, 'complete the relation', in which subjects must pick a choice stimulus that when paired with another (e.g., the bridge stimulus), completes the presented analogy (Gillan et al., 1981). The advantage to this approach is that the problems are inherently second-order, a requirement for assessing analogical reasoning. Previous studies examining analogical reasoning with crows and

parrots used a RMTS paradigm in which pairs of stimuli were presented together on single cards, rather than having each stimulus of a pair side-by-side on individual cards (Obozoba et al., 2015; Smirnova et al., 2015). This design makes it possible that the birds could have perceived the pairs of stimuli as single entities, resulting in first-order matching. In response to this concern, the current paradigm's 'complete the relation' style questions resolved this issue by presenting all stimuli as individual entities to be paired together. Moreover, the aim of the current study was to determine the true capabilities and extent to which an avian species could reason analogically by progressively increasing the difficulty of the analogies (i.e., including more dimensions, thus increasing the number of *same* and *different* within pair relations). Specifically, this paradigm could test reasoning using complex problems that include pairs of stimuli that present three distinct dimension relations simultaneously; the extent to which has not yet been evaluated with crows and parrots. Indeed, these species were given separate tests using each dimension to assess analogical reasoning, rather than combining dimensions in a single test (Obozoba et al., 2015; Smirnova et al., 2015). An added advantage of the current paradigm, which coincides with the addition of multiple dimensions in a single question, is the ability to investigate which factors explain incorrect choice-making. Because the correctly and incorrectly represented dimensions are balanced across all choice stimuli for a single trial, when an incorrect choice is made, it is possible to discern which dimensions the subject was attending to (e.g., a specific type of dimension (colour), or a dimension that it has had more or less exposure to). This determines which factors most influence a subject's ability to learn the task. Although the current paradigm met the aforementioned objectives, the experimental procedures were not adequate for training the pinyon jays on the necessary information to permit evaluation of their ability to solve analogical problems.

The Current Analogical Reasoning Paradigm in Practice

The current analogical reasoning procedure was designed to train the birds on relational rule use (i.e., solving the task by recognizing that the within pair relationships of the primary pair and the bridge-choice pair should be identical). The pinyon jays' suspended progression during initial training provides evidence that the current approach was not sufficient to train the birds. However, the current results cannot confirm that pinyon jays are incapable of reasoning by analogy, nor that this ability cannot be tested under operant conditions. It is evident that the

current paradigm requires some procedural changes. This paradigm was originally designed to be tested by humans, with the intention that the analogical problems presented would be applicable for testing reasoning ability across multiple species. During the original study, the majority of human subjects were able to complete all phases of training and testing (Leonard et al., in prep). Conversely, the current study procedures failed to teach the pinyon jays the requirements to reach the advancement criterion for *Dimension 1 Training*. The pinyon jays did show the ability to distinguish *same* and *different* concepts in a previous match to sample (MTS) operant task (Vernouillet et al., in press). During this MTS task and the current study, the pinyon jays were presented with alternative *same* and *different* choices. During the MTS task, each trial required a bird to peck a choice stimulus that was identical to the sample if they were the same or a white rectangle if the sample and comparison stimuli were different. During the current study, each trial required a bird to both recognize what dimensions were the *same* and *different* among the stimuli, and to use this information to peck a choice stimulus that would complete the presented *same* or *different* relationship. The success in the MTS task (Vernouillet et al., in press) compared to the current task may have been a result of the difficulties of learning to solve the problems presented by either study. Despite greater difficulty during the current study, there was little difference in the number of *same* and *different* trials that were correct across the birds' sessions, which indicated that this was not the concept that the birds were struggling with. Therefore, it is evident that despite being adapted from the original human study, there are aspects of the current experimental procedure that may not facilitate learning by pinyon jays. This study may act to provide information (e.g., specific unsuccessful procedures and modifications) to future researchers who are exploring adjustments that should be made to the current experimental design to better accommodate how pinyon jays attend to and recognize relational similarity between stimuli.

Mistakes of the Current Study

Despite carefully considering how to design an experiment that would accurately test the extent of analogical reasoning ability across species, my primary focus should have been on clearly answering the fundamental question of whether my experiment was adequate for examining this ability by *pinyon jays*. My mistake was placing more attention on how a paradigm that successfully tested this ability by humans could be adapted to make direct cross-

species comparisons under operant conditions. Instead, I should have made a more cautious evaluation of whether my training procedure was suitable for facilitating concept learning. Consequently, the pinyon jays struggled to understand the task requirements, as these requirements for recognizing relational similarity were not made clear using my design. Perhaps the pinyon jays required the analogical problems and procedures to be simplified further than they were during the initial phases of the experiment. Furthermore, the procedural modifications that were implemented did not help to facilitate this understanding. These modifications might have been more impactful to the pinyon jay's learning had their incorporation into my design been considered prior to the birds beginning the Analogical Reasoning Task. Additionally, a large focus of my design was on raising the complexity of the task with progressive phases, thus much attention was given to ensuring that the phases exceeding *Dimension 1 Training* possessed no foreseeable flaws. In retrospect, my primary concern should have been ensuring that shaping procedures and *Dimension 1 Training* were designed such that these critical learning phases would have been successful in promoting concept learning. If I were to fix this experimental design then I would consider revising these early training procedures. In general, revising how the stimuli were presented and the trial progression. For instance, this might involve some deconstruction of the analogical problem into its component stimuli and better training the birds on how to compare these stimuli before progressing to a fully structured analogical problem. Doing so could assist with training the birds to attend to and peck stimuli based on relevant features, and progressively relationships, as opposed to order of appearance. The notable error of my task was that it was designed to construct and display analogical problems, however, it fell short of teaching the birds to recognize relational similarity between the stimuli composing those problems. Thus, the birds were not taught the necessary precursor that could have led to analogical reasoning. Revising how the stimuli are first introduced during shaping and early training, for the purpose of improving how the birds are taught to recognize relational similarity, would make my task better suited to answering my research question.

