Behavior Research Methods (2019) 51:639–650
https://doi.org/10.3758/s13428-018-1065-0

CrossMark

An assessment of touchscreens for testing primate food preferences and valuations

Lydia M. Hopper[1] · Crystal L. Egelkamp[1] · Mason Fidino[2] · Stephen R. Ross[1]

Published online: 8 June 2018

Abstract
Typically, animals' food preferences are tested manually, which can be both time-consuming and vulnerable to experimenter biases. Given the utility of ascertaining animals' food preferences for research and husbandry protocols, developing a quick, reliable, and flexible paradigm would be valuable for expediting many research protocols. Therefore, we evaluated the efficacy of using a touchscreen interface to test nonhuman primates' food preferences and valuations, adapting previously validated manual methods. We tested a nonhuman primate subject with four foods (carrot, cucumber, grape, and turnip). Preference testing followed a pairwise forced choice protocol with pairs of food images presented on a touchscreen: The subject was rewarded with whichever food was selected. All six possible pairwise combinations were presented, with 90 trials per pairing. Second, we measured how hard the subject was willing to work to obtain each of the four foods, allowing us to generate demand curves. For this phase, a single image of a food item was presented on the touchscreen that the subject had to select in order to receive the food, and the number of selections required increased following a quarter-log scale, with ten trials per cost level (1, 2, 3, 6, 10, and 18). These methods allowed us to ascertain the subject's relative preferences and valuations of the four foods. The success of this touchscreen protocol for testing the subject's food preferences, from both a practical and a theoretical standpoint, suggests that the protocol should be further validated with other foods with this subject, with other subjects, and with other test items.

Keywords Demand curve · Preference · Touchscreen · Economic substitutability · Valuation

Considerable experimental effort has been applied to understanding nonhuman primates' food preferences, with such tests being run in both zoos (e.g., Fernandez, Dorey, & Rosales-Ruiz, 2004; Finestone et al., 2014) and laboratories (e.g., Baxter, Parker, Linder, Izquierdo, & Murray, 2000; Kemnitz & Francken, 1986). Knowledge of primates' food preferences has been used to inform experimental research paradigms (e.g., Brosnan & de Waal, 2003; Hopper, Kurtycz, Ross, & Bonnie, 2015), as well as to enhance husbandry and training efficacy (e.g., Gaalema, Perdue, & Kelling 2011). To assess primates' preferences, researchers typically evaluate which of two items subjects reach for first or select manually (e.g., Fernandez et al., 2004; Finestone et al., 2014; Tanaka, 2003, 2007); which they move toward first (e.g., McDermott & Hauser, 2004); or which they look at first or fixate on for longer, coded either manually from video footage (e.g., Adams & MacDonald, 2018; Kuwahata, Adachi, Fujita, Tomonaga, & Matsuzawa, 2004) or via an eye-tracking device (e.g., Howard, Wagner, Woodward, Ross, & Hopper, 2017; Wallis & Miller, 2003). All of these methods can be time-consuming, either to test or to code. It is also worth noting that food preference testing is often conducted as a precursor to a test, in order to determine reward values for different experimental conditions (e.g., to ascertain differently valued foods for tests of inequity aversion: Hopper, Lambeth, Schapiro, & Brosnan, 2014; or to create differently valued payoffs in tests of social learning strategies: Vale et al., 2017), or is used as a test of decision-making behavior under different experimental manipulations (e.g., Wallis & Miller, 2003), rather than as the primary focus of the experiment. Therefore, developing a quick, reliable, and flexible paradigm to test primates' preferences would be valuable for expediting many research protocols beyond those directly interested in primates' (food) preferences.

The food preferences of primates or other captive species are typically tested manually: A researcher or member of care

✉ Lydia M. Hopper
lhopper@lpzoo.org

[1] Lester E. Fisher Center for the Study and Conservation of Apes, Lincoln Park Zoo, Chicago, Illinois, USA

[2] Urban Wildlife Institute, Lincoln Park Zoo, Chicago, Illinois, USA

staff presents two or more food items to an animal subject that receives the items for which it reaches first (e.g., Brosnan & de Waal, 2003; Finestone et al., 2014; Wobber, Hare, & Wrangham, 2008). These pairwise trials are repeated, and if the animal chooses one option significantly above chance, it is recorded as a preferred food item (Hopper, Schapiro, Lambeth, & Brosnan, 2011). However, such manual methods for testing preferences are potentially vulnerable to cueing effects (Beran, 2012). For example, the researcher may (inadvertently) hold one food item closer to the animal than the other or may (unintentionally) provide a verbal cue as the animal selects a certain food item. Such biases, even if generated subconsciously by the experimenter, can impact the behavioral responses of the subjects and invalidate or bias interpretations of their behavior (e.g., Pfungst, 1911/2015). Additionally, food items are known to be highly potent stimuli; animals may reach for a food so quickly that they do not evaluate its worth in comparison to other options. Indeed, in cognitive research with primates, food items have often been covered (e.g., Claidière et al., 2015) or represented with symbols (e.g., Boysen, Berntson, Hannan, & Cacioppo, 1996) in order to overcome subjects' impulsiveness when presented with real food items, to allow for more nuanced responses to emerge.

Using a touchscreen interface to test animals' food preferences offers a potential solution to the biases and confounds introduced via manual testing. If both food items are presented simultaneously on a touchscreen, and the experimenter is blind to the subject's selection until after it has been made, then the opportunity for the experimenter to cue the animal is reduced, or is eliminated entirely. An additional advantage of using touchscreens is that preferences for stimuli that may not be feasible to be presented manually or in a controlled manner can be tested. For example, instead of necessitating the presentation of physical stimuli, a touchscreen interface allows researchers to show subjects images of group mates, enrichment devices, or people (e.g., familiar vs. unfamiliar people; see Tanaka, 2003, 2007, for examples). Thus, validating the ability to test primates' preferences for food items of known relative values (e.g., foods) opens the door to testing an array of stimuli in the future.

Touchscreen computers have been used to test human food preferences (e.g., Joseph, Egli, Koppekin, & Thompson, 2002) but are not typically used to test animals' food preferences (although see Judge, Kurdziel, Wright, & Bohrman, 2012). Despite the paucity of data demonstrating the efficacy of using touchscreens to assess primates' food preferences, touchscreens have been used to test many aspects of primate cognition (e.g., memory: Inoue & Matsuzawa, 2007; numerical understanding: Cantlon & Brannon, 2007; cooperation: Martin, Biro, & Matsuzawa, 2014; metacognition: Brown, Templer, & Hampton, 2017; and how cognition is influenced by emotion: Allritz, Call, & Borkenau, 2016). Given this fact, it is reasonable to assume that touchscreens could offer an effective method for testing primates' food preferences. Therefore, in this study we evaluated the efficacy of touchscreen computers for conducting forced choice pairwise food preference testing of primates. Additionally, since touchscreens have also been used to test various cognitive abilities in a variety of species (e.g., *Canis familiaris*: Zeagler et al., 2016; *Columba livia*: Fortes, Case, & Zentall, 2017; *Helarctos malayanus*: Perdue, 2016; *Ursus americanus*: Vonk & Beran, 2012), the methods being evaluated in this study also have the potential to be applied to nonprimate species.

