Friends or food: comparing social and food-based numerical cognition in captive pinyon jays

London M. Wolff[1]

[1] Department of Psychology, Center for Brain, Biology & Behavior, University of Nebraska, Lincoln, Lincoln, NE, USA

Abstract

Animals must often discriminate different quantities of objects in their environment, from numbers of food items to conspecifics. Yet we know little about how this numerical cognition compares across different types of objects. Based on past research, we would expect individuals to use both numerical ratio and numerical difference to choose between two numerical options. This study investigates whether numerical ratio and difference predict numerical preference in pinyon jays (*Gymnorhinus cyanocephalus*) for two types of stimuli, food items and conspecifics. Subjects (N=8 for food condition, N=10 for social condition) chose between two options using every paired combination of food item or group size number between 1 and 6. In both conditions, the pinyon jays showed an overall preference for the larger option over the smaller option. For the food condition, pinyon jays preferred numbers of items with higher numerical differences and lower numerical ratios. However, numerical difference did not influence preference independently ratio. For the social condition, neither difference nor ratio predicted the subjects' choices. Thus, the pinyon jays showed different effects of numerical difference and ratio on preference across food items and conspecifics. One rationale for these results are pinyon jays use different strategies when deciding between numbers of food items and flock mates. While number is important for selecting food items, other factors such as flock mate identity may be more important for selecting social groups to join. Thus, in numerical preference situations, the type of objects offered drive the numerical strategies that animals use.

Keywords: Numerical preference, Social cognition, Avian cognition
Word count: 7336

Introduction

When a hungry pinyon jay (*Gymnorhinus cyanocephalus*) fly's up to a grove of pine trees, should they choose the tree with just a few pine cones or the tree with several dozen? Later in the season when they are looking for a mate, should they join a flock of 10 birds or 50? Animals often must make choices between different numbers of options, and this occurs across a wide range of contexts, including food and social partners. Yet little is known about how corvids use numerical information outside of the foraging domain.

Many animal species have demonstrated the ability to quantify objects in their environment, including bees (Dacke & Srinivasan, 2008), fish (Agrillo & Dadda, 2007; Agrillo et al., 2008, 2011), amphibians (Uller et al., 2003), birds (Xia et al., 2001; Emmerton & Renner, 2006, 2009), and mammals (Vonk & Beran, 2012) but especially primates (Call, 2000; Beran, 2001; Nieder, 2018). As these animals are not closely related phylogenetically, some diverging hundreds of millions of years apart from each other (Nieder, 2018), quantification skills must have strong adaptive value that incur fitness benefits. These adaptive benefits can be broken down into two main subgroups: benefits for survival and reproduction. Survival benefits include but are not limited to navigation, predator avoiding, territory defense, and cooperative hunting/foraging (Yang & Chiao, 2016; Agrillo et al., 2017; Nieder, 2020). Reproduction benefits of numerical competence include courtship, mating, and avoiding brood parasitism (Arak, 1983; White et al., 2009; Carazo et al., 2012). For example, wolves are more likely to capture a buffalo, if they cooperatively hunt in groups of 9-13 (MacNulty et al., 2014). Having some form of numeric competence must be an important cognitive skill that incurs fitness benefits for survival and reproduction.

Two distinct cognitive systems, the object tracking system (OTS) and the approximate number system (ANS), are thought to describe how animals represent numbers (Hyde, 2011; Ditz & Nieder, 2016). The object tracking system (OTS) is a visual mechanism that tracks up to four objects precisely while the approximate number system (ANS) is a mental system that supports the estimation of numerical quantity without relying on language or symbols (Nieder, 2020). The most important difference to highlight is that the OTS works on small (3 or 4) but precise numbers while the ANS works on approximate quantities of larger sizes (Feigenson et al., 2004; Nieder, 2020). Current literature demonstrates that animals will use the absolute number of items and therefore the OTS when assessing numerical stimuli in some cases(Cantlon & Brannon, 2006; Emmerton & Renner, 2006; Dacke & Srinivasan, 2008). While animals use other cues such as surface area covered (Gómez-Laplaza & Gerlai, 2013; Stancher et al., 2015), density (Bertamini et al., 2018; Gómez-Laplaza et al., 2019), or size of items (Stevens et al., 2007; Emmerton & Renner, 2009) instead, of counting each object individually to estimate approximate quantities via the ANS.

Within the ANS, estimates are approximated through the use of numerical ratios and

London M. Wolff, ⬤ https://orcid.org/0000-0001-8359-2619.

Correspondence concerning this article should be addressed to London M. Wolff, B83 East Stadium, University of Nebraska, Lincoln, Lincoln, NE, USA 68588. ORCID 0000-0003-2375-1360. E-mail: lmwolff3@ gmail.com

differences through the *numerical magnitude effect* and the *numerical distance effect* (Dehaene et al., 1998; Ditz & Nieder, 2016). A numerical ratio is the mathematical quotient between two numbers: 2/4 has a ratio of 0.5. A numerical difference is the mathematical difference between two numbers: 4-2 has a difference of 2. The *numerical magnitude effect* asserts that, at a given numerical difference, discrimination worsens with increasing magnitude, which is equivalent to a decreasing numerical ratio. Essentially, it becomes harder to discriminate as the numerical ratio approaches 1. The *numerical distance effect* asserts that number discrimination improves with increasing numerical difference between two values. Essentially, it becomes easier to discriminate as the options become more dissimilar. Taken together, these two effects describe Weber's Law (Nieder, 2020). Quantity discriminations that follow Weber's law are a clear signature of the internal ANS and data showing difference and ratio effects illustrate the use of approximate amounts rather than precise numbers.

Past animal research investigating numerical cognition has found evidence that number is a relevant stimulus cue to which animals are sensitive across a number of contexts (Agrillo & Beran, 2013; Agrillo & Bisazza, 2014). Most of these numerical discrimination tasks use food or objects as stimuli (Call, 2000; Beran, 2001; Scarf et al., 2011; Rugani et al., 2013; Kelly, 2016). In previous studies, accuracy decreased as the ratio between the values approached 1, a direct confirmation of the *numerical magnitude effect* (Cantlon & Brannon, 2006; Hanus & Call, 2007; Evans et al., 2009; Merten & Nieder, 2009; Ditz & Nieder, 2016). The second half of Weber's Law has been confirmed showing animals discriminate between a pair of numbers with more accuracy at larger numerical differences, *numerical distance effect*, (Ditz & Nieder, 2015, 2016; Tornick et al., 2015; Kelly, 2016). However, a subset of studies looking at fish use conspecifics (members of the same species) to assess numerical cognition (Buckingham et al., 2007; Agrillo et al., 2008; Gómez-Laplaza & Gerlai, 2016). Many species prefer to be in larger groups, presumably because this dilutes their probability of being captured by predators (Krause & Ruxton, 2002; Silk et al., 2014). Though some studies show an effect of both difference and ratio on social preference (Agrillo & Dadda, 2007; Agrillo et al., 2008), others only show an effect of ratio (Buckingham et al., 2007; Gómez-Laplaza & Gerlai, 2011). In all four of these conspecific item studies, subjects were better at discriminating ratios above one half than below. This is in accordance with the numerical magnitude effect, with only one of the four providing evidence that fish were capable of discriminating ratios below 1:2. Research shows ratio and difference effects in animals across contexts, yet no one has looked at context effects on numerical preference within a single test population of any one species.

Pinyon jays are well-suited for examining effects of context on numerical cognition because of their diet and social habits. Pinyon jays are a highly social species of North American corvid (Balda & Kamil, 1998). They live in flocks ranging from 50 to 500 birds with fission-fusion dynamics in which members of a community form frequently changing subgroups (Balda & Kamil, 1998; Wiggins, 2005; Lehmann et al., 2007). Fission-fusion group living reduces predation risk and improves foraging success (Lehmann et al., 2007; Dange et al., 2021). This is relevant when thinking about number as birds in fission-fusion groups must often choose between breaking off into a smaller sub-group or rejoining the larger colony. One of the largest motivators of this decision-making process is foraging

benefit (Silk et al., 2014). Pinyon jays forage for protein-rich pine nuts, which they cache to retrieve in the winter. 70 to 90% of their winter diet is made up of cached seeds, even nestlings eat 10 to 32% pine seeds (Balda & Kamil, 1998). The need to retrieve cached food sources places strong selection pressure on numerical cognition, as they need to store as many pine seeds as possible to survive the winter. Pinyon jays rely on the quantity of food available for foraging decisions and the number of birds in a flock for social living decisions.

