Friends Aren't Food: Pinyon Jays (*Gymnorhinus cyanocephalus*) Show Context-Dependent Quantity Preference

London M. Wolff, Katherine Carey, and Jeffrey R. Stevens[*]

Department of Psychology, Center for Brain, Biology & Behavior, University of Nebraska-Lincoln

*Corresponding author (Email: jeffrey.r.stevens@gmail.com)

Abstract – Animals must often choose between different quantities of objects in their environment, from food items to conspecifics. Yet we know little about how quantitative cognitive abilities compare across different types of objects. Previous research shows individuals use both the numerical difference (large − small) and numerical ratio (small/large) between two quantities to choose between them. This study investigated whether numerical difference and ratio predict preferences for quantities of food items and conspecifics in pinyon jays (*Gymnorhinus cyanocephalus*) using quantity preference tasks. In two replications of the food experiment ($N = 12$), pinyon jays chose larger quantities of mealworms more when numerical differences were large and numerical ratios were small. However, numerical difference did not influence food choice independently of ratio. In two replications of the social experiment ($N = 20$), when choosing between groups of conspecifics, pinyon jays did not prefer the larger over smaller group sizes and did not show numerical difference or ratio effects. Therefore, pinyon jays may use quantity information differently when deciding between quantities of food items and conspecifics. Whereas quantity was important for selecting food items, additional factors such as individual identity may be more important for selecting social groups to join. Thus, the type of objects offered can influence how animals use quantity information to choose among quantities.

Keywords – Context dependence, Corvid, Quantity, Quantitative cognition, Preference

_____

When a hungry pinyon jay (*Gymnorhinus cyanocephalus*) flies 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 quantities of options across a wide range of contexts, such as foraging for food and selecting social partners. Yet little is known about how animals use information about quantity across these contexts.

Quantification skills have strong adaptive value for survival and reproduction (Nieder, 2018), playing roles in navigation, predator avoidance, territory defense, foraging, courtship, and mating (Arak, 1983; White et al., 2009; Carazo et al., 2012; Yang & Chiao, 2016; Agrillo et al., 2017; Nieder, 2020). And many animal species have demonstrated the ability to quantify objects in their environment, including arthropods (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 (Call, 2000; Beran, 2001; Vonk & Beran, 2012; Nieder, 2018). These abilities have been studied across a range of objects and contexts, using several different experimental techniques (Agrillo & Beran, 2013; Agrillo & Bisazza, 2014). In food preference tasks, for instance, animals are presented with two different quantities of food and are allowed to consume the chosen quantities. Animals should choose the larger rewards when they (1) can discriminate between the two quantities and (2) are motivated to choose and

consume more food. Importantly, a lack of preference does not mean an inability to discriminate (Mendelson et al., 2016). *Discrimination* requires an animal to recognize two things as different, whereas *preference* signifies a desire for one choice over the other. A preference implies discrimination, but lack of a preference does not imply an absence of discrimination. Most studies of quantification conflate these two types of tasks, but they are important to distinguish (Agrillo & Bisazza, 2014).

Most of the tasks designed to measure quantitative cognition use food or computer stimuli as quantifiable objects (Call, 2000; Beran, 2001; Scarf et al., 2011; Rugani et al., 2013; Kelly, 2016; Potrich et al., 2022). Yet other objects are also important to quantify. Quantities of conspecifics, for example, are important for avoiding predators by diluting the probability of being captured (Hamilton, 1971). And animals are sensitive to conspecific quantities in the wild, preferring to join larger over smaller groups (Emmerton, 2001; Krause & Ruxton, 2002; Silk et al., 2014), or comparing their current group size to the size of a rival group during competition (McComb et al., 1994; Wilson et al., 2001; Kitchen, 2004). In social preference tasks, animals (primarily fish) have been given the choice between joining groups of different quantities of conspecifics to assess quantification, and they are sensitive to conspecific quantity (Buckingham et al., 2007; Agrillo et al., 2008; Potrich et al., 2015; Gómez-Laplaza & Gerlai, 2016; Messina et al., 2022).

Despite existing work on quantifying food and conspecifics, little research has directly compared quantitative cognition across different types of objects to assess whether animals use quantity information similarly across these contexts. One of the key cognitive processes proposed for quantification is the *approximate number system*, which involves the estimation of quantity without relying on language or symbols (Feigenson et al., 2004; Nieder, 2020). The approximate number system is characterized by two key effects (Dehaene et al., 1998; Ditz & Nieder, 2016). For the *numerical distance effect*, discrimination improves with increasing numerical difference between two values (i.e., mathematical difference between two numbers: $4 - 2$ has a difference of 2). Discrimination becomes easier as the difference increases and the options become more dissimilar. For the *numerical ratio effect*, discrimination improves as the numerical ratio (i.e., mathematical quotient between two numbers: 2/4 has a ratio of 0.5) between values decreases. Discrimination becomes easier as the numerical ratio moves away from 1. The numerical ratio effect has been formalized as Weber's Law, which quantifies the perceived change between two stimuli and predicts that change detection depends on the stimuli's ratio (Fechner, 1860). Animals typically discriminate food quantities and group sizes better when there are larger numerical differences (Agrillo & Dadda, 2007; Agrillo et al., 2008; Kelly, 2016) and smaller numerical ratios (Cantlon & Brannon, 2006; Hanus & Call, 2007; Evans et al., 2009; Tornick et al., 2015; Potrich et al., 2015; Bisazza & Santacà, 2022), supporting the numerical distance and ratio effects respectively.

Data showing distance and ratio effects illustrate the use of approximate amounts rather than precise numbers. But difference and ratio are not independent; they are highly correlated (Lyons et al., 2015). As difference increases, ratio decreases. Thus, it is possible for animals to use either difference or ratio or both to quantify objects. For instance, quantity tasks in some species show effects of ratio but not difference (Cantlon & Brannon, 2006; Buckingham et al., 2007; Gómez-Laplaza & Gerlai, 2011; Tornick et al., 2015), whereas other studies show effects of both difference and ratio, with the other held constant (Agrillo et al., 2007; Kelly, 2016; Bisazza & Santacà, 2022). This is important because Weber's Law predicts that only ratio should drive discrimination—difference should not influence discrimination above and beyond ratio (Cantlon & Brannon, 2006). Therefore, an independent contribution of difference suggests an additional quantitative process beyond the ratio-dependence of Weber's Law.

Present Study

Pinyon jays, a highly social species of North American corvid (Balda & Kamil, 1998), are well-suited for examining effects of context on quantitative cognition because of their diet and social habits. They live in flocks ranging from 50 to 500 birds with fission-fusion dynamics in which members of a community form frequently changing subgroups (Marzluff & Balda, 1992). Fission-fusion group living reduces predation risk and improves foraging success (Lehmann et al., 2007; Dange et al., 2021). This is

relevant to quantification, 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 and rely on in the winter (Marzluff & Balda, 1992). The need to retrieve cached food sources may place strong selection pressure on quantitative cognition, as they need to store as many pine seeds as possible to survive the winter. Thus, pinyon jays rely on assessing the quantity of birds in a flock for social living decisions and the quantity of food items available for foraging decisions.

The primary aim of the present study was to investigate how pinyon jays use 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 quantities of either food items or conspecifics. To address our research question, we tested three hypotheses. Our first hypothesis posits that pinyon jays will, on average, prefer larger over smaller quantities. An animal is more likely to survive if they consume more food and live in larger rather than smaller groups. Our second hypothesis posits that pinyon jays will show stronger preferences for more items when the numerical pairs have higher numerical differences and lower numerical ratios. As differences and ratios move away from 1, discrimination becomes easier. Our third hypothesis posits that both numerical difference and ratio will influence preference independently of each other. Whereas discrimination tasks in primates show only ratio effects (Cantlon & Brannon, 2006), numerical judgment tasks in humans show independent effects of difference and ratio (Stevens & Soh, 2018). We suggest that the preference tasks used here are more comparable to those human numerical judgment tasks than discrimination tasks used in primates; therefore, we predict that results from our preference tasks will mirror results found in human numerical judgment tasks.