Although the importance of determining animals' preferences has been well demonstrated, the efficacy of pairwise forced choice tests has been questioned. As Schwartz, Silberberg, Casey, Paukner, and Suomi (2016) noted, pairwise preference tests do not account for economic substitutability, and therefore the conclusions about animals' preferences that can be drawn from such tests are limited. In this study, therefore, we also wished to test whether primates' valuations of foods could be measured with touchscreens. To do so, we adapted the manual methods used previously by Schwartz et al. with capuchin monkeys (*Cebus apella*) for use with touchscreens. When testing what cost capuchins would pay to obtain different foods, Schwartz et al. used differently weighted baskets that the monkeys had to lift to receive food rewards. From the monkeys' responses, Schwartz et al. were able to generate demand curves to describe the monkeys' valuations of the three foods tested (*sensu* Hursh & Silberberg, 2008). In the present study, we substituted "number of touches made on a touchscreen" for "weight" in the metric of cost. Simply put, we measured how many times a primate was willing to select a photograph of a food item, presented on the touchscreen, in order to receive that food item.

Our aim was to determine whether primate food preferences can be assessed using a touchscreen interface via a traditional pairwise paradigm (Exp. 1) and a cost-based paradigm (Exp. 2). For the pairwise paradigm, our aim was to test the efficacy of presenting photographs of food rewards, rather than symbols representing different rewards (e.g., Wallis & Miller, 2003). Since one study has already shown success with members of a macaque species (*Macaca silenus*; Judge et al., 2012) who were shown photographs of food items on a touchscreen, and other research has demonstrated primates' ability to recognize photographic images (e.g., Parron, Call, & Fagot, 2008), we expected that a touchscreen interface would be well-suited to testing primates' preferences. Additionally, and by modifying methods used by Schwartz et al. (2016), we wished to extend the pairwise preference-testing methods typically used to also generate demand curves of primates' food preferences, by measuring how hard animals are willing to work for different food rewards. This would allow us to develop a more nuanced understanding of primate food preferences while also demonstrating the potential to use touchscreen computers to run such evaluations.

Experiment 1

Method

This study was approved by the Lincoln Park Zoo Research Committee, which is the governing body for all animal research at the institution. The food substances, amounts, and frequency of provision were reviewed and approved by Lincoln Park Zoo veterinary and nutrition staff prior to the start of the project. No modifications were made to the standard animal care routines. This research adhered to legal requirements in the United States of America and to the American Society of Primatologists' Principles for the Ethical Treatment of Nonhuman Primates.

Subject and housing The subject for this evaluation was a 12-year-old male gorilla (*Gorilla gorilla gorilla*) who lived in an all-male group comprising four gorillas (average age: 10.8 years) at the Lincoln Park Zoo, Chicago, USA. As with many of the apes at the Lincoln Park Zoo, prior to this study the subject had had several years of experience using touchscreen computers, specifically with serial-learning tasks (e.g., Wagner, Hopper, & Ross, 2016), but had never before been shown photographs of real-life objects on the touchscreen, nor had he participated in preference testing using the touchscreen. Previous preference testing, in which the subject was manually offered two foods simultaneously in a forced choice paradigm (*sensu* Hopper et al., 2015), had revealed that this gorilla preferred grapes over pieces of carrot, but no other food preferences had been formally tested.

All touchscreen testing was run in the gorilla's home enclosure, and participation was voluntary. The four gorillas lived in a large and complex enclosure with both indoor and outdoor areas (total area: 1,932 m^2). The indoor enclosure had a deep mulch floor and numerous vines and climbing structures, and it was connected directly to the outdoor enclosure via floor-to-ceiling glass sliding doors that also ensured plenty of natural light in the indoor enclosure.

Apparatus and stimuli We used an HP ProDesk 400 G2 connected to a 55-cm ViewSonic touchscreen monitor (model: TD2220, 1,920 × 1,080 resolution) to run the touchscreen test sessions. A second ViewSonic monitor was set up so that the experimenter could review the subject's responses in real time. The custom software used to run the pairwise preference tests and reward–cost task was the Zenrichment ApeTouch software suite (Martin, 2017).

For stimuli, we used photographs of the four foods that were offered to the subject: a grape, a piece of carrot, a piece of turnip, and a piece of cucumber (Fig. 1). We photographed all the food images using a Canon Powershot S110 digital camera, keeping the camera's distance and height from the food items constant for every photo, as well as the zoom setting. To ensure consistent lighting for all photographs, we photographed the food items in a LimoStudio tabletop photography box (41 × 41 cm) using white Bristol paper as the background in order to eliminate shadows and enhance contrast. We then resized the photographs in Adobe Photoshop to 300 × 300 pixels for use with the ApeTouch software. When the stimuli were presented on the touchscreen, the screen background was black, and the white square stimuli showing the food photographs were presented in random locations on the screen for each trial.

Protocol: Paired preference test With Experiment 1 we aimed to test the efficacy of using touchscreens to evaluate the subject's relative preferences for four food items by presenting the items in a pairwise manner following a forced choice paradigm. For each food pairing, the subject was first presented with 50 training trials, in each of which a single food image was shown on the screen, and the subject was rewarded with that same food when he selected the image on the screen (Fig. 2A). The subject was presented with every possible pairing of the four foods (i.e., six pairings).

Following Judge et al. (2012), the aim of the training trials was to offer the subject the opportunity to learn the association between the photograph of a food item and receiving that same food. Across the 50 training trials, equal numbers of the two food items in the test pair were presented (i.e., 25 training trials per food type). In each trial, the location of the food image on the screen changed with the random placement of the image determined by the ApeTouch software. Training trials for each pairing were run over one or two sessions.

After the subject had completed the 50 training trials for a given food pairing, he was then tested on 90 test trials, with 30 trials run per session. In a test trial, two images were shown on the touchscreen, one for each of the two foods in a pairing (Fig. 2B). Following a classic forced choice paradigm, when the subject selected one of the two foods (by pressing the image on the screen), both images disappeared and the subject was rewarded with the food he had selected. The locations of the two food images were random and scattered around the screen, such that the subject had to move his hand between trials in order to reselect the same image (as in the training trials, the location of each stimulus was randomly generated by the ApeTouch software).

The subject completed 50 training trials and 90 test trials for each pair of foods until he had received all of the six possible pairwise combinations of the four foods. For this experiment, the randomly determined order of the food pairs tested was grape versus carrot, then carrot versus turnip, then

