Fractionating choice: A study on reward discrimination, preference and relative valuation in the rat (*Rattus norvegicus*)

Joshua M. Ricker, **Justin D. Hatch**, **Daniel D. Powers**, and **Howard C. Cromwell**

J.P. Scott Center for Neuroscience, Mind and Behavior and Department of Psychology, Bowling Green State University, Bowling Green Ohio, 43403

Abstract

Choice behavior combines discrimination between distinctive outcomes, preference for specific outcomes and relative valuation of comparable outcomes. Previous work has focused on one component (i.e., preference) disregarding other influential processes that might provide a more complete understanding. Animal models of choice have been explored primarily utilizing extensive training, limited freedom for multiple decisions and sparse behavioral measures constrained to a single phase of motivated action. The present study used a paradigm that combines different elements of previous methods with the goal to distinguish among components of choice and explore how well components match predictions based on risk-sensitive foraging strategies. In order to analyze discrimination and relative valuation, it was necessary to have an option that shifted and an option that remained constant. Shifting outcomes among weeks included a change in single-option outcome (0 to 1 to 2 pellets) or a change in mixed-option outcome (0 or 5 to 0 or 3 to 0 or 1 pellets). Constant outcomes among weeks were also mixedoption (0 or 3 pellets) or single-option (1 pellet). Shifting single-option outcomes among weeks led to better discrimination, more robust preference and significant incentive contrast effects for the alternative outcome. Shifting multi-options altered choice components and led to dissociations among discrimination, preference, and reduced contrast effects. During extinction, all components were impacted with the greatest deficits during the shifting mixed-option outcome sessions. Results suggest choice behavior can be optimized for one component but suboptimal for others depending upon the complexity of alterations in outcome value between options.

Keywords

There are a wide range of experimental designs used to investigate choice in animals and humans (Kahneman and Tversky, 1984; Rachlin et al., 1986; Real, 1991; de Matta et al., 2012). A majority of paradigms focus on preference and primarily examine either appetitive, instrumental behavior or consummatory measures separately (Kacelnik and Mouden, 2013). Appetitive behavior is defined as the behavior involved in acquiring the outcome while

Correspondence to: Howard Casey Cromwell, PhD, Department of Psychology, Bowling Green State University, Bowling Green, OH, 43403, Tel: (419) 372-9408, Fax: (419) 372-6013, ; Email: hcc@bgsu.edu.

consummatory behavior involves intake or terminal action with the outcome. In general, previous work neglects how different components of choice might interact, diverge, or converge during motivated action and lacks an ability to understand how different phases of motivation are influenced during the same choice. A major objective of the present study was to examine different components of choice while animals experience a relatively more open environment, one that simulates 'foraging choice' (Zabludoff et al., 1988). The present work focused on crucial, diverse components including discrimination, preference and relative reward valuation. Discrimination requires identifying a difference between alternatives. Discrimination is an essential part of work in sensation and perception but the process is equally critical in the production of motivated action including choice behavior (Watanabe et al., 2001; Peterson and Trapold, 1982; Michals, 1957). Discrimination becomes more demanding as stimulus or outcome properties converge (Shepard, 1987; Lawrence, 1949). Another key component is preference and for many, it is synonymous with choice. Preference builds on discrimination and inherently depends on value as a product of outcome dimensions (Hull, 1934; Spence, 1952). Preference becomes more equal as reward value converges (e.g., isohedonia can lead to equivalence in choice over time; see Guttman, 1954). Finally relative valuation is a component of value updating based upon external and internal factors (Clayton, 1964; Craft, Church, Rohrbach, & Bennett, 2011; Morgado, Marques, Silva, Sousa, & Cerqueira, 2014). Incentive contrast is one process that basically updates valuation of reward to more positive when reward value upshifts and negative when reward value downshifts (Flaherty, 1996). Animals are very proficient at reward discrimination, preference and relative valuation (Tolman, 1938; Horridge, 2005; Rivalan, Valton, Seriès, Marchand, & Dellu-Hagedorn, 2013; Wikenheiser, Stephens, & Redish, 2013). Ideally, these components work together to produce optimal choice that arises from evaluating short and long-term memories for reward value that are contextually dependent (Vestergaard and Schultz, 2015). Studied separately, the components can be reduced or absent. For example, incentive contrast is reduced due to alterations in physiological state (Panksepp and Trowill, 1971) or environmental shifts to outcome prediction or instrumental effort (Binkley et al., 2014; Webber et al., 2015). It is clear that animals can learn alternatives and express effective choice behavior (McMillan et al., 2015; Montes et al., 2015; Robbins, 2002; Linwick and Overmier, 2006); however, how these different components work together to produce choice is not well known especially in more unrestrained environmental settings.

One goal of the present study was to examine how enhanced ability to pace and to sequence responses during decision-making in the rat influences different components of choice. Determining the role of context in animal models of choice is important because it could enable a more complete understanding for decision-making during self-paced, sequential choice (Kacelnik and Bateson, 1996). Most experimental designs limit the pacing and the sequential expression of choice behavior (Weber, Shafir & Blais, 2004). Specifically, experimental research has focused on behavioral responses as a measure of preference such as pecks, lever presses or nosepokes (Baum, 1974; Bouton, Todd, Miles, León, & Epstein, 2013), one of two forced choices in a T-maze (Logan, 1965; Moustgaard & Hau, 2009), or licking / orofacial behavior (Flaherty & Rowan, 1986; Frutos, Pistell, Ingram, & Berthoud, 2012). Other paradigms examining decision-making utilize apparatuses such as single

operant boxes or runways restricting an animal's ability to pace activity and locomote (Marshall & Kirkpatrick, 2013; McClure, Podos, & Richardson, 2014). For example, an operant box is a single chamber with different devices within it that the animal can respond typically following extensive multi-week training and involving the learning of associations between different stimuli and the outcome delivery. The present paradigm combines aspects of this previous work on choice (i.e., lever response + alleyway + conditioned place) to create an environment with different temporal contingencies that produce behavior that resembles more free action to stay or leave and pattern responses in different ways. Recent work by Blanchard and Hayden (2015) has shown when comparing between a free-choice procedure and an intertemporal choice procedure, rhesus macaques show more patience when performing in the task with greater freedom. This performance in a choice task with more self-pacing suggests that monkeys are better at obtaining maximum reward than previous work using intertemporal tasks shows (Kim, Hwang, & Lee, 2008). A similar effect has been demonstrated in humans (Fisher, Thompson, Piazza, Crosland, & Gotjen, 1997). Moreover, non-human primates have shown significant preferences for a freechoice task compared to a forced choice condition (Suzuki, 1999). Pigeons are able to discriminate between human feeders dependent upon the physical characteristics and attitudes of the feeder (Belguermi et al., 2011). Comparable to human behavioral results (Fisher et al., 1997), free choice is also preferred to forced choice in avian models (Catania & Sagvolden, 1980). It is problematic that results could differ depending upon the nature of the choice environment (i.e., forced versus free) yet the bulk of work has been done using one of the designs while the other has been significantly neglected in experimental work.

A second goal of the present study was to ensure that the amount and form of training reduced behavioral autonomy and enabled valuation to potently influence different aspects of motivated action (Dickinson, 1985). Many designs require overtraining of animals due to either difficult or convoluted behavioral procedures (Funamizu, Ito, Doya, Kanzaki, & Takahashi, 2015) or complex discrimination between outcomes (Morgado, Marques, Silva, Sousa, & Cerqueira, 2014). This extensive training is necessary for animals to adequately discriminate between options and then preference or reward updating is measured (De La Piadad et al., 2006; Gibbon et al., 1988). The influence of reward value on choice can be reduced or absent because animals have reduced sensitivity to incentive value after extensive training (Adams, 1982). We have a relatively rapid training procedure most likely enabling reward value embedded in the more or less riskier contexts to persistently influence choice. In addition, the use of diverse, disparate levels of reward produced from multiple outcome dimensions reduces the production of behavioral autonomy (Kosaki and Dickinson, 2010).

We have developed a new, three-box reward-seeking paradigm (Figure 1) that integrates diverse features from previous methods used to study reward choice and decision making (Powers et al., 2011; Ricker et al., 2014). It creates a task that combines the advantageous aspects of commonly used choice paradigms. These include the exploration of an open field task, the discrimination of a T-maze, the environment of a conditioned place preference task, and the option to include more standard operant tasks (i.e., lever press). This paradigm overcomes the limitation of reduced response selection by providing the opportunity to measure multiple dependent variables. These include measures of reward approach (i.e., food cup checks, latencies to retrieve pellets & the number of times the animal chooses to

enter each box) and measures of outcome consumption (i.e., total reward received and trials performed) as well as measures of place conditioning (i.e., the average amount of time spent in each box per entry and total time spent in each box). By expanding the testing environment, we provide a more expansive area for the animal to roam as opposed to that utilized in a standard operant box. This straightforward improvement provides an opportunity for a diverse set of behaviors to be examined. This paradigm also overcomes the limitation of extensive training with a simplification of the actionoutcome relationship by making subsequent outcome deliveries contingent on more naturalistic responses during food outcome retrieval.