Method

We conducted experiments to investigate quantification of both food and conspecifics. Each experiment was replicated with two sets of birds, and most subjects experienced both the food and social experiment. Example videos of test sessions are available in the supplementary materials at https://osf.io/g45nk/.

Ethics Statement

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 Institutional Animal Care and Use Committee (protocols #1709 and #2059), and all procedures conformed to the ASAB/ABS Guidelines for the use of animals in research.

Subjects

Our study population of 21 pinyon jays (*Gymnorhinus cyanocephalus*) was wild born and locally housed. We intended to use all 35 pinyon jays in the colony, all of which completed some stages of training. The resulting 21 pinyon jays used in the study were the only birds that passed our training criteria. 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. At the end of the experiments, they ranged between 10-19 years old. All subjects completed prior cognitive and behavioral experiments and were handled by humans extensively during their time in the lab.

This project consisted of two replicates of each experiment. Within a replicate, the food and social experiments ran concurrently with separate birds as subjects. After the first replicate, most birds switched experiments for the second replicate. For replicate 1, 8 pinyon jays (1 female) completed all rounds of the food experiment from February 2021 to June 2021, and 10 jays (4 female) completed all rounds of the social experiment from February 2020 to February 2021. A further 17 jays (6 female) were used as conspecifics in the social experiment. Two jays were dropped from the social experiment due to unrelated

health concerns. For replicate 2, 4 pinyon jays (1 female) completed all rounds of the food experiment from October 2021 to March 2022, and 10 jays (1 female) completed all rounds of the social experiment from October 2021 to February 2022 (Table S1). A further 8 jays (4 female) from the colony were used as conspecifics in the social experiment. For social replication 2, we removed the three most popular birds from social repetition 1 to better control for individual preference across conspecifics. Four birds (1 female) served as both subjects and conspecific birds in separate replicates of the social experiment. A bird was never both a subject and a conspecific during the same replicate. See Table S1 for the order of experiments for each subject bird.

The jays in the food experiment were housed in pairs, whereas the subjects in the social experiment were individually housed and the conspecifics were group housed. Upon completion of the first phase of experiments birds were changed to the opposite housing type to stay consistent. Subjects were not food restricted in either experiment.

Numerical Pairs

We used the same set of all 15 possible numerical pairs between 1 and 6 (e.g., 1 vs 2, 1 vs 3, 1 vs 4, etc.) for all four experiments except in the second social experiment, which used all but four of the pairs (Table 1). We chose these numerical pairs because they systematically include all numerical pairs for numbers 1-6, which provides a good distribution of differences and ratios. Within these 15 pairs are the differences 1 through 5 and ratios of 0.17, 0.20, 0.25, 0.33, 0.40, 0.50, 0.60, 0.67, 0.75, 0.80, 0.83. To reduce the number of birds needed for the second social experiment, we removed the four numerical pairs that required nine or more conspecific birds.

Table 1

Factorial Pair Combinations

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

* Pairs not used in Social 2.
Note. Table used with permission under a CC-BY4.0 license: Wolff et al. (2022); available at https://doi.org/10.31234/osf.io/kxgwt.
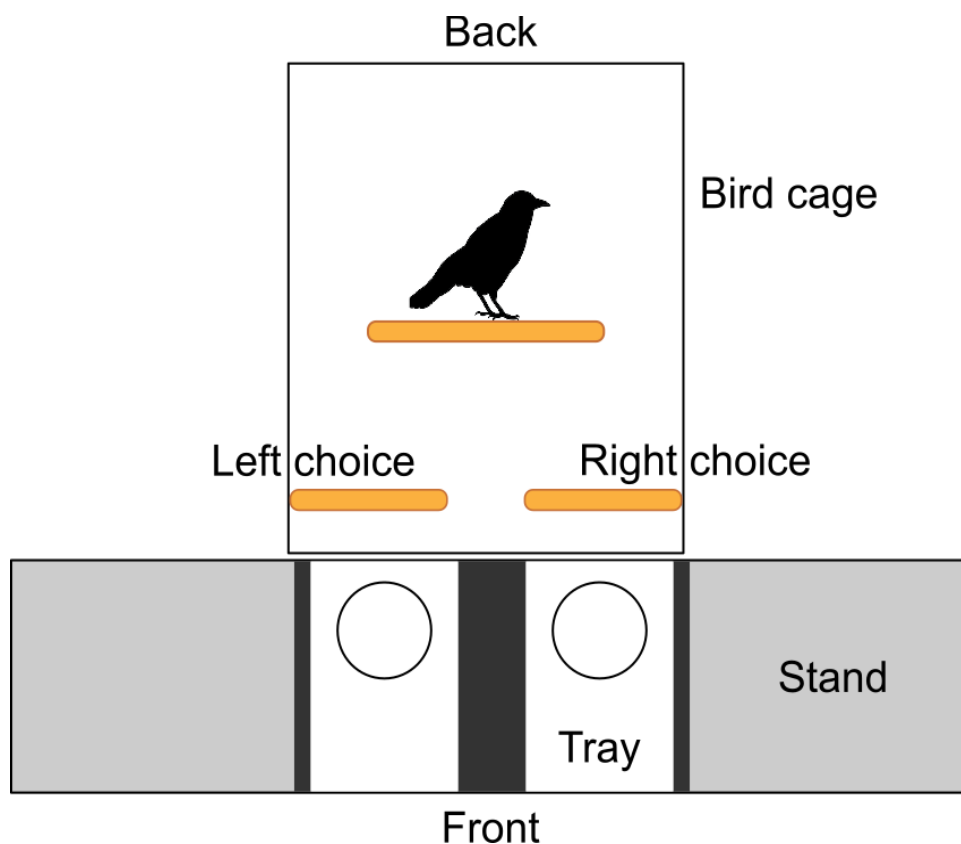
Food Experiment

The apparatus for the food experiment included a bird cage (72 × 48 × 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 mealworm viewing. The stand included two channels that guided Plexiglas trays. Each Plexiglas tray had a standard petri dish placed 1.5 cm away from the front. Mealworms were placed in the front two-thirds of the petri dish so the subjects could easily reach them, and they were evenly

distributed across the available area. The cage included three perches: one larger free-standing perch toward the back of the cage and two smaller perches attached to each side of the front of the cage. Subjects started each trial perched on the back perch, and they chose an option by landing on one of the front perches. The unchosen tray was slid out of reach, and the subjects then consumed the mealworms on the chosen side (see Supplementary Video).

Figure 1

Food Experiment Apparatus

Note. Overhead view. Subjects started the trial on the back perch and hopped forward to one of the front perches to signal choice. The experimenter sat on the other side of the stand and slid the opposite tray out of reach, allowing the subject to consume their chosen mealworms. Figure used with permission under a CC-BY4.0 license: Wolff et al. (2022); available at https://doi.org/10.31234/osf.io/kxgwt.

Habituation and Training

Prior to experimental sessions, all birds were habituated to the experimental room, cage, and stand and then experienced training. Depending on the bird, this took between 9-16 weeks, as each bird ran once a day between 11:00-15:00 CT. Subjects experienced four training phases in total, each meant to teach the bird a different piece of the paradigm.

The first phase (rear cup habituation) habituated subjects to the apparatus. 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.

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. Front dish habituation was similar to rear cup habituation, except three mealworms were placed in both dishes on the Plexiglas 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.

The third phase (moving dish training) introduced the subjects to the fact that the dish on the unchosen side moved out of reach. 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 fourth 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. The mixed reward training was identical to the moving dish training, except one of the dishes held no mealworms whereas the other held three mealworms. Subjects completed six of these 30-second trials per session.

Subjects progressed to a new phase after consuming at least 70% of the mealworms offered in their current phase for three consecutive days. Subjects could also regress 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 11:00-15:00 CT. The subjects were not on a restricted diet and were fed for the day directly after completing their respective test trials. 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 triggered de-bias training. If they succeed, they continued to the experimental trials.