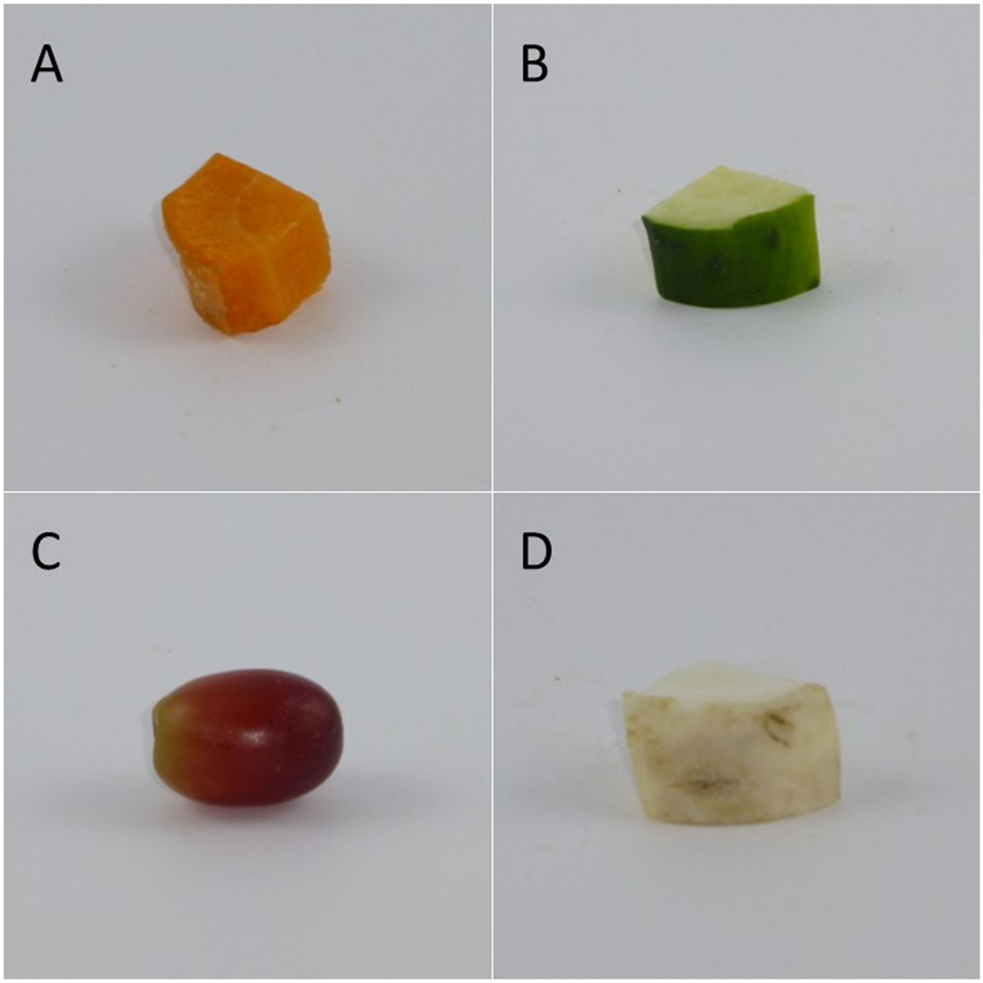

Fig. 1 The stimuli: Photographs of (A) a piece of carrot, (B) a piece of cucumber, (C) a grape, and (D) a piece of turnip

grape versus cucumber, then carrot versus cucumber, then turnip versus cucumber, and finally grape versus turnip.

No audio cues were provided to the subject (i.e., no chime sounded when he made a selection) in either the training or

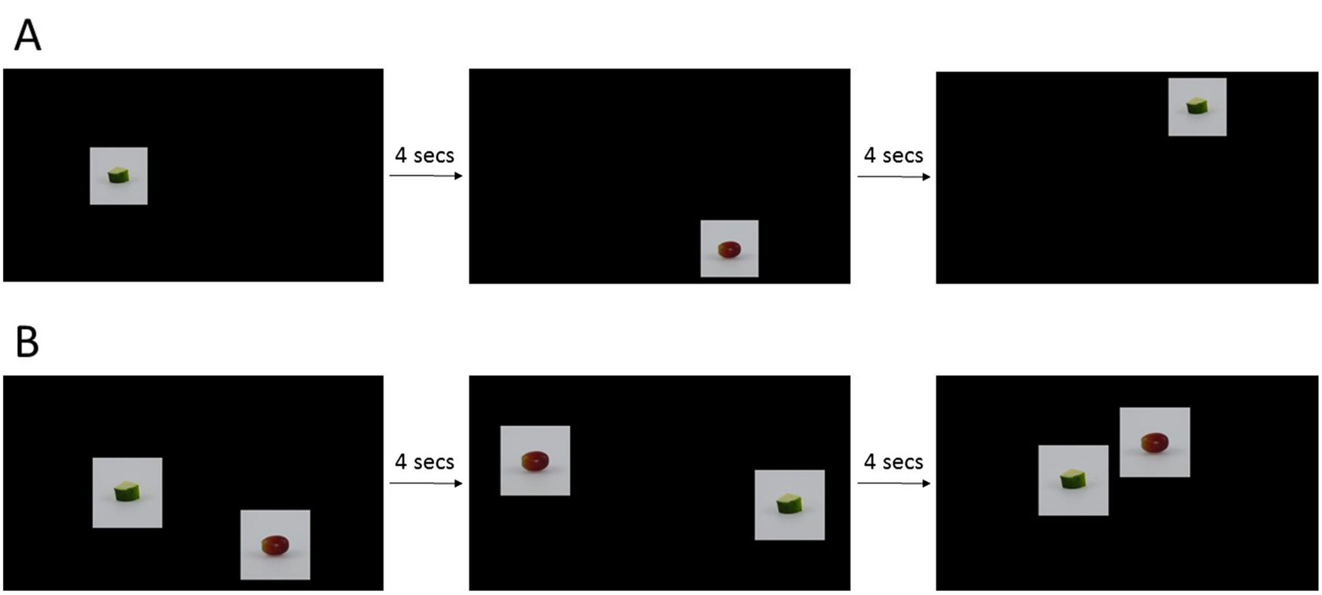

Fig. 2 Screenshots of the training trials (A) and testing trials (B) in Experiment 1, showing the grape and cucumber pairing in this example

test trials, as we did not wish to add a secondary reinforcer to the gorilla's choices. Additionally, for both training and test trials, the trials were separated by 4-s intertrial intervals. For all testing, no more than one test session was run in a single day, with sessions typically being run at 1:30 pm on weekdays. For both the training and test trials, the experimenter rewarded the subject by dropping the food reward into a PVC chute attached to the subject's cage. Regardless of the food type, the food pieces were all cut to the same size (i.e., the size of a grape), approximately 2–2.5 cm^3, and each weighed between 5 and 7 g.

After the conclusion of all the training and test trials for each of the six unique food pairings (Phase 1), we retested the subject's preferences again for the same foods (Phase 2), to ascertain the stability of his preferences over time. For this second round of testing, the training trials were omitted and the subject was simply tested using the paired testing trials. Again, the subject completed 90 test trials for each pairing, and the food pairs were presented in the same order as in the first phase of testing. The testing for Phase 1 took place from October to December 2016, and the testing for Phase 2 took place between December 2016 and April 2017.

Analyses To test the gorilla's relative preferences for the four foods presented in the paired comparisons, we fit a log-linear Bradley–Taylor model (LLBT) with the prefmod package in the R programming language (Hatzinger, 2015). Made specifically for paired-comparison testing, the LLBT model estimates a subject's relative "worth" values for each choice on a preference scale that sums to 1 (Hatzinger & Dittrich, 2012). A higher worth value indicates greater preference for an item. In a standard Bradley–Terry model, the probability that a subject will prefer object k to object j, or vice versa, is

$$p\left(y_{jk}\right) = c_{jk}\left(\frac{\sqrt{\pi_j}}{\sqrt{\pi_k}}\right)^{y_{jk}},$$

where π_j and π_k are the worth values for each object on the preference scale; y_{jk} is a response to the comparison of j to k, which takes the value of 1 if $j > k$ and the value of – 1 if $k > j$; and c_{jk} is a normalizing constant. For objects j and k, the LLBT model assumes that the observed number of selections for object j instead of k (or vice versa) follows a multinomial distribution conditional on a fixed number of trials. Thus, the expected number of selections for a given object over a paired trial is

$$m\left(y_{jk}\right) = n_{jk}p\left(y_{jk}\right),$$

where n_{jk} is the number of trials conducted for a given pair and $m(y_{jk})$ is the expected number of selections (Hatzinger & Dittrich, 2012). Since this is a log-linear model, the linear predictor of the LLBT is then

$$\ln m\left(y_{jk}\right) = u_{jk} + y_{jk}\left(\beta_j - \beta_k\right).$$

Here, u_{jk} is a nuisance parameter, whereas β_j and β_k correspond to the worth values on the preference scale, such that ln $\pi = 2\beta$ (Hatzinger & Dittrich, 2012).