Present Study

The primary aim of the present study was to investigate how pinyon jays use numerical information, specifically numerical difference and ratio, to choose between different quantities of food items or conspecifics. To address this aim, we offered pinyon jays a series of choices between smaller and larger numbers of items: either high calorie treats (mealworms) or conspecifics.

To test our research question, we tested three hypotheses. Our first hypothesis (H1) posits that pinyon jays will, on average, prefer larger over smaller numbers of food items and conspecifics. An animal is more likely to survive, increasing their fitness outcome, if they consume more food. Similarity, an animal is more likely to survive by living in larger rather than smaller groups because they are decreasing the chance they will be killed by a predator. We hypothesize that pinyon jays would choose the larger flock as it increases your chances of finding a suitable mate and avoiding predators. Our second hypothesis (H2) posits that pinyon jays will show strong preferences for items with higher numerical differences and lower numerical ratios. As differences and ratios approach 1, the discrimination difficulty decreases. Discriminating between five and six mealworms is more difficult than discrimination between one and six. Our third hypothesis (H3) posits that both numerical difference and ratio will influence preference independently of each other. This distinction is important because difference and ratio are highly correlated with each other: as difference increases, ratio decreases. It is important to understand whether pinyon jays use difference, ratio, or both in a preference task to better understand how corvids make decisions across foraging and social domains.

Methods

Subjects

Our study population of 20 Pinyon Jays (*Gymnorhinus cyanocephalus*) were wild born and locally housed. Researchers captured these birds in either Arizona or California (United States Fish and Wildlife permit MB694205) between 2006 and 2011. At capture, they were estimated to be between one and three years of age. The colony has an age range of 12 and 17 years with a mean of 14.45years . The University of Nebraska-Lincoln Institutional Animal Care and Use Committee approved this project (protocol number 1867 and 2059), and all procedures conformed to the ASAB/ABS Guidelines for the use of animals in research. All subjects have completed prior cognitive and behavioral experiments in their tenure with the lab. The colony is handled by humans extensively.

Eight Pinyon Jays (one female) completed all rounds of the food experiment and 10 Pinyon Jays (four female) completed all rounds of the social experiment. The Pinyon Jays in the food item experiment were housed two to a double cage, while the jays in the social experiment were individually housed. A further 17 Pinyon Jays (six female) from the colony were used as stooge conspecifics in the social experiment. This number includes six birds that had to be transferred out of use as stooge birds due to unrelated colony matters. Stooge birds were primarily housed socially and always in a separate room from the testing subjects.

Two Pinyon Jays were dropped from the social experiment due to unrelated health concerns.

Food Experiment

Apparatus. The apparatus for the food experiment included a bird cage (72 x 48 x 48 cm) abutting a plastic stand with sliding trays that contained mealworms (Figure 1). The stand was set at a 15 degree angle tilted toward the subject to facilitate viewing the number of mealworms. The stand had two pvc pipe lined channels the same width as our plexiglass trays. Each plexiglass tray had a standard petri dish placed half an inch away from the front of each tray. Mealworms were placed in the front two thirds of the petri dish so the subjects could easily reach the mealworms and they were evenly distributed across the available area. Within the cage were three perches. One large free standing perch stood in the middle of the cage, while two small perches attached to each side of the cage a few inches back from the front. Subjects started each trial perched on the large free standing perch, and they chose an option by landing on the smaller left or right perches. They then consumed the mealworms associated with their decision while the unchosen side was removed immediately.

Food Stimuli Habituation and Training. Prior to experimental sessions, all birds had to be habituated to the experimental room, cage, and stand and then experience training. Depending on the bird this took between 9-16 weeks as each bird was ran once a day between 11:00-13:00 CST. Subjects experienced four training phases in total; each aimed at teaching the bird a different piece of the paradigm. The first phase (rear cup habituation) habituated subjects to the apparatus. The second phase (front dish habituation) encouraged the birds to place their heads through holes at the front of the cage to consume mealworms from the dishes on the trays. The third phase (moving dish training) introduced the subjects to the fact that the dish on the unchosen side moved out of reach. The forth phase (mixed reward training) taught the subjects to quickly make a choice between zero or three mealworms, eat, and then set up for another trial in quick succession.

For rear cup habituation, the experimenter brought a subject from their home cage and released them into the test cage. Five mealworms were placed in each of the two feeding bowls into slots at the back of the cage. After three minutes, the experimenter counted the number of mealworms consumed in each food dish and returned the subject to its home cage.

For front dish habituation, the experimenter brought a subject from their home cage
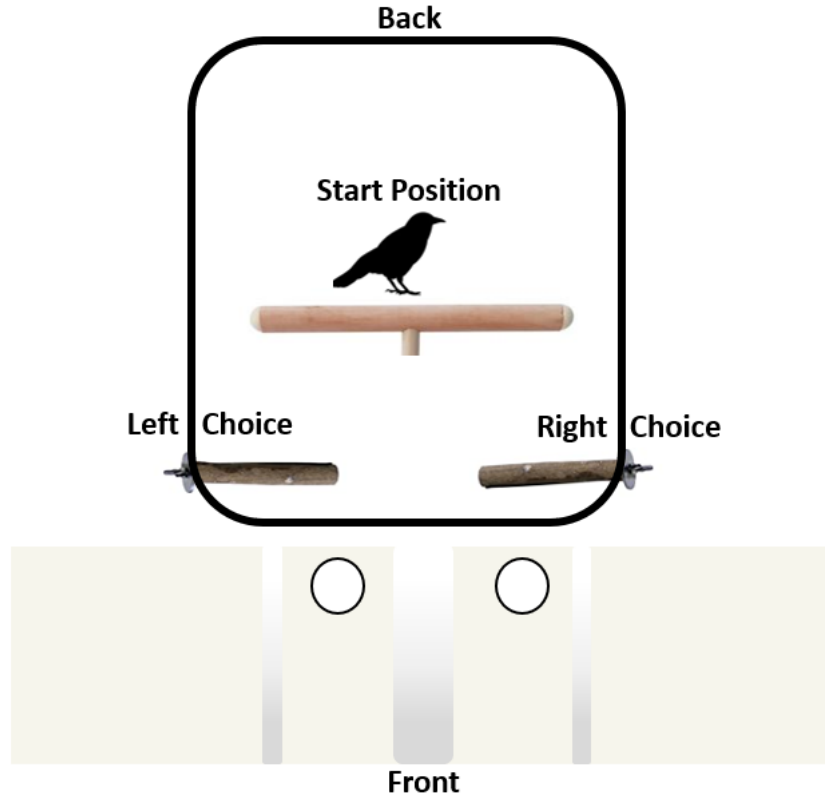
Figure 1. Food experiment apparatus (overhead view).

into the test cage. Three mealworms were placed in both dishes on the Plexiglass trays. The experimenter pushed the trays forward to present the dishes to the subject in one swift and smooth motion. After three minutes, the experimenter counted the number of mealworms eaten in each dish and returned the subject to their home cage.

Moving dish training was identical to front dish habituation, except the experimenter pulled back the tray in front of the unchosen perch. If subjects ate all three mealworms before three minutes expired, we repeated this process. If the bird did not finish their mealworms we waited the rest of the three minutes before returning them to their home cage.

The mixed reward training was identical to the moving dish training, except one of the dishes held no mealworms while the other held three mealworms and subjects completed six of these 30 second trials per session.

Subjects progressed to a new phase when they successfully consumed at least 70% of the mealworms offered in their current phase for three consecutive days. Subjects could also move to a previous phase if they consumed less than 25% of the mealworms offered on five out of seven days of training.

Experimental Procedure. All experimental sessions ran between 11:00-13:00 CST to keep the birds' hunger levels constant throughout the experiment. The subjects were not food restricted and were fed three hours prior to the start of the experiment. One

experimenter conducted each session. The first trial of the session consisted of one round of mixed reward training. If they failed this check, the experimenter completed two more rounds of mixed reward training. If they failed two out of three of these trials this triggers de-bias training. If they succeed, they continued to the experimental trials. For these trials, the experimenter placed the appropriate number of mealworms in each of the dishes, with mealworms placed 2.5 cm away from each other. The subject then started the trial on the back perch and hopped forward to one of the front perches to signal choice. The experimenter then removed the opposite dish and the subject had up to three minutes to consume their mealworms. If the subject did not make a choice and/or finish all mealworms within 3 minutes, we ended the session. Once the subject consumed all mealworms, we immediately started the next trial. The first trial of the day where the subject did not finish all their chosen mealworms triggered a stop on that days session. Subjects completed on average 4 trials per session.