At the beginning of each trial, the experimenter placed the appropriate number of mealworms 2.5 cm apart in each dish out of sight from the birds. 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 the mealworms within the chosen dish (see Supplementary Video). If the subject did not make a choice and/or finish all mealworms within three minutes, we ended the session. Once the subject consumed all mealworms, we immediately started the next trial. The first trial in which the subject did not finish all their chosen mealworms triggered a stop for that day's session. Subjects completed on average 3.25 trials per session for replicate 1 and 2.75 trials per session in replicate 2 with one session per day. Subjects ran in a randomized order each day. Each bird experienced 10 repetitions for each of the 15 numerical pairs between 1 and 6 (Table 1). The pairs were organized into blocks with one instance of each pair per block and order was randomly shuffled within each block.

Side-Bias Protocol

During habituation, three consecutive days of no choices and/or not eating mealworms on one 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. De-biasing training consisted of three mealworms placed in the dish the subject avoided and no mealworms in the side they preferred. The subject was allowed 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 removed the trays from 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 for the social experiment was a Y-maze formed out of chicken wire, plastic sheets, and Plexiglas (Figure 2). 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, transparent guillotine-style doors were closed after the bird walked or flew past them, thus making a choice between the option on the left or right. At the end of each arm, was a large bird cage (72 × 48 × 48 cm) housing the conspecific birds. Each cage had two lengthwise perches for the conspecific birds to use and one small perch hanging from the top.

Figure 2

Social Experiment Apparatus
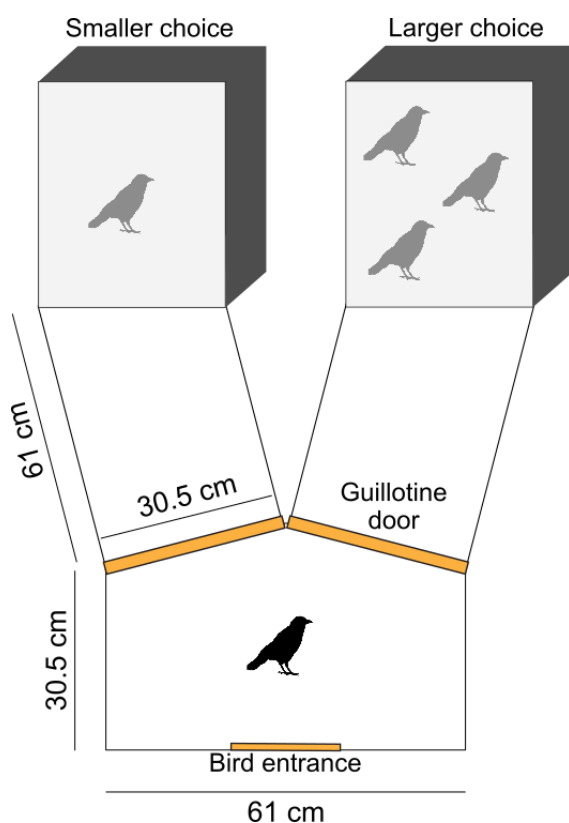
Note. Subjects were held inside the main chamber of the Y maze and shown each group of conspecifics for six seconds a piece. They were then released into the chamber to make a choice. Once the subject crossed under a door, the doors were closed and the bird stayed with their chosen conspecific(s) for three minutes. Figure used with permission under a CC-BY4.0 license: Wolff et al. (2022); available at https://doi.org/10.31234/osf.io/kxgwt).

Habituation and Training

Prior to experimentation, we habituated all birds to both the experimental room and the apparatus to ensure that neither the handling of the birds nor anything within the room would cause our subjects stress. For habituation, we attached a food 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, opened 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 closed behind them, and the bird could explore the chosen arm and consume 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 one habituation session per day for five 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 signs of a side bias. Depending on the bird, this took between 4-6 weeks.

Experimental Procedure

All experimental sessions ran between 09:00-17:00 CT, with birds being run once or twice a day depending on personnel. The subjects were not food restricted. During replicate 1, two experimenters were present at each session: the 'handler' handled the subject, while the 'recorder' handled the camera and the guillotine doors. The experimenter placed the conspecific 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 (see Supplementary Video). Once the subject crossed the threshold of one of the doors, the recorder closed both doors. 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. During replicate 2, the experimenter completed the duties of both handler and recorder.

For replicate 1, each subject experienced five repetitions for each of the 15 numerical pairs between 1 and 6 (Table 1). The side of the larger option was pseudo-randomized with no left or right runs longer than three in a row. The pairs were organized into blocks with one instance of each pair per block and pairs randomly shuffled within each block. The order in which the subjects ran in a particular day was also randomly shuffled. Subjects experienced one trial per session and one session per day. Conspecific birds participated in a mean of 308.8 and a median of 350.0 sessions across the replicate.

For replicate 2, each bird experienced 10 repetitions for each of the 15 numerical pairs between 1 and 6 except for the four pairs that required more than eight conspecific birds (4:5, 3:6, 4:6, 5:6). This was done in an effort to better account for individual bird preference among the subjects for certain conspecific birds. Randomization was the same as replicate 1. Subjects experienced one trial per session and one to two sessions per day. Conspecific birds participated in a mean of 825.0 and a median of 825.0 sessions across the replicate.

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 conspecific 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.

Inter-Rater Reliability

All data were live coded during sessions and also video recorded. JRS recoded choices from video recordings for about 5% of the sessions for each of the four studies. Although not naive to the hypotheses tested, JRS did not collect any of the original data and obscured the stimuli such that he could not view how many items were on each side when recoding choices. We used Cohen's kappa to assess the inter-rater

reliability of the binary response variable for the side of choice (right or left). For the food data, the recoded trials were in complete agreement with the original data ($\kappa = 1.00$, $N = 76$). For the social data, the agreement was very good ($\kappa = 0.94$, 95% CI [0.87, 1.00], $N = 71$).

Data Analysis

Data were analyzed and processed for the project using R (Version 4.3.2; R Core Team, 2023) and the R-packages *BayesFactor* (Version 0.9.12.4.6; Morey & Rouder, 2024), *bayestestR* (Version 0.13.1; Makowski et al., 2019), *ggcorrplot* (Version 0.1.4.1; Kassambara, 2023), *here* (Version 1.0.1; Müller, 2020), *lme4* (Version 1.1.35.1; Bates et al., 2015), *papaja* (Version 0.1.2; Aust & Barth, 2023), *patchwork* (Version 1.1.3; Pedersen, 2024), *performance* (Version 0.10.8; Lüdecke et al., 2021), *psych* (Version 2.3.12; Revelle, 2023), and *tidyverse* (Version 2.0.0; Wickham et al., 2019). The manuscript was created using *rmarkdown* (Version 2.25, Xie et al., 2018) and *papaja* (Version 0.1.2, Aust & Barth, 2023). Data, analysis scripts, supplementary materials, and reproducible research materials are available at the Open Science Framework (https://osf.io/g45nk/).

We drew inferences based on Bayes factors (BF_{10}), which offer the ratio of evidence for the alternative hypothesis (H_1) over evidence for the null hypothesis (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 three times more evidence for H_0 than H_1. We interpreted Bayes factors based on Wagenmakers et al. (2018), where a $BF_{10} > 3$ is moderate evidence for the alternative hypothesis, $BF_{10} < 1/3$ is moderate evidence for the null hypothesis, and $1/3 < BF_{10} < 3$ indicate neither hypothesis has sufficient evidence supporting it (suggesting the sample size is too small to draw conclusions). In addition, Bayes factors of 10 (1/10), 30 (1/30), and 100 (1/100) represent strong, very strong, and extreme evidence, respectively.