Importantly, this model assumes that the responses to paired comparisons are independent. Therefore, to test for independence within our data set, we performed a bootstrap with 1,000 iterations using the boot package for R (Canty & Ripley, 2017). This revealed no significant difference between our full data set and the bootstrapped model (residual df = 163.84, deviance = 0), suggesting independence between trials.

Additionally, we used the gnm function (Turner & Firth, 2015) to analyze the subject's relative preference for the foods, both within a phase and across the two phases of testing, to evaluate the consistency of his preferences. The prefmod package uses the multinomial-Poisson transformation (Baker, 1994) to fit the LLBT model through maximum likelihood. This requires the family of the gnm function to be specified as Poisson (Hatzinger & Dittrich, 2012).

All analyses were run in R version 3.3.2 (R Core Team, 2017) using RStudio (RStudio Team, 2016). The data and R script are available here: https://drive.google.com/drive/folders/1VOqttUUcQv-QQ9I2En5h7K5vf7UVZ6OC.

Results

The gorilla's responses in Experiment 1 revealed a strong preference for grape over the other three foods (Fig. 3). He significantly selected grape over cucumber (Z = 62.89, p < .001), carrot (Z = 8.65, p < .001), and turnip (Z = 14.39, p <

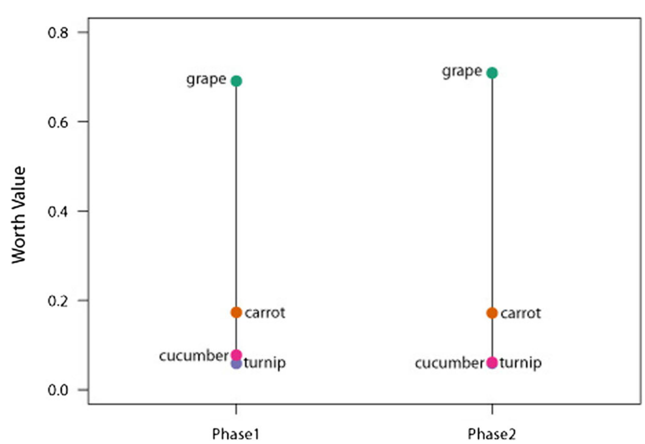

Fig. 3 The gorilla's relative preferences (worth values) for four foods across two rounds of testing following a forced choice paradigm presented in a series of paired comparisons. Phase 1 included training trials, whereas Phase 2 included only the paired testing trials

.001). He also preferentially selected carrot over both cucumber (Z = 12.95, p < .001) and turnip (Z = 6.66, p < .001) but showed no difference in the rates at which he selected cucumber and turnip (Z = 1.97, p = .073). Furthermore, his relative preferences for the four foods remained stable over the two testing phases (i.e., there was no difference across phases: χ^2 = 1.74, p = .629; Fig. 3).

Discussion

The gorilla's responses revealed differential selection rates of the four food options, suggesting a strong preference for grape over the other three foods (cucumber, carrot, and turnip) and that turnip was his least-preferred food. His relative preference for grape over carrot, as demonstrated through his selections via the touchscreen, reflected his earlier choices measured through a classic manual forced choice testing paradigm, providing validation for the results of the touchscreen method. Next, we ran a second experiment to evaluate the efficacy of using touchscreens to test primates' valuations of different food items. To do so, we adapted methods used by Schwartz et al. (2016) that required the subject to exert increasing effort to obtain each of the four foods presented in Experiment 1. In this way, we could generate demand curves to see the subject's relative willingness to work for the foods and to determine whether his responses to this test would reflect the relative preferences he had revealed in Experiment 1.

Experiment 2

Method

Subject and housing The subject, apparatus, and stimuli for Experiment 2 were the same as those for Experiment 1, with the exception that the photographs of the food items were resized to 290 × 290 pixels for use with an updated version of the ApeTouch software. Experiment 2 was run after the completion of Experiment 1, such that the subject had prior experience with each of the four food photograph stimuli before starting this second experiment.

Protocol: Reward–cost task In Experiment 2, rather than choosing between two foods presented on the touchscreen, as in Experiment 1, the subject had to select a single food image n times in order to be receive that same food item. The number of times the subject had to touch the image within a trial increased over phases, with ten trials for each of the four foods presented in each phase (i.e., 40 trials total per phase). After Bentzley, Fender, and Aston-Jones (2013), the phases were run such that the cost (number of required touches) increased following a quarter-log scale: 1, 2, 3, 6, 10, and 18. Only trials of the same cost were run in a single session, with no more than ten trials per test session and no more than one session per day. In every session, each of the four foods was presented to the subject at least once. Thus, rather than completing all of the trials for one food and then moving on to all of the trials for a second food, the subject completed all of the trials for all four foods for a certain cost, before moving on to the next cost phase (following the methods of Casey, Silberberg, Paukner, & Suomi, 2014; Schwartz et al., 2016). The order of presentation of the four foods within a test session was randomly determined.

The maximum length of time that a food image remained on the screen within a trial was 30 s. If the subject did not select an image within 30 s, the trial was considered incomplete. If the subject did select the image the required number of times within a trial (selecting each image in that trial in under 30 s), the experimenter gave the subject the corresponding food reward. Thus, a "refusal" was recorded for trials in which the subject did not select the image the required number of times, by "timing out." For trials in which the subject had to press the food image more than once within a trial, there was a 3-s interval after the subject selected the image before the presentations of the next image. With each presentation of the image within a trial (and between trials), its location on the screen changed due to the random placement of the image determined by the ApeTouch software (Fig. 4). In this way, these trials looked exactly the same as the training trials in Experiment 1, except that the subject now had to make more than one selection before being rewarded (in the ≥2-cost trials). All testing took place from June to December 2017.

Analyses To evaluate the gorilla's valuations of the four foods, as measured by the effort he was willing to exert to obtain them (i.e., number of presses on the touchscreen), we generated demand curves for each of the foods following the methods of Schwartz et al. (2016). To do so, we normalized our data in line with the procedure proposed by Hursh and Winger (1995; see also Schwartz et al., 2016) both for the quantity Q_f of food f eaten,

$$\varphi_f = \log_{10}\left(\left(\frac{Q_f}{Q_{0f}}\right)100\right),$$

and for the cost C (i.e., the number of touches on the touchscreen required within a trial to receive the food item) associated with each phase of testing for obtaining the food f,

$$\phi_f = \log_{10}\left(\frac{\left(CQ_{0f}\right)}{100}\right),$$

where Q_{0f} is the estimated quantity of food eaten with minimal cost. Next, and again following Schwartz et al. (2016), we

