Mice optimize switching decisions through cost-benefit evaluation

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, and Yongsoo Kim 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. To measure optimal cost-benefit switching decisions in mice, we developed a novel behavioral paradigm with two possibilities of obtaining reward. Mice initially prefer the easier method, but as its effort increases, their preference switches. We quantified these switching decisions using an indifference point where the values of the choices became equivalent. Manipulation of required effort and reward volume in a two-dimensional parameter space successfully shifted 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 parametric shift of the switching decisions. Taken together, our behavioral paradigm provides a quantitative platform to investigate neural circuits underlying cost-benefit assessment.

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 & Krebs, 1955). 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 (Levy & Glimcher, 2012; Montague & Berns, 2002). 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 (Kolling et al., 2012; Sugrue et al., 2004; Walton et al., 2006), 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. In addition, while reward processing in the brain has been a major focus in neuroscience research (Berridge et al., 2009; Schultz et al., 1997; Sutton, 1998), little is known about the neural circuit mechanisms underlying cost processing aside from identification of a few brain regions involved (Kolling et al., 2012; Mar et al., 2011; Rudebeck et al., 2006; Salamone et al., 1991)

Our long-term goal is to tackle this challenge by dissecting specific neural circuit mechanisms involved in neuroeconomic decision making. As a first step, here we established a quantitative behavioral paradigm that allows us to identify optimal cost-benefit switching behavior for mice in order to systematically and quantitatively investigate the cost-benefit interaction on the valuation process. Inspired by several past experiments on decision-making, effort, and motivated behavior (Atalayer & Rowland, 2009; Bailey et al., 2015; Hursh et al., 1988; Kepecs et al., 2008; Kolling et al., 2012; Padoa-Schioppa, 2009; Padoa-Schioppa & Assad, 2006; Rudebeck et al., 2006; Salamone et al., 2018; Salamone et al., 1991; 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). Using this behavioral paradigm, we tested the hypothesis that mice optimize switching decisions based on cost-benefit assessment. By independently modulating behavioral parameters associated with benefits and costs, we were able to quantify the moment of preference shift in mice. This finding is further validated with a theoretical framework of optimization called the marginal value theorem

(MVT; Charnov, 1976). Finally, this quantification of the preference shift is influenced by different internal states of motivation.

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.

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

105 between conditions. For instance, by fixing one parameter (e.g., reward amount) and
106 changing the other (e.g., lever press requirement), the contribution of cost and benefit to
107 decision-making can be evaluated semi-independently within the same task, a feat which
108 has required two different behavioral paradigms in past experiments. For our task, four
109 combinations of parameters were chosen. The ratio of large reward to small reward and
110 the fixed number of presses are used to denote the combination of parameters. For
111 instance, 2xFR12 means the volume of large reward, 6 μL, is twice as much as that of
112 the small reward, 3 μL, and the FR requirement is 12 presses. The four combinations of
113 parameters can be denoted as 2xFR6, 2xFR12, 5xFR6, and 5xFR12 (Fig. 1c).

114

115 **Mice adjust switching decisions proportionate to the values of session parameters**

116 We hypothesized that changes in the relative values of given choices would yield
117 quantifiable differences in the switching decisions of mice. To test our hypothesis, we ran
118 10 mice on our optimal switching task with the parameters 2xFR6, 2xFR12, 5xFR6, and
119 5xFR12. We evaluated an equal number of male and female mice and did not see
120 significant differences (Supplementary Fig. 1), therefore we combined the data for further
121 analysis. Four sessions with each parameter setting were collected (160 sessions total:
122 4 sessions x 4 parameter pairs x 10 mice). The PR and FR sides and the parameter pairs
123 were selected pseudo-randomly at the start of each session. Consistent with our initial
124 observation, in all parameter conditions mice initially preferred the PR with large reward
125 and switched their preference to the FR with small reward as the PR increased (Fig. 1b).
126 Taken together, we established a quantitative behavioral paradigm and mice were able

127 to make switching decisions by accurately evaluating the cost-benefit relationship of the
128 given choices.

129 The adjustment of switching decisions can be visible even from the plot of
130 decisions over time (Fig. 1b). The relative value of the PR with large reward is lowest at
131 2xFR6 because the reward is low and effort cost for the FR is also low. Alternatively, the
132 PR value is highest at 5xFR12 because the reward is high and the alternative choice
133 requires more effort. 2xFR12 and 5xFR6 are somewhere in between the two. Reflecting
134 these relative values, the number of PR choices increased and spread more as the value
135 increased. These measures were quantified in percentage and deviation of the PR
136 choices (Fig. 2a-b, Supplementary Fig. 2). The percentages of choosing the PR with large
137 reward were 6.5 ± 1.4%, 14 ± 3%, 7.8 ± 1.5%, and 19 ± 4% at 2xFR6, 2xFR12, 5xFR6,
138 and 5xFR12, respectively. The percentage of the FR choices were 86%, 69%, 84%, and
139 63%, respectively (Fig. 2a).