Most importantly and different from previous work, the new paradigm enables the analysis of diverse components of choice (see Table 1). Discrimination can be investigated precisely by monitoring how rats are able to identify choices that change in outcome magnitude over time. Preference is monitored by measuring how rats choose between two alternatives within a single testing session. Finally, relative reward valuation can be measured by seeing how animals choose between reward outcomes that remain constant between weeks, and that are compared to multiple alternative rewards (see Table 1). Previous experience with alternative rewards can influence an animal's choice. The subjective value that an animal assigns to a reward changes as the animal engages with new rewards with changing objective values (Crespi, 1942). A difference in performance to identical outcomes after exposure to a reward that differs in either magnitude or quality is thought of as a relative reward effect (Webber et al., 2015). We hypothesize that animals will change preference based on the shift in reward magnitude in each box but only when the higher level of magnitude offsets the increase in delay to outcome. Our preliminary data (Powers et al., 2011; Ricker et al., 2014) supported this rule, and it follows the general framework of risk-sensitive foraging. A point of indifference has been obtained using a 50% higher reward outcome to offset the longer delay (Weeks 2 and 5). Other weeks with similar disparity are not equivalent because the location with higher magnitude actually has the shorter delay making it attractive on both dimensions. Preference is reversed within a three week testing period and discrimination and contrast demonstrate 'scaling' according to this proposed interaction. Scaling is observed as shifts in responding that are proportional to shifts in reward outcome value (e.g., magnitude). If the interaction is upheld, it demonstrates the power of the risk-sensitive foraging model to predict not only preference but also discrimination (e.g., absolute preference) and dynamic valuation that arises from between session incentive relativity (e.g., positive and negative contrast derived from comparisons to the baseline weeks with equivalent responding).

Method

Subjects

Nine male Sprague-Dawley rats (*Rattus norvegicus*) weighing 291 – 490 grams were used in this study. The animals used in this study come from a larger sample used from a study involving quiniolinic acid lesions. The current animals are from the control group that received sham surgeries using a 0.9 M phosphate buffered saline solution. Animals were housed in $65 \times 24 \times 15$ cm cages with corncob bedding. Animals were food-deprived to no less than 85% of their free-feeding, baseline weight. They had *ad libitum* or open access to

food (Harlan Teklad Rat Chow #8604) from the end of testing Friday until approximately 24 hours prior to the beginning of testing on Monday. Water was available *ad libitum* in their home cages as well as throughout testing. They were maintained on a 12-h reverse light/dark cycle (lights off at 8:00 a.m.). The colony room was maintained at 70 degrees Fahrenheit and approximately 56% humidity. All procedures were approved by the Bowling Green State University Institutional Animals Care and Use Committee (Protocol 12-012). All efforts were made to keep animal suffering to a minimum.

Experimental Apparatus

The three-box reward-seeking paradigm consists of three $25.40 \times 30.48 \times 40.64$ cm cast acrylic boxes. A door is located on the front of the middle "decision" box. On the right and left wall of the box, there is a cast acrylic tunnel (9 cm diameter) that connects to the other two boxes that are approximately 117 cm apart. Each box is located in a separate closet and is connected to the box next to it via tunnels that traveled through the walls of the closets. Infrared (IR) beams are located at the entrance to each box to record entry / exit times. A food cup is located directly across from the tunnel in each of the "reward" boxes, and a pellet dispenser is connected to each food cup to dispense reward (45 mg dustless plain sucrose pellet; Bio-Serve, NJ). Infrared beams are located at the bottom of the food cup to record any time the rat checks for reward. A lever is located next to each food cup to record any superstitious behaviors. Guillotine doors are located just before the entrance to each reward box for training days and are triggered by IR beam breaks. Water nozzles are located to the left of the door of each reward box for *ad libitum* access throughout testing. Each reward box is placed within a sound-attenuating chamber and a white noise generator is used to mask outside noises. Both reward boxes are equipped with houselights that are only on during training until the rat enters the box. Behavioral measures are obtained using computer-controlled hardware with custom written programs (Med. Associates Inc, VT). Video cameras are suspended above each reward box to monitor each rat's behavior.

Schedule of Outcome Shifts

The current experiment was set up to expose rats in one box to reward magnitudes that changed over the course of two three-week sessions. During the first three-week session, the pellet number in the single-outcome box varied between weeks. In week (W) 1, no pellets were delivered in the single-outcome box. This pellet number increased to one pellet in W2. It increased again in W3 to two pellets. The reward number remained consistent throughout all three weeks in the mixed-outcome box. Each week, either no pellets or three pellets (0/3) were dispensed. The behavioral control program was arranged so that no more than two consecutive deliveries of the same pellet magnitude could occur; therefore, half of the time no pellets would be delivered, and half of the time three pellets would be delivered. During the second three-week session, the pellet magnitude in the mixed-outcome box varied between weeks. In W4, either zero or five pellets (0/5) were delivered. In W5, zero or three pellets (0/3) were delivered. In W6, zero or one pellet (0/1) was delivered. Each outcome had a 50% chance of being delivered. For the latter 3 week series, 1 pellet was delivered in the single-outcome box over the entire three-week session (W4–W6).

Outcome Exposure

Rats were run through behavioral training or outcome exposure on Monday and Tuesday of each week. This provided the animals with an equal time experience in each reward box to be able to express choice and demonstrate preference later in the week during the open testing days. Rats were placed in the decision box directly after the MED-PC program had been started. Once in the box, they had the option to go into the left or right tunnel. Once the IR beam to the reward box had been broken, the guillotine door lowered and a ten-minute timer started. Five seconds after entry, the pellet reward would be delivered to the food cup. Once five seconds had passed from an initial break of a food cup IR beam (the animal retrieves the reward), pellets were delivered. Pellets would then only be delivered again if five more seconds had passed and the IR beam was broken again (fixed-interval (FI) 5 schedule of reinforcement). After ten minutes, the guillotine door would lift, and the rat would be able to leave the box. When the IR beam entering the decision box was broken, that guillotine door would shut, forcing the rat to enter the other box while prohibiting it from re-entry into the already experienced box. Training would be complete once this same process occurred in the other reward box. This outcome exposure is labeled 'forced choice' but differs from other forced choice paradigms in that the animals are exposed to only one alternative in sequence with the order based on their box preference.

Free Choice Tests

Rats were tested on free choice preference on Wednesday (Free Choice 1) and Thursday (Free Choice 2) of each week. Two days were used with the initial experience as an acclimation to open testing and the second day used for data collection and comparison. During free choice testing, rats were placed in the middle, decision box immediately after trial initiation. Rats were then free to roam the apparatus for thirty minutes. Reward would be delivered to each food cup five seconds after the initial IR beam break to the reward box. Pellets were then delivered on an FI-5 second schedule when the food cup IR beam was broken. At the end of the thirty minute session, the guillotine doors would lower to keep the rat in the final box.

Extinction

Extinction sessions were ran on Friday each week to examine the power of the outcome to mediate behavior without primary reward delivery. The same procedures were followed on extinction days as on free choice testing with the exception that no pellets were delivered during the extinction phase of testing.

Behavioral Measures

The three-box reward seeking paradigm allows for multiple measures to be recorded to help fractionate the multiple components of choice behavior. We refer to these as measures of consumption (trials and total reward), measures of place conditioning (total time in box and average time in box), and measures of approach (foodcup checks, latencies, and entries). Trials refers to the number of times each animal retrieved food on the FI-5 schedule. Total reward refers to the total number of pellets a rat had eaten during the testing session. Data for trials is presented throughout the text as the indicator for consummatory action because

of the significant positive correlations between the two variables (significant values for r ranged from .874 – .995, $p < .000$). The amount of time the animal spent in each box was recorded and is labeled as total time in box. This time was averaged by dividing it by the number of entries made into that box over the session to see how much time on average the rat would spend in the box in a single visit. This measure was labeled average time in box. Once again, due to high positive correlations (significant values for r ranged from .632 – .957, $p < .020$ - .000), we limit reporting to total time in box. There were no significant correlations among our measures of appetitive motivation. These include foodcup checks (the number of times the IR beam in the food receptacle was broken); latencies to obtain food outcome (the duration of time (in milliseconds) between pellets being dispensed and IR beam breaks in the foodcup); and entries into food cup (the number of times the subject entered a specific box).

Data Analysis

Predictions for the different components of choice are listed in Table 1. All statistics were run using IBM SPSS statistics version 20. To assess reward discrimination, a one-way analysis of variance (ANOVA) was performed on the single-outcome box over the first three-week session and the mixed-outcome box over the second three-week session and pairwise tests between weekly sessions (see Table 1).

We assessed preference in two ways. First, we used a preference score to examine the percentage of optimal choice using only the trial numbers completed in each box. This score was a measure for the proportional choice for one alternative (the number of trials performed in the box with the most advantageous outcome divided by the total number of trials performed) and is used often in studies of choice, risk and decision-making. We completed a comparison among outcome exposure, open choice and extinction using 20-minutes of each session. A Friedman ANOVA test was performed to assess differences, and Wilcoxon signed ranks tests were then performed based upon obtaining significant main effects. Second, we completed an analysis of variance with box (2 levels) and week (3 levels) as the two variables (see Table 1). Pairwise comparisons were completed for each week providing a measure of preference for one outcome versus the other during a single context of choice (e.g., serial outcome exposure/training or the more open testing days).

To assess the relative reward effect, a one-way analysis of variance was performed on the stable outcome among weeks including the mixed-outcome box over the first three-week session and the single-outcome box over the second three-week session (see Table 1). Significant main effects or interactions were analyzed using pair-wise comparisons. P-values were adjusted within the 3-week sessions using Bonferroni corrections. Significance was determined for discrimination and preference (both comparing between 3 weeks) if $p < .017$, for relative reward effects (2 comparisons) using $p < .025$.