Prior to analysis, we transformed the left and right choice variable from each trial into a binary outcome, 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 numerical pair by subtracting the larger number from the smaller ($6 - 1 = 5$), as well as created the ratio by dividing the smaller by the larger number ($1/6 = 0.17$). Our hypotheses explore the relationship between our binary outcome variable (choice of the larger or smaller stimuli) and which possible numerical relationships (difference or ratio) predict subjects' 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-test of preference for larger numbers. Therefore, we calculated the mean *percent preference for larger quantities* for each subject across all pairs and used the t-test to compare the subject means to 50. We performed 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 & Rouder, 2024) 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, as the response variable was dichotomous and our subjects repeatedly made decision on the same numerical pairs. We used the trial-level choices for either the larger or smaller option available in the numerical pair as the response variable. To investigate our hypotheses, we used generalized linear (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. After comparing those models, we selected the final best-fitting model for each data set overall using Bayes factors comparing each random effects model to the intercept only model.

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) numerical pair as a random effect (to account for each bird repeatedly seeing each pair multiple times), and (4) both subject and numerical pair

as random effects. For example, the model with both subject and pair as random effects ran using the glmer() function from the *lme4* package (Bates et al., 2015) with the following structure: glmer(choice ~ (1|subject)+ (1|pair), family = binomial) (Table S2). 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 *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 used the following structure: choice ~ difference * ratio (Table S2). We calculated Bayes factors using the test_performance() function from the *performance* package (Lüdecke et al., 2021), which estimates Bayes factors by comparing model BIC values to the intercept only model BIC values using Wagenmakers' (2007) equation. The model with the highest Bayes factor is the best-fitting model. We then tested the best-fitting model for outliers and for a linear relationship between the predictor and the logit of the response variable. All models presented here satisfy these assumptions.

We tested the second hypothesis by assessing whether the fixed effect models with difference only or ratio only outperformed the intercept only model. That is, we inferred that difference or ratio influences choice if the Bayes factors for their respective models exceed 3. We tested the third hypothesis by assessing whether the difference + ratio or difference * ratio models outperformed all other models. If pinyon jays choose more items based on differences and ratios independently of each other, then the best-performing model should include both difference and ratio. Evidence in support of either of these models would indicate that both difference and ratio predict choice independently.

Results

Food Experiment

Our first hypothesis predicted that subjects ($N_{rep1} = 8$, $N_{rep2} = 4$) would on average choose the larger quantity of mealworms over the smaller quantity across all the numerical pairs in the food preference task. Data for both replicates satisfied the Shapiro-Wilk normality test (Replicate 1: $W = 0.97$, $p = .905$, Replicate 2: $W = 0.84$, $p = .202$), so we proceeded to use parametric statistics. One-sample t-tests provided moderate to strong evidence that preferences were above chance (50%) in both replicate 1 ($M = 60.8$, 95% CI [55.5, 66.0], $t(7) = 4.8$, $p = .002$, $BF_{10} = 24.3$) and replicate 2 ($M = 62.4$, 95% CI [55.5, 69.3], $t(3) = 5.7$, $p = .011$, $BF_{10} = 6.7$).

Our second and third hypotheses were investigated by comparing generalized linear models that included difference and/or ratio as predictors. Model comparisons allow us to investigate both the second hypothesis, that subjects would show stronger preferences for more items when the numerical pairs have higher differences and lower ratios, and the third hypothesis, that difference and ratio will influence preference independently of each other. For both replicates, the best-fitting random effect structure was no random effect structure, so we used no random effects in the subsequent fixed effect analysis. For the fixed effect model comparison, the model with only the main effect of ratio best fit both data sets (Replicate 1: $BF_{10} = 1.9 \times 10^3$, Replicate 2: $BF_{10} = 3.8$). Thus, subjects in the food experiment used the ratio between the two numbers of mealworms to choose between options, with stronger preferences for larger options at smaller ratios (Figure 3). The model with the main effect of difference showed evidence supporting stronger preferences for larger options at larger differences for replicate 1 but not replicate 2 (Replicate 1: $BF_{10} = 459.7$, Replicate 2: $BF_{10} = 0.37$). Consequently, this only partially supports our second hypothesis. Additionally, our third hypothesis was not supported, as the models including both difference and ratio were outperformed by models only including ratio (see Table S3 for full results).

Figure 3

Food Experiment Effects of Difference and Ratio on Choice
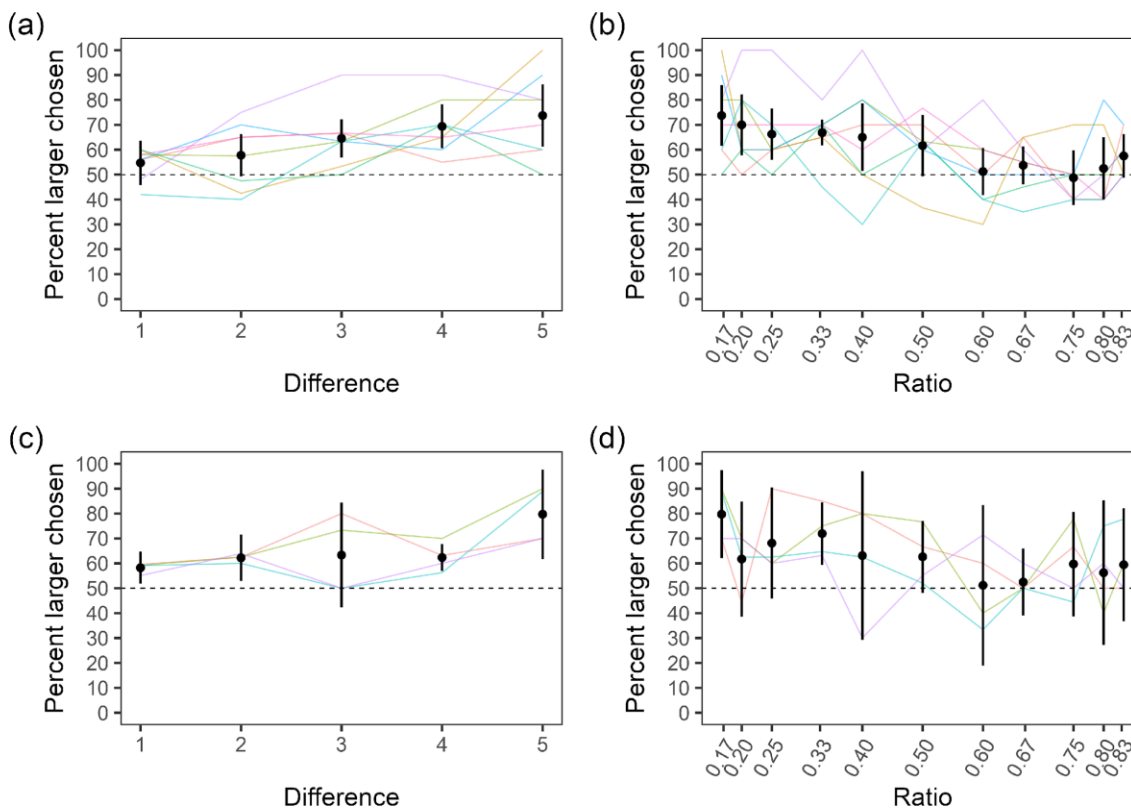
Note. Mean percent preference for the larger option is shown on the y-axis with the numerical difference or ratio values on the x-axis. (a) Preference for larger per difference in replicate 1. (b) Preference for larger per ratio in replicate 1. (c) Preference for larger per difference in replicate 2. (d) Preference for larger per ratio in replicate 2. Dots represent mean values across subjects and trials. Error bars represent 95% within-subject confidence intervals. Lines connect individual subject means. Figure used with permission under a CC-BY4.0 license: Wolff et al. (2022); available at https://doi.org/10.31234/osf.io/kxgwt.