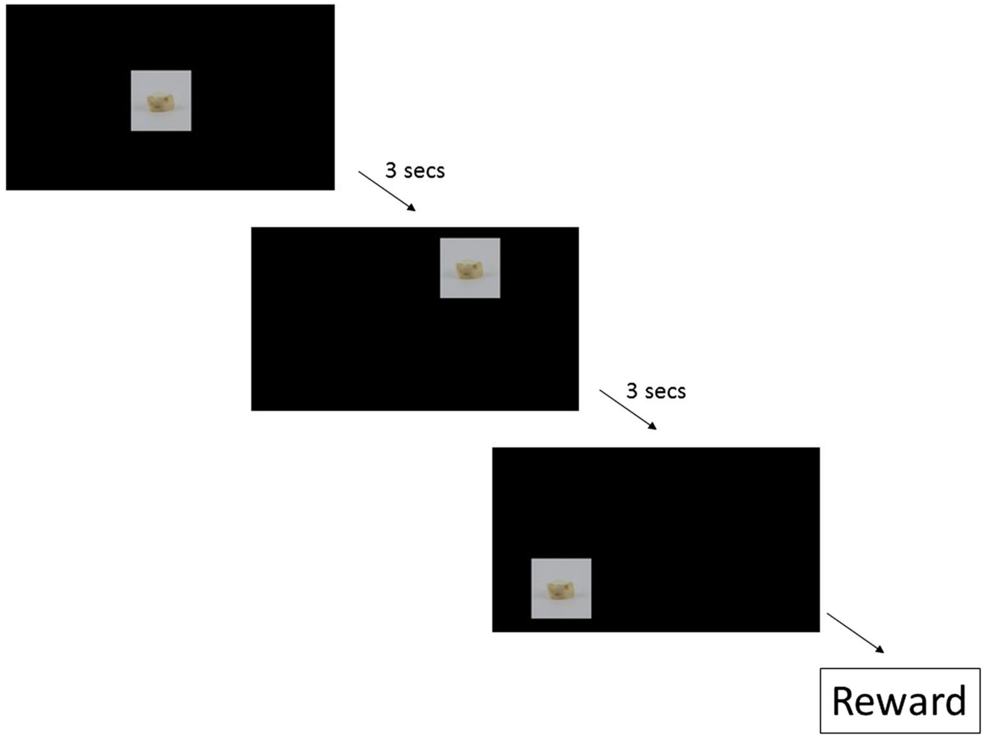

Fig. 4 Screenshots of trials in Experiment 2 showing the 3-cost protocol as the example, in which the subject had to select an image of a food reward three times before receiving that reward (in this case, a piece of turnip)

fitted the logarithm of normalized consumption as a function of the logarithm of normalized price with the exponential demand equation proposed by Hursh and Silberberg (2008):

$$\psi_f = \log_{10} Q_{0f} + k\left(e^{-aQ_{0f}\phi_f} - 1\right),$$

where ψ_f is the dependent variable (i.e., the log-normalized quantity of each food eaten at a given cost), k is the theoretical (estimated) range of ψ_f in logarithmic units (for this study, we set k at 1), and α is a parameter that determines the rate of decline in relative consumption (log consumption) with increases in price. Here it becomes clear that Q_0, the estimated quantity of food eaten with zero cost, becomes the intercept value of this linear equation.

In developing this equation, Hursh and Silberberg (2008), building on the earlier work of Allen (1962), proposed that the parameter α is negatively related to value and allows for direct comparisons to be drawn across the rewards offered to subjects, or across the same reward offered to multiple subjects. To estimate Q_0 and α, we applied a nonlinear least-squares estimate using the nls function for R (Baty et al., 2015).

To compare the gorilla's preferences for the four foods via the two methods (Exp. 1 and Exp. 2), we correlated his essential values (i.e., $1/\alpha$) for the four foods against his worth values for them (*sensu* Schwartz et al., 2016). Since a Schapiro–Wilk test revealed that the worth values were not normally distributed, we used a Spearman's correlation test for this comparison.

As for Experiment 1, all analyses were run in R version 3.3.2 (R Core Team, 2017) with RStudio (RStudio Team, 2016), The data and R script are available here: https://drive.google.com/drive/folders/1VOqttUUcQv-QQ9I2En5h7K5vf7UVZ6OC.

Results

Demand curves for the four foods revealed that as the cost required to obtain the foods increased, the subject was less willing to exert the required effort (Fig. 5). Comparing the demand curves in Figure 5 also shows the subject's relative preferences for the foods. The subject's essential values ($1/\alpha$) and r^2 values are presented in Table 1, again highlighting his differential valuations of the four foods, with grape as his most highly valued food item and cucumber as the lowest-valued item. There was not a significant correlation between the subject's worth values and essential values (rho = .4, r^2 = .88, p = .75; Fig. 6), although we note that the correlation coefficient (.32) suggests only a medium effect size, likely due to the small sample size (N = 4).

Discussion

To ascertain a subject's true valuation of an item, it is important to determine what cost that individual will "pay" to obtain it. For example, laboratory-housed mice will exert effort to obtain additional space in their home cage (Sherwin &

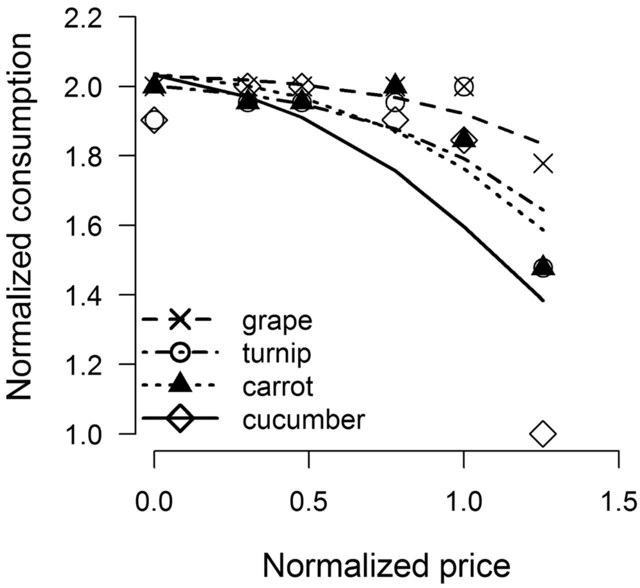

Fig. 5 Normalized numbers of food items consumed by the subject as a function of the normalized number of presses on the touchscreen that had to be made to obtain a food reward (effort). All four food types are presented together so that the different demand curves for each food type may be compared visually

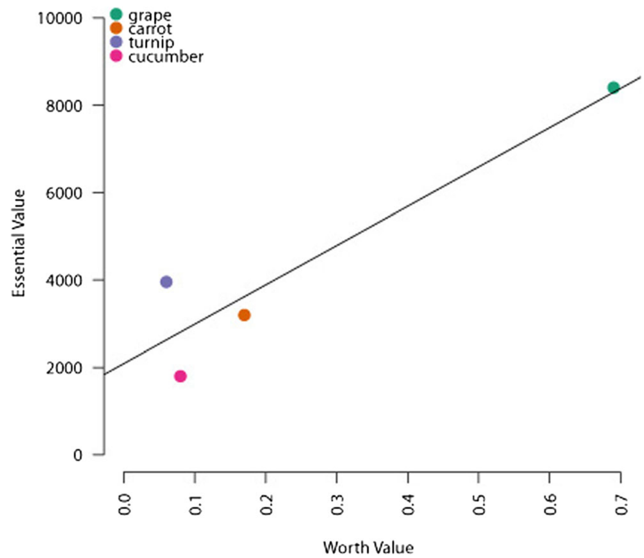

Fig. 6 The essential value of each of the four food rewards as a function of the subject's worth values for the same foods

Nicol, 1996); zoo-housed animals perform better in operant-training sessions when rewarded with preferred foods (Gaalema et al., 2011); and captive primates will walk farther to obtain preferred foods than they will for lower-valued foods (Hopper et al., 2015; Stevens, Rosati, Ross, & Hauser, 2005). In our study, as the cost to obtain the four foods increased, the subject's willingness to complete the required number of touchscreen presses to obtain the food rewards decreased. However, his responses, as shown by the demand curves, indicate that his willingness to work for the foods differed. The close resemblance of the demand curves generated from the subject's responses in Experiment 2, when tested with a touchscreen paradigm, to those of the capuchin monkeys tested via a manual task by Schwartz et al. (2016) provides support for the efficacy of using touchscreens to evaluate primates' food valuations.