Results

Outcome Exposure

Trials were measured during training sessions. Data presented are from outcome exposure day 2 (Figure 2). Taking data from this day provides an adequate amount of time for the rats to learn the outcomes associated with each box. A 3 (week) by 2 (box) ANOVA performed over the first three-week session revealed a main effect of week, $F(2,16) = 17.27$, $p < .01$, as well as a box by week interaction, $F(2,16) = 12.71$, $p < .01$. Post-hoc analyses revealed that the rats performed significantly more trials in the mixed-outcome box in W1 (0/3 pellets > 0 pellets: $t(8) = 2.46$, $p < .05$, 95% CI [1.16, 35.50], $d = .82$) and switched preference in W3 (2 pellets > 0/3 pellets: $t(8) = 2.33$, $p < .05$, 95% CI [.14, 26.52], $d = .78$). A 2×3 ANOVA performed on the second three-week session revealed a main effect of box, $F(1,7) = 31.56$, $p < .01$, and a main effect of week, $F(2,14) = 25.27$, $p < .01$. Post-hoc analyses showed that the rats performed significantly more trials in the mixed-outcome box in W4 (0/5 pellet > 1 pellets: $t(8) = 3.46$, $p < .01$, 95% CI [10.21, 51.12], $d = 1.15$) and this pattern of preference shifted to the single-outcome box in W6 (1 pellet > 0/1 pellet: $t(7) = 4.92$, $p < .01$, 95% CI [21.67, 61.83], $d = 1.74$).

Reward Discrimination

Predictions were for a scaling in responding dependent upon magnitude 'single-outcome' boxes (2>1>0) during the first-three week session and 'mixed outcome boxes' (0/5>0/3>0/1) during the second three-week session. Predictions were partially met for the single outcome discrimination (W1 < W2 < W3) but not observed for the second three-week session. In this latter series, responding during weeks 4 (0/5) and 5 (0/3) was nonsignificantly different leading to divergence from the predicted pattern (see Table 1).

Measure of consumption—The hypothesized behavior was observed in the first three-week session with trials (see Table 2). There was a main effect of week, $F(2,16) = 25.03$, $p < .01$, with significant differences between W1 and W2, as well as W1 and W3 (1 pellet > 0 pellets: 95% CI [58.45, 131.99], $d = 1.99$; 2 pellets > 0 pellets: 95% CI [60.96, 119.26], $d = 2.38$). There was a main effect of week, $F(2,16) = 15.32$, $p < .01$, during the second three-week session, with significant differences between W4 and W6 (0/5 pellets > 0/1 pellets: 95% CI [18.17, 46.72], $d = 1.75$).

Measures of approach—Analysis of food cup checks revealed a main effect of week, $F(2,16) = 16.28$, $p < .01$. There were significant differences between W1 and W2, as well as W1 and W3 (see Table 2; 1 pellet < 0 pellets: 95% CI [215.06, 389.17], $d = 2.67$; 2 pellets < 0 pellets: 95% CI [284.10, 727.01], $d = 1.75$). Finally, we found a main effect of week for latencies, $F(2,16) = 6.69$, $p < .01$, with a significant difference between W1 and W2, but not W1 andW3 or W2 and W3 (1 pellet >0 pellets: $t(8) = 3.59$, $p < .01$, 95% CI [123.56, 568.81], $d = 1.20$). We observed a main effect of week for food cup checks during the second three-week session, $F(2,16) = 7.52$, $p < .01$. Pairwise t-tests found significant differences between W4 and W6, as well as W5 and W6 (0/5 pellets > 0/1 pellet: 95% CI [187.74, 550.48], $d = 1.56$; 0/3 pellets > 0/1 pellet 95% CI [54.23, 674.22], $d = .90$).

Measure of place conditioning—There was a main effect of week found for total time in box, $F(2,16) = 18.69$, $p < .01$, with significant differences between W1 and W2, as well as W1 and W3 (see Table 2; 1 pellet > 0 pellets: 95% CI [272.45, 672.26], $d = 1.82$; 2 pellets > 0 pellets: 95% CI [351.79, 867.66], $d = 1.82$). There was also a main effect of week during the second three-week session for total time in box, $F(2,16) = 8.29$, $p < .01$, with significant differences between W4 and W6, as well as W5 and W6 (0/5 pellets > 0/1 pellet: 95% CI [327.76, 580.48], $d = 2.76$; 0/3 pellets > 0/1 pellet: 95% CI [51.10, 688.21], $d = .89$).

Preference

Preference scores across choice contexts—A preference score was obtained for each rat during Outcome Exposure (i.e., Forced Choice), Free Choice, and Extinction sessions (Figure 3). This score reflects the percentage of trials an animal performed that were considered advantageous out of all trials performed for that session. A Friedman's ANOVA revealed main effects for all weeks (W1: $\chi^2(2) = 14.22$, $p < .01$; W2: $\chi^2(2) = 10.89$, $p < .01$; W3: $\chi^2(2) = 14.00$, p < .01; W4: $\chi^2(2) = 10.67$, $p < .01$; W5: $\chi^2(2) = 13.56$, p < .01; W6: $\chi^2(2) = 16.22$, $p < .01$). Further analyses revealed significant differences between the outcome exposure day or forced choice sessions and extinction at every week (W1: $Z = -2.31$, $p < .05$; W2: $Z = -2.67$, $p < .01$; W3: $Z = -2.67$, $p < .01$; W4: $Z = -2.31$, $p < .05$; W5: $Z = -2.67$, $p < .01$; W6: $Z = -2.67$, $p < .01$), as well as significant differences between free choice and extinction at every week (W1: $Z = -2.67$, $p < .05$; W2: $Z = -2.55$, $p < .01$; W3: $Z = -2.67$, $p < .05$; W4: $Z = -2.67$, $p < .05$; W5: $Z = -2.67$, $p < .05$; W6: $Z = -2.67$, $p < .05$). We also found significant differences between forced choice and free choice sessions on W1 and W6 (W1: $Z = -2.13$, $p < .05$; W6: $Z = -2.20$, $p < .05$).

Preference during free choice testing day

Measure of consumption—Rats' responses closely followed our predictions (see Table 1 and 2). There was a main effect of week, $F(2,16) = 11.26$, $p < .01$, and a box by week interaction, $F(2,16) = 28.69$, $p < .01$. Significant differences between boxes were found in W1 (95% CI [81.55, 130.45], $d = 3.33$) and W3 (95% CI [20.22, 102.23], $d = 1.15$). For the second three-week session, we found a main effect of box, $F(1,8) = 18.49$, $p < .01$, a main effect of week, $F(2,16) = 12.81$, $p < .01$, and a box by week interaction, $F(2,16) = 11.39$, $p < .01$. A significant difference between boxes was found in W6 (95% CI [114.53, 211.69], $d = 2.58$).

Measures of approach—Food cup checks also fit the profile predicted during the first three-week session, but this pattern did not persist in the second three-week session (see Tables 1 and 2). A main effect of box, $F(1,8) = 14.18$, $p < .01$, and a box by week interaction, $F(2,16) = 18.68$, $p < .01$, were observed for food cup checks over the first three-week session. Significant differences between boxes were observed in W1 (95% CI [595.95, 1203.83], $d = 2.28$) and W3 (95% CI [51.53, 548.03], $d = .93$). There was a box by week interaction for the second three-week session, $F(2,16) = 12.75$, $p < .01$. Pairwise t-tests found significant differences between boxes in W6 (95% CI [431.70, 1042.53], $d = 1.86$).

There was a box by week interaction, $F(2,16) = 6.06$, $p < .05$, for latencies. Pairwise comparisons found significant differences between boxes in W1 ($t(8) = 3.56$, $p < .01$, 95% CI [124.54, 581.70], $d = 1.19$) and W2 ($t(8) = 2.31$, $p < .05$, 95% CI [.49, 270.92], $d = .77$).

Measure of place conditioning—Our measure of place conditioning, total time in box, was significantly different between boxes during multiple weeks, but did not precisely follow our predictions (see Tables 1 and 2). For the first three-week session, we found a box by week interaction, $F(2,16) = 20.69$, $p < .01$. Significant differences were found between boxes in W1 (95% CI [544.47,1027.30], $d = 2.5$) There was also a box by week interaction, $F(2,16) = 9.03$, $p < .01$, during the second three-week session, W6 was the only week with significant differences between boxes (95% CI [522.43,1078.19], $d = 2.21$).

Relative Reward Effect

Over the first three-week session, the magnitude of reward in the mixed-outcome box does not change objectively, but it should decrease subjectively as the absolute magnitude in the single-outcome box increases. Contrast effects were found for both positive contrast (W1>W2) and negative contrast (W3<W2) for several measures. For the second series, only positive contrast was obtained (W6>W5) for the 1 pellet outcome and few negative contrast indicators because of the parity for the outcomes in weeks 4 and 5.

Measure of consumption—We found a main effect of week for trials, $F(2,16) = 27.55$, $p < .01$, with significant differences between weeks (see Table 2; W1 > W2: 95% CI [25.31, 68.69], $d = 1.67$; W2 > W3: 95% CI [2.52, 57.71], $d = .84$). There was a main effect of week for trials over the second three-week session, $F(2,16) = 15.32$, $p < .01$, with significant differences between W5 and W6 (W6 > W5: $t(8) = 7.22$, $p < .01$, 95% CI [63.92, 123.86], $d = 2.41$) with no significant difference for the 4 and 5 combination.

Measures of approach—A main effect of week was also found for food cup checks, $F(2,16) = 15.72$, $p < .01$, with significant differences between W1 and W2 (W1 > W295% CI [160.46, 679.54], $d = 1.24$). During the second three-week session, food cup checks revealed a slightly different pattern of reward evaluation. There was a main effect of week, $F(2,16) = 11.91$, $p < .01$. Significant differences were found between W4 and W5, as well as W5 and W6 (W5 > W4: 95% CI [267.12, 717.09], $d = .88$; W6 > W5: 95% CI [110.84, 765.83], $d = 1.03$).

Measure of place conditioning—There was also a main effect of week for total time in box, $F(2,16) = 19.89$, $p < .01$. Pairwise t-tests found a significant difference between W1 and W2 (W1 > W2: 95% CI [244.33, 511.52], $d = 2.17$). There was also a main effect of week for total time in box for the second three-week session, $F(2,16) = 8.19$, $p < .01$, with significant differences between W5 and W6 (W6 > W5: 95% CI [47.63, 770.19], $d = .87$).

