A flexible behavior task allows quantitative analysis of cost-benefit evaluation in mice

Nathan A. Schneider, John Lisman, Hyun Jae Pi *

Affiliations:

Volen Center for Complex Systems,

Department of Biology,

Brandeis University, Waltham, Massachusetts 02453, USA.

*Correspondence: Hyun Jae Pi (hyunpi@brandeis.edu)

Keywords: decision-making, neuroeconomics, cost-benefit analysis, valuation, optimization, mouse behavior, switching decision

Acknowledgements

This work is supported by NARSAD Young Investigator Grant and NIH R01MH110391. The authors have no conflicting interests. We thank the Lisman laboratory for assistance, Drs. Don Katz, Sachin Ranade, Yongsoo Kim, Geoffrey Schoenbaum, and Andrew Wikenheiser for comments on the manuscript, and Drs. Shantanu Jadhav, Michael Halassa, and Michael Hasselmo for suggestions and ideas for experiments.

Abstract

Cost-benefit analysis is a key determinant of decision-making, yet little is known about the underlying neural circuit mechanisms, and investigating this concept using laboratory animals is challenging without quantitative behavioral readouts and theoretical frameworks. In order to tackle this challenge, here we took a theory-based approach and created an experimental design in order for mice to perform a task as a rationale agent and optimize decisions. Our novel behavioral paradigm offers two possibilities of obtaining reward. Mice initially prefer the easier method, but as its cost increases, their preference switches. We quantified these switching decisions using an indifference point where the values of the choices became equivalent. A novel component of our task design is the implementation of a systematic and flexible variation of cost and benefit parameters in two-dimensional parameter space. This allows researchers to choose a wide range of parameter combinations and quantify the shift of the switching decisions. Our results demonstrated that indeed the parametric manipulation successfully influenced the switching decisions relative to the given parameter values, suggesting that mice perform the task as rational agents. A theoretical framework based on the optimization principle further confirmed the switching decisions of mice were optimal. Furthermore, we demonstrated that although the preference shift was influenced by different internal states of motivation, the valuation process of mice was intact regardless of motivational states. Considering the arsenal of tools available for mice, our theory-centered approach provides a quantitative

and flexible platform to investigate the neural circuits underlying cost-benefit assessment in mice.

Introduction

When confronted with an uncertain environment, successful animal behavior maximizes the benefits and minimizes the costs associated with, for instance, acquiring food and mates. Across species, the correct assessment of cost and benefit is a key determinant for survival and long-term well-being (Herrnstein, 1961; Stephens and Krebs, 1985). This assessment is continually influenced by the animal's fluctuating intrinsic state and a constantly changing external environment; consequently, the best decision in one context is not always the best in another context. For instance, pursuing a low calorie diet may be a healthy choice for people in an advanced economy, but it is not a wise action for those who are uncertain about when they will have their next meal. Therefore, it is important to understand how such context-dependent factors adaptively interact with the process of cost-benefit analysis.

How is this decision process handled in the brain? One decision model suggests that the brain integrates the costs and benefits of a decision into a common neuronal currency, thus generating values for each choice that are comparable despite differences in the specific costs and benefits (Montague and Berns, 2002; Levy and Glimcher, 2012). However, this stand-alone integrator system cannot solve problems adaptably. Additional systems should exist in the brain that isolate the contributions of each decision variable (e.g. costs and benefits), update their values separately as intrinsic state and external stimuli fluctuate (Sugrue et al., 2004; Walton et al., 2006; Kolling et al., 2012), and feed this information back to the integrator in order to reach the

best decision. In this way, the brain could adapt to changing environments and internal states and guide decisions that maximize the net utility.

Unfortunately, investigating abstract concepts such as cost in non-verbal laboratory animals is challenging especially without quantitative behavioral readouts and theoretical frameworks. Although progress has been made in our understanding of multiple forms of cost processing (O'Doherty et al., 2001; Rudebeck et al., 2006; Kim et al., 2008; Ballard and Knutson, 2009; Amemori and Graybiel, 2012; Leathers and Olson, 2012; Hosokawa et al., 2013; Kolling et al., 2016; Azab and Hayden, 2018), the underlying neural circuit mechanisms still remain elusive. Considering the arsenal of tools available for mice, which is a popular model for circuit studies in contemporary neuroscience, we took a theory-centered approach and devised a novel quantitative behavior where mice act as a rational agent and optimize decisions to maximize the reward obtained.

Inspired by several past experiments on decision-making, effort, and motivated behavior (Hursh et al., 1988; Salamone et al., 1991; Padoa-Schioppa and Assad, 2006; Rudebeck et al., 2006; Kepecs et al., 2008; Atalayer and Rowland, 2009; Padoa-Schioppa, 2009; Kolling et al., 2012; Bailey et al., 2015; Salamone et al., 2018; Sweis et al., 2018), our behavioral task combines lever pressing on progressive and fixed ratios (PR; FR) with different amounts of water reward associated to each (Hodos, 1961), and the shift of the side preference was quantified by indifference points where the values of the two sides become equivalent. To validate our behavioral paradigm, we tested the hypothesis that mice optimize switching decisions based on cost-benefit

assessment. A novel component of our task design is not simply utilizing the PR and indifference points, but the implementation of systematic and flexible variation of cost and benefit parameters in two-dimensional parameter space. This allows researchers to choose a wide range of parameter combinations and quantify the shift of the switching decisions. Applying a theoretical framework of optimization called the marginal value theorem (MVT) to our behavioral paradigm (Charnov, 1976), we demonstrated that mice indeed make optimal decisions and the shift of the preference followed the relative value of the parameters. Finally, we demonstrated that although the preference shift was influenced by different internal states of motivation, the valuation process of mice was intact regardless of motivational states.

Materials and Methods

Animals

Five female and five male mice of the strain C57BL/6J were used in this study. These animals were bred on site from mice purchased from Charles River Laboratories (Wilmington, MA). The mice tested were between the ages of 8 and 13 months. All mice were kept on a 12hr/12hr light-dark cycle. All experiments were performed according to the guidelines of the Institutional Animal Care and Use Committee at Brandeis University.

Behavioral Setup

Experiments were conducted in a rectangular acrylic testing chamber (length, 19.3 cm; width, 14.0 cm; height, 13.9 cm, SanWorks, Stony Brook, NY) with grated flooring. A metal tray was placed underneath the chamber to collect waste. The chamber contained three nose-pokes with infrared LED/infrared phototransistor pair (Digikey, Thief River Falls, MN) to detect responses. A white LED (Digikey) inside the nose-pokes was used to cue trial availability, lever pressing progress, and reward availability. Only the center port was used for reward delivery. The two end ports were used only as lights and were covered with a snuggly-fit clear plastic cap. Plastic levers were custom designed for us by SanWorks for either side of the nose-pokes. The levers were also equipped with an infrared sensor to capture lever presses. A food pellet was placed inside the testing chamber at the start of each session to allow mice to eat in between

trials. The testing chamber was situated inside of a custom-built noise-reducing box (length, 42 cm; width, 39 cm; height, 39 cm). Water reward was delivered through a solenoid valve inside the nose-pokes (Lee Valve Co, Westbrook, CT). Water was supplied by a 60 mL syringe barrel mounted to the inside wall of the box and connected to the valve with silicone tubing (1/16” x 3/16”, Saint-Gobain Tygon, Malvern, PA). The syringe was refilled after every session to maintain water pressure. Two computer speakers (AmazonBasics, Seattle, WA) were placed inside the box on either side of the testing chamber to deliver punishment sounds. An infrared camera (Logitech, Binghamton, NY) was attached to the top of the inside of the box to allow observation during sessions. The testing chamber was connected to a Bpod state machine (SanWorks). Trial events were triggered through Matlab (MathWorks, Natick, MA).

Training

Prior to training, animals were water-restricted for 24 hours. Water was given daily to maintain 85-90% of their free-drinking body weight. Training occurred in 5 phases and took about 2-3 weeks. For all stages, mice were able to move to the next phase of training on the following day if they performed ~80 or more successful trials within an hour. First, mice were placed into the testing chamber to acclimate and could enter the center nose-poke (indicated by the center nose-poke being lit) for a small water reward (4μL). The light turned off when the mouse entered the port and collected the reward. After the mouse exited the port, there was a one second delay before the next trial

began. Second, mice had to press the right lever once (indicated by the right nose-poke being lit) in order for the center nose-poke to light up and provide reward. The lever had to be depressed for at least 100 ms to register as a press. Third, mice repeated the second phase but on the left side. Fourth, the animal had to press either the left or right lever once in order to obtain reward. The trials were pseudo-randomized and the correct side to obtain reward was indicated by the corresponding nose-poke being lit. Finally, mice repeated the fourth phase, but with an increasing number of presses required each day, from 2 presses all the way up to 10 presses. Once mice completed these stages of training, they were subjected to the optimal switching task and were allowed to familiarize themselves with the task for 3-5 sessions before data collection began.