The essential values for the foods revealed that grape was the food that the subject valued most highly, mirroring the results of Experiment 1. Experiment 1 revealed that his least-preferred foods of the four presented were turnip and cucumber. Similarly, his responses in Experiment 2 indicated that he valued cucumber lowest of the four foods. However, there was no significant relationship between the subject's worth values (Exp. 1) and essential values (Exp. 2) for the foods, although the two values were positively correlated. This lack of correlation across the two measures matches the findings of Schwartz et al. (2016), who also found no correlation between the capuchin's food preference rankings and their essential values for the same foods; however, our limited sample size (one subject, four foods) prevents us from making definitive conclusions about the relationship of the two metrics.

Table 1 Foods, essential values ($1/\alpha$), and r^2 values depicted in Figure 5

Food	Essential Value ($1/\alpha$)	r^2
Grape	8,403.36	.846
Turnip	3,952.57	.744
Carrot	3,194.89	.902
Cucumber	1,798.56	.829

General discussion

This assessment of testing primates' food preferences and valuations using a touchscreen interface demonstrated that this was an efficient and reliable method. The preference trials (Exp. 1) could be run quickly and without experimenter bias, and the subject readily participated in the test sessions because every trial was rewarded (although no secondary reinforcer was given). The success of the preference-testing protocol in testing the subject's relative food preferences, from both a practical and a theoretical standpoint, suggests that it should be further validated with other foods with this subject, with other subjects, and with other test items. Such validation will be key to evaluating the method's applied success. Additionally, Experiment 2 proved the efficacy of using touchscreens to test primates' food valuations by providing a method to change the required cost to obtain the different foods. Although this was a more time-consuming protocol than the preference-testing method, the results may be more theoretically informative (see Schwartz et al., 2016, for a discussion). Again, further testing is warranted to validate this technique.

Although the subject's preferences and valuations of the foods were not significantly correlated, there was a clear positive correlation between the two measures. It will be important to determine whether this relationship would be found to be significant if more subjects were tested on this paradigm. Schwartz et al. (2016) did not find a correlation between their subjects' essential values and food preference scores. However, rather than using a nuanced measure of preference, as we did, Schwartz and colleagues only recorded the monkeys' preferences as rank scores, which does not reveal relative degrees of preference across foods. We propose that using worth values as a preference score provides a more detailed picture of an animal's relative preferences than does using rank scores, and we propose that worth scores might be as informative as essential values.

Moving forward, it is likely that these methods will be applicable for use with any species that has good visual acuity and can use a touchscreen. Offering primates and other captive animals the option to make selections as part of their daily routine is inherently beneficial to enhancing their welfare (Kurtycz, 2015; Schapiro & Lambeth, 2007) and can impact their performance in cognitive testing. For example, a recent study showed that providing macaques with a choice over the order in which they completed a suite of cognitive tasks improved their performance as compared to yoked sessions in which the same test order was presented, but in a predetermined manner (Perdue, Evans, Washburn, Rumbaugh, & Beran, 2014). In this way, the paradigm that we evaluated could be used, not only to test preferences generally, but also provide captive animals with the opportunity to make "in the moment" choices, including for test conditions.

Despite the efficacy of the preference-testing protocol (Exp. 1), it is possible that this method might be vulnerable to subjects with strong side biases, in which they habitually select an image on one side of the touchscreen. It is likely that the randomization of stimulus presentation would mitigate this problem (and indeed we did not observe such side-bias patterns in our test subject), or that this protocol might actually help eliminate a subject's side biases if the subject was motivated to obtain more-preferred rewards. Additionally, without additional control conditions we cannot ascertain with the methods we have used here whether the subject perceived the stimuli as photographic representations of the food or simply as cues that he associated with the different rewards, learned via the initial training trials. Given primates' demonstrated ability to make cross-modal associations (e.g., from 2-D to 3-D stimuli; Davenport, Rogers, & Russell, 1975; Leighty, Menzel, & Fragaszy, 2008; Parron et al., 2008; see Person, 2008, for a review), such associations likely explain the success of the gorilla in our experiment.

Furthermore, we identified two key limitations to the methodology that we used in Experiment 2 to assess the value of learning the ape's valuation of the four foods. First, because we took advantage of existing software that we adapted for this study, there was a 3-s interval between the presentation of the stimuli *within* trials. Thus, the cost varied across phases not only through the number of selections (image presses) that the ape had to make, but also through the trial time. For example, at the 3-cost level (Fig. 4), the trial duration was three selections separated by two 3-s stimulus presentation intervals (minimally, 6 s total), whereas at the 6-cost level, the trial duration was six selections separated by five 3-s stimulus presentation intervals (minimally, 15 s total). Thus, we created a test of delayed gratification, although it is notable that the subject was actively engaged throughout the task, which may have aided his performance (Evans & Beran, 2007). Elimination of the interstimulus presentation time might result in a decrease in the total trial duration across cost levels, perhaps facilitating an increase in the number of trials per session.

Second, because the subject had to press a single image that was repeatedly presented within a trial, the subject was unable to assess the required cost to obtain the food at the beginning of each trial. Given animals' preference for discriminative over nondiscriminative stimuli (e.g., Roper & Zentall, 1999), we propose a modification to our methods that would provide subjects with more concrete information about the cost of each trial. In Schwartz et al. (2016), monkeys were required to lift a weighted tray in order to obtain a reward. Thus, the subjects could sample the cost required at the beginning of a trial to determine whether they wished to exert the required effort to obtain the reward. We propose that our touchscreen paradigm could be modified so that all the stimuli that the subject must select to obtain the food reward would be presented simultaneously on the screen at the beginning of the trial. In this way, at the start of each trial the subject would be shown the cost of a given reward, represented by the number of stimuli on the screen. Given primates' (Beran, 2001; Judge, Evans, & Vyas, 2005; Santos, Barnes, & Mahajan, 2005) and other species' (e.g., *Pterophyllum scalare*: Gomez-Laplaza & Gerlai, 2011; *Tursiops truncates*: Jaakkola, Fellner, Erb, Rodriguez, & Guarino, 2005; *Ursus americanus*: Vonk & Beran, 2012) counting and quantity estimation skills, such an approach should be applicable across species. Furthermore, since subjects' refusals (due to cost relative to reward value) would occur at the beginning of the trial, rather than midway through a trial (as occurred with our present protocol), more trials could be run in a shorter time, conferring practical benefits to this proposed modification of our protocol.

Building on previous work testing primates' preferences via both manual and eye-gaze protocols, we have evaluated two touchscreen methods that can be successfully applied to testing primates' preferences for, and valuations of, different food items. Our study adds to the growing number of measures that have been developed and evaluated to assess preferences (e.g., Clark, Howard, Woods, Penton-Voak, & Neumann, 2018). Future work will be required in order to test the applicability of these methods for use with other species, as well as the potential concord between worth values and essential values, by assessing these protocols with more subjects, but we believe that they have the potential to be broadly applicable. It would also be informative to ascertain the vulnerability of individuals' preferences to satiation, which is especially important when considering reinforcer values (e.g., Baxter et al., 2000) or in relation to motivation and effort (e.g., Pace, Ivancic, Edwards, Iwata, & Page, 1985; Svartdal & Mortensen, 1993), and which could be tested using paradigms akin to the protocol we used in Experiment 2. These methods could also be used to test how animals' selection behavior is modulated by relative preferences for the paired stimuli (*sensu* Sánchez-Amaro, Peretó, & Call, 2015) or their preferences for nonfood items (e.g., group mates, enrichment devices; *sensu* Adams & MacDonald, 2018). Knowing not just what animals prefer, but how motivated they are to obtain certain items, is key when working in a captive context, in terms of both maximizing welfare and enhancing experimental protocols (e.g., Fay & Miller, 2015).