Extinction

Rats were exposed to an extinction period at the end of every test week and received no food pellets for an entire thirty-minute session. All data presented for extinction is for the

dependent variable of numbers of trials because other measures followed this similar pattern of results.

Discrimination—For the first three-week session, there was a main effect of week, $F(2,16)$ = 24.61, p <.01. Significant differences were found between all weeks (1 pellet > 0 pellets: 95% CI [10.02,23.53], d = 1.91; 2 pellets > 0 pellets: 95% CI [3.35,11.98], d = 1.37; 1 pellet > 2 pellets: 95% CI [3.89,14.33], d = 1.34). For the second three-week session, there was no main effect found.

Preference—For the first three-week session, there was a main effect of week, $F(2,16)$ = 5.90, p < .05, and a box by week interaction, $F(2,16)$ = 36.55, p < .01. Significant differences were found between boxes only in W1 (0/3 pellets > 0 pellets: 95% CI [17.71,44.73], d = 1.77). For the second three-week session, there was a main effect of box, $F(1,8)$ = 6.57, p < .05. Significant differences were only found between boxes in W6 (1 pellet > 0/1 pellets 95% CI [3.34,11.33], d =1.41).

Relative Reward Effect—For the first three-week session, there was a main effect of week, $F(2,16)$ = 16.87, p <.01. Significant differences were found between W1 and W2 (W1 > W2: 95% CI [8.58,29.64], d = 1.40). There were no main effects for the second three-week session.

Discussion

In the current study, we used a paradigm that combines features of other commonly used choice paradigms to examine choice behavior in the rat and predicted that the animals would show optimal responding for reward in a lawful, parametric fashion. The basis for our predictions arises from the notion that animals will optimize outcomes but take into account energy costs and risks when performing different instrumental behaviors (Weber, Shafir & Blais, 2004; Shapiro, Siller & Kacelnik, 2008). Our animals weigh the costs of travel between two reward 'patches' and have one option that is riskier due to uncertain reward delivery with longer delays for higher magnitude outcomes. It is a basis for decisions used often in risk-sensitive foraging research (Kacelnik and Bateson, 1996) but typically in rodent work, the decisions are made in standard operant boxes and the subject is not in complete charge of the pace or sequence of choices. Our behavioral results indicated that choice varied depending upon the type of action and the feature of decision-making under scrutiny. Lawful relationships were upheld primarily for consummatory actions but not for appetitive behaviors. Most importantly, the findings show that this context of more open choice had a powerful impact on key components of motivation including discrimination, preference, and relative valuation. The findings show clearly that measuring discrimination, preference, and incentive contrast from the same set of outcomes over a series of sessions leads to clear dissociations among these measures with disruption of certain component(s) while sparing others.

We predicted discrimination among different magnitudes of outcome would show 'scaling' from one week to another. Animals can produce parametric responses that 'scale' to magnitude or the shifts in magnitude (Hutt, 1954). This has been shown using an

autoshaping (i.e., the ability of the animal to self-train via experience in the apparatus over several learning sessions) paradigm in rats for both anticipatory (Papini & Pellegrini, 2006) and consummatory (Pellegrini & Papini, 2007) behavior. Magnitude is clearly a potent reinforcer and has a powerful impact on learning with better performance using higher magnitude outcomes (Rose, Schmidt, Grabemann, & Güntürkün, 2009). Our prediction was upheld for the first three weeks in that we did find scaled output for the numbers of rewards consumed. The animals showed output of consumption proportional to the magnitude of the reward (0 vs. 1 vs. 2 pellets). This scaling of consumption disappeared for the second three weeks. A reason for this is the more complex outcome computations required in the second 3-week set when outcome value is derived from both magnitude plus a variable delay to reinforcement. This result is very important in terms of demonstrating how animals lose the ability to perform critical 'scaling' when value determination becomes more complicated. These differences should not hinge on the degree of magnitude disparity because this did not vary between the different set of 3 week sessions. It is clear that having a 0 reward reference point made a large impact on consummatory measures and reference points certainly enable animals to build an associative network for outcome value more efficiently (Marshall & Kirkpatrick, 2015).

For both the first and second weeks of testing, approach behaviors lacked scaling between weeks. For example, food cup latencies and entries into the chambers had only one difference for the 1st 3 weeks and no significant differences in t he 2nd set of 3 weeks. These anticipatory measures may reflect exploration during the task and exploratory acts can be self-reinforcing and lead to bursts of responding for the outcome intermixed with outcome switching. Support for this idea comes from the high intensity food cup checking (range: 571 to 936 checks during rewarded 30 min session). This compulsive checking at the highest rate of once every 2 s in one box or the other suggests that animals are searching and moving rapidly between chambers. This result could reflect a positive reinforcing aspect of the behavior per se in the setup with animals appearing to 'win-shift' because shifting from one outcome location to another per se holds incentive value. Wheel-running has been found to be reinforcing (Belke & Heyman, 1994) and diverse animals express 'contrafreeloading' by choosing to expend more effort for outcomes when provided with more or less effortful options (de Jonge,;Tilly, Baars, & Spruijt, 2008; Lindqvist & Jensen, 2009; McGowan, Robbins, Alldredge, & Newberry, 2010). This result is highly important because it demonstrates that having animals in a large environment self-directing choice behavior can lead to choice itself being reinforcing.

Choice as preference is typically measured as the proportional shift for one outcome over others and these preference scores provide a way to determine risk, learning, and how value can be relatively weighted. We examined preference using the percentage of optimal choice measure and examined these scores among our different daily experiences including 1) serial outcome exposure (forced choice or training day), 2) simultaneous outcome exposure with self-pacing (free choice or testing day) and 3) extinction. We found that the context has a substantial impact on how the animal chooses over an extended session. The self-paced environment had the highest level of optimal choice in all weeks when preference was predicted to significantly vary between options. This was most significant when one option was the lowest magnitude or 0. When animals had back-to-back experiences with 2 options,

they made significantly less optimal choices despite clear discriminable alternatives. For example, between 0 option and the 0/3 mixed-option animals chose the outcome with food only 58% of the choices whereas the open test day this optimal choice rose to nearly 90% of the responses. Extinction in a more open arena obliterated choice for the optimal outcome leading to higher activity in the non-preferred outcome in most cases. This could arise because of a 'frustration' effect when responding to the higher value outcome and a shift in responding directed to the previously lower value alternative.

Preference as the choice for one alternative over another during a single, daily session was mainly framed by the difference in delays for reward outcome between options. Magnitude was comparatively less influential. Animals were averse to waiting on average an additional 2.5 s in order to obtain a 1.5 pellet advantage. This is in contrast to previous work that showed rats will wait for up to 25 s in order to obtain 3 pellets when a 1 pellet/5 s outcome is the alternative (Logan, 1965). Logan (1965) used a forced choice paradigm with two alleyways for the animals to choose from in order to obtain each outcome. Our subjects had a strong aversion to the delay in the open or free foraging test sessions. Overall, it was consistent as animals reliably chose the box with the 5 s delay even when the magnitude was higher in the opposite box. The only time the animals chose the outcome with the 7.5 s delay on average was when it was pitted against the no-reward outcome. High levels of delay discounting (i.e., the reduction in value for an outcome that is relatively delayed over time) have been presented in basic and clinical science work (da Matta, Gonçalves, Leyser, & Lisiane, 2012; Heyman, 2003). Animals or humans with high levels of delay discounting have been shown to have poor behavioral and emotional control (Takayuki, Maguire, Henson, France, 2014; Odum, 2011; Kollins, 2003) and human work has connected higher discounting to diverse psychological disorders (Madden, Francisco.; Brewer, & Stein, 2011;). The results of the present work provide a unique perspective on the aversion to the delay because it may be arising from the preference for exploration as greater than the aversion to waiting. Adaptive foraging must take into account 'pursuit times' in order to decide whether or not delays are too costly (Kagel, Green & Caraco, 1986). In some cases, higher levels of gratification can be obtained following a longer delay possibly making the outcome value higher (Reynolds, de Wit, & Richards, 2002). Overall, the findings suggest that delays can have different incentive values or be influenced by different properties of the environment. Future work must explore preference and other aspects of choice across choice contexts including serial and simultaneous choice as well as 'forced' choice when options are limited during responding but available as a grouped set of alternatives over shorter and longer periods of time.

Relative valuation comparisons were examined using baseline activity during the 2nd and 5th weeks of testing. This strategy resembles intraschedule contrast that utilizes a period of outcome equivalence as a baseline period from which to explore contrast (McSweeney & Norman, 1979). Positive contrast would be the outcome upshift in week 1 or week 6 and negative contrast would be the outcome downshift in weeks 3 and 5. Comparing weeks 5 and 6, positive contrast was expressed for all measures except latencies to retrieve food pellet and entries into the box. Negative contrast was found in only a few consummatory measures during the initial 3 weeks. Positive contrast was found even during extinction but only for the comparison of the 0/3 outcomes between weeks 1 and 2. This positive contrast

effect occurs due to the low level of outcome in the comparable box in week 1. Typically, negative contrast is more robust compared to positive contrast (Mellgren, 1972; Flaherty, 1996). This is true especially for consummatory measures of licking rate monitored during rapid presentation or simultaneous experiences (Flaherty & Rowan, 1986). Negative contrast is more pervasive and instrumental successive negative contrast has been a reliable finding using the runway and other appetitive measures (Binkley et al., 2014). Positive contrast is typically obscured by ceiling effects or the obstruction of inducing behavioral effects because of limits in increasing behavior (Campbell, Crumbaugh, Knouse & Snodgrass, 1970). Animals respond at a high level to greater levels of incentive and are not able to increase responding. This confound was reduced in the present study of 'foraging choice' using a larger environmental setting and examining responses between distinct locations.