Future Procedural Considerations

The current procedures were not able to train the pinyon jays to solve single dimension analogical problems, despite humans learning to do so using nearly identical procedures. It may be that a certain aspect(s) of the training procedure resulted in the birds' difficulties with learning

the task. To examine how subjects acquire concepts in operant tasks, it is necessary to consider at least three aspects of operant designs, one or more of which may have affected the pinyon jays' success: (1) perceptual influences, such as the arrangement of stimuli in the display, (2) the training procedure, and (3) elements that control the progression through a session (e.g., CP, FR, and WS).

Stimuli Arrangement and Presentation

An important aspect of this paradigm was to ensure that trials presented second-order questions that promoted a between pair relational comparison. For instance, the primary stimuli needed to be spatially separated to ensure items were recognized as two separate entities, yet still represented as a pair. Thus, their primary presentation and proximity to one another (isolated side-by-side within their own frame) was designed to promote a within-pair comparison. The bridge stimulus, isolated in its own frame, indicated that it required pairing. Through training, the choice stimuli were to be recognized as matching options. Adult humans appear to understand this layout. Throughout development, humans typically learn analytical skills that may allow them to recognize relational similarities later in life, for example, when critical thinking is nurtured in youth and problem based learning is employed (for a review see Masek & Yamin, 2011). In fact, the development of relational reasoning ability is empirically linked to achievements and improved performance by students in mathematics, science, reading and writing (for a review see Dumas et al., 2013). Therefore, experience with learning and using relational logic may have aided performance by human participants (Leonard et al., in prep) in this task. Thus, these human participants may have possessed an advantage for understanding the nature of the relational problems as they were presented. Alternatively, in the case of pinyon jays, it is possible that the arrangement of this display may have created difficultly in making connections between how the stimuli should be regarded and the relationships they represented. This may be the result of the proximity of the choice stimuli from the remaining stimuli and/or the locations of the stimuli on the screen. Fagot and Parron (2010) showed that in an operant task, baboons were able to accurately match same/different relations between pairs of colour patches, when both elements of a pair were separated by a 2 pixel gap. When the gap sizes between the sample pair elements were increased, performance began to deteriorate as it became more difficult for the baboons to match relations when the sample pair's elements were more

spatially distant compared to the comparison pair elements. However, when *trained* to match relations as the sample and comparison gap sizes were progressively increased, accurate matching of the relational pairs was maintained, even when tested with a different set of colour patches. This suggested that a spatially distant pair of elements can be perceived as relating to one another if the mental connection was made between the elements at a point in time (e.g., initial training). It is possible that in the current experiment, having a gap size between the primary stimuli that was relatively large, may have made it less perceivable that both primary stimuli (while separate entities) should be regarded together as a pair. This lack of mental connectivity between stimuli may extend to the bridge-choice pairs, as these stimuli maintained relatively large distances from one another. A solution would be to minimize the intermediate distances between both the primary stimuli and the bridge-choice stimuli, thus helping the subject to recognize each as a pairing.

Considering that learning could be better facilitated by reducing the amount of screen space occupied by the choice stimuli, and increasing their proximity to the bridge stimulus, it may also be appropriate to optimize the locations that stimuli should appear on the screen. It is common when testing various species in the same/different learning paradigm for stimuli to be either centered on the display (e.g., Wright et al., 2017), or for the sample pair in the RMTS paradigm to be located at the center-top of the display and the comparisons to be to the lower left and right of the sample (Flemming et al., 2008; Flemming et al., 2013). Additionally, all pairs or groupings of stimuli are relatively close in proximity (Flemming et al., 2008; Flemming et al., 2013; Wright et al., 2017). As these previous designs have been shown to be adequate for training subjects with same/different matching or match-the-relation problems, a similar design may also be adequate for the complete-the-relation problems of the current paradigm. Gillan et al. (1981) presented a complete-the-relation analogical reasoning task to a previously token trained chimp in which a pair of cardboard shapes were arranged vertically on the left side of a tray and a single comparison shape was placed on the right side. Between the tray and the subject were two choice shapes, and at the center of the tray was a symbol that represented the concept *same*, which served as a cue for locating the relational match. The subject was successful at completing relations using this design. Therefore, the stimulus placement of this design can be adapted to the current paradigm, by having the primary and bridge stimuli centrally positioned on screen, with the options of the choice stimuli more proximal to the bridge stimulus. This may

assist in training the birds to understand that the primary and bridge stimuli together represent a single problem with a missing piece, and that one of the choice stimuli located directly below or adjacent to the problem represents that missing piece.

Initial Training Procedure

Not reaching the criterion for advancing from training to testing causes concern for whether the training procedure was adequate for teaching the pinyon jays the task requirements. In the current paradigm, the first *Dimension 1 Training* session given to the birds started with 30 single choice second-order matching trials, followed by 30 similar binary choice trials (with all following training sessions containing binary choice trials). Although previous studies have investigated the analogical reasoning abilities of primates while only presenting second-order relational matching trials throughout their experiments (Fagot & Parron, 2010; Flemming et al. 2008; Thompson et al. 1997), there are a number of primate and bird analogy studies that precede relational matching trials with first-order match-to-sample training or testing (Flemming & Kennedy, 2011; Kennedy & Fragaszy, 2008; Obozoba et al., 2015; Smirnova et al., 2015; Vonk 2003). Training subjects on first-order matching-to-sample problems, in which the correct choice stimulus is *identical* to the primary (sample) stimulus, may help to establish an understanding of sameness as well as promote the comparison of the primary stimuli to the choice stimuli. This experience with first-order matching may help to facilitate the recognition of relational similarities between stimuli as second-order problems are later presented. This preparation for recognizing similar relational structure, following experience with physically similar structures (i.e., the *progressive alignment effect*), was observed in children whose within dimensional physical matching experience facilitated improved recognition of higher-order similarities between stimuli during a test of cross-dimensional relational matching (Kotovsky & Gentner, 1996). Despite this evidence and the aforementioned hypothesis, presentation of a first-order problem in the current paradigm would be potentially problematic given that the primary stimuli are presented as a pair, thus requiring the presentation of two alternative comparison pairs (e.g., primary pair: two circles, correct comparison: two circles, incorrect comparison: two squares). Firstly, this example would not meet the requirements of a first-order problem. Secondly, this example presents much like a traditional RMTS problem, rather than a complete-the-relation problem. Finally, this example problem may cue subjects to consider featural