Optimal switching task

Our behavioral task involved combining a progressive ratio (PR) and fixed ratio (FR) schedule of lever pressing. The PR was associated with a large volume of water (either 6μL or 15μL) and the FR was associated with a small volume of water (3μL). In addition, the FR could either be 6 presses or 12 presses. Similar to the training phases, levers had to be depressed for 100 ms to count as a press and there was a one second delay between trials. The PR and FR sides as well as the parameter pairs were pseudo-randomly chosen at the start of each session. At the start of a trial, both the left and right nose-pokes were dimly lit, indicating that the mouse could choose either the left or right side. Once a mouse chose a side, the corresponding nose-poke would get

increasingly brighter with each press until the required number of presses was met and the center port lit up to indicate reward availability. Mice were able to freely choose either side; however, if they decided to switch sides in the middle of a trial before completing the number of presses on the initially chosen side, a punishment sound of white noise would play and the trial would end. That trial was then classified as an incomplete trial. Furthermore, if a mouse began pressing a lever but then stopped for more than 10 seconds to groom itself, eat, etc., the trial would end and that trial would be considered incomplete. A session could be anywhere from 1 to 3 hours long. A session ended if a mouse did not press either lever for a period of time longer than 5 minutes. If this did not occur within 3 hours, the session was ended by the operator. In this way, we can ensure that the mouse is well sated by the end of the session and that we capture the highest number of trials the mouse is willing to perform without keeping the animal in the chamber too long. At the end of each session, the mouse was weighed and additional water was given at the end of each session if necessary to maintain the animal's weight at 85-90% of their free-drinking body weight.

Biological vs. technical replication

We evaluated an equal number of 5 male and 5 female mice and did not see significant differences (Supplementary Fig. 1); therefore, we combined the data for further analysis. Four sessions with each parameter setting were collected (160 sessions total:

4 sessions x 4 parameter pairs x 10 mice). Therefore, our biological replicates are 10 mice and technical replicates are 4 sessions for each parameter settings.

Data analysis

All data analysis was carried out using built-in and custom-built software in Matlab (Mathworks). All data files and analysis codes are available on our website (https://sites.google.com/view/lismanlaboratoryupdated/internal?authuser=0). The first approach of the identification of indifference points was done by fitting our binary data to a sigmoid function, the Boltzmann function.

$$f(x) = \frac{1}{1+e^{\frac{(x-x_0)}{\tau}}}$$

where x_0 and τ are a 50% threshold and a slope. By assigning the value of the PR choice = 1 and that of the FR = 0 and assuming that the curve started from the PR and ended to the FR, the fitting curve was generated. An indifference trial number was estimated where the sigmoid curve crossed the midline and the number of lever presses required at the trial on the PR side was extracted from the data. This required number of lever presses provided the PR requirement at the indifference point of the session data. In the second approach, the median values of the choices were utilized. The median trial numbers of the PR and FR were first calculated and these two medians were connected with a line. An indifference trial number was defined at the intersections where this line connecting medians crosses the midline. An indifference point was then calculated by identifying the number of presses required on the PR side at this

intersection. The third approach was done with the marginal value theorem. First, we defined the net gain as the ratio between reward obtained and the number of presses (e.g. reward/effort) in a trial. The cumulative distribution of the total net gains of the PR was generated and fitted with a double-exponential function because it captured the sudden change of slope in the PR data. The cumulative distribution of the total net gains of the FR was fitted linearly, which can be interpreted as the long-term rate of gain. Then, the tangent point between two curves was identified and the indifference point was calculated by identifying the PR requirement at this tangent trial number. Statistical significance among four parameter pairs was first tested with the two-factor nonparametric Scheirer-Ray-Hare test and the post-hoc pairwise comparison was done with Wilcoxon Rank-sum test with Bonferroni correction for 4 comparison pairs. For single pair comparisons, a nonparametric Wilcoxon Rank-sum test was used. A *P value* cutoff of 0.05 was used for significance testing.

Results

A novel economic choice behavior for mice

In order to quantitatively evaluate the factors that influence economic decision-making, we developed a novel behavioral paradigm that allowed us to test our hypothesis that mice optimize switching decisions by assessing the cost and benefit of the given choices.

Water-restricted mice were required to press levers to collect water reward (Fig. 1a). Mice could freely choose either the fixed ratio (FR) lever or the progressive ratio (PR) lever. The FR lever provided a small volume of water reward when mice completed a fixed (i.e., unchanging) number of presses (Supplementary Video 1); the PR lever, meanwhile, provided a large volume of water but the required number of presses increased progressively (e.g., 2, 3, 4 …) with each visit to the side (Supplementary Video 2). At the beginning of the session, mice preferred the PR lever which afforded them a large reward with little effort compared to the FR (Fig. 1b). As the session progressed and the PR requirement became higher, however, the cost (i.e., the required number of presses) overwhelmed the benefit of the large reward. As a result, mice switched their preference to the FR with small reward, supporting the hypothesis that mice evaluate the values of given choices.

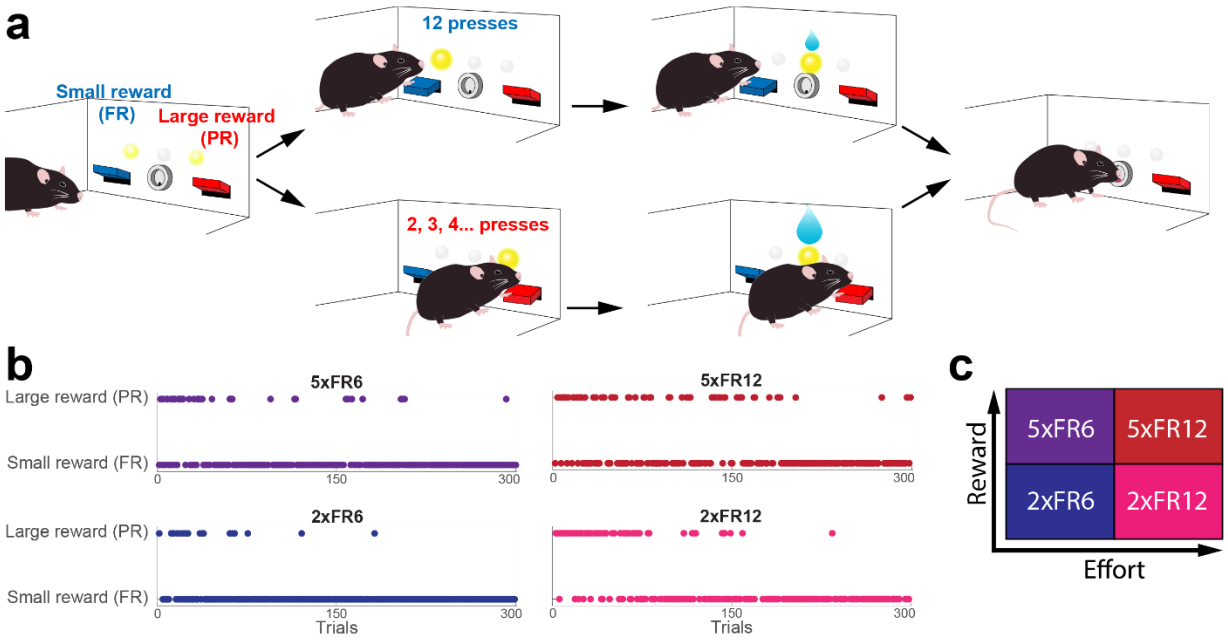

Figure 1. Optimal Switching Behavior Task. a, Schematic of task design. **b**, Example sessions for each parameter pair. Each point is a single trial within the session. Only the first 300 trials of each session are shown. All sessions are from the same animal. **c**, Grid showing increasing reward and effort for each type of parameter pair.

We wondered how the changes in relative reward size and the number of lever presses modulated the switching decisions of mice. Our design allows us to exploit behavioral data in a two-dimensional parameter space of large reward size and FR requirement (Fig. 1c), which provides a more quantitative and sensitive comparison between conditions. For instance, by fixing one parameter (e.g., reward amount) and changing the other (e.g., lever press requirement), the contribution of cost and benefit to decision-making can be evaluated semi-independently within the same task. In addition, our behavioral design provides users flexibility in choosing a wide range of parameter combinations. In this study, four combinations of parameters were chosen. The ratio of

large reward to small reward and the fixed number of presses are used to denote the combination of parameters. For instance, 2xFR12 means the volume of large reward, 6 μL, is twice as much as that of the small reward, 3 μL, and the FR requirement is 12 presses. The four combinations of parameters can be denoted as 2xFR6, 2xFR12, 5xFR6, and 5xFR12 (Fig. 1c).