The idea that animals can form a relative incentive comparison between weeks is not novel, and previous work has shown that animals retain and express a contrast effect over a 2 week interval between the training and testing sessions (Gordon, Flaherty & Riley, 1973). One way that the present contrast analysis is different from traditional contrast analyses is that an outcome comparison week with two different outcomes was used as the baseline. In order to explore how this difference in baseline could impact contrast, future work could incorporate a testing session in which the animals are exposed to identical outcomes in each location. Previous work using instrumental responses to study relative reward did find that different instrumental actions were differentially sensitive to contrast (Webber et al., 2015). Also, incentive contrast was only one of several relative reward effects observed to impact instrumental action (Webber et al., 2015). The present experimental paradigm could be used easily to investigate other relative effects including positive induction (Weatherly, Nurnberger, & Hanson, 2006) or variety effects (Bouton, Todd, Miles, León, Epstein, 2013) and note their relationship to other components of motivation. These relative reward effects are powerful and positive induction can be defined as the increase in response to an outcome of lower value proportional to the value of an alternative. It basically means that one outcome of lower value induces stronger responses during anticipation for an outcome of higher value. Variety on the other hand is dependent upon the overall context, and is defined as the impact of responding that depends upon the diversity of the set of alternatives available. A baseline rate of responding should be obtained with zero variety which is real-life might be impossible but is often used in experimental work (Bouton et al., 2013).

Discrimination, preference and contrast analyses coalesced as strong measures for choice behavior during the test and extinction sessions; however, these components also dissociated from one another. Making predictions that follow strict parametric relationships that incorporate variability of reward pitted against constant reward may not fit the data when animals have the ability to 'change one's mind' at multiple time points. This means that theoretical frameworks such as risk-sensitive foraging must take into account the simultaneous availability of one or even more options (Schuck-Paim and Kacelnik, 2007). Future research using rewards such as drugs of abuse or manipulating brain regions involved in reward processing can utilize these findings in order to make stronger predictions and extend existing theories.

An example in which reward preference has been divided includes the work on the divisible components of "liking" and "wanting" (for review, see Berridge & Kringelbach, 2013) In the present work, the dependent measure of total reward could be an indicator of overall preference for that reward or how much the animal may "like" the reward. Food cup checks resemble the compulsive nature that has developed from experiencing the reward or how much the animal "wants" the reward. Investigating components involved in "liking" or "wanting" of rewards within a context of free choice could help fill the gaps of knowledge still left in disorders such as addiction (Koob, 2015). Multiple studies have recently observed contrast effects while failing to find discrimination between outcomes (Binkley, Webber, Powers, & Cromwell, 2014; Webber, Chambers, Kostek, Mankin, & Cromwell, 2015). Typically, this is measured in different sessions but can be found with the same behaviors in the same sessions (Papini & Pelligrini, 2007). Wanting without liking and vice versa occurs following specific brain manipulations or drug exposure (Berridge & Robinson, 1993). The underlying neural regions involved in magnitude discrimination (Cromwell, Hassani, & Schultz, 2003) and relative reward processing (Cromwell, Hassani, Schultz 2005; Funamizu, Ito, Doya, Kanzaki, & Takahashi, 2015) overlap but could rely on different neural ensembles to encode motivational properties (Ito & Doya, 2015). Determining the physiological basis for distinct components of choice may reveal special vulnerabilities or computational abilities not otherwise uncovered (Berridge & Cromwell, 1990; Berridge, 2012).

Conclusion

Our findings show that animals on a negative energy budget respond in a risk-averse fashion yet express seeking or exploration as a choice irrespective of the differential values of options. This alters 'optimal' preference and rescales discrimination and relative valuation. The strict parametric relationships for appetitive and consummatory actions mostly fell short because of the variability of choice when animals have the ability to 'change one's mind' at multiple time points. This means that theoretical frameworks such as risk-sensitive foraging must take into account simultaneous availability of multiple options as well other psychological processes involved in choice besides preference (Abarca & Fantino 1982; MacDonall, 2009; Schuck-Paim and Kacelmik, 2007; Aw, Vasconcelos and Kacelnik, 2011). It should be noted that our experiment used food-restricted animals. This could be a major limitation of the current study due to the lack of generalizability to the normal status of humans and other animals. Future work would want to address this issue and explore components of choice as they interact with or are influenced by internal physiological state changes.

The results suggest the importance of dissociations among phases of motivation and components of choice. They suggest that acquiring reward at one point in time as a preference does not necessarily reflect either previous discrimination or contrast or anticipated outcome valuation computed from short or long-term experience. Foraging choice could be optimized if these components converge to evaluate risk most effectively. This form of interactive choice with 'scaled' responding of all components across weeks was observed only for consummatory measures and when animals worked for more 'simple' constant and certain outcomes. Future work will have to explore the necessity of components of choice for optimal foraging and decision-making. The present work provides

a method and a framework that can be used for future exploring and for developing ways to meet the challenge of understanding what factors are involved in optimizing and controlling choice and decision-making in comparative, experimental research.

Acknowledgments

References

Abarca N, Fantino E. Choice and foraging. Journal of Experimental Analysis of Behavior. 1982; 38(2): 117–123.

Adams CD. Variations in the sensitivity of instrumental responding to reinforcer devaluation. Quarterly Journal of Experimental Psychology. 1982; 34(2):77–98.

Aw JM, Vasconcelos M, Kacelnik A. How costs affect preferences: Experiments on state dependence, hedonic state and within-trial contrast in starlings. Animal Behaviour. 2011; 81(6):1117–1128.

Baum WM. Choice in free-ranging wild pigeons. Science. 1974; 185(4145):78–79. [PubMed: 17779288]

Belguermmi A, Bovet D, Pascal A, Prévot-Julliard A, Jalme M, Rat-Fischer L, Leboucher G. Pigeons discriminate between human feeders. Animal Cognition. 2011; 14(6):909–914. [PubMed: 21647649]

Belke T, Heyman GM. A matching law analysis of the reinforcing efficacy of wheel running in rats. Animal Learning & Behavior. 1994; 22:267–274.

Berridge KC, Kringelbach ML. Neuroscience of affect: Brain mechanisms of pleasure and displeasure. Current Opinion in Neurobiology. 2013; 23(3):294–303. [PubMed: 23375169]

Berridge KC, Robinson TE, Aldridge JW. Dissecting components of reward: 'liking', 'wanting', and learning. Current Opinion in Pharmacology. 2009; 9(1):65–73. [PubMed: 19162544]

Berridge KC. From prediction error to incentive salience: mesolimbic computation of reward motivation. European Journal of Neuroscience. 2012; 35(7):1124–1143. [PubMed: 22487042]

Berridge KC, Cromwell HC. Motivational-sensorimotor interaction controls aphagia and exaggerated treading after striatopallidal lesions. Behavioral Neuroscience. 1990; 104(5):778–795. [PubMed: 2244985]

Bhatti M, Jang H, Kralik JD, Jeong J. Rats exhibit reference-dependent choice behavior. Behavioural Brain Research. 2014; 267:26–32. [doi]. [PubMed: 24657593]

Bindra D. A motivational view of learning, performance, and behavior modification. Psychological Review. 1974; 81:199–213. [PubMed: 4424766]

Binkley KA, Webber ES, Powers DD, Cromwell HC. Emotion and relative reward processing: An investigation on instrumental successive negative contrast and ultrasonic vocalizations in the rat. Behavioural Processes. 2014; 107(0):167–174. doi:http://dx.doi.org/10.1016/j.beproc.2014.07.011. [PubMed: 25150068]

Blanchard TC, Hayden BY. Monkeys are more patient in a foraging task than in a standard intertemporal choice task. PloS One. 2015; 10(2):e0117057. [doi]. [PubMed: 25671436]

Bouton ME, Todd TP, Miles OW, León SP, Epstein LH. Within- and between-session variety effects in a food-seeking habituation paradigm. Appetite. 2013; 66(0):10–19. doi:http://dx.doi.org/10.1016/j.appet.2013.01.025. [PubMed: 23434973]

Braun S, Hauber W. Striatal dopamine depletion in rats produces variable effects on contingency detection: task-related influences. European Journal of Neuroscience. 2012; 35(3):486–495. [PubMed: 22277016]

Calvert AL, Green L, Myerson J. Delay discounting of qualitatively different reinforcers in rats. Journal of the Experimental Analysis of Behavior. 2010; 93(2):171–184. [doi]. [PubMed: 20885809]

Campbell PE, Crumbaugh CM, Knouse SB, Snodgrass ME. A test of the "ceiling effect"hypothesis of positive contrast. Psychonomic Science. 1970; 20:17–18.

Catania AC, Savgolden T. Preference for free choice over forced choice in pigeons. Journal of the Experimental Analysis of Behavior. 1980; 34(1):77–86. [PubMed: 16812181]

Clayton KN. T-maze choice learning as a joint function of the reward magnitudes for the alternatives. Journal of Comparative and Physiological Psychology. 1964; 58:333–338. [PubMed: 14241043]

Craft BB, Church AC, Rohrbach CM, Bennett JM. The effects of reward quality on risk-sensitivity in rattus norvegicus. Behavioural Processes. 2011; 88:44–46. [PubMed: 21801819]

Crespi LP. Quantitative variation of incentive and performance in the white rat. The American Journal of Psychology. 1942; 55(4):467–517.