Each bird experienced 10 repetitions for each of the 15 numerical pairs between 1 and 6 (e.g., 6 vs 5, 6 vs 4, 6 vs 3, etc.). The side of the larger option was pseudo randomized with no left or right runs longer than three in a row. We organized the pairs into 10 blocks. Each block was randomized within itself and had one instance of each factorial pair. The order the subjects ran in a particular day was also randomized.

Side-Bias Protocol. During habituation, three consecutive days of no choices and/or no eating of mealworms on one particular side triggered side de-bias training. During experimental sessions, there were two triggers for de-bias training: either not completing the first 0 vs. 3 practice test trial correctly or when a bird chose the same side for 10 consecutive trials.

For de-biasing training, we placed three mealworms in the dish associated with the side that the subject avoided and no mealworms in the side they preferred. We allowed the subject up to one minute to select the dish that contained mealworms. If the bird did not make a choice or eat any mealworms in the allotted time, we slide both trays all the way out of the stand, placed them out of view of the bird, waited a few seconds, then began the trial over. If they ate within the minute, we reset as soon as they ate until they had five total opportunities to eat. The bird returned to habituation or experimental sessions once they successfully chose the avoided side immediately and ate at least 60% of the mealworms provided.

Social Experiment

Apparatus. The apparatus (Figure 2) took the form of a Y maze formed out of chicken wire, plastic sheets, and Plexiglas. The subject entered a large chamber at the base of the maze before choosing one of two arms of the Y maze. At the entrance to both arms, a guillotine style door was closed after the bird walked or flew past it, thus making a choice between the option on the left or right. At the end of each arm, was a large bird cage housing the stooge birds. Each cage had two lengthwise perches for the stooge birds to use and one small perch hanging from the ceiling.
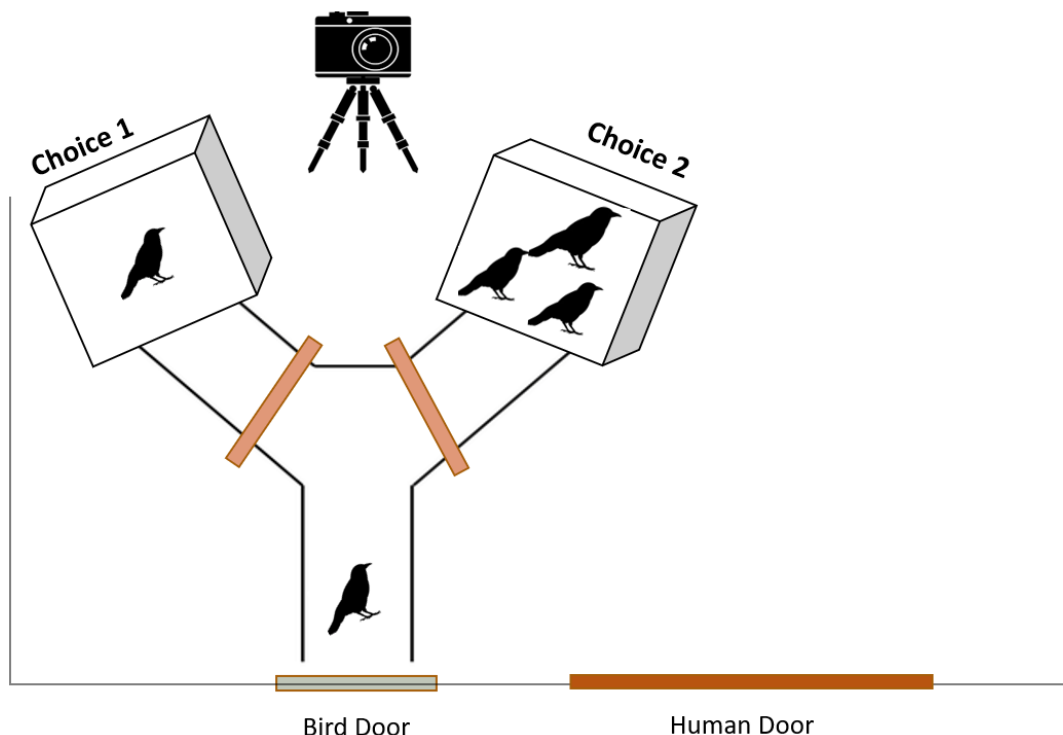
Figure 2. Social experiment apparatus (overhead view).

Habituation.   Prior to experimentation, we habituated all birds to both the exper-
imental room and the apparatus. For habituation, we attached a stainless steal cup to the
front of each bird cage. For a habituation session, the experimenter placed five mealworms
in each of the cups. The experimenter then brought the subject into the room, pulled up
both doors, and showed the subject each arm of the maze for six seconds, randomizing
between subjects the side shown first. The subject was then gently placed on the bottom of
the testing chamber as close to the center as possible with the bird facing away from their
options. Once the subject crossed the threshold of a door, the door was quietly and swiftly
closed behind them and the bird explored the chosen arm and consumed the mealworms for
two minutes. After the two minutes expired, the experimenter removed the subject from
the apparatus and recorded the number of mealworms consumed.

Subjects experienced 1 habituation session per day for 5 days a week. They completed
habituation once they consistently consumed at least 80% of the mealworms offered to them
in both arms of the apparatus and had no significant signs of a side bias. Depending on the
bird, this took between 4-6 weeks.

Experimental Procedure.   All experimental trials ran between 09:00-11:00 CST.
The subjects were not food restricted. Two experimenters were present at each session:
the 'handler' handled the subject, while the 'recorder' handled the camera, whiteboard,
and the guillotine doors. The experimenter placed stooge birds in their respective cages
and allowed them to acclimate to the room for 10 minutes before experimentation. The
handler then placed the subject inside the apparatus and showed them each option for six

seconds (counter balancing which was shown first) before releasing the subject into the chamber. Once the subject crossed the threshold of one of the doors, the recorder closed *both* doors gently but swiftly. After three minutes elapsed, the handler collected the subject and returned them to their home cage. These steps repeated until all birds had run through the experiment.

Each subject experienced 5 trials for each of the 21 numerical pairs between 0 and 6 (e.g., 6 vs 5, 6 vs 4, 6 vs 3, etc.). The side of the larger option was pseudo randomized with no left or right runs longer than 3 in a row. The pairs were organized into blocks with one instance of each pair per block and pairs randomized within each block. The order in which the subjects ran in a particular day was also randomized.

Side-Bias Protocol.   If any subject chose either the left or right side for six consecutive sessions in either habituation or experimentation, they experienced side de-biasing. For side de-biasing, only one door was open in the apparatus, the door leading to the side the subject avoided. We placed five mealworms in the food cup at the end of that arm with no stooge birds present. The bird had up to five minutes to walk/fly past the door into the correct side and three minutes once the door shut behind them to eat the mealworms. Subjects experienced five consecutive trials in a de-biasing session. The subject returned to habituation or experimental sessions once they successfully choose the avoided side immediately upon release and ate at least 60% of the mealworms provided.

Data Analysis

Data were analyzed and processed for the project using R [Version 4.1.0; R Core Team (2021)] and the R-packages *BayesFactor* [Version 0.9.12.4.2; Morey and Rouder (2018)], *bayestestR* [Version 0.11.5; Makowski et al. (2019)], *BMA* [Version 3.18.15; Raftery et al. (2021)], *broom* [Version 0.7.10; Robinson et al. (2021)], *coda* [Version 0.19.4; Plummer et al. (2006)], *dplyr* [Version 1.0.7; Wickham et al. (2021)], *forcats* [Version 0.5.1; Wickham (2021a)], *formattable* [Version 0.2.1; Ren and Russell (2021)], *ggplot2* [Version 3.3.5; Wickham (2016)], *here* [Version 1.0.1; Müller (2020)], *inline* [Version 0.3.19; Sklyar et al. (2021)], *knitr* [Version 1.36; Xie (2015)], *leaps* [Version 3.1; Fortran code by Alan Miller (2020)], *lme4* [Version 1.1.27.1; Bates et al. (2015)], *Matrix* [Version 1.3.4; Bates and Maechler (2021)], *papaja* [Version 0.1.0.9997; Aust and Barth (2020)], *patchwork* [Version 1.1.1; Pedersen (2020)], *performance* [Version 0.8.0; Lüdecke et al. (2021)], *purrr* [Version 0.3.4; Henry and Wickham (2020)], *readr* [Version 2.1.0; Wickham and Hester (2021)], *readxl* [Version 1.3.1; Wickham and Bryan (2019)], *robustbase* [Version 0.93.9; Todorov and Filzmoser (2009a)], *rrcov* [Version 1.6.0; Todorov and Filzmoser (2009b)], *stringr* [Version 1.4.0; Wickham (2019)], *survival* [Version 3.2.13; Terry M. Therneau and Patricia M. Grambsch (2000)], *tibble* [Version 3.1.6; Müller and Wickham (2021)], *tidyr* [Version 1.1.4; Wickham (2021b)], and *tidyverse* [Version 1.3.1; Wickham et al. (2019)]. Both the social and food stimuli data sets were analyzed using the same pre-registered analyses.