Mice adjust switching decisions proportionate to the values of session parameters

In order to validate our behavioral paradigm, we tested the hypothesis that changes in the relative values of given choices would yield quantifiable differences in the switching decisions of mice. We ran 10 mice on our optimal switching task with the parameters 2xFR6, 2xFR12, 5xFR6, and 5xFR12. We evaluated an equal number of male and female mice and did not see significant differences (Supplementary Fig. 1); therefore, we combined the data for further analysis. Four sessions with each parameter setting were collected (160 sessions total: 4 sessions x 4 parameter pairs x 10 mice). The PR and FR sides and the parameter pairs were selected pseudo-randomly at the start of each session. Consistent with our initial observation, in all parameter conditions mice initially preferred the PR with large reward and switched their preference to the FR with small reward as the PR increased (Fig. 1b). Taken together, we established a quantitative behavioral paradigm and mice were able to make switching decisions by accurately evaluating the cost-benefit relationship of the given choices.

The adjustment of switching decisions can be visible even from the plot of decisions over time (Fig. 1b). The relative value of the PR with large reward is lowest at 2xFR6 because the reward is low and effort cost for the FR is also low. Alternatively, the PR value is highest at 5xFR12 because the reward is high and the alternative choice requires more effort. 2xFR12 and 5xFR6 are somewhere in between the two. Reflecting these relative values, the number of PR choices increased and spread more as the value increased. These measures were quantified in percentage and deviation of the PR choices (Fig. 2a-b, Supplementary Fig. 2). The percentages of choosing the PR with large reward were 6.5 ± 1.4%, 14 ± 3%, 7.8 ± 1.5%, and 19 ± 4% at 2xFR6, 2xFR12, 5xFR6, and 5xFR12, respectively. The percentage of the FR choices were 86%, 69%, 84%, and 63%, respectively (Fig. 2a).

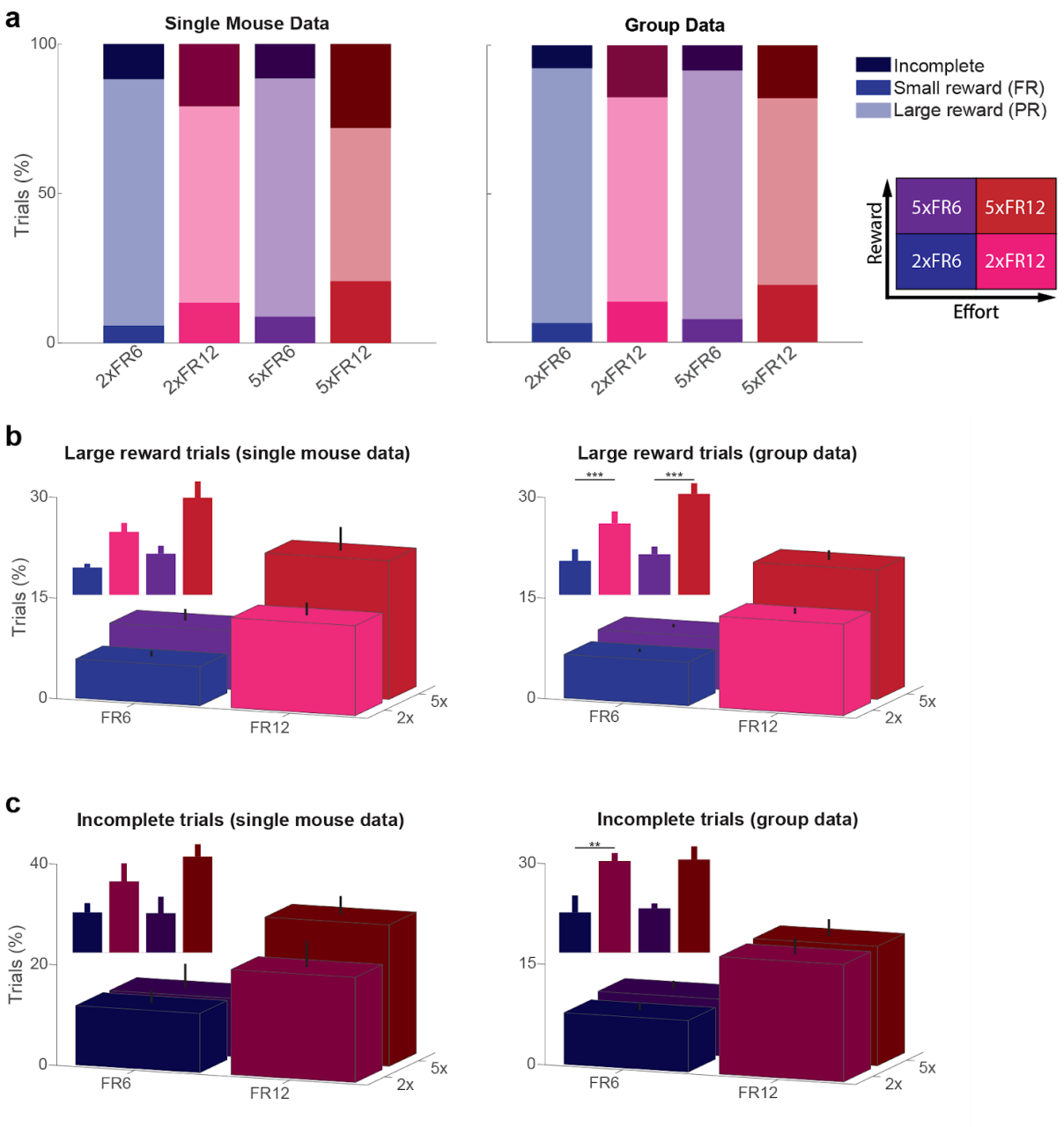

Figure 2. Trial type ratios reflect evaluation of changing cost and benefit between session parameters. a, Stacked bar graph showing trial distributions for each session (left, n = 4 sessions; right, n = 10 mice). **b**, Percent of trials where large reward (PR) was collected (left, n = 4 sessions; right, n = 10 mice). Error bars reflect the standard error of the mean. A two-way Scheirer-ray-Hare test indicated a significant effect in the FR requirement (H_1 = 28.98, p = 7 x 10^{-8}). Insert shows the same data with significant pairs notated (post hoc Wilcoxon Rank-sum test with Bonferroni correction, ***P <

0.001). **c**, Percent of trials where reward collection was failed (left, n = 4 sessions; right, n = 10 mice). Error bars reflect the standard error of the mean. A two-way Scheirer-Ray-Hare test indicated a significant effect in the FR requirement (H_1 = 16.46, p = 5 x 10^{-5}). Insert shows the same data with significant pairs notated (post hoc Wilcoxon Rank-sum test with Bonferroni correction, ***P* < 0.01).

The percentage of incomplete trials showed a similar trend as the percentage of large reward choices. A trial is considered incomplete if a mouse stopped pressing the lever in the middle of a trial for more than ten seconds or if the mouse pressed the opposite lever after initiating a trial. If the relative values of session parameters were higher, mice tried and failed more (Fig. 2c) and the failure was more frequent on the PR side (Supplementary Fig. 3-4). The incomplete trial percentages were 8 ± 3%, 17 ± 7%, 8 ± 4%, and 18 ± 8% for 2xFR6, 2xFR12, 5xFR6, and 5xFR12 respectively. Changing the cost parameter (e.g. number of FR presses) contributed more to the incomplete trials. The parameters 2xFR6 and 5xFR6 have a difference of only 0.7%, yet between 2xFR6 and 2xFR12 there is a difference of 9% (p < 0.01, Wilcoxon Rank-sum test with Bonferroni correction). The cost change also exerted a stronger effect on the task performance as indicated in the total number of trials and water collected (Supplementary Table 1). Consistent with our hypothesis, these results suggest that mice can differentiate the relative values of the session parameters and adjust their switching decisions accordingly.

Estimation of switching decisions by indifference points

When is the right moment for mice to switch their preference from the PR to the FR to maximize the gain? What is an optimal strategy to maximize gains? In our behavioral design, these questions can be explored by examining “indifference points.” Theoretically, these indifference points occur when the subjective values of each side become equal. Initially, the PR with large reward is more valuable than the FR with small reward. As the PR increases, however, the value on the PR side starts to decrease while the FR side remains fixed. At some point, the subjective value of the PR becomes equivalent to that of the FR:

$$V_{PR} = V_{FR}$$

An indifference point provides a quantitative and behavioral readout of how mice evaluate cost and gain. In our behavioral paradigm, the number of PR lever presses needed at the indifference point (PR requirement) can also be interpreted as mice’s willingness to work to collect large reward.