140 The percentage of incomplete trials showed a similar trend as the percentage of
141 large reward choices. A trial is considered incomplete if a mouse stopped pressing the
142 lever in the middle of a trial for more than ten seconds or if the mouse pressed the
143 opposite lever after initiating a trial. If the relative values of session parameters were
144 higher, mice tried and failed more (Fig. 2c) and the failure was more frequent on the PR
145 side (Supplementary Fig. 3-4). The incomplete trial percentages were 8 ± 3%, 17 ± 7%,
146 8 ± 4%, and 18 ± 8% for 2xFR6, 2xFR12, 5xFR6, and 5xFR12 respectively. Changing the
147 cost parameter (e.g. number of FR presses) contributed more to the incomplete trials.
148 The parameters 2xFR6 and 5xFR6 have a difference of only 0.7%, yet between 2xFR6
149 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

171 PR and FR trials; and 3) developing a measure based on an optimization theory called
172 the marginal value theorem (Charnov, 1976; Stephens & Krebs, 1955).

173 The first approach for the estimation of the PR requirement at indifference points
174 was fitting the data with a function that represented the distribution of the choices.
175 Because our data was binary (e.g. two choices), session data were fit with a sigmoid
176 function (Boltzmann function). Sigmoid fitting curves captured the profiles of mouse
177 decisions, showing the transition from the PR to the FR (Fig. 3a-b). The indifference trial
178 number was first estimated where the sigmoid curve crossed the midline. Then the
179 number of lever presses required for the PR at that trial was extracted from the data,
180 which provided the PR requirement at the indifference point of the session data. Figure
181 3a also shows where the PR requirement lies on the cumulative distribution of the PR.
182 The results showed that both at a single mouse level and at the animal average, the
183 estimated PR requirement at indifference points was lowest at 2xFR6, highest at 5xFR12
184 and those of 2xFR12 and 5xFR6 lied in between (20 ± 4 for 2xFR6, 28 ± 4 for 2xFR12,
185 29 ± 4 for 5xFR6, and 45 ± 7 for 5xFR12; Fig. 3c). This shifting pattern of the PR
186 requirement proportional to the relative values of session parameters suggests that mice
187 perform the task as a rational agent, which is a basic premise of economic decision-
188 making theory (Levy & Glimcher, 2012).

189 Although binary data is conventionally fit with a sigmoid function, this approach did
190 not capture the spread of the PR choices. Even after passing an indifference point, mice
191 occasionally pursued the PR side with large reward despite the high number of required
192 lever presses (Supplementary Fig. 3). When the relative value of the PR was higher, mice
193 attempted to collect large reward more often (Supplementary Fig. 4). In order to capture

194 the spread in estimation of indifference points, we utilized median values of the choices
195 (Fig. 4a). In this approach, the median trial numbers of the PR and FR were first calculated
196 and these two medians were connected with a line. An indifference trial number was
197 defined at the intersections where this line connecting medians crosses the midline. The
198 PR requirement at the indifference point was then calculated by identifying the number of
199 presses required on the PR side at this intersection.

200 The PR requirement estimated using median values showed the same trend as
201 those estimated by the sigmoid function. 2xFR6 had the lowest value of the PR
202 requirement, 5xFR12 had the highest, and 2xFR12 and 5xFR6 had values in between
203 (30 ± 3 for 2xFR6, 38 ± 4 for 2xFR12, 33 ± 4 for 5xFR6, and 44 ± 5 for 5xFR12). Although
204 the trend was similar, the estimated PR requirements from two different approaches
205 conferred slightly different values. The standard deviations of PR requirements estimated
206 by medians were smaller, but the relationship between the PR water volumes was not
207 significant. In summary, the two different estimations of the PR requirement at
208 indifference points provide a similar trend in the relative values of session parameters:
209 the higher the value of the PR, the greater the PR requirement at the indifference point.

210

211 **Estimation of indifference based on a normative model**

212 Switching decisions in our behavioral task can be viewed as an optimization
213 problem. Our question can be rephrased as such: how do mice optimize their switching
214 decisions to maximize the reward gained and minimize the effort? We addressed this
215 question with the marginal value theorem (MVT), a normative model developed to explain

optimal foraging behavior (Charnov, 1976; Stephens & Krebs, 1955). 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 & Krebs, 1955). 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).