Cromwell HC, Hassani OK, Schultz W. Relative reward processing in primate striatum. Experimental Brain Research. 2005; 162:520–525. [PubMed: 15754177]

Cromwell HC, Schultz W. Influence of the expectation for different reward magnitudes on behavior-related activity in primate striatum. Journal of. Neurophysiology. 2003; 89:2823–2838. [PubMed: 12611937]

da Matta A, Gonçalves FL, Bizarro L. Delay discounting: concepts and measures. Psychology & Neuroscience. 2012; 5(2):135–146.

de Jonge FH, Tilly S, Baars AM, Spruijt BM. On the rewarding nature of appetitive feeding behaviour in pigs (*Sus scrofa*): Do domesticated pigs contrafreeload? Applied Animal Behaviour Science. 2008; 114(3–4):359–372.

de la Piedad X, Field D, Rachlin H. The influence of prior choices on current choice. Journal of the Experimental Analysis of Behavior. 2006; 85(1):3–21. [PubMed: 16602373]

Dias-Ferreira E, Sousa JC, Melo I, Morgado P, Mesquita AR, Cerqueira JJ, Sousa N. Chronic stress causes frontostriatal reorganization and affects decision-making. Science. 2009; 325(5940):621–625. [doi]. [PubMed: 19644122]

Dickinson A. Actions and habits: the development of behavioural autonomy. Phil. Trans. Royal Society of London. 1985; 308:67–78.

Evenden JL, Robbins TW. Win-stay behaviour in the rat. The Quarterly Journal of Experimental Psychology. 1984; 36(1):1–26.

Feja M, Hayn L, Koch M. Nucleus accumbens core and shell inactivation differentially affects impulsive behaviors in rats. Progress in Neuropsychopharmacology & Biological Psychiatry. 2014; 54:31–42.

Fisher WW, Thompson RH, Piazza CC, Crosland K, Gotjen D. On the relative reinforcing effects of choice and differential consequences. Journal of Applied Behavior Analysis. 1997; 30(3):423–438. [PubMed: 9316257]

Flaherty, CF. Incentive relativity. Cambridge, U.K: Cambridge University Press; 1996.

Flaherty CF, Rowan GA. Successive, simultaneous, and anticipatory contrast in the consumption of saccharin solutions. Journal of Experimental Psychology.Animal Behavior Processes. 1986; 12(4): 381–393. [PubMed: 3772302]

Frutos MG, Pistell PJ, Ingram DK, Berthoud HR. Feed efficiency, food choice, and food reward behaviors in young and old fischer rats. Neurobiology of Aging. 2012; 33(1):206.e41–206.e53. [doi]. [PubMed: 20970890]

Funamizu A, Ito M, Doya K, Kanzaki R, Takahashi H. Condition interference in rats performing a choice task with switched variable- and fixed-reward conditions. Frontiers in Neuroscience. 2015; 9:27. [doi]. [PubMed: 25741231]

Gibbon J, Fairhurst S, Church RM, Kacelnik A. Scalar expectancy theory and choice between delayed rewards. Psychological Review. 1988; 95:102–114. [PubMed: 3353474]

Gordon WC, Flaherty CF, Riley EP. Negative contrast as a function of the interval between preshift and postshift training. Bulletin of the Psychonomic Society. 1973; 1(1):25–27.

Guttman N. Equal reinforcement values for sucrose and glucose solutions prepared with equal sweetness values. Journal of Comparative and Physiological Psychology. 1954; 47:358–361. [PubMed: 13221680]

Green L, Myerson J. A discounting framework for choice with delayed and probabilistic rewards. Psychological Bulletin. 2004; 130(5):769–792. [PubMed: 15367080]

Herrnstein RJ. Relative and absolute strength of response as a function of frequency of reinforcement. Journal of the Experimental Analysis of Behavior. 1961; 4:267–272. [PubMed: 13713775]

Heyman GM. The remarkable agreement between people and pigeons concerning rewards delayed: Comments on Suzanne Mitchell's paper. Choice, Behavioral Economics, and Addiction. 2003:358–362.

Horridge A. What the honeybee sees: a review of the recognition system of. Apis mellifera. Physiological Entomology. 2005; 30(1):2–13.

Hull CL. The concept of the habit-family hierarchy and maze learning: Part 1. Psychological Review. 1934; 41:33–54.

Hutt PJ. Rate of bar pressing as a function of quality and quantity of food reward. Journal of Comparative and Physiological Psychology. 1954; 47:235–239. [PubMed: 13163261]

Ito M, Doya K. Distinct neural representation in the dorsolateral, dorsomedial, and ventral parts of the striatum during fixed- and free-choice tasks. The Journal of Neuroscience : The Official Journal of the Society for Neuroscience. 2015; 35(8):3499–3514. [doi]. [PubMed: 25716849]

Ivory NJ, Kambouropoulos N, Staiger PK. Cue reward salience and alcohol cue reactivity. Personality and Individual Differences. 2014; 69(0):217–222. doi:http://dx.doi.org/10.1016/j.paid. 2014.06.005.

Jensen GD. Preference for bar pressing over "freeloading" as a function of number of rewarded presses. Journal of Experimental Psychology. 1963 May; 65(5):451–454. 1963, http://dx.doi.org/ 10.1037/h0049174. [PubMed: 13957621]

Kacelnik A, Bateson M. Risky Theories—The effects of variance on foraging decisions. Integrative and Comparative Biology. 1996; 36(4):402–434.

Kacelnik A, El Mouden C. Triumphs and trials of the risk paradigm. Animal Behaviour. 2013; 86(6): 1117–1129.

Kagel JH, Green L, Caraco T. When foragers discount the future: constrint or adaptation? Animal Behavior. 1986; 34:271–283.

Kahneman D, Tversky A. Choices, values, and frames. American Psychologist. 1984; 39(4):341–350.

Kim S, Hwang J, Daeyeol L. Prefrontal coding of temporally discounted values during intertemporal choice. Neuron. 2008; 59(1):161–172. doi:http://dx.doi.org/10.1016/j.neuron.2008.05.010. [PubMed: 18614037]

Kim S, Hwang J, Lee D. Prefrontal coding of temporally discounted values during inter-temporal choice. Neuron. 2008; 59(3):161–172. doi:http://dx.doi.org/10.1016/j.neuron.2008.05.010. [PubMed: 18614037]

Kollins SH. Delay discounting is associated with substance use in college students. Addictive Behaviors. 2003; 28(6):1167–1173. [PubMed: 12834659]

Koob GF. The dark side of emotion: The addiction perspective. European Journal of Pharmacology. 2015; 753:73–87. [PubMed: 25583178]

Kosaki Y, Dickinson A. Choice and contingency in the development of behavioral autonomy during instrumental conditioning. J Exp. Psych. Anim. Beh. Proc. 2010; 36(3):334–342.

Lawrence DH. Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. Journal of Experimental Psychology. 1949; 39:779–784.

Lerman DC, Kelley ME, Van Camp CM, Roane HS. Effects of reinforcement magnitude on spontaneous recovery. Journal of Applied Behavior Analysis. 1999; 32:197–200. [PubMed: 10396772]

Lindqvist C, Jensen P. Domestication and stress effects on contrafreeloading and spatial learning performance in red jungle fowl (*Gallus gallus*) and White Leghorn layers. Behavioural Processes. 2009; 81(1):80–84. [PubMed: 19429200]

Linwick DC, Overmier JB. Associatively activated representations of food events resemble food outcome expectancies more closely than they resemble food-based memories. Learning and Behavior. 2006; 34(1):1–12. [PubMed: 16786879]

Ludvig EA, Balci F, Spetch ML. Reward magnitude and timing in pigeons. Behavioural Processes. 2011; 86(3):359–363. [PubMed: 21244854]

MacDonall JS. Reinforcing staying and switching while using a changeover delay. Journal of the Experimental Analysis of Behavior. 2003; 79(2):219–232. [PubMed: 12822688]

MacDonall JS. The stay/switch model of concurrent choice. Journal of the Experimental Analysis of Behavior. 2009; 91(1):21–39. [PubMed: 19230510]

Madden GJ, Francisco MT, Brewer AT, Stein JS. Delay discounting and gambling. Behavioural Processes. 2011; 87(1):43–49. [PubMed: 21352902]

Marshall AT, Kirkpatrick K. The effects of the previous outcome on probabilistic choice in rats. Journal of Experimental Psychology.Animal Behavior Processes. 2013; 39(1):24–38. [doi]. [PubMed: 23205915]

Marshall AT, Kirkpatrick K. Relative gains, losses and reference points in probabilistic choice in rats. PLOS One. 2015; 10(2):1–33.

McClure J, Podos J, Richardson HN. Isolating the delay component of impulsive choice in adolescent rats. Frontiers in Integrative Neuroscience. 2014; 8:3. [doi]. [PubMed: 24478644]

McGowan RTS, Robbins CT, Alldredge JR, Newberry RC. Contrfreeloading in grizzly bears: implications for captive foraging enrichment. Zoo Biology. 2010; 29(4):484–502. [PubMed: 19816856]

McMillan N, Kirk CR, Roberts WA. Pigeon (*columba livia*) and rat (*rattus norvegicus*) performance in the midsession reversal procedure depends upon cue dimensionality. Journal of Comparative Psychology. 2014; 128(4):357–366. [PubMed: 25111628]

McMillan N, Sturdy CB, Spetch ML. When Is a Choice Not a Choice? Pigeons Fail to Inhibit Incorrect Responses on a Go/No-Go Midsession Reversal Task. Journal of Experimental Psychology-Animal Learning and Cognition. 2015; 41(3):255–265. [PubMed: 25915749]

McSweeney FK, Norman WD. Defining behavioral contrast for multiple schedules. Journal of Experimental Analysis of Behavior. 1979; 32:457–461.