Throughout this paper, we will draw inferences based on Bayesian statistics (*BF* values), as opposed to frequentist stats (*p* values). While frequentist statistics are more prevalent in the current literature, we use Bayes statistics because they offer bidirectional information about both the alternative and the null hypotheses. That is, Bayes factors are

the ratio of evidence for H_1 over evidence for H_0 (Wagenmakers, 2007; Wagenmakers et al., 2010). Therefore, a Bayes factor of 3 indicates three times more evidence for H_1 than H_0, whereas a Bayes factor of 1/3 (the reciprocal of 3) indicates 3 times more evidence for H_0 than H_1. We interpreted Bayes factors based on Wagenmakers et al. (2018), where a *BF* > 3 is sufficient evidence for the alternative hypothesis, *BF* $< 1/3$ is sufficient evidence for the null hypothesis, and $1/3 < BF < 3$ indicate neither hypothesis has evidence supporting it (suggesting the sample size is too small to draw conclusions).

Prior to analysis, we transformed the left and right choice variable from each trial into a binary operator, with 1 representing a choice for the larger option and 0 representing a choice for the smaller option. We also created variables with the numerical difference between each number pair by subtracting the larger number from the smaller (6-1 = 5), as well as creating the ratio by dividing the smaller by the larger number (1/6 = 0.16). Our hypotheses explore the relationship between our binary outcome variable, choice of the larger or smaller stimuli, and which possible mechanism, difference or ratio, subjects use to make choices when presented with either food or social items.

Our first hypothesis investigated whether pinyon jays prefer larger over smaller numbers of food items and conspecifics. To test this, we conducted a one sample t-tests of preference for larger numbers. Therefore, we calculated the mean *percent preference for larger numbers* for each subject and used the t-test to compare the subject means to 50. We perform both frequentist and Bayesian t-tests, with inferences based on Bayes factors. Bayes factors for t-tests were calculated using the `ttestBF` function from the *BayesFactor* R package (Morey et al., 2021) using default, noninformative priors.

Our second hypothesis investigated whether numerical difference and ratio predict preferences between smaller and larger options and the third hypothesis investigated whether difference and ratio predicted preferences *independently*. To test these hypotheses, we used generalized linear mixed-effects modeling because the response variable was dichotomous and our subjects repeatedly made decision on the same number pairs. We used the trial-level choices for either the larger (coded as 1) or smaller (coded as 0) option available in the number pair as the response variable. To investigate our hypotheses, we used generalized logistic models to compare which combination of random (subject, pair, or both) and fixed (ratio, difference, or a combination of both) effects best describe each data set (food and social). We first found the best-fitting random effect structure, then added this random structure to all of the possible fixed effect structures, leaving us with the final best fitting model for each data set overall.

To explore random effect structure, we included models with no fixed effect and either (1) no random effects (intercept only), (2) subject as a random effect, (3) number pair as a random effect (to account for each bird repeatedly seeing each pair multiple times), and (4) both subject and number pair as random effects. For example, the model with both subject and pair as random effects ran using the `glmer()` function with the following structure: `r glmer(choice ~ (1|subject)+ (1|pair), family = binomial)` (Figure *A1* a). We then used Bayes factors to select the model with the best-fitting random effect structure. We added the chosen random effect structure to our fixed effects to find the best-fitting model for the data set overall. The five fixed effects models were: (1) no fixed

effects (intercept only), (2) ratio as a fixed effect, (3) difference as a fixed effect, (4) both difference and ratio as a fixed effects but *without* an interaction, and (5) both difference and ratio as fixed effects *with* an interaction. The model with both difference and ratio as fixed effects with an interaction term ran using the `glm()` function and the following structure: `r glm(choice ~ difference * ratio, family = binomial)` (Figure *A1* b). We calculated Bayes factors using the `test_performance()` function from the *performance* package (*Performance*, 2021), which estimates Bayes factors from model BIC values using Wagenmakers' (Wagenmakers, 2007) equation. The best fitting model has the highest Bayes factor. The second hypothesis will be supported if difference and ratio are both included as main effects in the best fitting model. The third hypothesis will be supported if the interaction between difference and ratio is included in the best fitting model.

Results

Food Experiment

Our first hypothesis predicted that subjects would choose the larger number of mealworms over the smaller number of mealworms within any given pair across all sessions. On average, our subjects in the food study chose the larger option 60.75% of the time with a standard deviation of 6.31. Results from a one sample t-test provided strong evidence that our hypothesis was supported ($M = 60.75$, 95% CI [55.48, 66.02], $t(7) = 4.82$, $p = .002$, $BF_{10} = 24.28$).

To investigate our other two hypotheses, we used model selection with mixed effect models. Hypothesis two predicts that (H2) pinyon jays prefer higher differences and lower ratios. If difference and ratio are both included as main effects in the best fitting model than hypothesis two will be supported. Hypothesis three predicts that (H3) difference and ratio are independent. If the interaction between difference and ratio is included in the best fitting model than this hypothesis three will be supported. We used a two-stage model selection approach that first selected the best random effects structure before investigating fixed effects. To determine the best random-effects structure, we compared models with all possible random effect combinations: subject, pair, and both subject and pair against the intercept only model. The model with the largest Bayes factor was assumed to have the best fit.

The best-fitting random effect structure for the food item data was the intercept only model, or no random effect structure, by rule of parsimony. The inclusion of only subject as a random effect ($BF_{10} = 0.15$) was under 0.33, providing evidence that this random effect structure was not supported by the data. The inclusion of pair ($BF_{10} = 0.31$), or both subject and pair ($BF_{10} = 0.06$), also did not improve model fit. Therefore, we included no random effect structure in the second stage of comparisons looking at fixed effects. The five fixed effect model structures we compared included the (1) intercept only model against the model with (2) ratio as a fixed effect, (3) difference as a fixed effect, (4) both difference and ratio as a fixed effects but *without* an interaction, and (5) both difference and ratio as fixed effects *with* an interaction. The model with only the main effect of ratio ($BF_{10} = 1.87 \times 10^3$)

374 best fit the data out of the five models that we compared. The model with the second-
375 best fit was the difference only model. To ensure that the ratio only model was the best
376 fitting model we compared the ratio only and difference only models to find 4.08 times the
377 evidence for the ratio only model as the difference only model. Thus, subjects in the food
378 study used the ratio between the two numbers of mealworms to choose between options,
379 with stronger preferences for larger options at smaller ratios (Figure 3 a). Consequently,
380 this only partially supports our second hypothesis, since difference was not included in the
381 best fitting model (Figure 3 b). Additionally, our third hypothesis was not supported, as
382 the interaction term between difference and ratio was not included in the best fitting model
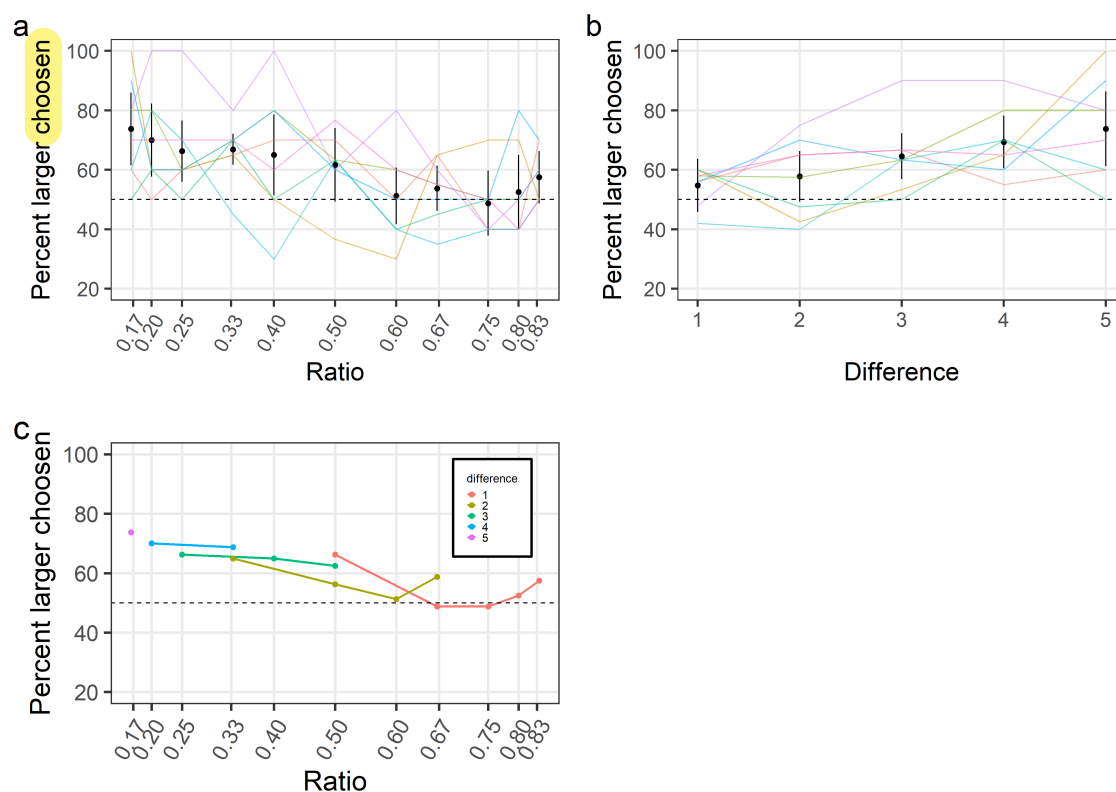383 (Figure 3 c).