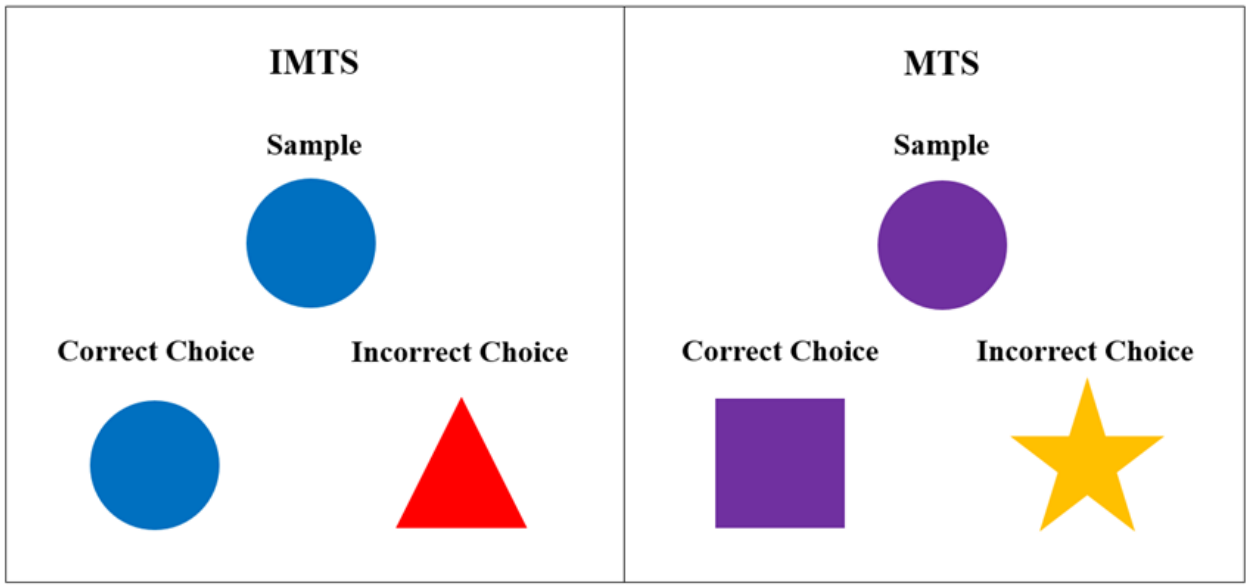
information (e.g., circles match circles) rather than conceptual information (e.g., the correct comparison pair are the *same* as the primary pair), thus promoting the use of perceptual cues to solve the task.

The critical component of success in the current training procedure may be the properties of the task which will foster relational rule learning by the birds. Recently, Kroupin and Carey (2021) evaluated whether successful RMTS training procedures for humans and non-human animals were a result of task properties that provided subjects with learning experiences either aiding in the *generation* of mental representations of sameness and difference, or aiding in *identifying* these same preexisting representations, which serve to facilitate correct matching. The latter experience occurs when there are properties of a task that would change a subject's inductive biases (i.e., considering which of several stimulus features are relevant to the current problem). Inductive biases allow subjects to link their current concepts of sameness and difference to correct matching. Kroupin and Carey used Obozoba et al. (2015) and Smirnova et al. (2015) as examples of studies in which parrots and crows, respectively, were successful on RMTS testing following experience with *both* identity-match-to-sample (IMTS) training (i.e., the single item sample stimulus and its correct comparison stimulus are identical across all exemplars) and progressive alignment trials (e.g., similar to IMTS except each stimulus is composed of two items). Importantly, progressive alignment can assist in producing conceptual representations. For instance, progressive alignment trials require comparisons of physical features, but these comparisons are made between pairs of items. These pairs may possess two dimensions, thus the subject must identify which comparison pair matches the sample pair along the relevant dimension. Although, the relevant dimension will be featurally identical between the pairs (e.g., blue colour), there is also relational similarity between the stimuli (i.e., same colour relationship). Hence, this process can facilitate the recognition of relational similarity, which may result in the production of representations of sameness or difference. Conversely, match-to-sample training can only effect preexisting representations. For instance, the MTS procedure only requires the comparison of physical features. Therefore, MTS trials do not require the subject to use representations of sameness or difference, as would be needed for relational matching, thus it does not produce such representations. Kroupin and Carey recognized that the success of parrots and crows could have been facilitated by the contribution of both experiences, but suggest that progressive alignment alone could not have guaranteed success in the task, as

even the rate of success for progressive alignment and RMTS trials during the first test session were relatively equal (i.e., if progressive alignment trials helped to facilitate success in RTMS trials, it would be expected that success in progressive alignment trials would be lower initially). Thus, Kroupin and Carey evaluated whether first-order MTS training was sufficient for success on a task in which adult humans' inductive biases would require shifting towards relational responding as the basis for matching. Participants were given either IMTS training or MTS training (i.e., trials in which only the relevant exemplar is identical between the sample and correct comparison stimuli) on one of three dimensions (colour, size, or number; see Figure 9 for an example). Only two types of MTS training (not IMTS) facilitated relational matching. This finding showed that the sole cause for this increase in relational responding was due to changes in inductive biases. During MTS problems, subjects must determine which exemplar is relevant thus acting on their representations of sameness and difference, however this is not the case in IMTS problems, as there is no distinguishing which exemplars are relevant when they are all the same between the sample- and correct choice stimulus. Kroupin and Carey replicated this study (with minor changes) using four-year-old children (a population that performs similarly to non-human animals on standard RMTS tasks). These children exhibited the same trend as the adults. These findings suggest that when evaluating relational reasoning via RMTS tasks, training procedures that effectively alter inductive biases can result in increased relational responding. Therefore, this process can help to promote an understanding of relational similarity. The current paradigm's training procedure was not adequate for teaching subjects to recognize relational comparisons and thus may benefit from incorporating a component that can alter inductive biases. However, Kroupin and Carey's task utilizes a match-the-relation style problem, whereas the current paradigm using a complete-the-relation style problem. This poses challenges to incorporating this component, as transitioning from a two-choice matching structure to a A:A′::B:B′ structure during training would change how stimulus comparisons are understood and may negatively affect learning. Additionally, the properties of MTS problems that affect inductive biases may be based on distinguishing which of two or more dimensions are relevant among the sample and comparison stimuli. Whereas, stimuli could be composed of two dimensions during early shaping/training procedures, subjects may have difficulty transferring their inductive biases to *Dimension 1 Training* which presents stimuli composed of a single dimension for the first time. To design a suitable training procedure for the current paradigm,

future studies may need to determine if inductive biases can be changed using a complete-the-relation style problem, without the incorporation of first-order MTS problems. The current training procedure may not have induced changes to inductive biases, nor facilitated reasoning using relational similarity. An appropriate training procedure for this paradigm should serve to familiarize subjects with the comparative nature of the task by fostering the appropriate mental connections between the stimuli.