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

239 difference between MVT and the other approaches was that MVT generated the smallest
240 standard deviations for all four parameters (Table 1).

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

253

254 **Different motivational modulation of indifference points**

255 The valuation process is subjective and context-dependent. Motivation is one of
256 the internal states that can strongly influence the evaluation process. The same volume
257 of water has significantly different values to thirsty vs. sated mice. Assuming the level of
258 thirst as a proxy of motivational level, we looked at how different levels of motivation
259 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).

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.

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.

We checked whether this conjecture could be confirmed in our dataset and indeed it was (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 progresses (Fig 6e). These results suggest that the relative values of two choices are independent of the motivational levels. Taken together, our results indicate that while the level of motivation can influence the overall performance, the evaluation process of the relative values of given options is independent of the motivation.

Discussion

A major goal in this study is to develop a behavioral paradigm that quantitatively assesses the contribution of costs and benefits to the valuation process in economic choice behavior. Our task design was inspired by several previous studies, of which the topics include foraging behavior (Charnov, 1976; Kolling et al., 2012; Stephens & Krebs, 1955), decision-making (Atalayer & Rowland, 2009; Berridge et al., 2009; Herrnstein, 1961; Hursh et al., 1988; Kepecs et al., 2008; Padoa-Schioppa & Assad, 2006; Platt & Huettel, 2008; Sugrue et al., 2004; Walton et al., 2006), and motivation (Bailey et al., 2015; Berridge et al., 2009; Dayan & Balleine, 2002; Salamone et al., 2018). However, a novel contribution of our study is to provide a quantitative behavior focusing on effort cost and a theoretical framework based on a normative model. By systematically varying the amount of reward and effort cost in the two-dimensional parameter space, we were able to generate an indifference plane that is potentially a more sensitive measure to distinguish subtle behavioral differences. To the best of our knowledge, this is the first systematic investigation of the effect of different amounts of effort cost and reward benefit in mice.

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 (e.g. water reward and effort cost) 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

313 were not significant in our trial distribution analysis (both large reward and incomplete
314 trials) as well as the median fit indifference point estimations. It is possible that different
315 volumes of water are not as significant to mice as, for example, two different rewards of
316 water and sucrose, where a sucrose reward would likely have a much higher reward value
317 than plain water.

318 Indifference points were calculated with three different methods – 1) fitting data
319 with a sigmoid function (logistic regression), 2) utilizing median values of choices, and 3)
320 marginal value theorem, a theory for optimal foraging decisions. The trend of shift of the
321 side preference was consistent in all three measures. Importantly the values calculated
322 from MVT were very close to those from the sigmoid fit. This suggests that mice indeed
323 utilize the optimization principle in their decision-making. It will be important to see how
324 neuronal activity reflects this switching point and whether the neural activity supports
325 MVT.

326 In our optimal switching behavior, the pressing requirement is the main cost that
327 we wanted to implement. However, there are other forms of cost that are widely used in
328 in the laboratory setup. Probabilistic delivery of reward is one method that represents cost
329 associated with risk. Discount temporal delay is another one that accounts a temporal
330 aspect of cost and explains how the relative valuation of reward is placed differently at
331 different time points. For instance, we tend to assign more value on an immediate reward
332 than the same amount that can be received in the distant future. Therefore, the time delay
333 discounts the value of the reward. Although our behavioral task was designed not to
334 include risk and temporal delay but rather to focus on effort cost, we cannot exclude the
335 possibility that temporal discount may play a role in the valuation process. However, we

336 reason that the effect size of temporal discount is relatively small because the time scale
337 for temporal discount effect tends to be much larger than that of one trial used in our
338 experiments (e.g. hours, days, and months vs. seconds). Considering the time scale of
339 seconds can be consumed by grooming and exploration, the perception of a few seconds
340 delay to mice and its discounting factor is unlikely to play a significant role in our task.

341 The value placement is a subjective and relative process. A range of factors can
342 influence the process and assign different values to the item (Sugrue et al., 2004). For
343 instance, the same amount of water reward can have a very different value to the same
344 subject depending on its internal state. However, the large amount of water still has more
345 value than the small amount. Our data captured this aspect of valuation process. Although
346 different motivational levels impacted the overall task performance and other behavioral
347 readouts, the relative value of PR was very similar regardless of the level of thirst.
348 Although our dataset was limited to one parameter pair (2xFR12) for statistical evaluation,
349 future research should address whether this finding holds in different parameter pairs.