Exploratory Analyses

Because of potential variability between replicates caused by small sample sizes, we conducted an exploratory analysis in which we combined the replicate 1 and 2 data (N = 12) and ran the same analyses as described above. First, this combined data did not show evidence for or against differences in preferences between sexes ($M = -0.01$, 95% CI [-0.14, 0.12], $t(2.18) = -0.24$, $p = .831$, $BF_{10} = 0.55$), with data meeting assumptions of normality ($W = 0.99$, $p = .999$) and homogeneity of variance ($F(1) = 0.67$, $p = .433$). Second, this analysis corroborated replicate 1's analysis with preferences for larger options above chance ($M = 61.3$, 95% CI [57.7, 64.8], $t(11) = 7.0$, $p < .001$, $BF_{10} = 1.1 \times 10^3$), models with either difference or ratio outperforming models without them (difference: $BF_{10} = 82.9$; ratio: $BF_{10} = 3.3 \times 10^3$), and the ratio only model outperforming models with both difference and ratio included (Table S3). The only difference from replicate 1's results was that numerical pair was included as a random effect because it was the best-fitting random effects model.

Difference and ratio are necessarily correlated, and it can be difficult to statistically separate their effects since collinear predictors can affect coefficient estimates (Belsley et al., 1980). Because we were expressly interested in comparing these predictors, we could not use the typical solution for collinearity, which is to drop the predictor with the highest variance inflation factor (Belsley et al., 1980). Instead, as an exploratory analysis, we took a subsample of our numerical pairs that had the same ratio (.5) but various

differences (1, 2, 3) to determine whether difference affected preferences independent of ratio. Comparing a generalized linear model with binary error distribution (logistic regression) with preference as a function of difference to a comparable intercept-only model, we observe that the intercept-only model fit better (BF_{10} = 0.07). Therefore, difference does not account for preference when holding ratio constant in this subset of data.

Social Experiment

Hypothesis 1 predicted that subjects (N_{rep1} = 10, N_{rep2} = 10) would choose the larger quantity of conspecifics over the smaller in the social preference task. Data for both replicates satisfied the Shapiro-Wilk normality test (Replicate 1: W = 0.90, p = .233, Replicate 2: W = 0.88, p = .148), so we proceeded to use parametric statistics. One-sample t-tests provided evidence that our hypothesis was strongly supported in replicate 1 (M = 55.3, 95% CI [52.3, 58.3], $t(9)$ = 4.0, p = .003, BF_{10} = 16.7) but not supported in replicate 2 (M = 50.2, 95% CI [47.1, 53.3], $t(9)$ = 0.1, p = .896, BF_{10} = 0.31).

For hypotheses 2 (the presence of difference and ratio effects) and hypothesis 3 (independent difference and ratio effects), we again used model selection (Table S4). Models with no random effect structure performed best, so we included no random effects in the fixed effects analysis. For fixed effects, the intercept only model best fit the data (Replicate 1: BF_{10} = 0.07, Replicate 2: BF_{10} = 0.06), suggesting that neither ratio nor difference influenced choice (Figure 4). Because no model other than the intercept only had evidence suggesting that it was true, neither hypotheses 2 nor 3 was supported by the data.

Exploratory Analyses

We conducted an exploratory analysis in which we combined the replicate 1 and 2 data (N = 20) and ran the same analyses as described above. First, this combined data did not show evidence for or against differences in preferences between sexes (M = 0.03, 95% CI [-0.01, 0.07], $t(14.45)$ = 1.77, p = .099, BF_{10} = 0.74), with data meeting assumptions of normality (W = 0.95, p = .394) and homogeneity of variance ($F(1)$ = 1.98, p = .177). Second, this analysis found intermediate results between the two replicate analyses with insufficient evidence to support the presence or absence of an overall preference for larger options above chance (M = 52.8, 95% CI [50.5, 55.1], $t(19)$ = 2.5, p = .022, BF_{10} = 2.7). The results for hypothesis 2 and 3 corroborated the individual analyses with no effects of either difference or ratio on preference (difference: BF_{10} = 0.08; ratio: BF_{10} = 0.05; Table S4).

As another exploratory analysis mirroring what we did with the food experiment data, we assessed the effect of difference for a subset of the numerical pairs with the same ratio (0.5). Comparing our model with preference as a function of difference to a comparable intercept-only model, we observe that the intercept-only model fits better (BF_{10} = 0.06). Therefore, difference does not account for preference when holding ratio constant.

The complete lack of difference or ratio effects on preference for conspecifics surprised us. Therefore, upon completion of our previous confirmatory analyses, we conducted an exploratory analysis of subject bird preferences for individual conspecific birds. We began by determining the percentage of sessions that each subject chose the side associated with each conspecific bird. For example, the subject bird Juniper experienced 12 sessions in which Hermia was included in one of the groups, but Juniper only chose the group that included Hermia 3 times, resulting in a percentage of 25%. Similarly, Juniper chose the group with Bruno 3 out of 4 sessions that included Bruno, resulting in a percentage of 75%. Figure 5 illustrates these percentages with subject and conspecific sex indicated. We used more conspecific birds in replicate 1 than replicate 2, so birds met fewer times in replicate 1 but met more birds overall (Table S5).

Figure 4

*Social Experiment Effects of Difference and Ratio on Choice*
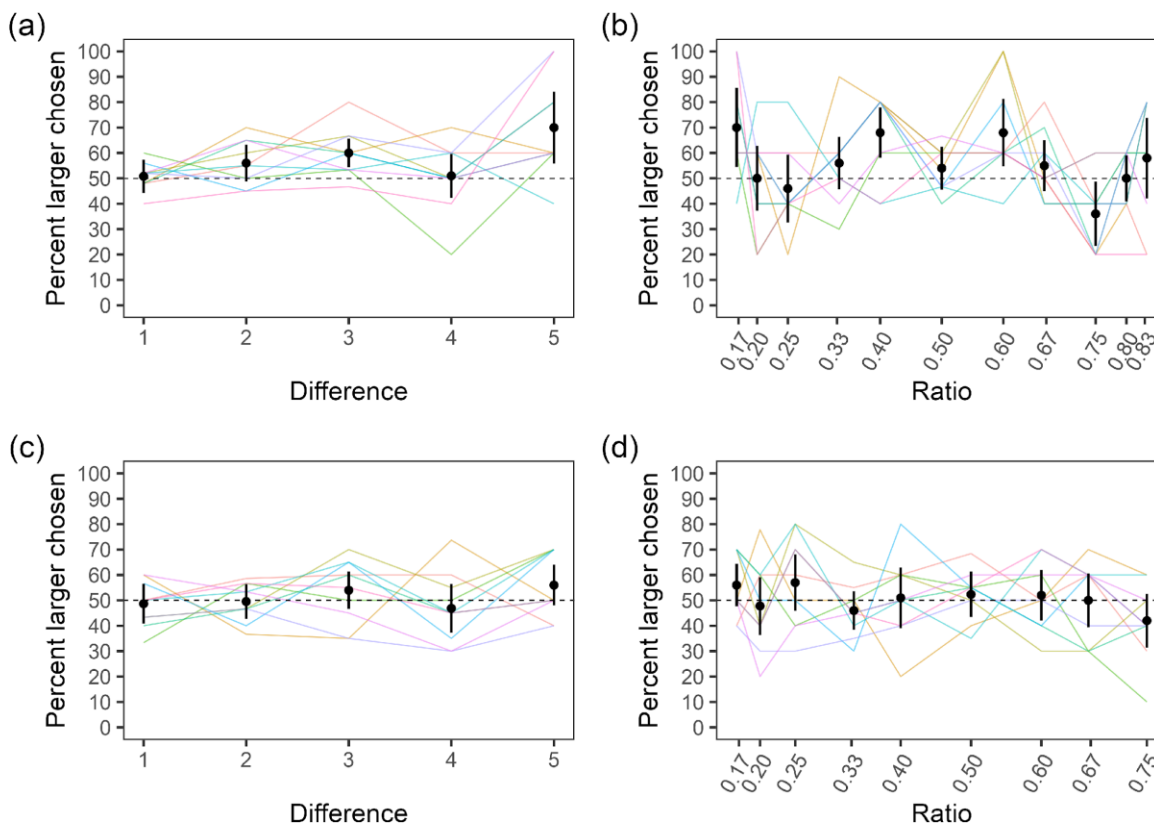
Note. Mean percent preference for the larger option is shown on the y-axis with the numerical difference or ratio values on the x-axis. (a) Preference for larger per difference in replicate 1. (b) Preference for larger per ratio in replicate 1. (c) Preference for larger per difference in replicate 2. (d) Preference for larger per ratio in replicate 2. Dots represent mean values across subjects and trials. Error bars represent 95% within-subject confidence intervals. Lines connect individual subject means. Figure used with permission under a CC-BY4.0 license: Wolff et al. (2022); available at https://doi.org/10.31234/osf.io/kxgwt.