We estimated the PR requirement at indifference points using three different approaches: 1) fitting data with a sigmoid function; 2) utilizing median values of completed PR and FR trials; and 3) developing a measure based on an optimization theory called the marginal value theorem (Charnov, 1976; Stephens and Krebs, 1985).

The first approach for the estimation of the PR requirement at indifference points was fitting the data with a function that represented the distribution of the choices. Because our data was binary (e.g. two choices), session data were fit with a sigmoid function (Boltzmann function). Sigmoid fitting curves captured the profiles of mouse decisions, showing the transition from the PR to the FR (Fig. 3a-b). The indifference trial

number was first estimated where the sigmoid curve crossed the midline. Then the number of lever presses required for the PR at that trial was extracted from the data, which provided the PR requirement at the indifference point of the session data. Figure 3a also shows where the PR requirement lies on the cumulative distribution of the PR. The results showed that both at a single mouse level and at the animal average, the estimated PR requirement at indifference points was lowest at 2xFR6, highest at 5xFR12 and those of 2xFR12 and 5xFR6 lied in between (20 ± 4 for 2xFR6, 28 ± 4 for 2xFR12, 29 ± 4 for 5xFR6, and 45 ± 7 for 5xFR12; Fig. 3c). This shifting pattern of the PR requirement proportional to the relative values of session parameters suggests that mice perform the task as a rational agent, which is a basic premise of economic decision-making theory (Levy and Glimcher, 2012).

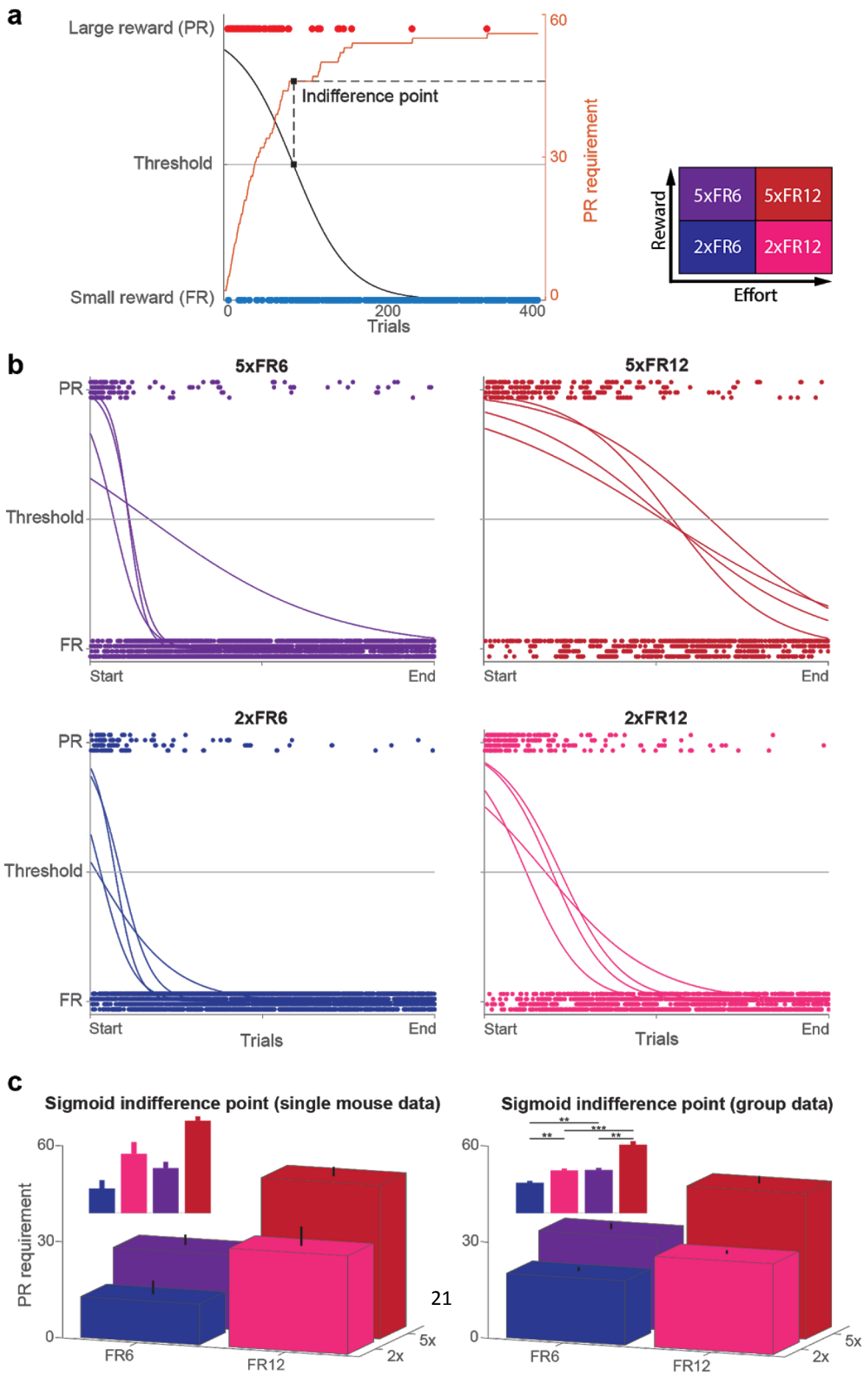
a
Large reward (PR)
Small reward (FR)
Threshold
Indifference point
PR requirement
60
30
0
0
200
400
Trials
Reward
Effort
5xFR6
5xFR12
2xFR6
2xFR12
b
5xFR6
5xFR12
2xFR6
2xFR12
PR
FR
Threshold
Start
End
Trials
c
Sigmoid indifference point (single mouse data)
Sigmoid indifference point (group data)
PR requirement
FR6
FR12
5x
2x

Figure 3. Estimating indifference points through a sigmoid curve shows variation in switching behavior between parameter conditions. a, Example session showing how the indifference point is estimated. Black curve is the sigmoid fit and orange curve is the cumulative PR. First, the data is fitted to a sigmoid curve. Then, where the curve crosses the threshold, a trial number is identified (black squares). Finally, the indifference point is determined by finding the number of presses required for the PR at this trial (dashed lines). **b**, Sigmoid curves for all 16 sessions of one mouse, sorted by session parameters. Trials are plotted as individual points normalized to the total number of trials in the session and each row of points is a different session. For sessions with higher reward or effort, the threshold intersection shifts farther to the right. **c**, Indifference points for a single mouse (left, n = 4 sessions) and the entire population (right, n = 10 mice). Error bars reflect standard error of the mean. A two-way Scheirer-Ray-Hare test indicated a significant effect in the FR requirement ($H_1 = 14.05$, $p = 0.0002$) and the PR reward volume ($H_1 = 15.72$, $p = 7 \times 10^{-5}$). Insert shows the same data with significant pairs notated (post hoc Wilcoxon Rank-sum test with Bonferroni correction, $^{**}p < 0.01$, $^{***}p < 0.001$).

Although binary data is conventionally fit with a sigmoid function, this approach did not capture the spread of the PR choices. Even after passing an indifference point, mice occasionally pursued the PR side with large reward despite the high number of required lever presses (Supplementary Fig. 3). When the relative value of the PR was higher, mice attempted to collect large reward more often (Supplementary Fig. 4). In order to capture the spread in estimation of indifference points, we utilized median values of the choices (Fig. 4a). In this approach, the median trial numbers of the PR and FR were first calculated and these two medians were connected with a line. An indifference trial number was defined at the intersections where this line connecting

medians crosses the midline. The PR requirement at the indifference point was then calculated by identifying the number of presses required on the PR side at this intersection.

The PR requirement estimated using median values showed the same trend as those estimated by the sigmoid function. 2xFR6 had the lowest value of the PR requirement, 5xFR12 had the highest, and 2xFR12 and 5xFR6 had values in between (30 ± 3 for 2xFR6, 38 ± 4 for 2xFR12, 33 ± 4 for 5xFR6, and 44 ± 5 for 5xFR12). Although the trend was similar, the estimated PR requirements from two different approaches conferred slightly different values. The standard deviations of PR requirements estimated by medians were smaller, but the relationship between the PR water volumes was not significant. In summary, the two different estimations of the PR requirement at indifference points provide a similar trend in the relative values of session parameters: the higher the value of the PR, the greater the PR requirement at the indifference point.