350 Cost-benefit analysis is a critical component in neuroeconomic decision-making,
351 yet little is known about neural circuit mechanisms underlying it. Due to the abstract nature
352 of cost and benefit, and lack of behavioral readouts, the mechanistic investigation at the
353 neural circuit level has been challenging, especially using an animal model system. Our
354 contribution is to provide a novel behavioral paradigm with quantitative readouts that are
355 modulated parametrically. The interpretation of the behavioral results in the context of the
356 optimization framework (e.g. MVT) conferred additional confirmation of our findings. We
357 believe that our approach with a quantitative behavior and theoretical framework will
358 guide the direction of future studies aimed at identifying underlying neuronal circuits.

359 Taking into account the stochastic nature of neuronal activity, it is difficult to expect how
360 cost information is encoded; however, quantitative behavioral readouts and a theoretical
361 framework may help us to identify the neuronal signature of the specific information. Since
362 dysfunction in cost-benefit assessment is a key phenotype in diverse psychopathology
363 such as addiction, schizophrenia and severe impulsivity, the application of our
364 quantitative behavior to disease models will be a useful tool to identify behavioral
365 differences and circuit dysfunction.

366 **Methods**

367 *Animals*

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

373

374 *Behavioral Setup*

375 Experiments were conducted in a rectangular acrylic testing chamber (length, 19.3 cm;
376 width, 14.0 cm; height, 13.9 cm, SanWorks, Stony Brook, NY) with grated flooring. A
377 metal tray was placed underneath the chamber to collect waste. The chamber contained
378 three nose-pokes with infrared LED/infrared phototransistor pair (Digikey, Thief River
379 Falls, MN) to detect responses. A white LED (Digikey) inside the nose-pokes was used
380 to cue trial availability, lever pressing progress, and reward availability. Only the center
381 port was used for reward delivery. The two end ports were used only as lights and were
382 covered with a snuggly-fit clear plastic cap. Plastic levers were custom designed for us
383 by SanWorks for either side of the nose-pokes. The levers were also equipped with an
384 infrared sensor to capture lever presses. A food pellet was placed inside the testing
385 chamber at the start of each session to allow mice to eat in between trials. The testing
386 chamber was situated inside of a custom-built noise-reducing box (length, 42 cm; width,
387 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) 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

411 obtain reward. The trials were pseudo-randomized and the correct side to obtain reward
412 was indicated by the corresponding nose-poke being lit. Finally, mice repeated the fourth
413 phase, but with an increasing number of presses required each day, from 2 presses all
414 the way up to 10 presses. Once mice completed these stages of training, they were
415 subjected to the optimal switching task and were allowed to familiarize themselves with
416 the task for 3-5 sessions before data collection began.

417

418 *Optimal switching task*

419 Our behavioral task involved combining a progressive ratio (PR) and fixed ratio (FR)
420 schedule of lever pressing. The PR was associated with a large volume of water (either
421 6μL or 15μL) and the FR was associated with a small volume of water (3μL). In addition,
422 the FR could either be 6 presses or 12 presses. Similar to the training phases, levers had
423 to be depressed for 100 ms to count as a press and there was a one second delay
424 between trials. The PR and FR sides as well as the parameter pairs were pseudo-
425 randomly chosen at the start of each session. At the start of a trial, both the left and right
426 nose-pokes were dimly lit, indicating that the mouse could choose either the left or right
427 side. Once a mouse chose a side, the corresponding nose-poke would get increasingly
428 brighter with each press until the required number of presses was met and the center port
429 lit up to indicate reward availability. Mice were able to freely choose either side; however,
430 if they decided to switch sides in the middle of a trial before completing the number of
431 presses on the initially chosen side, a punishment sound of white noise would play and
432 the trial would end. That trial was then classified as an incomplete trial. Furthermore, if a
433 mouse began pressing a lever but then stopped for more than 10 seconds to groom itself,

434 eat, etc., the trial would end and that trial would be considered incomplete. A session
435 could be anywhere from 1 to 3 hours long. A session ended if a mouse did not press
436 either lever for a period of time longer than 5 minutes. If this did not occur within 3 hours,
437 the session was ended by the operator. In this way, we can ensure that the mouse is well
438 sated by the end of the session and that we capture the highest number of trials the
439 mouse is willing to perform without keeping the animal in the chamber too long. At the
440 end of each session, the mouse was weighed and additional water was given at the end
441 of each session if necessary to maintain the animal's weight at 85-90% of their free-
442 drinking body weight.

443

444 *Biological vs. technical replication*

445 We evaluated an equal number of 5 male and 5 female mice and did not see significant
446 differences (Supplementary Fig. 1), therefore we combined the data for further analysis.
447 Four sessions with each parameter setting were collected (160 sessions total: 4 sessions
448 x 4 parameter pairs x 10 mice). Therefore, our biological replicates are 10 mice and
449 technical replicates are 4 sessions for each parameter settings.