Figure 3. Food study difference, ratio, and interaction results (a) Mean preference for the
larger option is shown on the y axis with the numerical difference options on the x axis.
(b) Mean preference for the larger option is shown on the y axis with the numerical ratio
options on the x axis. Dots represent mean values across all subjects and trials and error
bars represent 95% within-subject confidence intervals. Lines represent individual birds
(N=8). (c) Mean preference for the larger option is shown on the y axis numerical ratio
options on the x axis. The ratios were then grouped into the five numerical differences they
represent. Dots represent mean values across all subjects and trials within a study.

Social Experiment

Our first hypothesis predicted that subjects would choose the larger number of flock mates over the smaller number of flock mates within any given pair across all sessions. On average, our subjects in the social study choose the larger option 54.83% of the time with a standard deviation of 4.67. Results from a one sample t-test provided moderate evidence that our hypothesis was supported ($M = 54.83$, 95% CI [51.48, 58.17], $t(9) = 3.27$, $p = .010$, $\mathrm{BF}_{10} = 6.33$).

Results progressed similarly for the social study, we tested all possible random effects (subject, pair, and both) against the intercept only model to determine which, if any, of these random effects should be used in the full model.. Model testing revealed that the best-fitting random effect structure for the flock mate study no random effect structure. The inclusion of subject ($\mathrm{BF}_{10} = 0.04$), pair ($\mathrm{BF}_{10} = 0.22$), or both subject and pair ($\mathrm{BF}_{10} = 0.01$), did not improve the fit of the empty random effect model. All models had Bayes factors below $1/3$, which provides evidence for the null hypothesis of the intercept only model. The intercept only model with a Bayes factor of 1 is the chosen random effect structure by rule of parsimony as it is the least complex model, therefore no random effect structure was added to our fixed effect models for the social study.

The same fixed effect structures were used in the social as the food study: an empty model, a model with only difference, only ratio, both but no interaction, and one model with both difference and ratio and an interaction. The intercept only model ($\mathrm{BF}_{10} = 0.11$) best fit the data by rule of parsimony. Similarly, to the random effect model structures, our fixed effect models also had Bayes factors less than $1/3$. The model with the second best fit was the difference only model ($\mathrm{BF}_{10} = 0.11$), then the ratio only model ($\mathrm{BF}_{10} = 0.32$), both ($\mathrm{BF}_{10} = 0.01$), and the full model with an interaction ($\mathrm{BF}_{10} = 0.00$). Because ratio and difference failed to predict choices, neither of our hypothesis were supported by the data (Figure 4). Subjects did not use difference or ratio to make numerical decisions about conspecifics.

Discussion

We examined pinyon jays' numerical cognitive abilities in a food and social preference task. In both studies birds chose the larger of the two options on average across trials confirming our first hypothesis. In the food study, smaller numerical ratios but not larger numerical differences predicted the birds' choices partially confirming our second hypothesis. In the social experiment, neither ratio nor difference predicted choice showing that our second hypothesis was unsupported. In both the food and social experiments, difference and ratio did not independently predict choice indicating that our third hypothesis was also unsupported across studies. In neither context do pinyon jays use both ratio and difference together to choose between flock mates or food items

Results from the food study align with previous corvid research that birds prefer larger over smaller quantities and smaller ratios (Tornick et al., 2015; Ditz & Nieder, 2016; Kelly, 2016). This ratio effect supports our theory that preference for the larger option decreases as the ratio between the values approaches 1, confirming that birds acted in accordance with
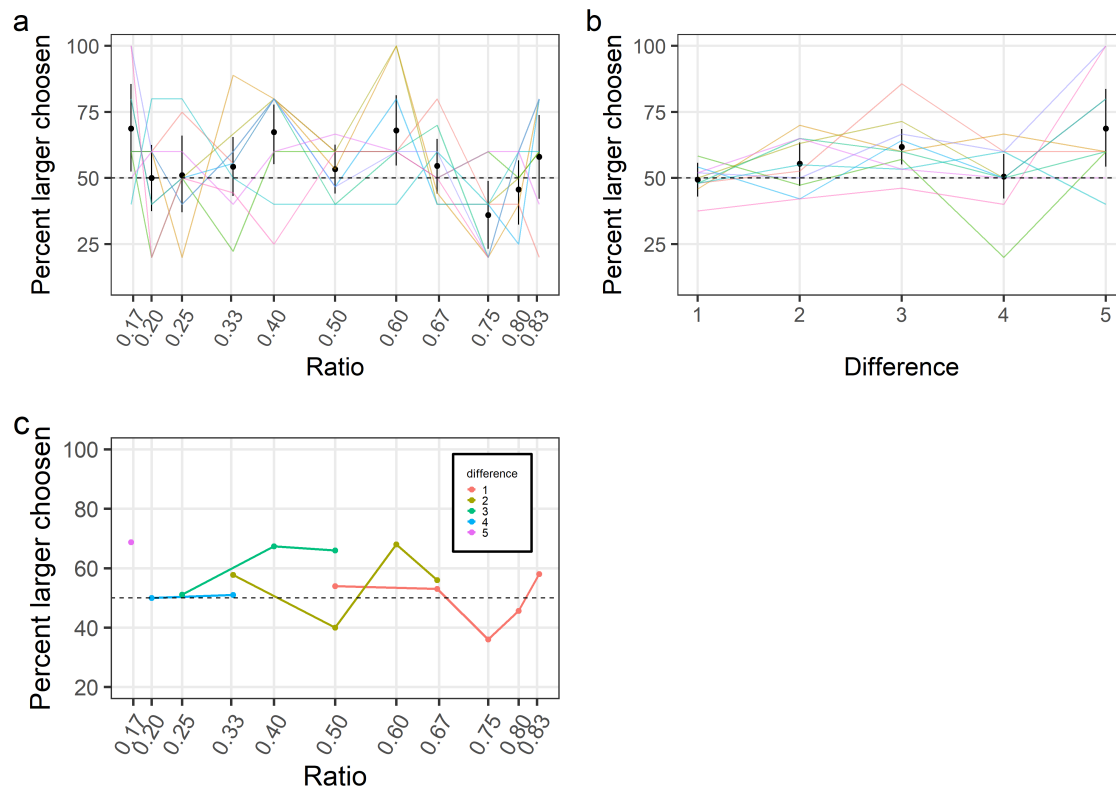
Figure 4. Social study difference, ratio, and interaction results (a) Mean preference for the larger option is shown on the y axis with the numerical difference options on the x axis. (b) Mean preference for the larger option is shown on the y axis with the numerical ratio options on the x axis. Dots represent mean values across all subjects and trials and error bars represent 95% within-subject confidence intervals. Lines represent individual birds (N=10). (c) Mean preference for the larger option is shown on the y axis numerical ratio options on the x axis. The ratios were then grouped into the five numerical differences they represent. Dots represent mean values across all subjects and trials within a study.

the *numerical magnitude effect*. Therefore, we have evidence supporting subjects using the approximate number system as a mechanism for quantity discrimination. Pinyon jays use amount of food, rather than precise number. However, the *numerical distance effect* was not present in our data as number preference did not increase as difference increased between two values. See limitations for further comment on this outcome.