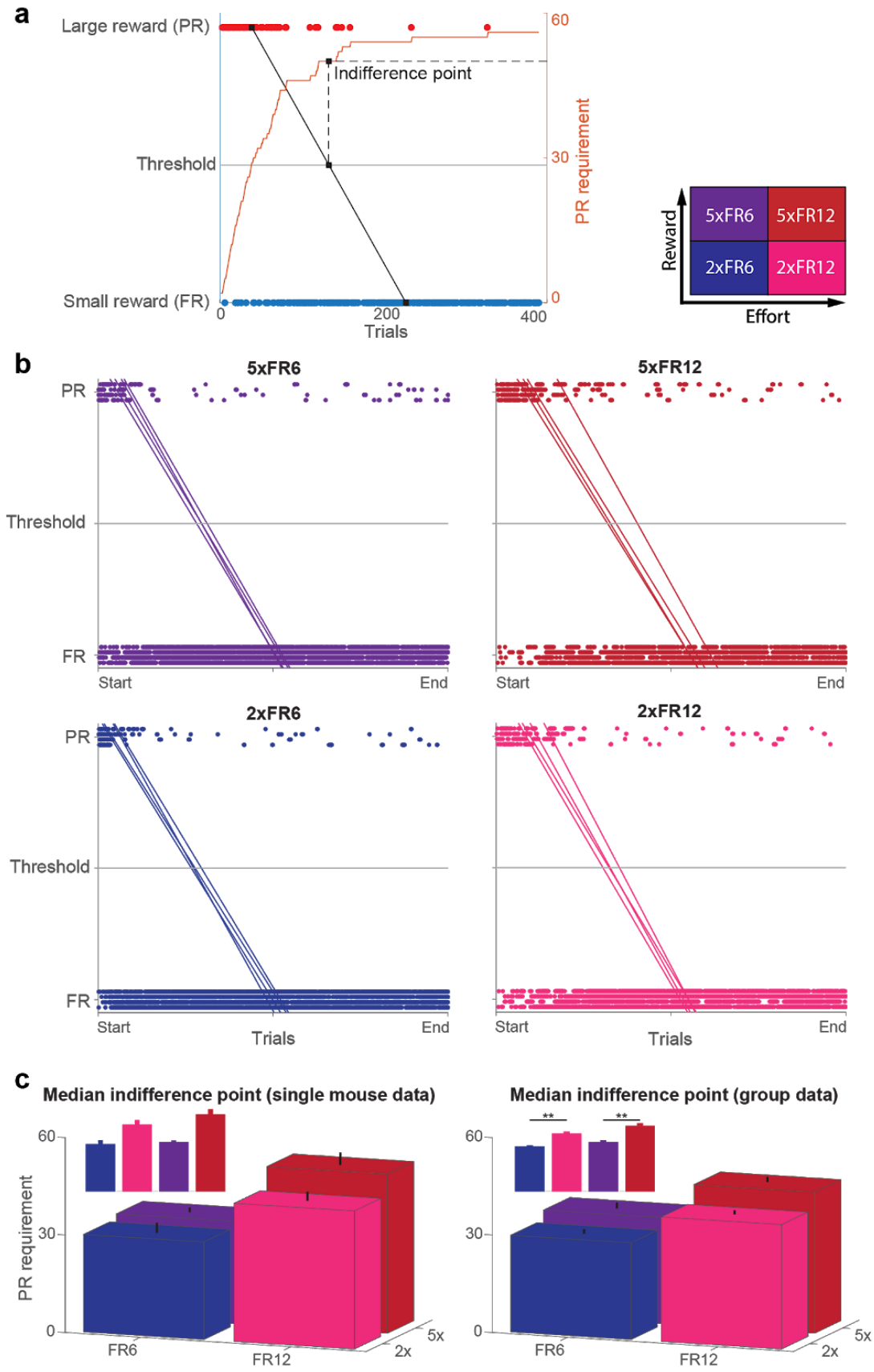
a
Large reward (PR)
Indifference point
Threshold
Small reward (FR)
60
30
0
PR requirement
0
200
400
Trials
Reward
Effort
5xFR6
5xFR12
2xFR6
2xFR12
b
5xFR6
5xFR12
2xFR6
2xFR12
PR
Threshold
FR
Start
End
Trials
c
Median indifference point (single mouse data)
Median indifference point (group data)
PR requirement
60
30
0
FR6
FR12
5x
2x
**

Figure 4. Calculating indifference points using median trial values displays variation in switching behavior between parameter conditions. a, Example session showing how the indifference point is estimated. Black line is the linear fit and orange curve is the cumulative PR. First, the median trial numbers for PR and FR trials are calculated and connected by a line. Then, where the line crosses the threshold, a trial number is identified (black squares). Finally, the indifference point is determined by finding the number of presses required for the PR at this trial (dashed lines). **b**, Median lines for all 16 sessions of one mouse, sorted by session parameters. Trials are plotted as individual points normalized to the total number of trials in the session and each row of points is a different session. **c**, Indifference points for a single mouse (left, n = 4 sessions) and the entire population (right, n = 10 mice). Error bars reflect standard error of the mean. A two-way Scheirer-Ray-Hare test indicated a significant effect in the FR requirement (H_1 = 22.70, p = 2 x 10^{-6}). Insert shows the same data with significant pairs notated (post hoc Wilcoxon Rank-sum test with Bonferroni correction, **p < 0.01).

Estimation of indifference based on a normative model

Switching decisions in our behavioral task can be viewed as an optimization problem. Our question can be rephrased as such: how do mice optimize their switching decisions to maximize the reward gained and minimize the effort? We addressed this question with the marginal value theorem (MVT), a normative model developed to explain optimal foraging behavior (Charnov, 1976; Stephens and Krebs, 1985). MVT addresses an optimal strategy of leaving an area with less food to find a better food area, thus maximizing the energy gained and minimizing the costs. In our behavioral

framework, a switching decision can also be interpreted as a leaving decision. When is the optimal moment to leave the PR side with large reward to maximize the gain?

According to MVT, the optimal strategy to maximize energy intake is to leave the food patch when the current intake rate becomes smaller than the long-term average intake rate that the predator experiences (Charnov, 1976; Stephens and Krebs, 1985). Applying this theorem to our decision-making behavior, the optimal solution is to leave the PR side when the reward/effort ratio becomes smaller than that of the FR. The optimal switching point can be identified where the slope of the FR intake rate is tangent to the curve of the PR intake rate (Fig. 5a).

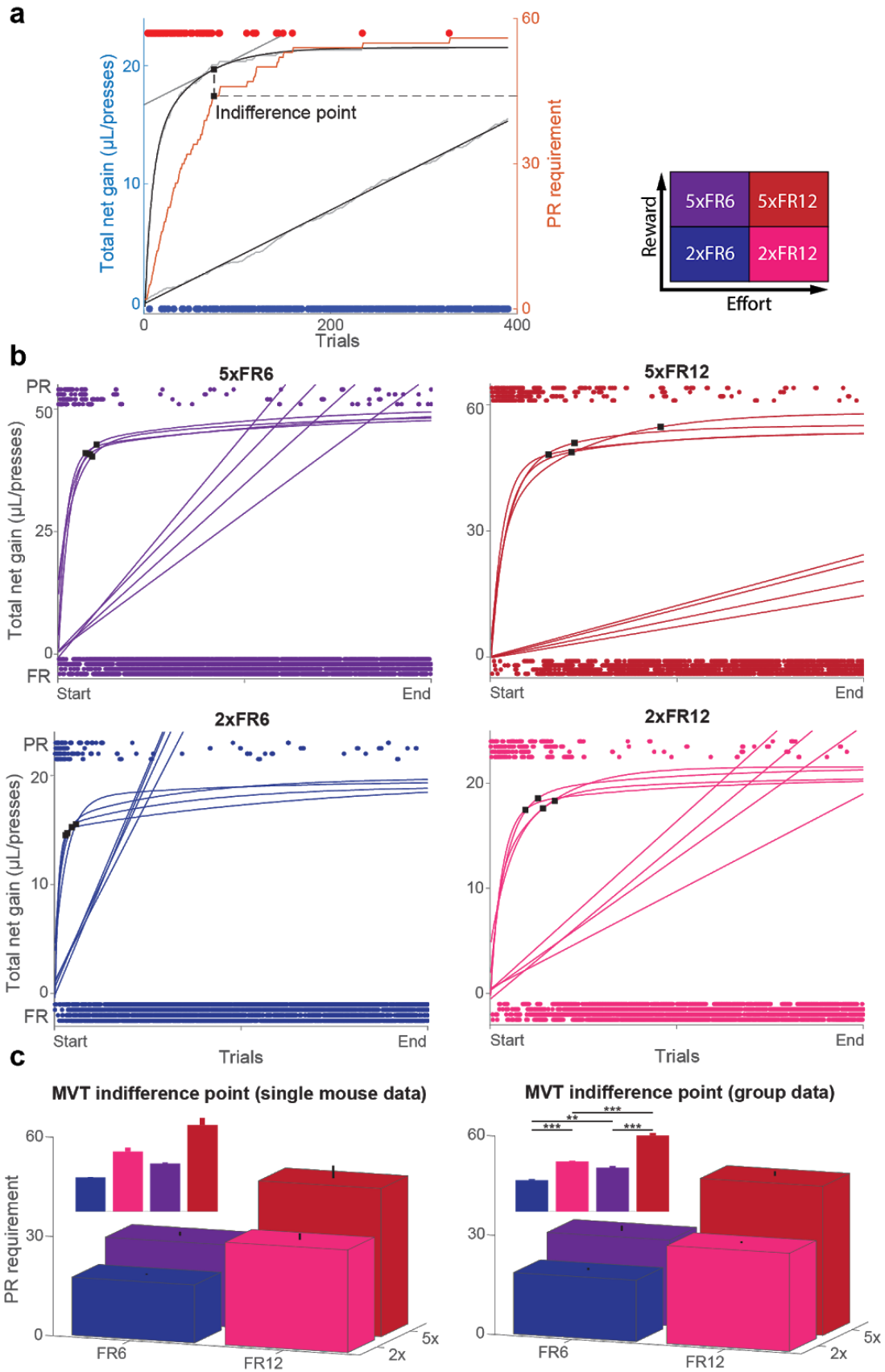
a
Total net gain (µL/presses)
PR requirement
Indifference point
Trials
5xFR6
5xFR12
2xFR6
2xFR12
Reward
Effort
b
5xFR6
5xFR12
2xFR6
2xFR12
PR
FR
Start
End
Trials
c
MVT indifference point (single mouse data)
MVT indifference point (group data)
FR6
FR12
5x
2x