Figure 9. Examples of identity-match-to-sample (IMTS) and match-to-sample (MTS) trials. The sample and correct choice match along all exemplars in IMTS trials, but only along the relevant dimension, in this case colour, in MTS trials.



Training Procedure Elements

Correction Procedures. Some of the elements of the training procedure that were meant to serve as learning aids, or help to draw awareness to the task, did not prove helpful during this critical acquisition period. For instance, regardless of the number of repeated attempts to solve the presented problem, correction trials were not largely beneficial during training, perhaps due to difficulties learning the task requirements. Alternatively, the birds simply did not understand the correction trials. Scheper et al. (2019) tested memory recall performance by young adults using a spatial search task which required them to learn the locations of objects on a virtual grid and to commit their locations to memory. During training, the number of errors that could be made during trials was pre-set so that clicking on a grid cell only revealed an object after a

certain number of attempts (i.e., 0, 2, 3, 4, and 5) were made across the grid. Participants were then tested to determine how their learning process might have been effected by error frequency. The amount of errors made during task acquisition had no significant effects on recall accuracy. In a similar way, during an analogical reasoning task, seeing the same problem repeatedly may not improve present or future performance, especially if a subject has not yet learned the requirements to solve the problem. Therefore, correction trials may not necessarily stop subjects from either guessing or fixating incorrectly on a choice that never results in reinforcement—even after making many consecutive pecks to that choice. However, during the learning process, correction trials can still aid in correcting a failed attempt to accurately determine relational similarity. For instance, Blaisdell and Cook (2005) gave pigeons a two-item same/different discrimination task. During training in the first experiment, accuracy on shape trials declined compared to colour trials when previous colour-shape stimuli were later trained with these dimensions independent. During a second experiment, training modifications were incorporated into the procedures in attempt to improve accuracy on shape-only training trials prior to testing with novel shapes. Of these modifications (i.e., stimulus size reduction, correction procedures, and removal of colour-shape trials in favor of colour- and shape-only trials) incorporation of a correction procedure aided the majority of pigeons in raising their shape choice accuracy from near-chance levels in experiment 1 to approximately 80%. In the modified procedures of the current study, neither removal of the CPs nor having all stimuli appear simultaneously during correction trials facilitated the subjects to make the necessary comparisons between stimuli. A possible solution to optimize the effect of CPs would be to limit the number of CPs given per trial. Doing so will prevent the unnecessary overexposure of subjects to previously incorrect trials and increase exposure to new trials (i.e., new examples of analogy). Contingent on performance, this also creates the opportunity for increasing the frequency of reward for successful choice-making throughout a session.

Fixed Ratio. It has been suggested that fixed ratios of five may be too small to detect effects on performance, whereas greater increases in FRs may result in significant learning improvements (Sacks et al., 1972). Indeed, pigeons are shown to learn concepts at a faster rate with increased FR requirements for sample stimuli (Katz et al., 2008; Wright, 1997). Typically, in match-to-sample operant studies, pecks are made to a single sample stimulus followed by a peck to a choice stimulus. In the current study, there are three stimuli that must be pecked before

the choice stimuli appear. This was part of the initial rationale for not making each individual stimulus FR too high during training, as it would result in making a greater number of pecks during each individual trial. Instead, a higher FR requirement for a sample was divided amongst the primary and bridge stimuli so that their cumulative FR was high and to avoid decreased motivation due to the amount of effort to complete each trial. However, if repeating the current study, it should be expected that raising the FR requirements of the primary and bridge stimuli should improve task acquisition, as was observed with pigeons (Katz et al., 2008; Sacks et al., 1972; Wright, 1997). However, consideration should be taken to how high the FR should be set for each stimulus in order to avoid an overabundance of pecks per trial. Furthermore, there are pigeon discrimination tasks that include raising the FR of choice stimuli to 10 pecks, however pecks to the choice were not preceded by pecks to a sample (Blaisdell & Cook, 2005). Therefore, the FR for the choice stimulus in the current study could remain low given that the birds' performance did not improve when it was raised. Further evidence may be needed to understand the learning effects of raising the FR requirement of choice stimuli above 1-2 pecks during trials that require pecks to multiple stimuli.

Warning Stimulus. Following completion of their 87th session, two of the pinyon jays were introduced to a WS that appeared before the beginning of each new trial. This stimulus functions as a signal to enhance "attentional preparation" (Hackley & Valle-Inclan, 2003; Périn et al., 2010) and also served as an aid for more easily distinguishing the initiation of a new trial from a correction trial. Throughout the subsequent sessions, there was no significant rise in the number of correct trials by either bird. With the suspension of training (due to the Covid-19 pandemic) shortly after these sessions began, it is undeterminable what the prolonged effects of the WS on the percentage of correct trials might have been. Observing no initial effect during training may suggest that the WS would not be useful in the task as it cannot be linked to improved concept learning. Katz et al. (2018) assessed how procedural modifications (e.g., the inclusion of a WS) effected stimulus control over mice when given stimulant drugs during an attentional task. Mice were placed in an operant chamber with five equally spaced holes, one of which was randomly illuminated across trials by a brief signal stimulus. This stimulus indicated the correct choice. During testing, the mice were given trials both with and without a WS (i.e., the chamber overhead light was extinguished for 0.5 s) preceding the signal stimulus, because the WS was believed to alter signal detection by the mice while they were under the influence of