Our lack of significant findings in the social study suggests that pinyon jays do not have just one mechanism they use across contexts. While ratio was shown to be an important mechanism to pinyon jays when foraging, neither difference nor ratio were the mechanism of choice when choosing between flock mates. This outcome is surprising as previous numeric discrimination tasks with conspecifics found effects of difference and ratio (Agrillo & Dadda, 2007; Agrillo et al., 2008). However, these experiments were completed in fish, not corvids. Corvids have a more complex system of sociality (Balda & Kamil, 1998). Possibly,

individual identity of birds overrides the implicit need to go to the smaller ratio and larger difference. Work in geese indicate that movement patterns of individuals may be shaped by the personality types present in the group (Kurvers et al., 2009). Furthermore, when testing Mexican Jay's on food item discrimination in the presence of flock mates, (Kelly, 2016) found no significant effect of ratio or difference as the birds were more likely to follow the previous bird, than pick the food item of a larger quantity. Possibly, this is because observing conspecifics provides important information, such as the safety of a food patch or resource availability (Krause et al., 2010; Handegard et al., 2012). Taken together, this suggests that birds do no view conspecifics in the same numeric terms as other objects and further investigation is needed to understand what internal systems animals are using to make numerical choices about conspecifics.

The differences in numerical preference between food and social contexts may be due to different selective pressures. Both flock size and foraging techniques have consequences for evolutionary fitness, but they tackle different adaptive problems. Food consumption acts primarily via natural selection by enhancing survival. Flock size, however, is integral to both natural and sexual selection: natural selection in the form of predator evasion, and sexual selection in the form of mate preference/reproduction. Joining a larger flock size allows an animal to dilute their chances of being eaten by predators (i.e., the dilution effect). Predator evasion acts on natural selection, allowing the animal to live long enough to reproduce, getting its genes into the next generation. This selection pressure could account for the confirmation of our first hypothesis or the subjects' preferences for larger over smaller flocks. Instead of using numerical difference and ratio to optimize predator evasion, birds are more likely focused on other biological imperatives like finding a suitable mate. Based on our current data it seems that when choosing between two flocks, birds are making decision not based on the number of animals on each side but rather the identity of the animals. For example, a female bird should go to the flock option with suitable males, aligning with sexual selection mechanisms, over optimizing predator evasion. It is these evolutionary differences that could account for the variation we see in numerical cognitive strategies.

It is important to highlight that these results correspond to a numerical *preference* task, not a discrimination task. If a subject discriminates between two objects, that means they recognize them as different. Preference denotes a liking of one option over the other. Meaning, the lack of a preference between two numbers of objects does not mean that the subject cannot discriminate the two objects. They just might not care about the difference. If a bird chooses indiscriminately between 5 or 6 mealworms it may not mean that they cannot discriminate between 5 and 6, but rather they are equally preferred. A clear preference implies discrimination, but lack of a preference does not imply an absence of discrimination.

To be clear, we are not making any claims about the method by which the birds make these preference-based numerical choices. That is, our study does not allow us to distinguish between the birds choosing larger *numbers* of food items or conspecifics and larger *amounts* of them. Previous studies have shown that instead of the absolute number, animals often make decisions based on the surface area or item size available (Gómez-Laplaza & Gerlai, 2013; Gómez-Laplaza et al., 2019). For example, in our study, subjects could have chosen an

option because they have a preference for 10% of a tray covered by mealworms over another with only 7% of the tray covered as opposed to preferring six over four mealworms. It is possible surface area, density, or quantity were all possible methods used by the subjects to make the numerical decisions discussed in the paper. For our purposes, the method of preference is not critical. In the food item study, total quantity is a reasonable criteria because the birds ultimate goal is to obtain as much food as possible to stay alive (e.g. two large mealworms are better than 3 small mealworms). Therefore, total calories or overall food intake, rather than absolute number, is a better evolutionary strategy. Future work is needed to tease apart which methods birds use when presented with food or conspecifics.

Limitations

Though numerical ratio predicted choice in the food study, it did not in the social study, and difference did not predict choice in either study. This result was unlike previous corvid research, which found both difference and ratio effects in food studies (Tornick et al., 2015; Ditz & Nieder, 2016; Kelly, 2016). This could be due to the limited range within our pair options. The smallest difference option we offered subjects was one (2-1=1) and the largest was five (6-1=5).Previous studies difference options ranged from 8 as the largest option offered to 30 (Tornick et al., 2015; Ditz & Nieder, 2016). Considering that pinyon jay flock sizes range from 50 to 500 (Wiggins, 2005), the limited range of difference options may not have provided enough opportunities to demonstrate effects of difference.

Additionally, our relatively low sample size with an aging study population could have also contributed to our findings. Considering our sample size (Nfood=8, Nsocial=10) we completed a within-subject study design and completed 10 trials per pair to account for within-subject variation. Furthermore, we are currently replicating these findings in our lab with another group of subjects. The 18 subjects that completed the experiment were around twice the age of a free ranging pinyon jay's average lifespan (Marzluff & Balda, 1992). Possibly, numerical cognitive abilities degrade with age. However, in a closely related corvid seed caching species there was no signs of age-related function loss in spatial cognitive performance in birds at age six (Heinen et al., 2021). Yet, our subjects have an average age of 15. Between the ages of six and 15 our animals could have incurred some age-related function loss.

No subject completed both the food item and the flock mate study, so we could not directly compare performance across contexts. To make a more robust claim about context-dependent numerical cognition, we would need to test the same subjects in both contexts. We are currently replicating these experiments to investigate context effects within subjects. Subjects from the previous food experiment are now experiencing the social study and vice versa, this will allow us to better understand how difference and ratio effects change depending on context.

Conclusion

The present research investigated how pinyon jays assess numbers of food items and conspecifics in a preference task. We showed that overall, pinyon jays chose the larger option

over the smaller in both contexts. In a food context, pinyon jays use numerical ratio but *not* numerical difference to make decisions. In a social context, subjects did not use difference or ratio to make decisions. Taken together, our findings suggest that while number or amount is important for selecting food items, other factors such as flock mate identity may be more important for selecting social groups to join. Thus, in numerical preference situations, the type of objects offered drive the numerical strategies that animals use.

Acknowledgments

Ethics approval

All procedures were conducted in an ethical and responsible manner, in full compliance with all relevant codes of experimentation and legislation and were approved by the UNL Internal Review Board (protocol # 17922) and Institutional Animal Care and Use Committee (protocol # 1621).

References

Agrillo, C., & Beran, M. J. (2013). Number without language: Comparative psychology and the evolution of numerical cognition. *Frontiers in Psychology.* https://www.academia.edu/6879466/Number_without_language_comparative_psychology_and_the_evolution_of_numerical_cognition

Agrillo, C., & Bisazza, A. (2014). Spontaneous versus trained numerical abilities. A comparison between the two main tools to study numerical competence in non-human animals. *Journal of Neuroscience Methods*, *234*, 82–91. https://doi.org/10.1016/j.jneumeth.2014.04.027

Agrillo, C., & Dadda, M. (2007). Discrimination of the larger shoal in the poeciliid fish Girardinus falcatus. *Ethology Ecology & Evolution*, *19*(2), 145–157. https://doi.org/10.1080/08927014.2007.9522574

Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2008). Do fish count? Spontaneous discrimination of quantity in female mosquitofish. *Animal Cognition*, *11*(3), 495–503. https://doi.org/10.1007/s10071-008-0140-9

Agrillo, C., Miletto Petrazzini, M. E., & Bisazza, A. (2017). Numerical abilities in fish: A methodological review. *Behavioural Processes*, *141*(Pt 2), 161–171. https://doi.org/10.1016/j.beproc.2017.02.001

Agrillo, C., Piffer, L., & Bisazza, A. (2011). Number versus continuous quantity in numerosity judgments by fish. *Cognition*, *119*(2), 281–287. https://doi.org/10.1016/j.cognition.2010.10.022

Arak, A. (1983). Vocal interactions, call matching and territoriality in a Sri Lankan treefrog, Philautus leucorhinus (Rhacophoridae). *Animal Behaviour*, *31*(1), 292–302. https://doi.org/10.1016/S0003-3472(83)80199-7

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown.* https://github.com/crsh/papaja

Balda, R. P., & Kamil, A. (1998). *The Ecology and Evolution of Spatial Memory in Corvids of the Southwestern USA: The Perplexing Pinyon Jay.* 37.