Figure 5. Using a normative model to estimate indifference points. a, Example session showing how the indifference point is estimated. Black lines are the MVT curves and orange line is the cumulative PR. First, the data for PR and FR are fit to a double-exponential and linear function, respectively (methods) A tangent point is calculated where the slope of the FR line (grey line) is tangent to the PR curve (black square). Finally, the indifference point is determined by finding the number of presses required for the PR at the trial where the tangent point occurred (dashed lines). **b**, MVT model curves for all 16 sessions of one mouse, sorted by session parameters. Tangent points are indicated by black squares. Trials are plotted as individual points normalized to the total number of trials in the session and each row of points is a different session. For sessions with higher reward or effort, the tangent points shift farther to the right. **c**, Indifference points for a single mouse (left, n = 4 sessions) and the entire population (right, n = 10 mice). Error bars reflect standard error of the mean. A two-way Scheirer-Ray-Hare test indicated a significant effect in the FR requirement ($H_1 = 25.20$, $p = 5 \times 10^{-7}$) and the PR reward volume ($H_1 = 9.35$, $p = 0.002$). Insert shows the same data with significant pairs notated (post hoc Wilcoxon Rank-sum test with Bonferroni correction, $^{**}p < 0.01$, $^{***}p < 0.001$).

We applied MVT to our data. First, the cumulative distribution of the net gain from the PR was generated and was fitted with a double-exponential function that captured the sudden change in slope of the PR net gain over time. The cumulative distribution of the net gain from the FR was fitted linearly, which can be interpreted as the long-term rate of gain. Then, the tangent point between the two curves was identified. This can be visualized by shifting the FR line upward (Fig. 5a-b). Finally, the PR requirement was calculated using this tangent point trial number. The PR requirements estimated by MVT showed the same trend as the other two measures, with requirements increasing

with the relative values of the PR in session parameters. The estimated PR requirements at indifference points were 19 ± 2, 30 ± 2, 26 ± 3, and 45 ± 5 at 2xFR6, 2xFR12, 5xFR6, and 5xFR12, respectively (See Supplementary Table 2 for individual mice). One notable difference between MVT and the other approaches was that MVT generated the smallest standard deviations for all four parameters (Table 1).

	2xFR6	2xFR12	5xFR6	5xFR12
Sigmoid Fit	20 ± 4	28 ± 4	29 ± 4	45 ± 7
Medians	30 ± 3	38 ± 4	33 ± 4	44 ± 5
MVT	19 ± 2	30 ± 2	26 ± 3	45 ± 5

Table 1. Summary of indifference points. Each average indifference point and standard deviation (n = 10 mice) are compiled for easy comparison of each parameter pair and indifference point estimation method.

The PR requirements at indifference points estimated by two data-driven approaches and one based on a theoretical framework are summarized in Table 1. Comparing the PR requirements from data-driven approaches, the values obtained by medians were higher than those by sigmoid fit. Considering the estimation with medians was devised to capture the spread of choices in the later part of a session, it explained the shift of the PR requirements to the right and the higher values. Furthermore, the indifference point values estimated by MVT were very close to those estimated by sigmoid fit. Similar to our theory of the subjective values of PR and FR, MVT shows that the rate of FR remains constant while PR decreases until the two are equal. Taken together, MVT supports the data-driven approach by sigmoid fit for estimating

indifference points and strengthens our claim that the indifference point occurs when the value of the two choices becomes equivalent.

Different motivational modulation of indifference points

The valuation process is subjective and context-dependent. Motivation is one of the internal states that can strongly influence the evaluation process. The same volume of water has significantly different values to thirsty vs. sated mice. Assuming the level of thirst as a proxy of motivational level, we looked at how different levels of motivation modulated the evaluation process in our decision-making behavior.

In the data presented so far, we tried to keep the level of motivation/thirst consistent by providing a limited amount of water daily. Mice collected some water during the task and additional water was given after the task if necessary to maintain the health of mice yet keep them thirsty enough to perform the task consistently. However, we noticed that in sessions where mice had access to free water the day before, the performance of mice was different, potentially because they were less thirsty and therefore less motivated (Fig. 6a). Among the dataset in these less motivated conditions, the parameter 2xFR12 happened to have enough sessions across mice and allowed us to compare the effect of two different levels of motivation statistically (e.g. high vs. low motivation).