Author note We are grateful to the editor and two anonymous reviewers, who provided useful feedback on an earlier draft of this article. We thank Katherine Cronin and Sarah Jacobson for helpful discussions on the design of this study. We also thank Maureen Leahy, Jill Moyse, and the animal care staff at Lincoln Park Zoo's Regenstein Center for African Apes, who provide the best possible care for the apes and who are so supportive of our research. This research was funded by the Chauncey and Marion Deering McCormick Foundation, the Leo S. Guthman Fund, and the Abra Prentice Foundation.

References

Adams, L. C., & MacDonald, S. E. (2018). Spontaneous preference for primate photographs in Sumatran orangutans (*Pongo abelii*). *International Journal of Comparative Psychology*, *31*. Retrieved from https://escholarship.org/uc/item/08t203bk

Allen, R. G. D. (1962). Mathematical analysis for economists. New York, NY: Macmillan.

Allritz, M., Call, J., & Borkenau, P. (2016). How chimpanzees (*Pan troglodytes*) perform in a modified emotional Stroop task. *Animal Cognition*, *19*, 435–449.

Baker, S. G. (1994). The multinomial-Poisson transformation. *Journal of the Royal Statistical Society: Series D*, *43*, 495–504. doi:https://doi.org/10.1002/0471667196.ess0683.pub2

Baty, F., Ritz, C., Charles, S., Brutsche, M., Flandrois, J.-P., & Delignette-Muller, M.-L. (2015). A toolbox for nonlinear regression in R: The package nlstools. *Journal of Statistical Software*, *66*, 1–21. doi:https://doi.org/10.18637/jss.v066.i05

Baxter, M. G., Parker, A., Linder, C. C. C., Izquierdo, A. D., & Murray, E. A. (2000). Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *Journal of Neuroscience*, *20*, 4311–4319.

Bentzley, B. S., Fender, K. M., & Aston-Jones, G. (2013). The behavioral economics of drug self-administration: A review and new analytical approach for within-session procedures. *Psychopharmacology*, *226*, 113–125.

Beran, M. J. (2001). Summation and numerousness judgments of sequentially presented sets of items by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *115*, 181–191.

Beran, M. J. (2012). Did you hear the one about the horse that could count? *Frontiers in Psychology*, *3*, 357. doi:https://doi.org/10.3389/fpsyg.2012.00357

Boysen, S. T., Berntson, G. G., Hannan, M. B., & Cacioppo, J. T. (1996). Quantity-based interference and symbolic representations in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, *22*, 76–86.

Brosnan, S. F., & de Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature*, *425*, 297–299.

Brown, E. K., Templer, V. L., & Hampton, R. R. (2017). An assessment of domain-general metacognitive responding in rhesus monkeys. *Behavioural Processes*, *135*, 132–144.

Cantlon, J. F., & Brannon, E. M. (2007). Basic math in monkeys and college students. *PLoS Biology*, *5*, e328. doi:https://doi.org/10.1371/journal.pbio.0050328

Canty, A., & Ripley, B. (2017). Bootstrap functions (Originally by Angelo Canty for S). Retrieved from https://CRAN.R-project.org/package=boot

Casey, A. H., Silberberg, A., Paukner, A., & Suomi, S. J. (2014). Defining reward by cross-modal scaling. *Animal Cognition*, *17*, 177–183.

Claidière, N., Whiten, A., Mareno, M. C., Messer, E., Brosnan, S. F., Hopper, L. M., … McGuigan, N. (2015). Selective and contagious prosocial resource donation in capuchin monkeys, chimpanzees and humans. *Scientific Reports*, *5*, 7631.

Clark, A. P., Howard, K. L., Woods, A. T., Penton-Voak, I. S., & Neumann, C. (2018). Why rate when you could compare? Using the "EloChoice" package to assess pairwise comparisons of perceived physical strength. *PLoS ONE*, *13*, e0190393. doi:https://doi.org/10.1371/journal.pone.0190393

R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.R-project.org

Davenport, R. K., Rogers, C. M., & Russell, I. S. (1975). Cross-modal perception in apes: Altered visual cues and delay. *Neuropsychologia*, *13*, 229–235.

Evans, T. A., & Beran, M. J. (2007). Chimpanzees use self-distraction to cope with impulsivity. *Biology Letters*, *3*, 599–602.

Fay, C., & Miller, L. J. (2015). Utilizing scents as environmental enrichment: Preference assessment and application with Rothschild giraffe. *Animal Behavior and Cognition*, *2*, 285–291.

Fernandez, E. J., Dorey, N. R., & Rosales-Ruiz, J. (2004). A two-choice preference assessment with five cotton-top tamarins (*Saguinus oedipus*). *Journal of Applied Animal Welfare Science*, *7*, 163–169.

Finestone, E., Bonnie, K. E., Hopper, L. M., Vreeman, V. M., Lonsdorf, E. V., & Ross, S. R. (2014). The interplay between individual, social, and environmental influences on chimpanzee food choices. *Behavioural Processes*, *105*, 71–78.

Fortes, I., Case, J. P., & Zentall, T. R. (2017). Pigeons, unlike humans, do not prefer near hits in a slot-machine-like task. *Behavioural Processes*, *138*, 67–72.

Gaalema, D. E., Perdue, B. M., & Kelling, A. S. (2011). Food preference, keeper ratings, and reinforcer effectiveness in exotic animals: The value of systematic testing. *Journal of Applied Animal Welfare Science*, *14*, 33–41.

Gomez-Laplaza, L. M., & Gerlai, R. (2011). Can Angelfish (*Pterophyllum scalare*) count? Discrimination among different shoal sizes follows Weber's law. *Animal Cognition*, *14*, 1–9.

Hatzinger, R. (2015). Package "prefmod": Utilities to fit paired comparison models for preferences. Retrieved from https:/CRAN.R-project.org/package=prefmod

Hatzinger, R., & Dittrich, R. (2012). premod: An R package for modeling preferences based on paired comparisons, rankings, or ratings. *Journal of Statistical Software*, *48*(10). doi:https://doi.org/10.18637/jss.v048.i10

Hopper, L. M., Kurtycz, L. M., Ross, S. R., & Bonnie, K. E. (2015). Captive chimpanzee foraging in a social setting: A test of problem solving, flexibility, and spatial discounting. *PeerJ*, *3*, e833.

Hopper, L. M., Lambeth, S. P., Schapiro, S. J., & Brosnan, S. F. (2014). Social comparison mediates chimpanzees' responses to loss, not frustration. *Animal Cognition*, *17*, 1303–1311.

Hopper, L. M., Schapiro, S. J., Lambeth, S. P., & Brosnan, S. F. (2011). Chimpanzees' socially maintained food preferences indicate both conservatism and conformity. *Animal Behaviour*, *81*, 1195–1202.

Howard, L. H., Wagner, K. E., Woodward, A. L., Ross, S. R., & Hopper, L. M. (2017). Social models enhance apes' memory for novel events. *Scientific Reports*, *7*, 40926.

Hursh, S. R., & Silberberg, A. (2008). Economic demand and essential value. *Psychological Review*, *115*, 186–198.