Bates, D., & Maechler, M. (2021). *Matrix: Sparse and dense matrix classes and methods.* https://CRAN.R-project.org/package=Matrix

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

Beran, M. (2001). Summation and numerousness judgments of sequentially presented sets of items by chimpanzees (Pan troglodytes). *Journal of Comparative Psychology (Washington, D.C. : 1983)*, *115*, 181–191. https://doi.org/10.1037//0735-7036.115.2.181

Bertamini, M., Guest, M., Vallortigara, G., Rugani, R., & Regolin, L. (2018). The effect of clustering on perceived quantity in humans (Homo sapiens) and in chicks (Gallus

gallus). *Journal of Comparative Psychology (Washington, D.C.: 1983)*, *132*(3), 280–293. https://doi.org/10.1037/com0000114

Buckingham, J. N., Wong, B. B. M., & Rosenthal, G. G. (2007). Shoaling decisions in female swordtails: How do fish gauge group size? *Behaviour*, *144*(11), 1333–1346. https://doi.org/10.1163/156853907782418196

Call, J. (2000). Estimating and operating on discrete quantities in Orangutans (Pongo pygmaeus). *Journal of Comparative Psychology*, *114*(2), 136–147. https://doi.org/10.1037/0735-7036.114.2.136

Cantlon, J. F., & Brannon, E. M. (2006). Shared System for Ordering Small and Large Numbers in Monkeys and Humans. *Psychological Science*, *17*(5), 401–406. https://doi.org/10.1111/j.1467-9280.2006.01719.x

Carazo, P., Fernández-Perea, R., & Font, E. (2012). Quantity Estimation Based on Numerical Cues in the Mealworm Beetle (Tenebrio molitor). *Frontiers in Psychology*, *3*, 502. https://doi.org/10.3389/fpsyg.2012.00502

Dacke, M., & Srinivasan, M. V. (2008). Evidence for counting in insects. *Animal Cognition*, *11*(4), 683–689. https://doi.org/10.1007/s10071-008-0159-y

Dange, P., Mhaisalkar, P., & Paranjpe, D. (2021). To group or not to group: Group size dynamics and intestinal parasites in Indian peafowl populations. *Acta Ethologica*, *24*(2), 107–117. https://doi.org/10.1007/s10211-021-00366-x

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences*, *21*(8), 355–361. https://doi.org/10.1016/S0166-2236(98)01263-6

Ditz, H. M., & Nieder, A. (2016). Numerosity representations in crows obey the Weber–Fechner law. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1827), 20160083. https://doi.org/10.1098/rspb.2016.0083

Ditz, H. M., & Nieder, A. (2015). Neurons selective to the number of visual items in the corvid songbird endbrain. *Proceedings of the National Academy of Sciences*, *112*(25), 7827–7832. https://doi.org/10.1073/pnas.1504245112

Emmerton, J., & Renner, J. C. (2006). Scalar effects in the visual discrimination of numerosity by pigeons. *Learning & Behavior*, *34*(2), 176–192. https://doi.org/10.3758/BF03193193

Emmerton, J., & Renner, J. C. (2009). Local rather than global processing of visual arrays in numerosity discrimination by pigeons (Columba livia). *Animal Cognition*, *12*(3), 511–526. https://doi.org/10.1007/s10071-009-0212-5

Evans, T. A., Beran, M. J., Harris, E. H., & Rice, D. F. (2009). Quantity judgments of sequentially presented food items by capuchin monkeys (Cebus apella). *Animal Cognition*, *12*(1), 97–105. https://doi.org/10.1007/s10071-008-0174-z

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Fortran code by Alan Miller, T. L. based on. (2020). *Leaps: Regression subset selection.* https://CRAN.R-project.org/package=leaps

Gómez-Laplaza, L. M., & Gerlai, R. (2011). Can angelfish (Pterophyllum scalare) count? Discrimination between different shoal sizes follows Weber's law. *Animal Cognition*, *14*(1), 1–9. https://doi.org/10.1007/s10071-010-0337-6

Gómez-Laplaza, L. M., & Gerlai, R. (2013). The role of body surface area in quantity discrimination in angelfish (Pterophyllum scalare). *PloS One*, *8*(12), e83880. https://doi.org/10.1371/journal.pone.0083880

Gómez-Laplaza, L. M., & Gerlai, R. (2016). Discrimination of large quantities: Weber's law and short-term memory in angelfish, Pterophyllum scalare. *Animal Behaviour*, *112*, 29–37. https://doi.org/10.1016/j.anbehav.2015.10.022

Gómez-Laplaza, L. M., Romero, L., & Gerlai, R. (2019). The role of item size on choosing contrasted food quantities in angelfish (Pterophyllum scalare). *Scientific Reports*, *9*(1), 15305. https://doi.org/10.1038/s41598-019-51753-1

Handegard, N. O., Boswell, K. M., Ioannou, C. C., Leblanc, S. P., Tjøstheim, D. B., & Couzin, I. D. (2012). The Dynamics of Coordinated Group Hunting and Collective Information Transfer among Schooling Prey. *Current Biology*, *22*(13), 1213–1217. https://doi.org/10.1016/j.cub.2012.04.050

Hanus, D., & Call, J. (2007). Discrete quantity judgments in the great apes (Pan paniscus, Pan troglodytes, Gorilla gorilla, Pongo pygmaeus): The effect of presenting whole sets versus item-by-item. *Journal of Comparative Psychology*, *121*(3), 241–249. https://doi.org/10.1037/0735-7036.121.3.241

Heinen, V. K., Pitera, A. M., Sonnenberg, B. R., Benedict, L. M., Branch, C. L., Bridge, E. S., & Pravosudov, V. V. (2021). Specialized spatial cognition is associated with reduced cognitive senescence in a food-caching bird. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1947), 20203180. https://doi.org/10.1098/rspb.2020.3180

Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools.* https://CRAN.R-project.org/package=purrr

Hyde, D. C. (2011). Two Systems of Non-Symbolic Numerical Cognition. *Frontiers in Human Neuroscience*, *5*, 150. https://doi.org/10.3389/fnhum.2011.00150

Kelly, E. M. (2016). Counting on your friends: The role of social environment on quantity discrimination. *Behavioural Processes*, *128*, 9–16. https://doi.org/10.1016/j.beproc.2016.03.019

Krause, J., & Ruxton, G. D. (2002). *Living in Groups.* OUP Oxford.

Krause, J., Ruxton, G. D., & Krause, S. (2010). Swarm intelligence in animals and humans. *Trends in Ecology & Evolution*, *25*(1), 28–34. https://doi.org/10.1016/j.tree.2009.06.016

Kurvers, R. H. J. M., Eijkelenkamp, B., Oers, K. van, Lith, B. van, Wieren, S. E. van, Ydenberg, R. C., & Prins, H. H. T. (2009). Personality differences explain leadership in barnacle geese. *Animal Behaviour*, *78*(2), 447–453. https://doi.org/10.1016/j.anbehav.2009.06.002

Lehmann, J., Korstjens, A. H., & Dunbar, R. I. M. (2007). Fission–fusion social systems as a strategy for coping with ecological constraints: A primate case. *Evolutionary Ecology*, *21*(5), 613–634. https://doi.org/10.1007/s10682-006-9141-9

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

MacNulty, D. R., Tallian, A., Stahler, D. R., & Smith, D. W. (2014). Influence of Group Size on the Success of Wolves Hunting Bison. *PLOS ONE*, *9*(11), e112884. https://doi.org/10.1371/journal.pone.0112884

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the bayesian framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Marzluff, J. M., & Balda, R. P. (1992). *The Pinyon Jay: Behavioral Ecology of a Colonial and Cooperative Corvid* (First Edition). Academic Press.