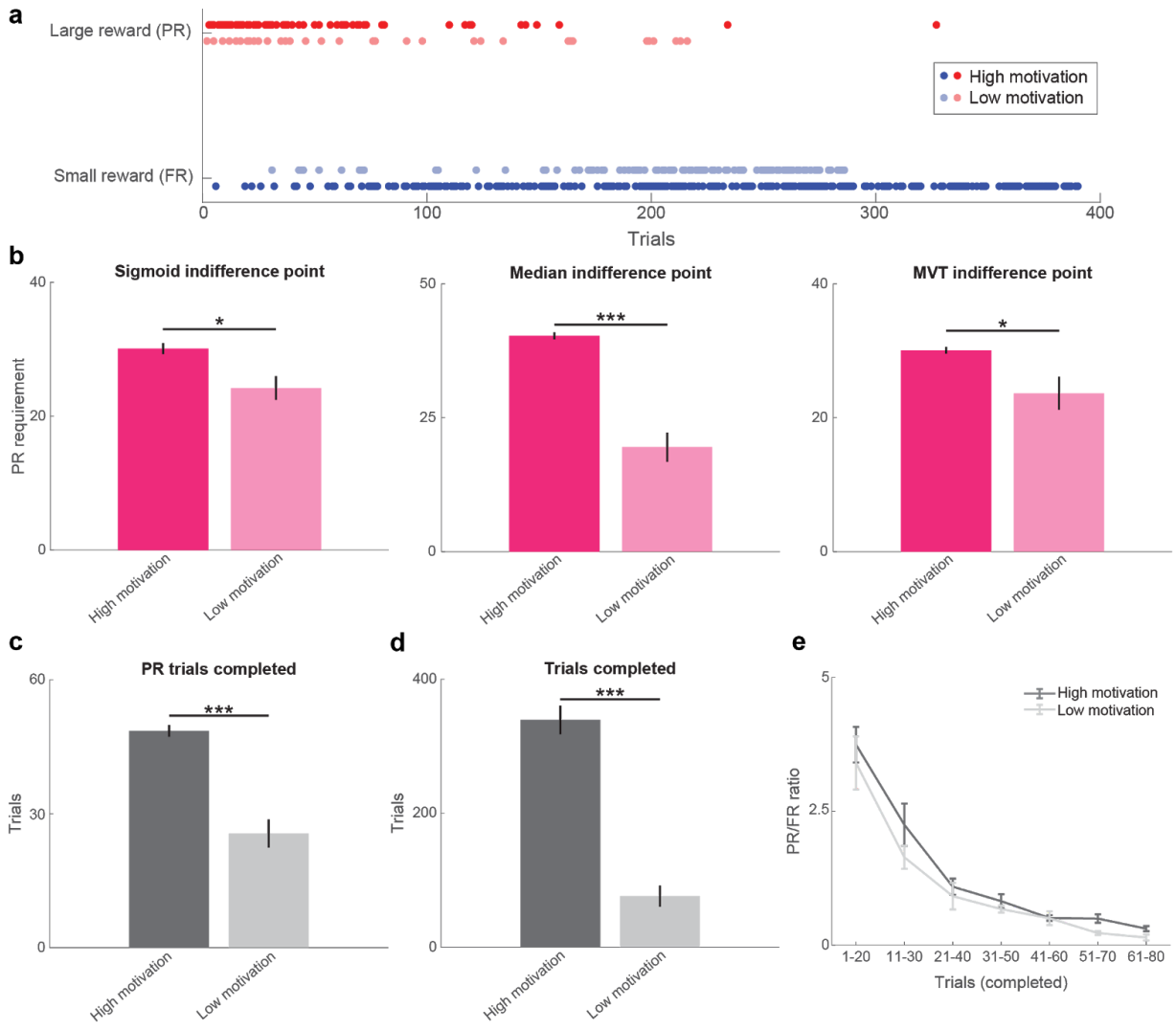

Figure 6. Motivation level effects choices in optimal switching task. a, Example sessions of a high and low motivational state from the same mouse. **b**, Mice switch from PR to FR earlier in a low motivational state (Wilcoxon Rank-sum test, *p < 0.05, ***p < 0.001, n = 7 mice). Error bars reflect the standard error of the mean. **c**, Mice complete more PR trials when in a high motivational state (Wilcoxon Rank-sum test, ***p < 0.001, n = 7 mice). Error bars reflect the standard error of the mean. **d**, Mice complete more trials when in a high motivational state (Wilcoxon Rank-sum test, ***p < 0.001, n = 7 mice). Error bars reflect the standard error of the mean. **e**, Comparison of high and low

motivation PR/FR ratios over time (n = 7 mice). Error bars reflect the standard error of the mean.

The differences were detected between low vs. high motivation conditions. The PR requirement at the indifference points, which represent willingness to work for the large reward, were significantly lower in sated mice, indicating the value of the PR was diminished in the low motivational condition (Fig. 6b; sigmoid fit: 30 ± 2 vs. 24 ± 5; median fit: 40 ± 2 vs. 20 ± 7; MVT: 30.1 ± 1.5 vs. 24 ± 7). Thirsty mice also tended to perform more trials from the PR with larger reward compared to sated mice (Fig. 6c, 49 ± 3 vs. 37 ± 8). Unfortunately, animals perform significantly fewer trials in a low motivated state (Fig. 6d, 340 ± 60 vs. 80 ± 40) resulting in a less pronounced switching behavior.

Valuation process is intact regardless of motivational states

Although different levels of motivation can shift mice's willingness to work for the large reward, they should not change the relative valuation of the PR compared to the FR. The relative value of the PR is still higher regardless of the motivational level. For instance, although a drop of water reward is not as valuable to sated mice as it is to thirsty mice, a large reward is still more valuable than a small reward regardless of thirst level. Indeed our behavioral results captured this conjecture (Fig. 6e). In the first 80 completed trials, where PR trials are most prevalent, we looked at the ratio of PR to FR trials. This PR to FR ratio changed consistently between the two groups as the session

progressed (Fig 6e). These results suggest that the relative values of two choices are independent of motivational levels. Taken together, our results indicate that while the level of motivation can influence the overall performance, the valuation process of the relative values of given options is independent of the motivation.

Conclusions

A major goal in this study is to develop a theory-centered behavioral paradigm for mice that quantitatively assesses the contribution of costs and benefits to the valuation process in economic choice behavior. Considering the arsenal of cutting edge tools available for mice in contemporary neuroscience (Lima et al., 2009; Katona et al., 2012; Chung et al., 2013; Roth, 2016; Deisseroth and Hegemann, 2017), the establishment of a quantitative behavior for mice is significant. These tools often expose an additional dimension in the information space that was inaccessible in the past. For instance, the neural correlates of a small population are often ignored. However, even a small population of neurons (e.g. 1~2%) can transfer significant information to the downstream neurons by divergently targeting them. Once the anatomical and physiological cell type is identified, the findings from these neurons can make a strong statement (Pi et al., 2013).

Our task design was inspired by several previous studies, of which the topics include foraging behavior (Charnov, 1976; Stephens and Krebs, 1985; Kolling et al., 2012), decision-making (Herrnstein, 1961; Hursh et al., 1988; Sugrue et al., 2004; Padoa-Schioppa and Assad, 2006; Walton et al., 2006; Kepecs et al., 2008; Platt and Huettel, 2008; Atalayer and Rowland, 2009; Berridge et al., 2009; Hosokawa et al., 2013), and motivation (Dayan and Balleine, 2002; Berridge et al., 2009; Bailey et al., 2015; Salamone et al., 2018). PR and indifference points have a long history that advanced our understanding of operant behavior (Pareto, 1919; Hodos, 1961; Salamone et al., 2018). A novel contribution of our study is not simply combining them,

but providing a quantitative and flexible paradigm that allows the manipulation of cost and benefit parameters independently and systematically. By systematically varying the amount of reward and cost in two-dimensional parameter space, we were able to generate an indifference plane that is potentially a more sensitive measure to distinguish subtle behavioral differences. Furthermore, the theoretical validation of our behavioral results suggests that mice indeed acted as a rational agent and that the switching decisions were optimal. Our paradigm can serve as a quantitative tool to investigate neural circuit mechanisms underlying cost using mice as a model system.

Across all mice tested, our results showed that the initial preference for the PR with large reward shifted to the FR with small reward as the PR got higher. The shift in preference was quantified by an indifference point where the values of the FR and PR became equivalent. The relative values of the PR and FR are different in different parameter pairs and indifference points should reflect these relative values. Surprisingly, our data seemed to show that mice placed more weight on the FR requirement during valuation. Changes in the PR large reward volume were not significant in our trial distribution analysis (both large reward and incomplete trials) as well as the median fit indifference point estimations. It is possible that different volumes of water are not as significant to mice as, for example, two different rewards of water and sucrose, where a sucrose reward would likely have a much higher reward value than plain water.

Indifference points were calculated with three different methods – 1) fitting data with a sigmoid function (logistic regression), 2) utilizing median values of choices, and

3) marginal value theorem, a theory for optimal foraging decisions. The trend of shift of the side preference was consistent in all three measures. Importantly the values calculated from MVT were very close to those from the sigmoid fit. This suggests that mice indeed utilize the optimization principle in their decision-making. It will be important to see how neuronal activity reflects this switching point and whether the neural activity supports MVT.

In our optimal switching behavior, the pressing requirement is the main cost that constitutes both effort and time cost. Effort cost is the requirement of physical pressing action that mice have to invest. The time cost is a temporal delay of reward due to the time delay proportional to the number of lever presses. This is better understood in the context of discount temporal delay that explains how the relative valuation of reward is placed differently at different time points (Roesch et al., 2006; Berns et al., 2007). For instance, most people would prefer to receive a hundred dollars now rather than a hundred dollars three months from now. Therefore, the time delay discounts the value of the reward. There are other forms of cost that are widely used in the laboratory setup. Probabilistic delivery of reward is one method that represents cost associated with risk. Because of the physically different nature of these different costs mentioned, they are thought to be processed in different circuits and networks in the brain (Rudebeck et al., 2006; Mar et al., 2011). Even though it is an important topic to investigate how different forms of cost are processed differently, and several studies have been devoted to distinguishing them, our current design did not aim to address this issue. Instead, we step back and see how cost, in general, is processed in the valuation process.

Considering the flexibility of our behavioral design, however, some of the issues can be addressed with a minor modification of the task.

The value placement is a subjective and relative process. A range of factors can influence the process and assign different values to the item (Sugrue et al., 2004; Amemori and Graybiel, 2012; Cai and Padoa-Schioppa, 2012; Lee, 2013; Yamada et al., 2013). For instance, the same amount of water reward can have a very different value to the same subject depending on its internal state. However, the large amount of water still has more value than the small amount. This aspect of the valuation process was captured in our data. Although different motivational levels impacted the overall task performance and other behavioral readouts, the relative value of PR was very similar regardless of the level of thirst. Because our dataset was limited to one parameter pair (2xFR12) for statistical evaluation, however, future research should address whether this finding holds in different parameter pairs. We validated our behavioral results in the context of the optimization framework which conferred additional confirmation of optimal decisions. However, our additional intention of introducing a theoretical framework is to guide studies into the neuronal circuits involved. Taking into account the stochastic nature of neuronal activity, it is difficult to expect how cost information is encoded; however, quantitative behavioral readouts and a theoretical framework may help us to identify the neuronal signature of the specific relevant information. In addition, theoretical frameworks can serve as a medium that connects inter-species studies (e.g. rodents, primates and humans). Therefore, it will

be interesting to see whether a theory-centered experimental design will pan out compared to the experiment-centered approaches.