Hursh, S. R., & Winger, G. (1995). Normalized demand for drugs and other reinforcers. *Journal of the Experimental Analysis of Behavior*, *64*, 373–384.

Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, *17*, R1004–R1005.

Jaakkola, K., Fellner, W., Erb, L., Rodriguez, M., & Guarino, E. (2005). Understanding of the concept of numerically "less" by bottlenose dolphins (*Tursiops truncatus*). *Journal of Comparative Psychology*, *119*, 286–303.

Joseph, B., Egli, M., Koppekin, A., & Thompson, T. (2002). Food choice in people with Prader–Willi syndrome: Quantity and relative preference. *American Journal on Mental Retardation*, *107*, 128–135.

Judge, P. G., Evans, T. A., & Vyas, D. K. (2005). Ordinal representation of numeric quantities by brown capuchin monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, *31*, 79–94.

Judge, P. G., Kurdziel, L. B., Wright, R. M., & Bohrman, J. A. (2012). Picture recognition of food by macaques (*Macaca silenus*). *Animal Cognition*, *15*, 313–325.

Kemnitz, J. W., & Francken, G. A. (1986). Characteristics of spontaneous obesity in male rhesus monkeys. *Physiology and Behavior*, *38*, 477–483.

Kurtycz, L. M. (2015). Choice and control for animals in captivity. *Psychologist*, *28*, 892–894.

Kuwahata, H., Adachi, I., Fujita, K., Tomonaga, M., & Matsuzawa, T. (2004). Development of schematic face preference in macaque monkeys. *Behavioural Processes*, *66*, 17–21.

Leighty, K. A., Menzel, C. R., & Fragaszy, D. M. (2008). How young children and chimpanzees (*Pan troglodytes*) perceive objects in a 2D display: Putting an assumption to the test. *Developmental Science*, *11*, 778–792.

Martin, C. F. (2017). ApeTouch software suite [Computer software]. Indianapolis, IN: Author.

Martin, C. F., Biro, D., & Matsuzawa, T. (2014). The Arena System: A novel shared touch-panel apparatus for the study of chimpanzee social interaction and cognition. *Behavior Research Methods*, *46*, 611–618. doi:https://doi.org/10.3758/s13428-013-0418-y

McDermott, J., & Hauser, M. (2004). Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primates. *Cognition*, *94*, B11–B21.

Pace, G. M., Ivancic, T., Edwards, G. L., Iwata, B. A., & Page, T. J. (1985). Assessment of stimulus preference and reinforce value with profoundly retarded individuals. *Journal of Applied Behavior Analysis*, *18*, 249–255.

Parron, C., Call, J., & Fagot, J. (2008). Behavioural responses to photographs by pictorially naive baboons (*Papioi abubis)*, gorillas (*Gorilla gorilla*) and chimpanzees (*Pan troglodytes*). *Behavioural Processes*, *78*, 351–357.

Perdue, B. M. (2016). The effect of computerized testing on sun bear behavior and enrichment preferences. *Behavioral Sciences*, *6*, 19.

Perdue, B. M., Evans, T. A., Washburn, D. A., Rumbaugh, D. M., & Beran, M. J. (2014). Do monkeys choose to choose? *Learning & Behavior*, *42*, 164–175. doi:https://doi.org/10.3758/s13420-014-0135-0

Persson, T. (2008). Pictorial primates: A search for iconic abilities in great apes (PhD thesis; Lund University Cognitive Studies 136). Retrieved May 6, 2018, from http://lup.lub.lu.se/record/938491

Pfungst, O. (2015). *Clever Hans (The horse of Mr. von Osten): A contribution to experimental animal and human psychology.* London, UK: Forgotten Books. (Original work published 1911)

Roper, K. L., & Zentall, T. R. (1999). Observing behavior in pigeons: The effect of reinforcement probability and response cost using a symmetrical choice procedure. *Learning and Motivation*, *30*, 201–220.

RStudio Team. (2016). RStudio: Integrated development for R. Boston, MA, USA: RStudio, Inc. Retrieved from www.rstudio.com/

Sánchez-Amaro, A., Peretó, M., & Call, J. (2015). Differences in between-reinforcer value modulate the selective-value effect in great apes (*Pan troglodytes, P. paniscus, Gorilla gorilla, Pongo abelii*). *Journal of Comparative Psychology*, *130*, 1–12. doi:https://doi.org/10.1037/com0000014

Santos, L. R., Barnes, J. L., & Mahajan, N. (2005). Expectations about numerical events in four lemur species *Eulemur fulvus*, *Eulemur mongoz*, *Lemur catta* and *Varecia rubra*. *Animal Cognition*, *8*, 253–262.

Schapiro, S. J., & Lambeth, S. P. (2007). Control, choice, and assessments of the value of behavioral management to nonhuman primates in captivity. *Journal of Applied Animal Welfare Science*, *10*, 39–47.

Schwartz, L. P., Silberberg, A., Casey, A. H., Paukner, A., & Suomi, S. (2016). Scaling reward value with demand curves versus preference tests. *Animal Cognition*, *19*, 631–641.

Sherwin, C. M., & Nicol, C. J. (1996). Reorganization of behaviour in laboratory mice, *Mus musculus*, with varying cost of access to resources. *Animal Behaviour*, *51*, 1087–1093.

Stevens, J. R., Rosati, A. G., Ross, K. R., & Hauser, M. D. (2005). Will travel for food: Spatial discounting in two new world monkeys. *Current Biology*, *15*, 1855–1860.

Svartdal, F., & Mortensen, T. (1993). Effects of reinforcer value on sensitivity to non-verbal operant contingencies in humans. *Quarterly Journal of Experimental Psychology*, *46A*, 347–364.

Tanaka, M. (2003). Visual preference by chimpanzees (*Pan troglodytes*) for photos of primates measured by a free choice-order task: Implication for influence of social experience. *Primates*, *44*, 157–165.

Tanaka, M. (2007). Development of the visual preference of chimpanzees (*Pan troglodytes*) for photographs of primates: Effect of social experience. *Primates*, *48*, 303–309.

Turner, H., & Firth, D. (2015). *Generalized nonlinear models*. Retrieved from http://go.warwick.ac.uk/gnm

Vale, G. L., Flynn, E. G., Kendal, J., Rawlings, B., Hopper, L. M., Schapiro, S. J., … Kendal, R. L. (2017). Testing differential use of payoff-biased social learning strategies in children and chimpanzees. *Proceedings of the Royal Society B*, *284*, 20171751.

Vonk, J., & Beran, M. (2012). Bears "count" too: Quantity estimation and comparison in black bears (*Ursus americanus*). *Animal Behaviour*, *84*, 231–238.

Wagner, K. E., Hopper, L. M., & Ross, S. R. (2016). Asymmetries in the production of self-directed behaviour by chimpanzees and gorillas during a computerized cognitive test. *Animal Cognition*, *19*, 343–350.

Wallis, J. D., & Miller, E. K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *European Journal of Neuroscience*, *18*, 2069–2081.

Wobber, V., Hare, B., & Wrangham, R. (2008). Great apes prefer cooked food. *Journal of Human Evolution*, *55*, 340–348.

Zeagler, C., Zuerndorfer, J., Lau, A., Freil, L., Gilliland, S., Starner, T., & Jackson, M. M. (2016). Canine computer interaction: Towards designing a touchscreen interface for working dogs. In ACI '16: Proceedings of the Third International Conference on Animal–Computer Interaction (Article 2). New York, NY: ACM Press. doi: https://doi.org/10.1145/2995257.2995384