We calculated how often each conspecific bird was chosen by aggregating over each sex of subject birds to illustrate how often male and female subject birds chose each conspecific bird (Table S5). For instance, the most preferred conspecific was the female, Sapphire, who was chosen 67.6% of the time she was on offer for male subject birds and 62.5% of the time for female subject birds. Whereas we see clear individual preferences, these choices do not differ depending on the sex of the conspecific. Male subjects do not differ in their preferences between male or female conspecifics ($M = -0.1$, 95% CI [-2.6, 2.3], $t(14) = -0.1$, $p = .909$, $BF_{10} = 0.26$). There were not enough female subjects to draw conclusions about possible sex differences ($M = 2.3$, 95% CI [-5.4, 10.0], $t(4) = 0.8$, $p = .450$, $BF_{10} = 0.52$). Thus, identity matters for individual conspecifics, but we do not have evidence for preferences based on the sex of potential flock mates.

Figure 5

Preference for Individual Conspecifics Per Subject
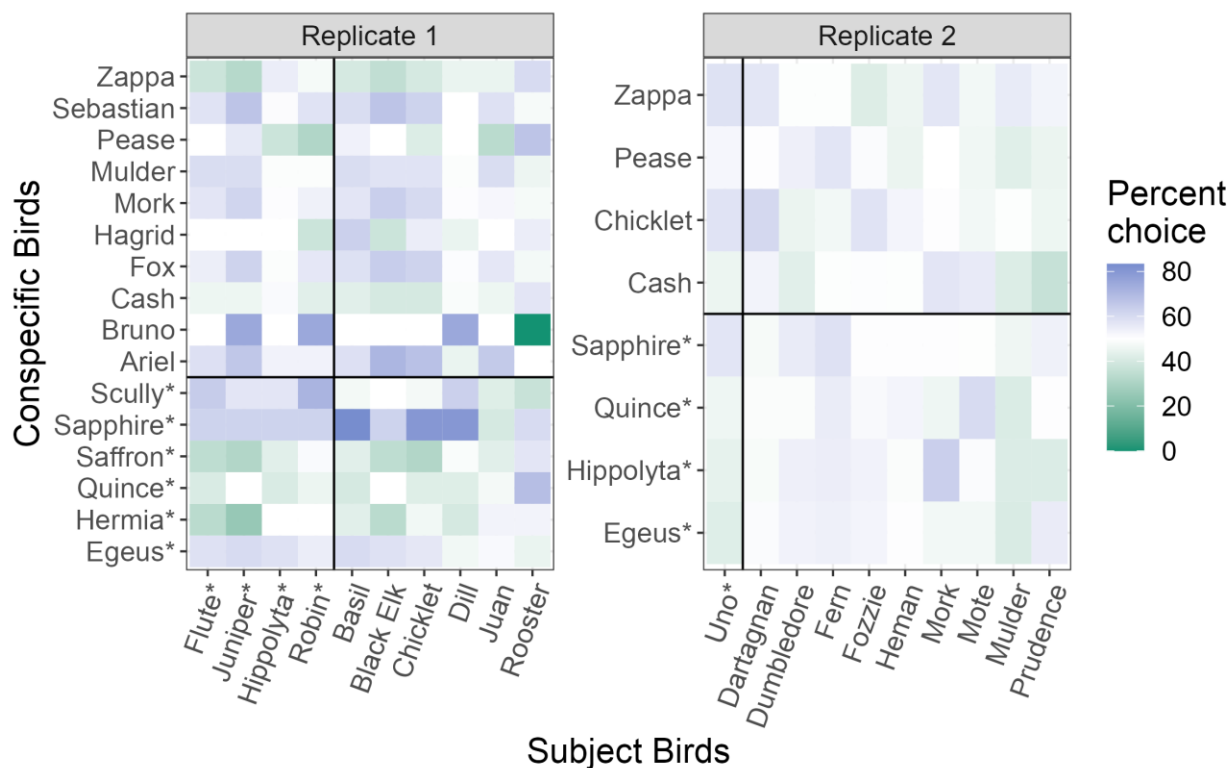
Note. Squares represent a subject/conspecific pairing, and the color illustrates the percentage of times the group including that conspecific was chosen by that subject. Blue represents a preference *for* a particular conspecific, green represents a preference *against* a particular conspecific, and white represents no preference. Asterisks note female subjects and conspecifics. Lines divide quadrants of male/female, male/male, female/female, female/male pairs. Figure used with permission under a CC-BY4.0 license: Wolff et al. (2022); available at https://doi.org/10.31234/osf.io/kxgwt.

Discussion

We examined pinyon jays' quantitative abilities to choose between different quantities of food items and conspecifics. Across all numerical pairs, birds chose the larger of the two options in the food experiment but not in the social experiment, partially confirming our first hypothesis. In the food experiment, smaller numerical ratios predicted the birds' choices for larger amounts, but larger differences predicted birds' choices only in the first replicate, partially confirming our second hypothesis. In the social experiment, neither ratio nor difference predicted choice, contradicting our second hypothesis. In both the food and social experiments, difference and ratio did not independently predict choice, contradicting our third hypothesis.

In the food experiment, our pinyon jays preferred larger over smaller quantities more as the numerical ratios decreased, which aligns with previous corvid research demonstrating a numerical ratio effect in both quantity preference tasks (Tornick et al., 2015; Kelly, 2016) and numerosity discrimination tasks (Ditz & Nieder, 2016). This provides evidence for pinyon jays using the approximate number system when judging food quantities. Whereas we did find evidence for the numerical distance effect in our first replicate, it seemed to be driven primarily by the numerical ratio effect, supporting the ratio-dependence of Weber's Law. When ratio was held constant, difference did not influence choice, matching findings by Tornick et al. (2015) on nutcrackers. Ditz and Nieder (2016) and Kelly (2016), however, found both difference and ratio effects.

In the social experiment, neither difference nor ratio predicted choice, suggesting that pinyon jays do not use quantity information in the same way across object types. This outcome is surprising, as previous quantification tasks with conspecifics in fish found effects of difference and ratio (Agrillo & Dadda, 2007; Agrillo et al., 2008). Our results suggest that how pinyon jays use quantity information to make decisions differs across contexts. Both foraging techniques and flock size 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 avoidance and sexual selection in the form of mate preference. Joining a larger flock allows an animal to dilute their chances of being eaten by predators (i.e., the dilution effect) but also provides a larger pool of potential mates. For food items and predation risk, quantity matters. But for mate preference or other social preferences, whereas quantity may be important (e.g., to increase the pool of potential partners), the *identity* of the partners also matters.

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 selection of one option over the other. Thus, the lack of a preference between two quantities 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.

One possible explanation for the lack of a ratio or difference effect for the social preference task is that individual identity of birds overrides the importance of quantity. That is, the birds may be able to *discriminate* between different quantities of conspecifics based on ratio and/or difference, but their *preferences* do not reflect this because additional factors come into play. An exploratory follow-up analysis of our data showed wide variation in preferences for groups that contained individual conspecific birds (Figure 5). Interestingly, choices did not differ depending on the sex of the conspecific (Table S5). Pinyon jays have complex, long-term bonds with other flock members and mates (Marzluff & Balda, 1992), which may make identity of group mates more important than sheer quantities. Moreover, the birds in our studies did not experience signals of predation danger during the experiment. Without pressure to dilute risk in larger groups, social partner identity may have overridden quantity to determine choice.