Merten, K., & Nieder, A. (2009). Compressed scaling of abstract numerosity representations in adult humans and monkeys. *Journal of Cognitive Neuroscience*, *21*(2), 333–346. https://doi.org/10.1162/jocn.2008.21032

Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of bayes factors for common designs.* https://CRAN.R-project.org/package=BayesFactor

Morey, R. D., Rouder, J. N., Jamil, T., Urbanek, S., Forner, K., & Ly, A. (2021). *BayesFactor: Computation of Bayes Factors for Common Designs.* https://CRAN.R-project.org/package=BayesFactor

Müller, K. (2020). *Here: A simpler way to find your files.* https://CRAN.R-project.org/package=here

Müller, K., & Wickham, H. (2021). *Tibble: Simple data frames.* https://CRAN.R-project.org/package=tibble

Nieder, A. (2018). Evolution of cognitive and neural solutions enabling numerosity judgements: Lessons from primates and corvids. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1740), 20160514. https://doi.org/10.1098/rstb.2016.0514

Nieder, A. (2020). The Adaptive Value of Numerical Competence. *Trends in Ecology & Evolution*, *35*(7), 605–617. https://doi.org/10.1016/j.tree.2020.02.009

Pedersen, T. L. (2020). *Patchwork: The composer of plots.* https://CRAN.R-project.org/package=patchwork

Performance: Assessment of Regression Models Performance. (2021). https://CRAN.R-project.org/package=performance

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, *6*(1), 7–11. https://journal.r-project.org/archive/

R Core Team. (2021). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Raftery, A., Hoeting, J., Volinsky, C., Painter, I., & Yeung, K. Y. (2021). *BMA: Bayesian model averaging.* https://CRAN.R-project.org/package=BMA

Ren, K., & Russell, K. (2021). *Formattable: Create 'formattable' data structures.* https://CRAN.R-project.org/package=formattable

Robinson, D., Hayes, A., & Couch, S. (2021). *Broom: Convert statistical objects into tidy tibbles.* https://CRAN.R-project.org/package=broom

Rugani, R., Cavazzana, A., Vallortigara, G., & Regolin, L. (2013). One, two, three, four, or is there something more? Numerical discrimination in day-old domestic chicks. *Animal Cognition, 16*(4), 557–564. https://doi.org/10.1007/s10071-012-0593-8

Scarf, D., Hayne, H., & Colombo, M. (2011). Pigeons on Par with Primates in Numerical Competence. *Science.* https://doi.org/10.1126/science.1213357

Silk, M. J., Croft, D. P., Tregenza, T., & Bearhop, S. (2014). The importance of fission–fusion social group dynamics in birds. *Ibis, 156*(4), 701–715. https://doi.org/10.1111/ibi.12191

Sklyar, O., Murdoch, D., Smith, M., Eddelbuettel, D., Francois, R., Soetaert, K., & Ranke, J. (2021). *Inline: Functions to inline c, c++, fortran function calls from r.* https://CRAN.R-project.org/package=inline

Stancher, G., Rugani, R., Regolin, L., & Vallortigara, G. (2015). Numerical discrimination by frogs (Bombina orientalis). *Animal Cognition, 18*(1), 219–229. https://doi.org/10.1007/s10071-014-0791-7

Stevens, J. R., Wood, J. N., & Hauser, M. D. (2007). When quantity trumps number: Discrimination experiments in cotton-top tamarins (Saguinus oedipus) and common marmosets (Callithrix jacchus). *Animal Cognition, 10*(4), 429–437. https://doi.org/10.1007/s10071-007-0081-8

Terry M. Therneau, & Patricia M. Grambsch. (2000). *Modeling survival data: Extending the Cox model.* Springer.

Todorov, V., & Filzmoser, P. (2009a). An object-oriented framework for robust multivariate analysis. *Journal of Statistical Software, 32*(3), 1–47. https://www.jstatsoft.org/article/view/v032i03/

Todorov, V., & Filzmoser, P. (2009b). An object-oriented framework for robust multivariate analysis. *Journal of Statistical Software, 32*(3), 1–47. https://www.jstatsoft.org/article/view/v032i03

Tornick, J. K., Callahan, E. S., & Gibson, B. M. (2015). An investigation of quantity discrimination in Clark's nutcrackers (Nucifraga columbiana). *Journal of Comparative Psychology (Washington, D.C.: 1983), 129*(1), 17–25. https://doi.org/10.1037/a0037863

Uller, C., Jaeger, R., Guidry, G., & Martin, C. (2003). Salamanders (Plethodon cinereus) go for more: Rudiments of number in an amphibian. *Animal Cognition, 6*(2), 105–112. https://doi.org/10.1007/s10071-003-0167-x

Vonk, J., & Beran, M. J. (2012). Bears 'count' too: Quantity estimation and comparison in black bears, Ursus americanus. *Animal Behaviour*, *84*(1), 231–238. https://doi.org/10.1016/j.anbehav.2012.05.001

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems ofp values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/BF03194105

Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. *Cognitive Psychology*, *60*(3), 158–189. https://doi.org/10.1016/j.cogpsych.2009.12.001

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

White, D. J., Ho, L., & Freed-Brown, G. (2009). Counting Chicks Before They Hatch: Female Cowbirds Can Time Readiness of a Host Nest for Parasitism. *Psychological Science*, *20*(9), 1140–1145. https://doi.org/10.1111/j.1467-9280.2009.02418.x

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string operations*. https://CRAN.R-project.org/package=stringr

Wickham, H. (2021a). *Forcats: Tools for working with categorical variables (factors)*. https://CRAN.R-project.org/package=forcats

Wickham, H. (2021b). *Tidyr: Tidy messy data*. https://CRAN.R-project.org/package=tidyr

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., & Bryan, J. (2019). *Readxl: Read excel files*. https://CRAN.R-project.org/package=readxl

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *Dplyr: A grammar of data manipulation*. https://CRAN.R-project.org/package=dplyr

Wickham, H., & Hester, J. (2021). *Readr: Read rectangular text data*. https://CRAN.R-project.org/package=readr

Wiggins, D. (2005). *Pinyon jay (Gymnorhinus cyanocephalus): A technical conservation assessment*. https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb5182043.pdf

Xia, L., Emmerton, J., Siemann, M., & Delius, J. D. (2001). Pigeons (Columba livia) learn to link numerosities with symbols. *Journal of Comparative Psychology (Washington, D.C.: 1983)*, *115*(1), 83–91. https://doi.org/10.1037/0735-7036.115.1.83

Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Chapman; Hall/CRC. https://yihui.org/knitr/

Yang, T.-I., & Chiao, C.-C. (2016). Number sense and state-dependent valuation in cuttlefish. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1837), 20161379. https://doi.org/10.1098/rspb.2016.1379

Appendix

Model Structures

Random Effects

a	Model	Formula
	Intercept Only Model	Choice~1
	Subject Only Model	Choice~(1|Subject)
	Pair Only Model	Choice~(1|Pair)
	Both Subject and Pair	Choice~(1|Subject)+(1|Pair)

Fixed Effects

b	Model	Formula
	Intercept Only Model	Choice~1
	Ratio Only Model	Choice~Ratio
	Difference Only Model	Choice~Difference
	Both Fixed Effects, No Interaction	Choice~Ratio+Difference
	Both Fixed Effects, With Interaction	Choice~Ratio*Difference

Figure A1. Random and fixed effect model structures (a) The four random effect models with their corresponding formulas. (b) The five fixed effect models with their corresponding formulas. Choice represents the binary dependent variable of larger or smaller item chosen by the subject.

Food Study

a

Random Effects

Model	AIC	BIC	BF
(1|Subject)	1608	1619	0.15
(1|Pair)	1607	1617	0.31
(1|Subject)+(1|Pair)	1605	1620	0.06

b

Fixed Effects

Model	AIC	BIC	BF
ratio	1589	1600	1874.2
difference	1592	1602	459.7
difference + ratio	1591	1606	72.4
difference * ratio	1593	1613	2.3

Social Study

c

Model	AIC	BIC	BF
(1|Subject)	1010	1020	0.04
(1|Pair)	1007	1016	0.22
(1|Subject)+(1|Pair)	1009	1023	0.01

d

Model	AIC	BIC	BF
ratio	1008	1018	0.11
difference	1006	1015	0.32
difference + ratio	1008	1022	0.01
difference * ratio	1008	1027	0.00

Figure A2. Bayes factor, BIC, and AIC values per model (a) Random effect models and Bayes factor, BIC, and AIC values for the food study (b) Fixed effect models and Bayes factor, BIC, and AIC values for the food study (c) Random effect models and Bayes factor, BIC, and AIC values for the social study. (d) Fixed effect models and Bayes factor, BIC, and AIC values for the social study. Reminder: $BF > 3$ show moderate evidence for the alternative hypothesis while $BF < 0.33$ show moderate evidence for the null hypothesis. A BF in between 3 and 0.33 is inconclusive.

Pair	Ratio	Difference
1/2	0.50	1
1/3	0.33	2
1/4	0.25	3
1/5	0.20	4
1/6	0.17	5
2/3	0.67	1
2/4	0.50	2
2/5	0.40	3
2/6	0.33	4
3/4	0.75	1
3/5	0.60	2
3/6	0.50	3
4/5	0.80	1
4/6	0.67	2
5/6	0.83	1

Figure A3. Factorial pair combinations used each experiment. We used every factorial pair from 1 to 6 resulting in 15 combinations. The second and third columns show the ratio and difference for each pair, respectively.