various stimulant drugs. The results showed that the effects of the drugs (either when increasing accuracy or having no effect on accuracy) didn't differ between the presence or absence of the WS. In fact, during the control condition (i.e., no drugs) accuracy decreased slightly with the addition of the WS. Most notably, a WS will minimize the need for constant vigilance when awaiting a signal. However, when considering accuracy performance, it appears that a WS will not necessarily have a profound or improved influence on the accuracy of a subject's signal response behaviour (Katz et al., 2018), nor perhaps its complete-the-relation matching performance, as shown in the current study. It is evident that the WS may not be essential to task acquisition, however incorporation of the WS at the beginning of training could better aid in the learning process. Currently, there is no cause to remove the WS from the training procedure as its presence would continue to serve the primary purpose of readying the birds for a distinctly new trial.

Learning Points and Contributions of the Current Study

I have learned that when determining if a task can be used as a valid, comparative measure of analogical reasoning ability, the procedures must teach necessary concepts in a way that is understandable to the species being evaluated. Therefore, it must be determined if there are aspects of a training procedure that may negatively affect the ease by which concepts are acquired. I have considered multiple aspects of my procedures (i.e., the current design and the modifications) that might have influenced performance. In regards to the modifications, I found that the pinyon jays did not respond to adjustments to procedural elements (e.g., CP, FR, or WS) as expected. Although intended to simplify the procedures, these modifications did not help the pinyon jays learn the task and thus it is perhaps best to avoid attempts to cultivate learning by implementing incremental shifts in procedures. Moreover, my Stage 2 shaping procedure was designed to progressively introduce the stimuli in the placement and order in which stimuli would later appear in *Dimension 1 Training*. Doing so was intended to show the birds that the primary stimuli are a pair and that the bridge stimulus required a match in order to create a relationally similar pair. The structure of these problems (A:A′::B:B′) appears to foster relational problem solving by humans (Leonard et al., in prep). However, it is possible that setting up Stage 2 shaping such that it introduced A, A′, B, and B′ in this progression did not facilitate relational learning by pinyon jays. This progression immediately introduced many ideas required to solve

the analogical problems. Starting by identifying that individual stimuli are composed of a dimension exemplar to progressively using the exemplar relationship between stimuli to make a between-pair comparison. These requirements for interpreting the stimuli perhaps created difficulty for acquiring a relational rule based strategy. For instance, the current procedures trained the birds using relational problems, rather than leading up to relational problems. Greater training success may occur if pinyon jays are given a procedure that familiarizes them with rudimentary steps necessary to construct the analogical problem. Perhaps these steps need to be more temporally isolated from one another, compared to the current procedures. Thus, each idea can be acquired individually and a preceding step can act as a scaffold for learning the following step.

My study was the first attempt at training pinyon jays using a new analogical paradigm providing complete-the-relation style problems. This problem is advantageous compared to the more commonly used match-the-relation style problems because the arrangement of the stimuli better reflect the presentation of an analogy. To my knowledge, my study was also the first attempt to train a non-human animal on complete-the-relation analogical problems without the use of tokens or prior language training. Although the analogical reasoning abilities of pinyon jays is inconclusive, my findings regarding the difficulties associated with my task provide a basis for future research into this question. Furthermore, this paradigm was designed for (1) the purpose of comparative investigations into analogical reasoning abilities across species, and (2) to minimize the risk of subjects developing a perceptual processing strategy, rather than an analogical reasoning strategy. However, my study showed that if this paradigm is to be suitable for evaluating the analogical reasoning abilities of pinyon jays, future researchers must adapt the procedures such that they are adequate for promoting relational reasoning. Moreover, the failure of my task demonstrates a need to explore training procedures that can successfully teach non-human animals how to solve complete-the-relation analogical problems. First, by incorporating adequate procedures for teaching the birds how the stimuli are related. Second, by avoiding the exact progression and arrangement of stimuli that I used during training. Finally, by facilitating the use of relational concepts, such as sameness and difference. For instance, this paradigm may benefit from training procedures that change inductive biases, such that subjects may infer how concepts of sameness and difference are necessary for the task.

Conclusions

The objective of this experiment was to test a new paradigm that was intended to provide an accurate assessment of the analogical reasoning abilities of pinyon jays. There is a demand in the field of comparative cognition to develop new methods to investigate analogy and relational learning across species. The learning difficulties of the pinyon jays in the current task provide indication of the importance of carefully regulating task procedures when using the same paradigm across different species. Although methods should be as similar as possible, there are multiple factors that should be considered for the species in question when adapting the operant task for their use. Specifically, how the placement of stimuli on the screen will affect whether a subject recognizes the relational similarity between them, whether or not training should begin with simpler problems that work to alter inductive biases and facilitate relational learning, and whether elements of the training procedure need adjustment between species. The latter includes limiting the number of correction trials per problem, beginning to raise stimulus FRs during the start of training, and including a WS at the start of training. Importantly, this study is useful for demonstrating that producing comparative paradigms requires more than an appropriate test of analogical reasoning, but also an experimental procedure that adequately promotes learning across different species. It is clear that careful and moderate adjustment of a procedure must include consideration of how a particular species is known to processes visual information and build mental connections between stimuli. The current paradigm offers a non-verbal, purely visual task suitable for testing multiple species. However, even with the added advantage of using an apparatus and same/different concepts that the pinyon jays were familiar with, they still had difficulty understanding what the task required. The current study showed that learning by pinyon jays is not facilitated by my procedure. However, my contribution to the field is the broadening of our understanding that further procedural considerations must be explored if the current paradigm or complete-the-relation problems are to be used for comparative investigations of analogical reasoning. Specifically, future studies that aim to create a cross-species approach to investigating analogical reasoning should consider whether the relational problems of the task are indeed appropriate, relevant, and presented such that analogical reasoning can be learned regardless of which species is the subject.