Limitations

Our study design does not allow us to pinpoint the exact features by which the birds make these quantitative choices. For the food preference tasks, the birds may choose larger *numbers* of individual food items or larger *amounts* of them. Using number involves tracking the quantity of individuated objects. However, in many cases, animals choose based on amount, which refers to other measures or proxies of quantity such as item size, surface area, volume, perimeter, and density (Menzel, 1960; Uller et al., 2003; Stevens et al., 2007; 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. For our purposes, the method of preference is not critical. In the food experiment, total quantity is a reasonable criterion because the birds' ultimate goal is to obtain as much food as possible to stay alive (e.g., two large mealworms may include more calories than three small mealworms). Therefore, choosing based on total calories or overall food intake, rather than absolute number, could be a better evolutionary strategy.

The birds also may not have tracked number per se for conspecifics but rather some kind of aggregate amount like area. Although quantity is probably more important than number for food, absolute number can be more important for aspects of social preferences. Assuming a predator attacks a single individual at random, each member of a group has a $1/N$ chance of being attacked. Thus, the absolute number of individuals in the group determines an individual's probability of being attacked. So, number itself is critical to choosing groups to survive in a predator avoidance situation. Similarly, when assessing group size for the purpose of mate selection, every additional conspecific increases the pool of potential

mates. Again, absolute number is relevant. However, though large group sizes are beneficial for predator avoidance and mate selection, they are detrimental to food and mate competition. Larger groups means more competition, and absolute number is relevant here because competition is a function of resource amount per individual. Thus, absolute number is important for both the benefits and costs of being in a large group. Yet if most individuals are roughly the same size—which tends to be the case for adult pinyon jays (Ligon, 1978)—then quantity may be a sufficient proxy for number. That is, given the more sophisticated cognition required to track number, tracking quantity may be a simpler (in terms of cognition) and sufficient (in terms of evolutionary fitness) way to assess group size. Future work is needed to tease apart which methods animals use when presented with food or conspecifics. This work requires carefully controlled stimuli that can account for item size, surface area, volume, perimeter, and density (Zanon et al., 2022).

None of our studies found difference effects independent of ratio, even in exploratory analyses in which we held ratio constant and tested effects of difference. 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 ($6 - 1 = 5$). In previous discrimination studies, differences ranged from 8 to 30 (Ditz & Nieder, 2016). Considering that pinyon jay flock sizes range from 50 to 500 (Marzluff & Balda, 1992), the limited range of difference options may not have provided enough opportunities to demonstrate effects of difference. Moreover, because difference and ratio are correlated, future tests of their independent contributions to preference should involve choosing numerical pairs with fixed differences and ratios.

Another reason we might have not found ratio or difference effects in the social experiment is the differing reinforcement levels. Possibly, time with conspecifics in a social setting is not as reinforcing as eating mealworms, leading to less cost for a bird's choice. Therefore, rather than the context of food or social interaction driving differences in preferences, the motivation to choose the larger might have differed. Additionally, the lack of results in the social experiment may result from subjects using the collective movement of conspecific groups as a cue. Subjects might have relied on amount of movement rather than the actual quantity of birds in each group for their choices. Previous studies on fish have demonstrated that the preference for shoal size is influenced by stimulus motion (Messina et al., 2022). Subjects might have relied on amount of movement rather than the actual quantity of birds in each group for their choices.

Additionally, our relatively low sample size with an aging study population could have also contributed to our findings. Considering our sample size ($N = 21$), we completed a within-subject study design offering multiple trials per pair to account for individual subject variation. Our subjects were about twice as old as free-ranging pinyon jays' average lifespan (Marzluff & Balda, 1992). Possibly, quantitative cognitive abilities degrade with age.

Conclusion

This research investigated how pinyon jays assess quantities of food items and conspecifics in preference tasks. For food items, numerical ratio predicted their choices in both replicates whereas numerical difference predicted choice in only the first replicate, but neither ratio nor difference predicted choices in the social experiment. Although quantity is important for selecting food items, additional factors such as conspecific identity may be more important for selecting social groups to join. Thus, in quantification situations, the type of objects to be quantified may drive how animals use different information to make quantity decisions. Furthermore, many adaptive problems beyond foraging require sensitivity to quantities, and we encourage further exploration of quantitative cognition of non-food objects.

Author Contributions: Wolff: Conceptualization, Data Curation, Formal Analysis, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft

Preparation, Writing-review & editing. Trevino: Investigation, Project Administration, Resources, Visualization, Writing – Review & Editing. Stevens: Conceptualization, Formal Analysis, Funding Acquisition, Methodology, Investigation, Software, Supervision, Validation, Visualization, Writing – review & editing.

Conflict of Interest: The authors declared that no conflicts of interest exist.

Data Availability: The data and analysis code are available at: https://osf.io/g45nk/.

References

Agrillo, C., & Beran, M. J. (2013). Number without language: Comparative psychology and the evolution of numerical cognition. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2013.00295

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., & Bisazza, A. (2007). Quantity discrimination in female mosquitofish. *Animal Cognition*, *10*(1), 63–70. https://doi.org/10.1007/s10071-006-0036-5

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*, 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. (2023). *papaja: Create APA manuscripts with R Markdown*. https://github.com/crsh/papaja

Balda, R. P., & Kamil, A. C. (1998). The ecology and evolution of spatial memory in corvids of the southwestern USA: The perplexing pinyon jay. In R. P. Balda, I. M. Pepperberg, & A. C. Kamil (Eds.), *Animal cognition in nature: The convergence of psychology and biology in laboratory and field* (pp. 29–64). Academic Press.

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

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. John Wiley & Sons, Ltd. https://doi.org/10.1002/0471725153.ch1

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

Bisazza, A., & Santacà, M. (2022). Zebrafish excel in number discrimination under an operant conditioning paradigm. *Animal Cognition*, *25*(4), 917–933. https://doi.org/10.1007/s10071-022-01602-y

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

Emmerton, J. (2001). Birds' judgments of number and quantity. In R. G. Cook (Ed.), *Avian visual cognition*. http://www.pigeon.psy.tufts.edu/avc/emmerton/

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

Fechner, G. T. (1860). *Elemente der psychophysik*. Breitkopf und Härtel.

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

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. (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

Hamilton, W. D. (1971). Geometry for the selfish herd. *Journal of Theoretical Biology*, *31*(2), 295–311. https://doi.org/10.1016/0022-5193(71)90189-5

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

Kassambara, A. (2023). *ggcorrplot: Visualization of a correlation matrix using "ggplot2"*. https://CRAN.R-project.org/package=ggcorrplot

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

Kitchen, D. M. (2004). Alpha male black howler monkey responses to loud calls: Effect of numeric odds, male companion behaviour and reproductive investment. *Animal Behaviour*, *67*, 125–139. https://doi.org/10.1016/j.anbehav.2003.03.007

Krause, J., & Ruxton, G. D. (2002). *Living in groups*. Oxford University Press.

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

Ligon, J. D. (1978). Reproductive interdependence of piñon jays and piñon pines. *Ecological Monographs*, *48*(2), 111–126. https://doi.org/10.2307/2937295

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

Lyons, I. M., Nuerk, H.-C., & Ansari, D. (2015). Rethinking the implications of numerical ratio effects for understanding the development of representational precision and numerical processing across formats. *Journal of Experimental Psychology: General*, *144*, 1021–1035. https://doi.org/10.1037/xge0000094

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*. Academic Press.