Mellgren RL. Positive and negative contrast effects using delayed reinforcement. Learning and Motivation. 1972; 3:185–193.

Michels KM. Response latency as a function of the amount of reinforcement. The British Journal of Animal Behaviour. 1957; 5:50–52.

Montes DR, Stopper CM, Floresco SB. Noradrenergic modulation of risk/reward decision making. Psychopharmacology. 2015; 232(15):2681–2696. [PubMed: 25761840]

Morgado P, Marques F, Silva MB, Sousa N, Cerqueira JJ. A novel risk-based decision-making paradigm. Frontiers in Behavioral Neuroscience. 2014; 8:45. [PubMed: 24596547]

Moustgaard A, Hau J. Induction of habits in rats by a forced-choice procedure in T-maze and the effect of pre-test free exploration. Behavioural Processes. 2009; 82(1):104–107. doi:http://dx.doi.org/10.1016/j.beproc.2009.04.011. [PubMed: 19409455]

Odum AL. Delay discounting: Trait variable? Behavioural Processes. 2011; 87:1. [PubMed: 21385637]

Paine TA, O'Hara A, Plaut B, Lowes DC. Effects of disrupting medial prefrontal cortex GABA transmission on decision-making in a rodent gambling task. Psychopharmacology. 2015; 232(10):1755–1765. [PubMed: 25420610]

Papini MR, Pellegrini S. Scaling relative incentive value in consummatory behavior. Learning and Motivation. 2006; 37:357–378.

Pellegrini A, Papini MR. Scaling relative incentive value in anticipatory behavior. Learning and Motivation. 2007; 38:128–154.

Peterson GB, Trapold MA. Expectancy mediation of concurrent conditional discriminations. American Journal of Psychology. 1982; 95:571–580. [PubMed: 7168456]

Powers, DD.; Cromwell, HC. Society for Neuroscience. New Orleans, La: 2012. The relative reward effect: A novel choice paradigm examining decisions in the rat model.

Rachlin H, Green L. Commitment, choice and self-control. Journal of the Experimental Analysis of Behavior. 1972; 17(1):15–22. [PubMed: 16811561]

Rachlin H, Logue AW, Gibbon J, Frankel M. Cognition and behavior in studies of choice. Psychological Review. 1986; 93(1):33–45.

Rachlin H. In what sense are addicts irrational? Drug and Alcohol Dependence. 2007; 90:92–99.

Real LA. Animal Choice behavior and the evolution of cognitive architecture. Science. 1991; 253:980–986. [PubMed: 1887231]

Reynolds B, de Wit H, Richards JB. Delay of gratification and delay discounting in rats. Behavioural Processes. 2002; 59(3):157–168. [PubMed: 12270518]

Ricker, JM.; Downey, C.; Hatch, JD.; Cromwell, HC. International Behavioral Neuroscience Society. Las Vegas, NV: 2014. Brain substrates for choice behavior in a novel 3-box paradigm.

Rivalan M, Valton V, Series P, Marchand AR, Dellu-Hagedorn F. Elucidating poor decision-making in a rat gambling task. PloS One. 2013; 8(12):e82052. [PubMed: 24339988]

Robbins TW. The 5-choice serial reaction time task: behavioural pharmacology and functional neurochemistry. Psychopharmacology. 2002; 163(3–4):362–380. [PubMed: 12373437]

Rose J, Schmidt R, Grademann M, Güntürkün O. Theory meets piegeons: The influence of reward-magnitude on discrimination-learning. Behavioural Brain Research. 2009; 198(1):125–129. [PubMed: 19041347]

Schuck-Paim C, Kacelnik A. Choice processes in multialternative decision making. Behavioral Ecology. 2007; 18(3):541–550.

Shapiro MS, Siller S, Kacelnik A. Simultaneous and sequential choice as a function of reward delay and magnitude. Journal of Experimental Psychology: Animal Behavior Processes. 2008; 34(1):75–93. [PubMed: 18248116]

Shepard RN. Toward a universal law of generalization for psychological science. Science. 1987; 237:1317–1323. [PubMed: 3629243]

Spence K. "Mathematical formulations of learning phenomena". Psychological Review. 1952; 59:152–160. [PubMed: 14920650]

Suzuki S. Selection of forced- and free-choice by monkeys (macaca fascicularis). Perceptual and Motor Skills. 1999; 88:242–250.

Tanno T, Maguire DR, Henson C, France CP. Effects of amphetamine and methylphenidate on delay discounting in rats: interactions with order of delay presentation. Psychopharmacology. 2014; 231(1):85–95. [PubMed: 23963529]

Tolman EC. The determinants of behavior at a choice point. Psychological Review. 1938; 45:1–41.

Uban KA, Rummel J, Floresco SB, Galea LM. Estradiol modulates effort-based decision making in female rats. Neuropsychopharmacology. 2012; 37(2):390–401. [PubMed: 21881567]

Vestergaard MD, Schultz W. Choice mechanisms for past, temporally extended outcomes. Proc R. Soc. B. 2015; 282(1810) pii: 20141766.

Volkert VM, Lerman DC, Vorndran CM. The effects of reinforcement magnitude on functional analysis outcomes. Journal of Applied Behavior Analysis. 2005; 38:147–162. [PubMed: 16033163]

Watanabe M, Cromwell HC, Tremblay L, Hollerman JR, Hikosaka K, Schultz W. Behavioral reactions reflecting differential expectations of outcomes in monkeys. Experimental Brain Research. 2001; 140:511–518. [PubMed: 11685405]

Weatherly JN, Nurnberger JT, Hanson BC. Investigating the procedural variables that determine whether rats will display negative anticipatory contrast or positive induction. Behavioral Processes. 2005; 70:10–18.

Webber ES, Chambers NE, Kostek JA, Mankin DE, Cromwell HC. Relative reward effects on operant behavior: Incentive contrast, induction and variety effects. Behavioural Processes. 2015; 116(0):87–99. doi:http://dx.doi.org/10.1016/j.beproc.2015.05.003. [PubMed: 25979604]

Weber EU, Shafir S, Blais A. Predicting risk sensitivity in humans and lower animals. Psychological Review. 2004; 111(2):430–445. [PubMed: 15065916]

Wikenheiser AM, Stephens DW, Redish AD. Subjective costs drive overly patient foraging strategies in rats on an intertemporal foraging task. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110(20):8308–8313. [doi]. [PubMed: 23630289]

Zabludoff SD, Wecker J, Caraco T. Foraging choice in laboratory rats: constant versus variable delay. Behavioral Processes. 1988; 16:95–110.
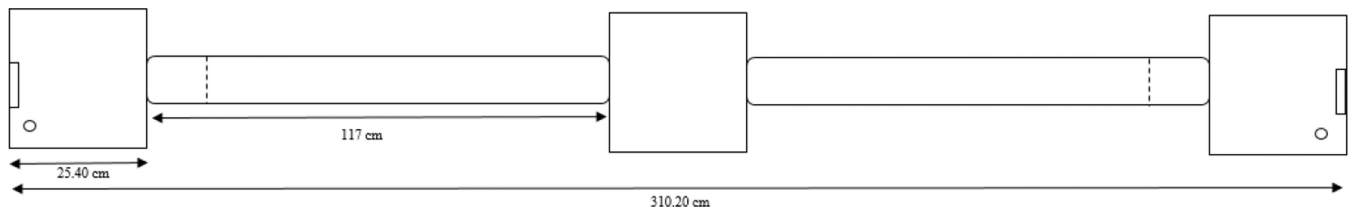
Figure 1. Representative schematic of the 3-box apparatus

The 3-box apparatus is depicted above. Rats are placed in the middle "decision" box at the start of the experiment, and given the option to enter a cast acrylic tube attached to either side of the box. Food receptacles (rectangles) are located directly opposite from the tunnel, and water nozzles (circles) are located next to the door of each reward box. During outcome exposure of each week, a guillotine door (dashed line) is lowered once IR sensors are broken just before entering one of the reward boxes. Rats are able to roam this expanded environment for the entire 30-minute session on test days.
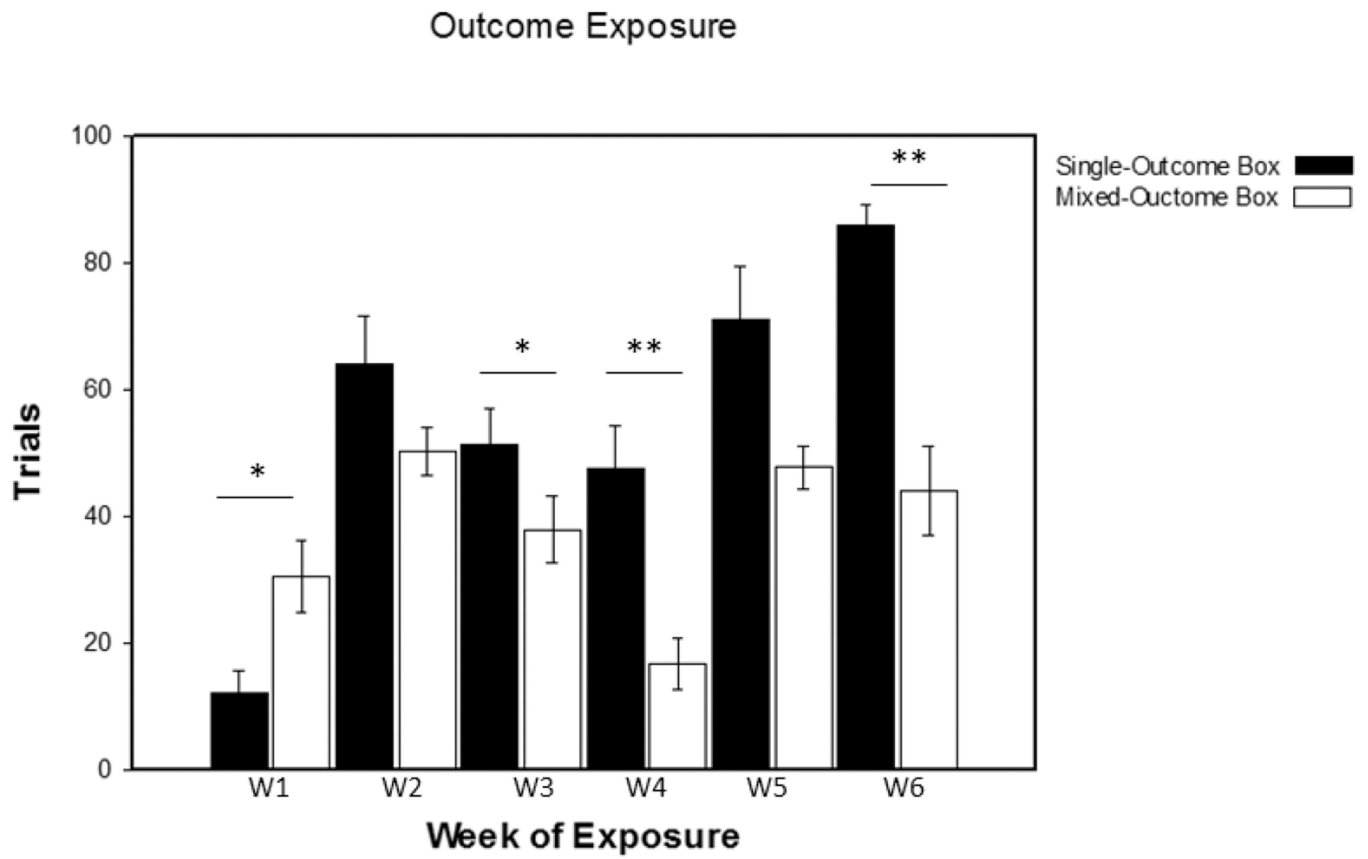
Outcome Exposure