450

451 *Data analysis*

452 All data analysis was carried out using built-in and custom-built software in Matlab
453 (Mathworks). The first approach of the identification of indifference points was done with
454 our binary data with a sigmoid function, the Boltzmann function.

455 $$f(x) = \frac{1}{1 + e^{(x - x_0)/\tau}}$$

456 where x_0 and τ are a 50% threshold and a slope. By assigning the value of the PR choice
457 = 1 and that of the FR = 0 and assuming that the curve started from the PR and ended to
458 the FR, the fitting curve was generated. An indifference trial number was estimated where
459 the sigmoid curve crossed the midline and the number of lever presses required at the
460 trial on the PR side was extracted from the data. This required number of lever presses
461 provided the PR requirement at the indifference point of the session data. In the second
462 approach, the median values of the choices were utilized. The median trial numbers of
463 the PR and FR were first calculated and these two medians were connected with a line.
464 An indifference trial number was defined at the intersections where this line connecting
465 medians crosses the midline. An indifference point was then calculated by identifying the
466 number of presses required on the PR side at this intersection. The third approach was
467 done with the marginal value theorem. First, we defined the net gain as the ratio between
468 reward obtained and the number of presses (e.g. reward/effort) in a trial. The cumulative
469 distribution of the total net gains of the PR was generated and fitted with a double-
470 exponential function because it captured the sudden change of slope in the PR data. The
471 cumulative distribution of the FR was fitted linearly, which can be interpreted as the long-
472 term rate of gain. Then, the tangent point between two curves was identified and the
473 indifference point was calculated by identifying the PR requirement at this tangent trial
474 number. Statistical significance among four parameter pairs was first tested with the two-
475 factor nonparametric Scheirer-Ray-Hare test and the post-hoc pairwise comparison was
476 done with Wilcoxon Rank-sum test with Bonferroni correction for 4 comparison pairs. For

477 single pair comparisons, a nonparametric Wilcoxon Rank-sum test was used. A *P value*

478 cutoff of 0.05 was used for significance testing.

References

Atalayer, D., & Rowland, N. E. (2009). Meal patterns of mice under systematically varying approach and unit costs for food in a closed economy. *Physiol Behav, 98*(1-2), 85-93. doi:10.1016/j.physbeh.2009.04.016

Bailey, M. R., Jensen, G., Taylor, K., Mezias, C., Williamson, C., Silver, R., . . . Balsam, P. D. (2015). A novel strategy for dissecting goal-directed action and arousal components of motivated behavior with a progressive hold-down task. *Behav Neurosci, 129*(3), 269-280. doi:10.1037/bne0000060

Bailey, M. R., Simpson, E. H., & Balsam, P. D. (2016). Neural substrates underlying effort, time, and risk-based decision making in motivated behavior. *Neurobiol Learn Mem, 133*, 233-256. doi:10.1016/j.nlm.2016.07.015

Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: 'liking', 'wanting', and learning. *Curr Opin Pharmacol, 9*(1), 65-73. doi:10.1016/j.coph.2008.12.014

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theor Popul Biol, 9*(2), 129-136. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1273796

Dayan, P., & Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron, 36*(2), 285-298. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12383782

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *J Exp Anal Behav, 4*, 267-272. doi:10.1901/jeab.1961.4-267

Hodos, W. (1961). Progressive ratio as a measure of reward strength. *Science, 134*(3483), 943-944. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/13714876

Hursh, S. R., Raslear, T. G., Shurtleff, D., Bauman, R., & Simmons, L. (1988). A cost-benefit analysis of demand for food. *J Exp Anal Behav, 50*(3), 419-440. doi:10.1901/jeab.1988.50-419

Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature, 455*(7210), 227-231. doi:10.1038/nature07200

Kolling, N., Behrens, T. E., Mars, R. B., & Rushworth, M. F. (2012). Neural mechanisms of foraging. *Science, 336*(6077), 95-98. doi:10.1126/science.1216930

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Curr Opin Neurobiol, 22*(6), 1027-1038. doi:10.1016/j.conb.2012.06.001

Mar, A. C., Walker, A. L., Theobald, D. E., Eagle, D. M., & Robbins, T. W. (2011). Dissociable effects of lesions to orbitofrontal cortex subregions on impulsive choice in the rat. *J Neurosci, 31*(17), 6398-6404. doi:10.1523/JNEUROSCI.6620-10.2011

Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron, 36*(2), 265-284. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12383781

Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. *J Neurosci, 29*(44), 14004-14014. doi:10.1523/JNEUROSCI.3751-09.2009