McComb, K., Packer, C., & Pusey, A. E. (1994). Roaring and numerical assessment in contests between groups of female lions, *Panthera leo*. *Animal Behaviour*, *47*(2), 379–387. https://doi.org/10.1006/anbe.1994.1052

Mendelson, T. C., Fitzpatrick, C. L., Hauber, M. E., Pence, C. H., Rodríguez, R. L., Safran, R. J., Stern, C. A., & Stevens, J. R. (2016). Cognitive phenotypes and the evolution of animal decisions. *Trends in Ecology & Evolution*, *31*(11), 850–859. https://doi.org/10.1016/j.tree.2016.08.008

Menzel, E. W. (1960). Selection of food by size in the chimpanzee, and comparison with human judgments. *Science*, *131*, 1527–1528. https://doi.org/10.1126/science.131.3412.1527

Messina, A., Potrich, D., Perrino, M., Sheardown, E., Miletto Petrazzini, M. E., Luu, P., Nadtochiy, A., Truong, T. V., Sovrano, V. A., Fraser, S. E., Brennan, C. H., & Vallortigara, G. (2022). Quantity as a fish views it: Behavior and neurobiology. *Frontiers in Neuroanatomy*, *16*. https://doi.org/10.3389/fnana.2022.943504

Morey, R. D., & Rouder, J. N. (2024). *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

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. (2024). *patchwork: The composer of plots*. https://CRAN.R-project.org/package=patchwork

Potrich, D., Sovrano, V. A., Stancher, G., & Vallortigara, G. (2015). Quantity discrimination by zebrafish (*Danio rerio*). *Journal of Comparative Psychology*, *129*(4), 388–393. https://doi.org/10.1037/com0000012

Potrich, D., Zanon, M., & Vallortigara, G. (2022). Archerfish number discrimination. *eLife*, *11*, e74057. https://doi.org/10.7554/eLife.74057

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Revelle, W. (2023). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University. https://CRAN.R-project.org/package=psych

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

Stevens, J. R., & Soh, L.-K. (2018). Predicting similarity judgments in intertemporal choice with machine learning. *Psychonomic Bulletin & Review*, *25*(2), 627–635. https://doi.org/10.3758/s13423-017-1398-1

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

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*, *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 of p 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., 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

Wilson, M. L., Hauser, M. D., & Wrangham, R. W. (2001). Does participation in intergroup conflict depend on numerical assessment, range location or rank for wild chimpanzees? *Animal Behaviour*, *61*, 1203–1216.

Wolff, L., Carey, K., & Stevens, J. R. (2022). *Friends aren't food: Pinyon jays show context-dependent numerical cognition*. PsyArXiv. https://doi.org/10.31234/osf.io/kxgwt

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

Xie, Y., Allaire, J. J., & Grolemund, G. (2018). *R markdown: The definitive guide*. Chapman; Hall/CRC. https://bookdown.org/yihui/rmarkdown

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

Zanon, M., Potrich, D., Bortot, M., & Vallortigara, G. (2022). Towards a standardization of non-symbolic numerical experiments: GeNEsIS, a flexible and user-friendly tool to generate controlled stimuli. *Behavior Research Methods*, *54*(1), 146–157. https://doi.org/10.3758/s13428-021-01580-y

Supplementary Materials

Table S1

Subject Bird Information

Subject	Sex	Age (years)	Food 1	Food 2	Social 1	Social 2
Uno	Female	12	X			X
Dumbledore	Male	11	X			X
Fern	Male	15	X			X
Fozzie	Male	12	X			X
He-man	Male	12	X			X
Mork	Male	12	X			X
Mote	Male	14	X			X
Prudence	Male	10	X			X
Mulder	Male	11				X
Dartagnan	Male	10		X		X
Saffron	Female	12		X		
Dill	Male	15		X	X	
Rooster	Male	12		X	X	
Flute	Female	14			X	
Hippolyta	Female	14			X	
Juniper	Female	15			X	
Robin	Female	14			X	
Basil	Male	15			X	
Black Elk	Male	10			X	
Chicklet	Male	12			X	
Juan	Male	19			X	

Note. Table used with permission under a CC-BY4.0 license: Wolff et al. (2022); available at https://doi.org/10.31234/osf.io/kxgwt.

Table S2

Random and Fixed Effect Model Structures Tested

Model	Formula
Random effects	
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	
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

Note. Table used with permission under a CC-BY4.0 license: Wolff et al. (2022); available at https://doi.org/10.31234/osf.io/kxgwt .

Table S3

Model Comparison for the Food Experiment

Model	AIC	BIC	BF
Replicate 1 random effects			
intercept only	1609.6	1614.7	1.00
(1|subject)	1608.4	1618.5	0.15
(1|pair)	1606.9	1617.1	0.31
(1|subject) + (1|pair)	1605.2	1620.4	0.06
Replicate 1 fixed effects			
intercept only	1609.6	1614.7	1.00
ratio	1589.5	1599.7	1,874.16
difference	1592.3	1602.5	459.68
difference + ratio	1590.9	1606.2	72.38
difference * ratio	1592.7	1613.0	2.34
Replicate 2 random effects			
intercept only	753.5	757.9	1.00
(1|subject)	755.5	764.2	0.04
(1|pair)	755.2	763.9	0.05
(1|subject) + (1|pair)	757.2	770.2	0.00
Replicate 2 fixed effects			
intercept only	753.5	757.9	1.00
ratio	746.5	755.2	3.79
difference	751.2	759.8	0.37
difference + ratio	747.8	760.8	0.23
difference * ratio	748.6	766.0	0.02

Table S4

Model Comparison for the Social Experiment

Model	AIC	BIC	BF
Replicate 1 random effects			
intercept only	1033.2	1037.8	1.00
(1|subject)	1035.2	1044.4	0.04
(1|pair)	1030.9	1040.1	0.32
(1|subject) + (1|pair)	1032.9	1046.7	0.01
Replicate 1 fixed effects			
intercept only	1033.2	1037.8	1.00
ratio	1033.7	1043.0	0.07
difference	1031.7	1040.9	0.21
difference + ratio	1033.1	1047.0	0.01
difference * ratio	1034.1	1052.6	0.00
Replicate 2 random effects			
intercept only	1524.1	1529.1	1.00
(1|subject)	1526.1	1536.1	0.03
(1|pair)	1526.1	1536.1	0.03
(1|subject) + (1|pair)	1528.1	1543.1	0.00
Replicate 2 fixed effects			
intercept only	1524.1	1529.1	1.00
ratio	1524.9	1534.9	0.06
difference	1525.6	1535.6	0.04
difference + ratio	1526.7	1541.7	0.00
difference * ratio	1528.7	1548.7	0.00

Note. Table used with permission under a CC-BY4.0 license: Wolff et al. (2022); available at
https://doi.org/10.31234/osf.io/kxgwt.

Table S5

Choice for Individual Conspecifics

Conspecific	Age	Number of trials	Female choice %	Male choice %	Overall choice %
Replicate 1					
Male					
Zappa	11	400	43.1	44.2	43.8
Cash	11	350	46.4	45.7	46.0
Pease	14	141	43.8	49.4	46.8
Hagrid	11	179	46.9	51.3	49.7
Bruno	19	40	62.5	45.8	52.5
Mulder	11	570	54.4	55.0	54.7
Mork	12	630	56.0	55.3	55.6
Fox	11	530	55.7	56.6	56.2
Comanche	10	21	62.5	40.0	57.1
Sebastian	19	464	58.3	57.4	57.8
Ariel	19	406	57.9	59.5	58.9
Female					
Saffron	12	350	40.0	42.9	41.7
Hermia	14	134	39.6	45.3	43.3
Quince	14	206	44.3	48.3	46.6
Scully	11	163	62.5	47.7	52.8
Egeus	14	600	57.9	53.1	55.0
Sapphire	12	66	62.5	67.6	65.2
Replicate 2					
Male					
Cash	11	770	45.5	48.5	48.2
Pease	14	818	52.4	49.2	49.5
Zappa	11	790	58.2	50.6	51.4
Chicklet	12	828	57.8	50.7	51.4
Female					
Egeus	14	828	42.2	50.3	49.5
Quince	14	848	45.9	50.6	50.1
Hippolyta	14	818	43.9	50.8	50.1
Sapphire	12	888	57.3	51.4	52.0