Future Directions

Future applications of the adapted experimental design should include as assessment of how multiple species respond to analogical relationships that incorporate higher degrees of incongruence between dimensions. For instance, relationships that include two (or more) relevant dimensions are not equally challenging. The same/different relationships that arise from combinations of two dimensions vary (see Table 3). The relationships within a pair of stimuli are considered the least complex when there is congruency in their dimensions, whereas incongruence is considered more complex, thus making incongruent relationships harder to disambiguate. Interpreting relationship congruency was first illustrated with one relevant dimension in Figure 3a and 3b, where both *same* and *different* relationships are presented. When considering trials with two relevant dimensions, understanding relationships requires more interpretation. For instance, relationships where both dimensions are congruent between stimuli (e.g., same shape, same colour) would be considered less challenging to comprehend than relationships where one dimension between the stimuli is incongruent (e.g., same shape, different colour or same colour, different shape). The most challenging relationship would be when two of three dimensions differ between stimuli (e.g., different shape, same colour, different pattern). Prior to the start of my study, I expected performance would decline as the number of incongruent dimensions increased among relationships that incorporated more than one dimension, as recognizing a correct choice should become more difficult when one dimension or more are incongruent than when dimensions are all congruent. This exact case was demonstrated using the current paradigm by human participants who advanced to trials that contained three dimensions. Humans performed at their highest accuracies when no dimensions differed in trials, and their accuracies progressively declined as more dimensions differed. They performed at their lowest accuracy when two dimensions differed, but interestingly their accuracies were equivalent when either one or three dimensions differed (Leonard et al., in prep.). The latter finding might suggest that among relationships with multiple (three) dimensions, an entirely incongruent relationship does not equate in difficulty to an entirely congruent relationship, but instead was shown to match in difficultly to a relationship that was only slightly less congruent (i.e., one of three dimensions differ). It would be informative to know whether this trend is consistent across different species, and if not, how species would perform when given relationships with three dimensions of varying congruency.

Literature Cited

Akagi, M. (2018). Rethinking the problem of cognition. *Synthese*, *195*(8), 3547-3570.

Arantes, J., & Berg, M. E. (2009). Intermodal transfer from a visual to an auditory discrimination using an errorless learning procedure. *Behavioural Processes*, *81*(2), 303-308.

Arantes, J., & Machado, A. (2011). Errorless learning of a conditional temporal discrimination. *Journal of the Experimental Analysis of Behavior*, *95*(1), 1-20.

Balda, R. P., & Kamil, A. C. (1989). A comparative study of cache recovery by three corvid species. *Animal Behaviour*, *38*(3), 486-495.

Barrett, L., Henzi, P., & Dunbar, R. (2003). Primate cognition: from 'what now?' to 'what if?'. *Trends in Cognitive Sciences*, *7*(11), 494-497.

Beck, J. (2012). Do animals engage in conceptual thought?. *Philosophy Compass*, *7*(3), 218-229.

Bednekoff, P. A., & Balda, R. P. (1996). Social caching and observational spatial memory in pinyon jays. *Behaviour*, *133*(11-12), 807-826.

Beran, M. J., Parrish, A. E., Perdue, B. M., & Washburn, D. A. (2014). Comparative cognition: past, present, and future. *International Journal of Comparative Psychology*, *27*, 3-30.

Blaisdell, A. P., & Cook, R. G. (2005). Two-item same-different concept learning in pigeons. *Animal Learning & Behavior*, *33*(1), 67-77.

Bovet, D., & Vauclair, J. (2001). Judgment of conceptual identity in monkeys. *Psychonomic Bulletin & Review*, *8*(3), 470-475.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 625-641.

Catania, A. C. (1984). The operant behaviorism of BF Skinner. *Behavioral and Brain Sciences*, *7*(4), 473-475.

Clary, D., & Kelly, D. M. (2011). Cache protection strategies of a non-social food-caching corvid, Clark's nutcracker (*Nucifraga columbiana*). *Animal Cognition*, *14*(5), 735-744.

Clary, D., & Kelly, D. M. (2016a). Graded mirror self-recognition by Clark's nutcrackers. *Scientific Reports, 6,* 36459.

Clary, D., & Kelly, D. M. (2016b). Clark's Nutcrackers (*Nucifraga columbiana*) Flexibly Adapt Caching Behavior to a Cooperative Context. *Frontiers in Psychology*, *7*, 1643.

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature, 395*(6699), 272-274.

Clayton, N. S., & Emery, N. J. (2007). The social life of corvids. *Current Biology*, *17*(16), R652-R656.

Clayton, N., Emery, N., & Dickinson, A. (2006). The rationality of animal memory: Complex caching strategies of western scrub jays. In S. Hurley & M. Nudds (Eds.). *Rational Animals?* (pp. 197-216). London: Oxford University Press.

Clayton, N. S., & Russell, J. (2009). Looking for episodic memory in animals and young children: Prospects for a new minimalism. *Neuropsychologia*, *47*(11), 2330-2340.

Cook, R. G. (2002). Same-different concept formation in pigeons. In M. Beckoff, C. Allen, & G. M. Burghart (Eds.), *The cognitive animal* (pp. 229-237). Cambridge, MA: MIT Press.

Cook, R. G., Katz, J. S., & Cavoto, B. R. (1997). Pigeon same–different concept learning with multiple stimulus classes. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*(4), 417-433.

Dally, J. M., Emery, N. J., & Clayton, N. S. (2005). Cache protection strategies by western scrub-jays, Aphelocoma californica: implications for social cognition. *Animal Behaviour*, *70*(6), 1251-1263.

Dere, E., Kart-Teke, E., Huston, J. P., & Silva, M. D. S. (2006). The case for episodic memory in animals. *Neuroscience & Biobehavioral Reviews*, *30*(8), 1206-1224.

Dumas, D., Alexander, P. A., & Grossnickle, E. M. (2013). Relational reasoning and its manifestations in the educational context: A systematic review of the literature. *Educational Psychology Review*, *25*(3), 391-427.

Emery, N. J., & Clayton, N. S. (2001). Effects of experience and social context on prospective caching strategies by scrub jays. *Nature*, *414*(6862), 443.

Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science*, *306*(5703), 1903-1907.