524 Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode
525 economic value. *Nature, 441*(7090), 223-226. doi:10.1038/nature04676
526 Platt, M. L., & Huettel, S. A. (2008). Risky business: the neuroeconomics of decision
527 making under uncertainty. *Nat Neurosci, 11*(4), 398-403. doi:10.1038/nn2062
528 Rudebeck, P. H., Walton, M. E., Smyth, A. N., Bannerman, D. M., & Rushworth, M. F.
529 (2006). Separate neural pathways process different decision costs. *Nat Neurosci,*
530 *9*(9), 1161-1168. doi:10.1038/nn1756
531 Salamone, J. D., Correa, M., Yang, J. H., Rotolo, R., & Presby, R. (2018). Dopamine,
532 Effort-Based Choice, and Behavioral Economics: Basic and Translational
533 Research. *Front Behav Neurosci, 12*, 52. doi:10.3389/fnbeh.2018.00052
534 Salamone, J. D., Steinpreis, R. E., McCullough, L. D., Smith, P., Grebel, D., & Mahan, K.
535 (1991). Haloperidol and nucleus accumbens dopamine depletion suppress lever
536 pressing for food but increase free food consumption in a novel food choice
537 procedure. *Psychopharmacology (Berl), 104*(4), 515-521. Retrieved from
538 https://www.ncbi.nlm.nih.gov/pubmed/1780422
539 Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and
540 reward. *Science, 275*(5306), 1593-1599. Retrieved from
541 https://www.ncbi.nlm.nih.gov/pubmed/9054347
542 Stephens, D. W., & Krebs, J. R. (1955). *Foraging theory. Monographs in behavior and*
543 *ecology*: Princeton Academic Press.
544 Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004). Matching behavior and the
545 representation of value in the parietal cortex. *Science, 304*(5678), 1782-1787.
546 doi:10.1126/science.1094765

Sutton, R. S. B., A. G. (1998). *Reinforcement lerning: an introduction*: A Bradford Book, The MIT Press.

Sweis, B. M., Abram, S. V., Schmidt, B. J., Seeland, K. D., MacDonald, A. W., 3rd, Thomas, M. J., & Redish, A. D. (2018). Sensitivity to "sunk costs" in mice, rats, and humans. *Science, 361*(6398), 178-181. doi:10.1126/science.aar8644

Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E., & Rushworth, M. F. (2006). Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. *Neural Netw, 19*(8), 1302-1314. doi:10.1016/j.neunet.2006.03.005

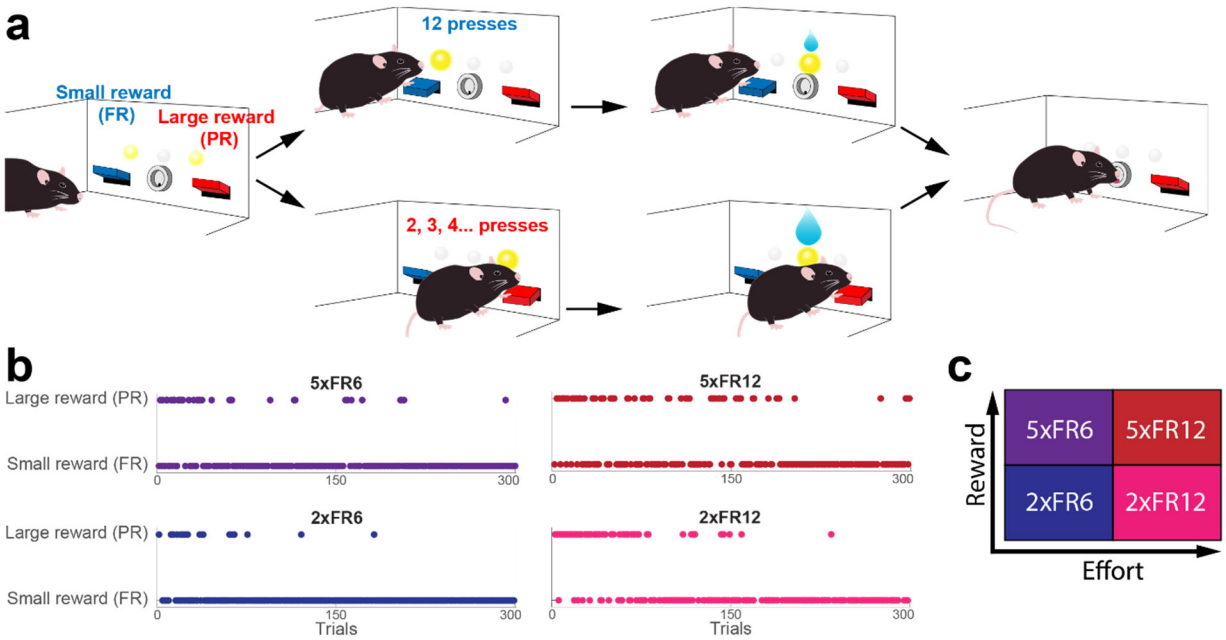

Figure 1. Optimal Switching Behavior Task. a, Schematic of task design. **b**, Example sessions for each parameter pair. Each point is a single trial within the session. Only the first 300 trials of each session are shown. All sessions are from the same animal. **c**, Grid showing increasing reward and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