Dysfunction in cost-benefit assessment is a key phenotype in diverse psychopathology including addiction, schizophrenia, depression, anxiety disorders, and severe impulsivity (Pasquier and Petit, 1997; Amemori and Graybiel, 2012; Hosokawa et al., 2013; Kolling et al., 2014; Amemori et al., 2015; Husain and Roiser, 2018). The findings of this study demonstrated that our behavioral paradigm can examine different aspects of the valuation process. For example, indifference points or an indifference plane can be used as a readout for the values proportional to the task parameters (e.g. 2xFR12), whereas the ratio between PR and FR conveys the relative values of the given choices (e.g. large vs. small). A single genetic mutation can cause several behavioral symptoms by dysregulating signaling cascades divergently in the downstream process. A recent study managed to link a specific circuit dysfunction that causes a specific behavioral phenotype in a genetic mouse model (Wells et al., 2016). It will be interesting to apply our behavioral readouts to mouse disease models in order to see whether different valuation processes dysfunction differently. Taken together, this study established a quantitative behavioral paradigm and theoretical framework that can be used to investigate the function and dysfunction of cost-benefit assessment at both behavioral and circuit levels.

References

Amemori K, Graybiel AM (2012) Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making. Nat Neurosci 15:776-785.

Amemori K, Amemori S, Graybiel AM (2015) Motivation and affective judgments differentially recruit neurons in the primate dorsolateral prefrontal and anterior cingulate cortex. J Neurosci 35:1939-1953.

Atalayer D, Rowland NE (2009) Meal patterns of mice under systematically varying approach and unit costs for food in a closed economy. Physiol Behav 98:85-93.

Azab H, Hayden BY (2018) Correlates of economic decisions in the dorsal and subgenual anterior cingulate cortices. Eur J Neurosci 47:979-993.

Bailey MR, Jensen G, Taylor K, Mezias C, Williamson C, Silver R, Simpson EH, Balsam PD (2015) A novel strategy for dissecting goal-directed action and arousal components of motivated behavior with a progressive hold-down task. Behav Neurosci 129:269-280.

Ballard K, Knutson B (2009) Dissociable neural representations of future reward magnitude and delay during temporal discounting. Neuroimage 45:143-150.

Berns GS, Laibson D, Loewenstein G (2007) Intertemporal choice--toward an integrative framework. Trends Cogn Sci 11:482-488.

Berridge KC, Robinson TE, Aldridge JW (2009) Dissecting components of reward: 'liking', 'wanting', and learning. Curr Opin Pharmacol 9:65-73.

Cai X, Padoa-Schioppa C (2012) Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. J Neurosci 32:3791-3808.

Charnov EL (1976) Optimal foraging, the marginal value theorem. Theor Popul Biol 9:129-136.

Chung K, Wallace J, Kim SY, Kalyanasundaram S, Andalman AS, Davidson TJ, Mirzabekov JJ, Zalocusky KA, Mattis J, Denisin AK, Pak S, Bernstein H, Ramakrishnan C, Grosenick L, Gradinaru V, Deisseroth K (2013) Structural and molecular interrogation of intact biological systems. Nature 497:332-337.

Dayan P, Balleine BW (2002) Reward, motivation, and reinforcement learning. Neuron 36:285-298.

Deisseroth K, Hegemann P (2017) The form and function of channelrhodopsin. Science 357.

Herrnstein RJ (1961) Relative and absolute strength of response as a function of frequency of reinforcement. J Exp Anal Behav 4:267-272.

Hodos W (1961) Progressive ratio as a measure of reward strength. Science 134:943-944.

Hosokawa T, Kennerley SW, Sloan J, Wallis JD (2013) Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex. J Neurosci 33:17385-17397.

Hursh SR, Raslear TG, Shurtleff D, Bauman R, Simmons L (1988) A cost-benefit analysis of demand for food. J Exp Anal Behav 50:419-440.

Husain M, Roiser JP (2018) Neuroscience of apathy and anhedonia: a transdiagnostic approach. Nat Rev Neurosci 19:470-484.

Katona G, Szalay G, Maak P, Kaszas A, Veress M, Hillier D, Chiovini B, Vizi ES, Roska B, Rozsa B (2012) Fast two-photon in vivo imaging with three-dimensional random-access scanning in large tissue volumes. Nat Methods 9:201-208.

Kepecs A, Uchida N, Zariwala HA, Mainen ZF (2008) Neural correlates, computation and behavioural impact of decision confidence. Nature 455:227-231.

Kim S, Hwang J, Lee D (2008) Prefrontal coding of temporally discounted values during intertemporal choice. Neuron 59:161-172.

Kolling N, Wittmann M, Rushworth MFS (2014) Multiple neural mechanisms of decision making and their competition under changing risk pressure. Neuron 81:1190-1202.
Kolling N, Behrens TE, Mars RB, Rushworth MF (2012) Neural mechanisms of foraging. Science 336:95-98.
Kolling N, Wittmann MK, Behrens TE, Boorman ED, Mars RB, Rushworth MF (2016) Value, search, persistence and model updating in anterior cingulate cortex. Nat Neurosci 19:1280-1285.
Leathers ML, Olson CR (2012) In monkeys making value-based decisions, LIP neurons encode cue salience and not action value. Science 338:132-135.
Lee D (2013) Decision making: from neuroscience to psychiatry. Neuron 78:233-248.
Levy DJ, Glimcher PW (2012) The root of all value: a neural common currency for choice. Curr Opin Neurobiol 22:1027-1038.
Lima SQ, Hromadka T, Znamenskiy P, Zador AM (2009) PINP: a new method of tagging neuronal populations for identification during in vivo electrophysiological recording. PLoS One 4:e6099.
Mar AC, Walker AL, Theobald DE, Eagle DM, Robbins TW (2011) Dissociable effects of lesions to orbitofrontal cortex subregions on impulsive choice in the rat. J Neurosci 31:6398-6404.
Montague PR, Berns GS (2002) Neural economics and the biological substrates of valuation. Neuron 36:265-284.
O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. Nat Neurosci 4:95-102.
Padoa-Schioppa C (2009) Range-adapting representation of economic value in the orbitofrontal cortex. J Neurosci 29:14004-14014.
Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. Nature 441:223-226.
Pareto V (1919) Manuale di Economia Politica — con una Introduzione alla Scienza Sociale. Piccola Biblioteca Scientifica 13.
Pasquier F, Petit H (1997) Frontotemporal dementia: its rediscovery. Eur Neurol 38:1-6.
Pi HJ, Hangya B, Kvitsiani D, Sanders JI, Huang ZJ, Kepecs A (2013) Cortical interneurons that specialize in disinhibitory control. Nature 503:521-524.
Platt ML, Huettel SA (2008) Risky business: the neuroeconomics of decision making under uncertainty. Nat Neurosci 11:398-403.
Roesch MR, Taylor AR, Schoenbaum G (2006) Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. Neuron 51:509-520.
Roth BL (2016) DREADDs for Neuroscientists. Neuron 89:683-694.
Rudebeck PH, Walton ME, Smyth AN, Bannerman DM, Rushworth MF (2006) Separate neural pathways process different decision costs. Nat Neurosci 9:1161-1168.
Salamone JD, Correa M, Yang JH, Rotolo R, Presby R (2018) Dopamine, Effort-Based Choice, and Behavioral Economics: Basic and Translational Research. Front Behav Neurosci 12:52.
Salamone JD, Steinpreis RE, McCullough LD, Smith P, Grebel D, Mahan K (1991) Haloperidol and nucleus accumbens dopamine depletion suppress lever pressing for food but increase free food consumption in a novel food choice procedure. Psychopharmacology (Berl) 104:515-521.
Stephens DW, Krebs JR (1985) Foraging theory. Monographs in behavior and ecology: Princeton Academic Press.

Sugrue LP, Corrado GS, Newsome WT (2004) Matching behavior and the representation of value in the parietal cortex. Science 304:1782-1787.

Sweis BM, Abram SV, Schmidt BJ, Seeland KD, MacDonald AW, 3rd, Thomas MJ, Redish AD (2018) Sensitivity to "sunk costs" in mice, rats, and humans. Science 361:178-181.

Walton ME, Kennerley SW, Bannerman DM, Phillips PE, Rushworth MF (2006) Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. Neural Netw 19:1302-1314.

Wells MF, Wimmer RD, Schmitt LI, Feng G, Halassa MM (2016) Thalamic reticular impairment underlies attention deficit in Ptchd1(Y/-) mice. Nature 532:58-63.

Yamada H, Tymula A, Louie K, Glimcher PW (2013) Thirst-dependent risk preferences in monkeys identify a primitive form of wealth. Proc Natl Acad Sci U S A 110:15788-15793.