Fagan, W. F., Lewis, M. A., Auger-Méthé, M., Avgar, T., Benhamou, S., Breed, G., LaDage, L., Schlägel, U. E., Tang, W., Papastamatiou, Y. P., Forester, J., & Mueller, T. (2013). Spatial memory and animal movement. *Ecology Letters*, *16*(10), 1316-1329.

Fagot, J., & Parron, C. (2010). Relational matching in baboons (*Papio papio*) with reduced grouping requirements. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*(2), 184-193.

Flemming, T. M., Beran, M. J., Thompson, R. K., Kleider, H. M., & Washburn, D. A. (2008). What meaning means for same and different: Analogical reasoning in humans (*Homo sapiens*), chimpanzees (*Pan troglodytes*), and rhesus monkeys (*Macaca mulatta*). *Journal of Comparative Psychology*, *122*(2), 176-185.

Flemming, T. M., & Kennedy, E. H. (2011). Chimpanzee (*Pan troglodytes*) relational matching: Playing by their own (analogical) rules. *Journal of Comparative Psychology*, *125*(2), 207-215.

Flemming, T. M., Thompson, R. K., & Fagot, J. (2013). Baboons, like humans, solve analogy by categorical abstraction of relations. *Animal cognition*, *16*(3), 519-524.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155-170.

Gentner, D., & Holyoak, K. J. (1997). Reasoning and learning by analogy: Introduction. *American Psychologist*, *52*(1), 32.

Gentner, D. & Maravilla, F. (2018). Analogical reasoning. L. J. Ball & V. A. Thompson (Eds.) *International Handbook of Thinking & Reasoning* (pp. 186-203). NY, NY: Psychology Press.

Gillan, D. J., Premack, D., & Woodruff, G. (1981). Reasoning in the chimpanzee: I. Analogical reasoning. Journal of Experimental Psychology: *Animal Behavior Processes*, *7*(1), 1-17.

Goto, K., & Watanabe, S. (2009). Visual working memory of jungle crows (*Corvus macrorhynchos*) in operant delayed matching-to-sample 1. *Japanese Psychological Research*, *51*(3), 122-131.

Hackley, S. A., & Valle-Inclán, F. (2003). Which stages of processing are speeded by a warning signal?. *Biological psychology*, *64*, 27-45.

Håstad, O., & Ödeen, A. (2014). A vision physiological estimation of ultraviolet window marking visibility to birds. *PeerJ*, *2*, e621.

Håstad, O., Victorsson, J., & Ödeen, A. (2005). Differences in color vision make passerines less conspicuous in the eyes of their predators. *Proceedings of the National Academy of Sciences*, *102*(18), 6391-6394.

Healy, S. D. (2019). The face of animal cognition. *Integrative Zoology*, *14*(2), 132-144.

Katz, J. S., Bodily, K. D., & Wright, A. A. (2008). Learning strategies in matching to sample: If-then and configural learning by pigeons. *Behavioural processes*, *77*(2), 223-230.

Katz, J. L., Kohut, S. J., & Soto, P. (2018). Vigilance Demand and the Effects of Stimulant Drugs in a 5-Choice Reaction-Time Procedure in Mice. *Behavioural Pharmacology*, *29*(8), 701-708.

Katz, J. S., & Wright, A. A. (2006). Same/different abstract-concept learning by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*(1), 80-86.

Katz, J. S., Wright, A. A., & Bachevalier, J. (2002). Mechanisms of same/different abstract-concept learning by rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, *28*, 358–368.

Kelly, D. M., Bischof, W. F., Wong-Wylie, D. R., & Spetch, M. L. (2001). Detection of glass patterns by pigeons and humans: Implications for differences in higher-level processing. *Psychological Science*, *12*(4), 338-342.

Kennedy, E. H., & Fragaszy, D. M. (2008). Analogical reasoning in a capuchin monkey (*Cebus apella*). *Journal of Comparative Psychology*, *122*(2), 167-175.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, *67*(6), 2797-2822.

Kroupin, I., & Carey, S. (2021). Population differences in performance on Relational Match to Sample (RMTS) sometimes reflect differences in inductive biases alone. Current Opinion in *Behavioral Sciences*, *37*, 75-83.

Leonard, K., Sepehri, P., Cheri, B., & Kelly, D. M. (in prep). Relational complexity influences analogical reasoning ability.

MacDonald, S. E., & Ritvo, S. (2016). Comparative cognition outside the laboratory. *Comparative Cognition & Behavior Reviews*, *11*, 49-61.

Magnotti, J. F., Katz, J. S., Wright, A. A., & Kelly, D. M. (2015). Superior abstract-concept learning by Clark's nutcrackers (*Nucifraga columbiana*). *Biology Letters*, *11*(5), 20150148.

Magnotti, J. F., Wright, A. A., Leonard, K., Katz, J. S., & Kelly, D. M. (2017). Abstract-concept learning in Black-billed magpies (*Pica hudsonia*). *Psychonomic Bulletin & Review*, *24*(2), 431-435.

Masek, A., & Yamin, S. (2011). The effect of problem based learning on critical thinking ability: a theoretical and empirical review. *International Review of Social Sciences and Humanities*, *2*(1), 215-221.

Obozova, T., Smirnova, A., Zorina, Z., & Wasserman, E. (2015). Analogical reasoning in amazons. *Animal Cognition*, *18*(6), 1363-1371.

Ödeen, A., & Håstad, O. (2003). Complex distribution of avian color vision systems revealed by sequencing the SWS1 opsin from total DNA. *Molecular Biology and Evolution*, *20*(6), 855-861.

Paz-y-Miño C. G., Bond, A. B., Kamil, A. C., & Balda, R. P. (2004). Pinyon jays use transitive inference to predict social dominance. *Nature*, *430*(7001), 778-781.

Pepperberg, I. M. (1987). Acquisition of the same/different concept by an African Grey parrot (*Psittacus erithacus*): Learning with respect to categories of color, shape, and material. *Animal Learning & Behavior*, *15*(4), 423-432.

Pepperberg, I. M. (2013). Abstract concepts: data from a Grey parrot. *Behavioural Processes*, *93*, 82-90.