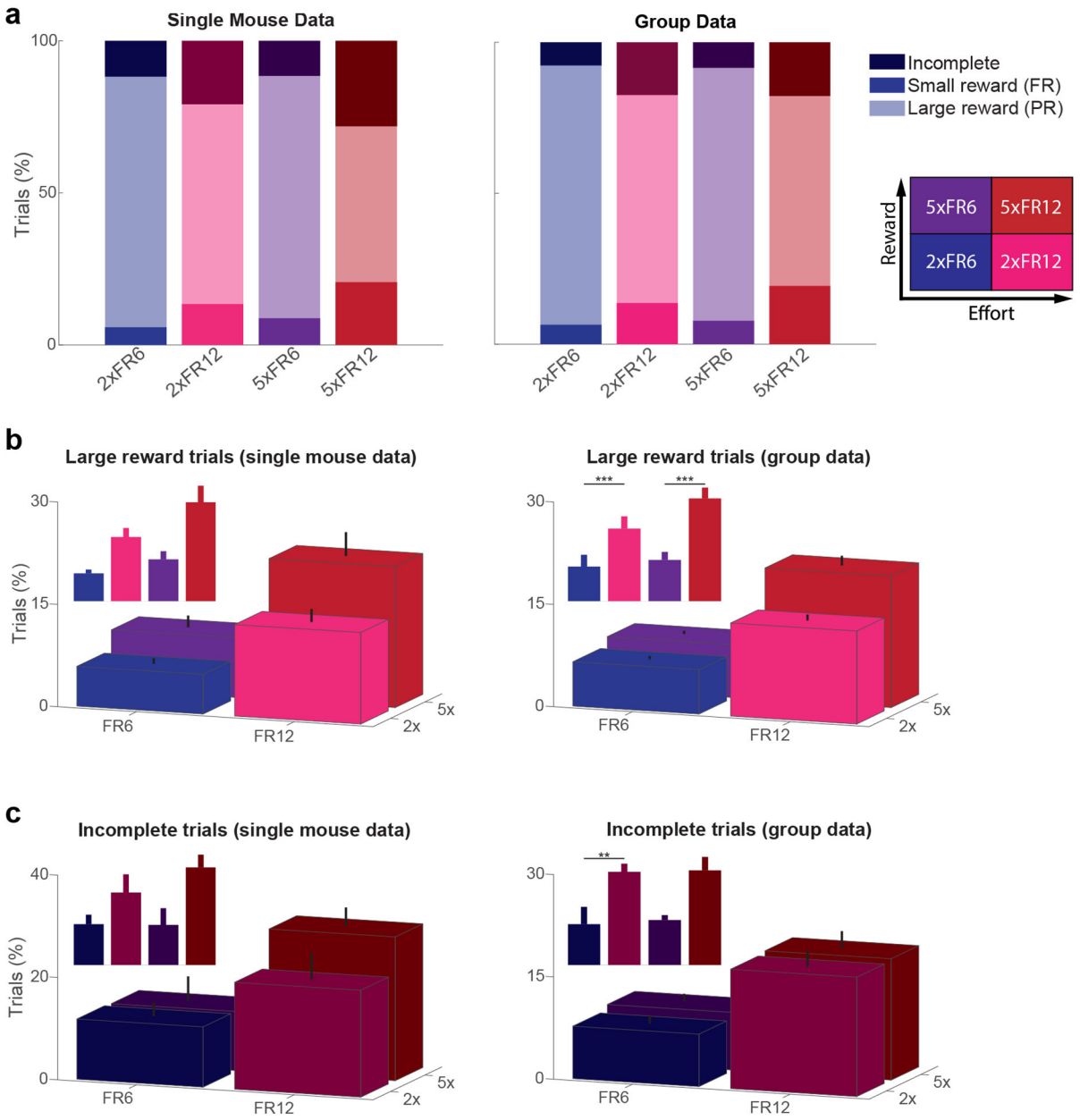
a
Single Mouse Data
Group Data
Trials (%)
100
50
0
2xFR6
2xFR12
5xFR6
5xFR12
Incomplete
Small reward (FR)
Large reward (PR)
Reward
Effort
5xFR6
5xFR12
2xFR6
2xFR12
b
Large reward trials (single mouse data)
Large reward trials (group data)
30
15
0
FR6
FR12
2x
5x
***
c
Incomplete trials (single mouse data)
Incomplete trials (group data)
40
20
**

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 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 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).