Figure 2. Trials performed during outcome exposure training
The number of trials performed in each box over both three-week sessions. Each week (W)
is listed on the x-axis with the single-outcome box (black bars) and mixed-outcome box
(white bars) represented above. Values are mean ± standard error. * = p < .05. ** = p < .01.
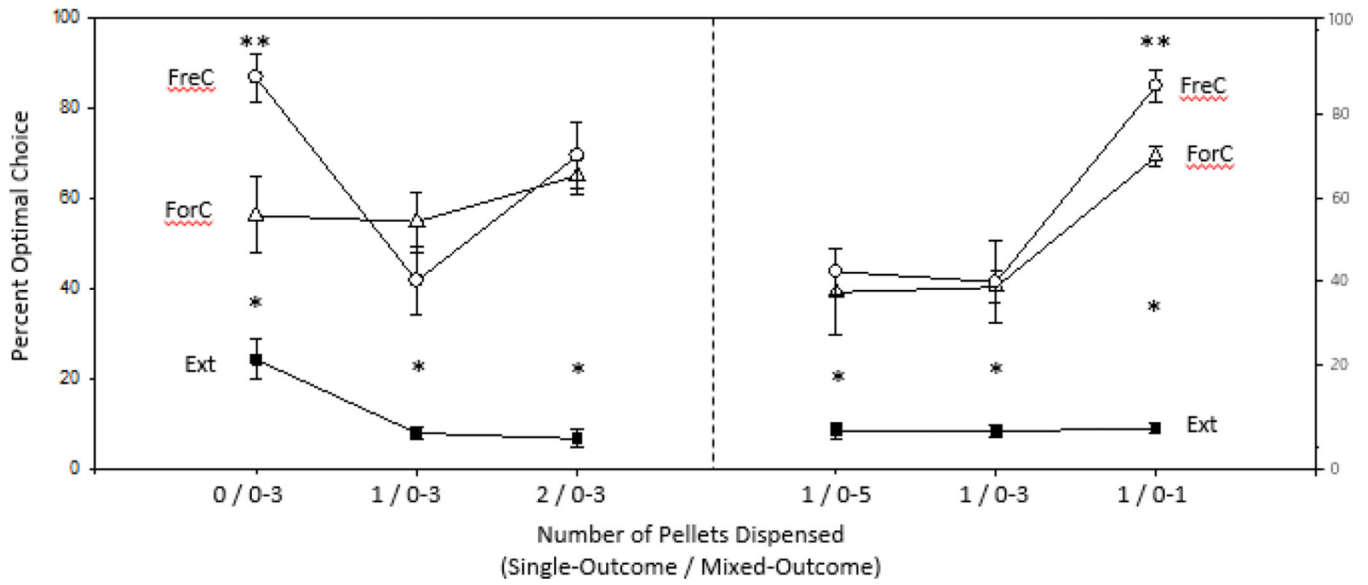
Figure 3. Comparative preferences across choice contexts
Percentage of optimal preference is compared across free-choice (FreC), forced-choice (ForC), and extinction (Ext) conditions. Numbers along the x-axis represent the number of pellets dispensed in the single-outcome box followed by the number of pellets dispensed in the mixed-outcome box for each week of testing. ** = significant differences between FreC and ForC. * = significant differences between FreC and Ext, as well as ForC and Ext.

Table 1

Predictions for the multiple components of choice behavior

Discrimination ------------------→			Preference ------------------→			Relative Reward Effect			
Comparisons		Predictions	Comparisons		Predictions	Comparisons		Predictions	Incentive Relativity
Week	Box	Magnitude Outcomes	Week	Box	Alternating Outcomes	Week	Box	Identical Outcomes	Direction of Contrast
1 vs 2	Single-Outcome	$W1 < W2$ $(0 < 1)$	1	S vs M	$M > S$ $(0/3 > 0)$	1 vs 2	$M_1 > M_2$	$0/3 > 0/3$	Positive
2 vs 3	Single-Outcome	$W2 < W3$ $(1 < 2)$	2	S vs M	$M = S$ $(0/3 = 1)$	2 vs 3	$M_3 < M_2$	$0/3 < 0/3$	Negative
1 vs 3	Single-Outcome	$W1 < W3$ $(0 < 2)$	3	S vs M	$S > M$ $(2 > 0/3)$	4 vs 5	$S_4 < S_5$	$1 < 1$	Negative
4 vs 5	Mixed-Outcome	$W4 > W5$ $(0/5 > 0/3)$	4	M vs S	$M > S$ $(0/5 > 1)$	5 vs 6	$S_6 > S_5$	$1 > 1$	Positive
5 vs 6	Mixed-Outcome	$W5 > W6$ $(0/3 > 0/1)$	5	M vs S	$S = M$ $(1 = 0/3)$				
4 vs 6	Mixed-Outcome	$W4 > W6$ $(0/5 > 0/1)$	6	M vs S	$S > M$ $(1 > 0/1)$				

Note. The three main components involved in choice behavior include discrimination (left-side), preference (middle) and relative reward effects (right-side). Predictions listed are based on what would be considered optimal choice. The relative reward effects are listed as incentive contrast effects based on a baseline level of activity during the middle week of testing when the two outcomes are purportedly equal between the two boxes.(S = single-outcome box and M = mixed-outcome box). The numbers in the cells and located in parentheses represent the food pellet number dispensed during the respective weeks.

Table 2

Components of Choice in the 3-box Apparatus

Week	Outcomes	Trials		Food Cup Checks		Total Time in Box (s)	
		Single - Outcome	Mixed - Outcome	Single - Outcome	Mixed - Outcome	Single - Outcome	Mixed - Outcome
1	0 vs 0/3	6.78 ± 1.08 **, ++	112.78 ± 10.49 ##	17.22 ± 3.51 **, ++	917.11 ± 130.62 ##	289.12 ± 32.71 **, ++	1075.00 ± 83.62 ##
2	1 vs 0/3	102.00 ± 16.38	65.78 ± 13.38	319.33 ± 36.87	497.11 ± 129.47	761.47 ± 91.65	697.08 ± 99.26
3	2 vs 0/3	96.89 ± 12.66 **, ++	35.67 ± 10.47 #	522.78 ± 94.95 *, ++	223.00 ± 62.79	898.84 ± 108.83 ++	522.09 ± 89.22
4	1 vs 0/5	93.00 ± 13.53	56.22 ± 8.24	346.78 ± 67.11	470.89 ± 82.35	647.59 ± 73.60	760.56 ± 82.74
5	1 vs 0/3	106.67 ± 20.46	60.78 ± 14.95	400.56 ± 87.58	466.00 ± 138.34 +	697.84 ± 120.19	676.09 ± 119.20 +
6	1 vs 0/1	186.89 ± 18.53 **, ##	23.78 ± 5.22 ++	838.89 ± 134.80 **, #	101.78 ± 22.34 ++	1106.75 ± 84.17 **, #	306.44 ± 49.13 ++

Note. Raw data for measures of consumption, approach, and place conditioning are represented. Data presented are for the number of trials each animal performed (Trials), the number of times they checked each food cup (Food Cup Checks), and the total amount of time (in seconds) spent in each box (Total Time in Box). Preference can be assessed by comparing the single-outcome to mixed-outcome in each week. Discrimination is assessed by comparing single-outcome from W1 – W3 and mixed-outcome from W4 – W6. Relative reward is assessed by comparing mixed-outcome from W1 – W2 and W2 – W3, as well as single-outcome from W4 – W5 and W5 – W6. Significant differences are indicated by * for preference, + for discrimination, and # for relative reward.

A single symbol represents $p < .05$, while double symbols represent $p < .01$ For example, *, +, or # represent $p < .05$. **, ++, or ## reflect $p < .01$. Values are mean ± standard error.