Périn, B., Godefroy, O., Fall, S., & de Marco, G. (2010). Alertness in young healthy subjects: an fMRI study of brain region interactivity enhanced by a warning signal. *Brain and Cognition*, *72*(2), 271-281.

Premack, D. (1988). Minds with and without language. In L. Weiskrantz (Ed.). *Thought without language* (pp. 46-65). New York: Oxford University Press.

Qadri, M. A., Leonard, K., Cook, R. G., & Kelly, D. M. (2018). Examination of long-term visual memorization capacity in the Clark's nutcracker (*Nucifraga columbiana*). *Psychonomic Bulletin & Review*, *25*(6), 2274-2280.

Raby, C. R., Alexis, D. M., Dickinson, A., & Clayton, N. S. (2007). Planning for the future by western scrub-jays. *Nature*, *445*(7130), 919.

Raver, C. C. & Blair, C. (2016). Neuroscientic insights: Attention, working memory, and inhibitory control. *Future of Children*, *26*(2), 95-118.

Roitblat, H. L. (1980). Codes and coding processes in pigeon short-term memory. *Animal Learning & Behavior*, *8*(3), 341-351.

Sacks, R. A., Kamil, A. C., & Mack, R. (1972). The effects of fixed-ratio sample requirements on matching to sample in the pigeon. *Psychonomic Science*, *26*(5), 291-292.

Scheper, I., de Bruijn, E. R., Bertens, D., Kessels, R. P., & Brazil, I. A. (2019). The impact of error frequency on errorless and errorful learning of object locations using a novel paradigm. *Memory*, *27*(10), 1371-1380.

Shettleworth, S. J. (2001). Animal cognition and animal behaviour. *Animal behaviour*, *61*(2), 277-286.

Smirnova, A., Zorina, Z., Obozova, T., & Wasserman, E. (2015). Crows spontaneously exhibit analogical reasoning. *Current Biology*, *25*(2), 256-260.

Snowdon, C. T. (1990). Language capacities of nonhuman animals. *American Journal of Physical Anthropology*, *33*(S11), 215-243.

Stout, D. (2010). The evolution of cognitive control. *Topics in Cognitive Science*, *2*(4), 614-630.

Sutton, J. E., & Shettleworth, S. J. (2008). Memory without awareness: Pigeons do not show metamemory in delayed matching to sample. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*(2), 266.

Tebbich, S., Griffin, A. S., Peschl, M. F., & Sterelny, K. (2016). From mechanisms to function: an integrated framework of animal innovation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1690), 20150195.

Terrace, H. S. (1963). Discrimination learning with and without "errors" 1. *Journal of the experimental analysis of behavior*, *6*(1), 1-27.

Thompson, R. K., & Oden, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science*, *24*(3), 363-396.

Thompson, R. K., Oden, D. L., & Boysen, S. T. (1997). Language-naive chimpanzees (*Pan troglodytes*) judge relations between relations in a conceptual matching-to-sample task. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*(1), 31-49.

Thornton, A., & Lukas, D. (2012). Individual variation in cognitive performance: developmental and evolutionary perspectives. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1603), 2773-2783.

Vernouillet, A., Casidsid, H., & Kelly. D.M. (2020) Conspecific presence, but not pilferage, influences pinyon jays' (*Gymnorhinus cyanocephalus*) caching behavior. *Learning & Behavior*.

Vernouillet, A., Clary, D., & Kelly, D.M. (in prep-a). Highly social corvids, but not non-social corvids, can determine heterospecific individuals as threat for their food stores.

Vernouillet, A., Leonard, K., Katz, J.S., Magnotti, J.M., Wright, A., & Kelly, D.M. (in press). Abstract-concept learning in two species of New World corvids, pinyon jays (*Gymnorhinus cyanocephalus*) and California scrub jays (*Aphelocoma californica*). *Journal of Experimental Psychology: Animal Learning & Cognition*.

Vernouillet, A., Huang, N., & Kelly, D.M. (in prep-b). Identifying an observing individual as a threat to caches in two corvid species.

Vernouillet, A., & Kelly, D.M. (2020). Individual exploratory responses are not repeatable across time or context for four species of food-storing corvid. *Scientific Reports, 10*, 394.

Vonk, J. (2003). Gorilla (*Gorilla gorilla gorilla*) and orangutan (*Pongo abelii*) understanding of first-and second-order relations. *Animal Cognition, 6*(2), 77-86.

Vonk, J. (2015). Corvid cognition: Something to crow about? *Current Biology, 25*(2), R69-R71.

Vonk, J., & MacDonald, S. E. (2002). Natural concepts in a juvenile gorilla (*Gorilla gorilla gorilla*) at three levels of abstraction. *Journal of the Experimental Analysis of Behavior*, *78*(3), 315-332.

Vonk, J., & MacDonald, S. E. (2004). Levels of abstraction in orangutan (*Pongo abelii*) categorization. *Journal of Comparative Psychology*, *118*(1), 3.

Williams, B. A. (1971). The effects of intertrial interval on discrimination reversal learning in the pigeon. *Psychonomic Science*, *23*(3), 241-243.

Wright, A. A. (1997). Concept learning and learning strategies. *Psychological Science, 8*(2), 119-123.

Wright, A. A., & Katz, J. S. (2006). Mechanisms of same/different concept learning in primates and avians. *Behavioural Processes, 72*(3), 234-254.

Wright, A. A., Magnotti, J. F., Katz, J. S., Leonard, K., & Kelly, D. M. (2016). Concept learning set-size functions for Clark's nutcrackers. *Journal of the Experimental Analysis of Behavior*, *105*(1), 76-84.

Wright, A. A., Magnotti, J. F., Katz, J. S., Leonard, K., Vernouillet, A., & Kelly, D. M. (2017). Corvids outperform pigeons and primates in learning a basic concept. *Psychological Science*, *28*(4), 437-444.

Wright, A. A., Rivera, J. J., Katz, J. S., & Bachevalier, J. (2003). Abstract concept learning and list-memory processing by capuchin and rhesus monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*, *29*, 184–198.

Zentall, T. (2018). The value of research in comparative cognition. *International Journal of Comparative Psychology*, *31*, 1–17.