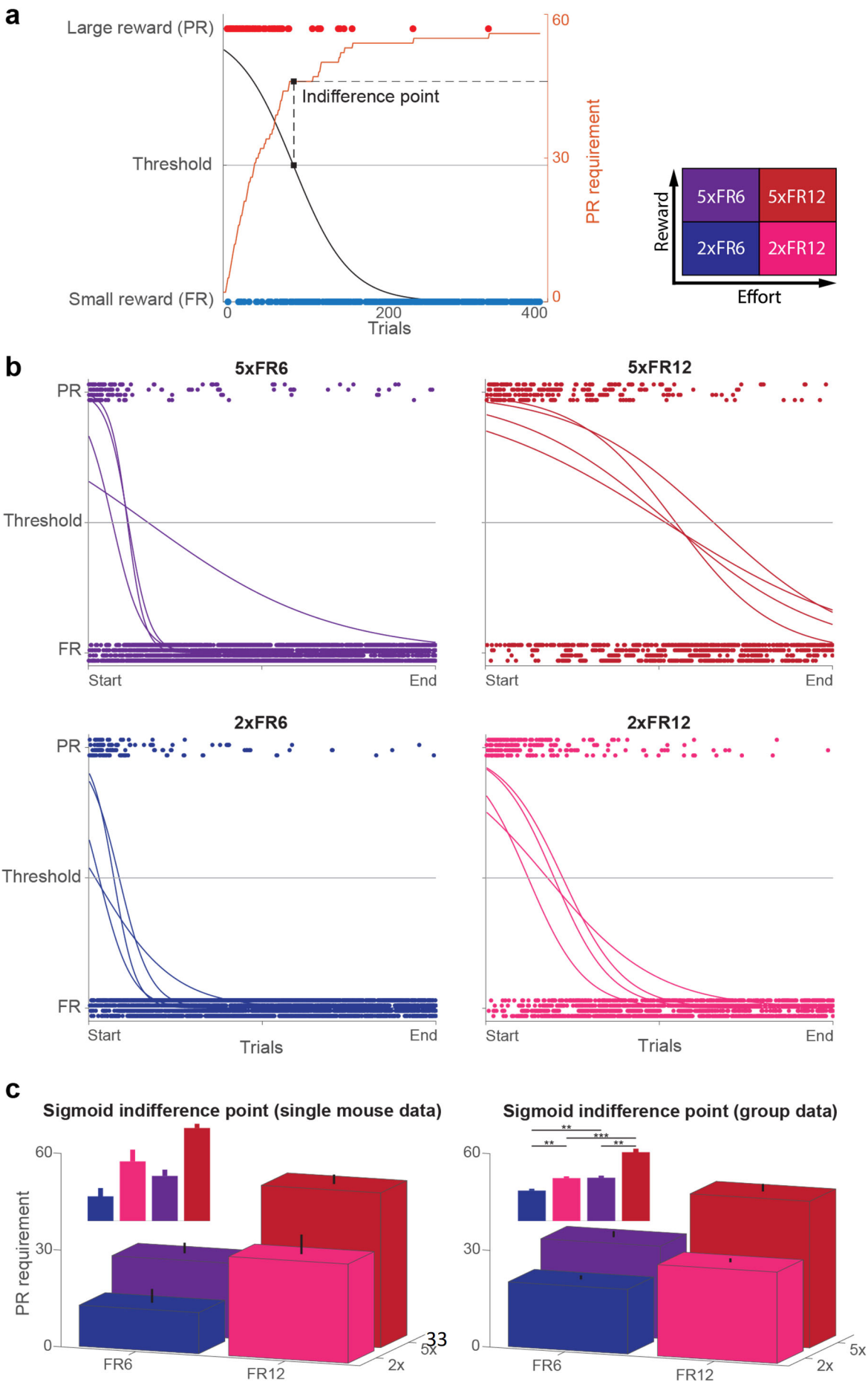
a
Large reward (PR)
Small reward (FR)
Threshold
Indifference point
PR requirement
60
30
0
0
200
400
Trials
5xFR6
5xFR12
2xFR6
2xFR12
Reward
Effort
b
5xFR6
5xFR12
2xFR6
2xFR12
PR
FR
Threshold
Start
End
Trials
c
Sigmoid indifference point (single mouse data)
Sigmoid indifference point (group data)
PR requirement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

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 x 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).

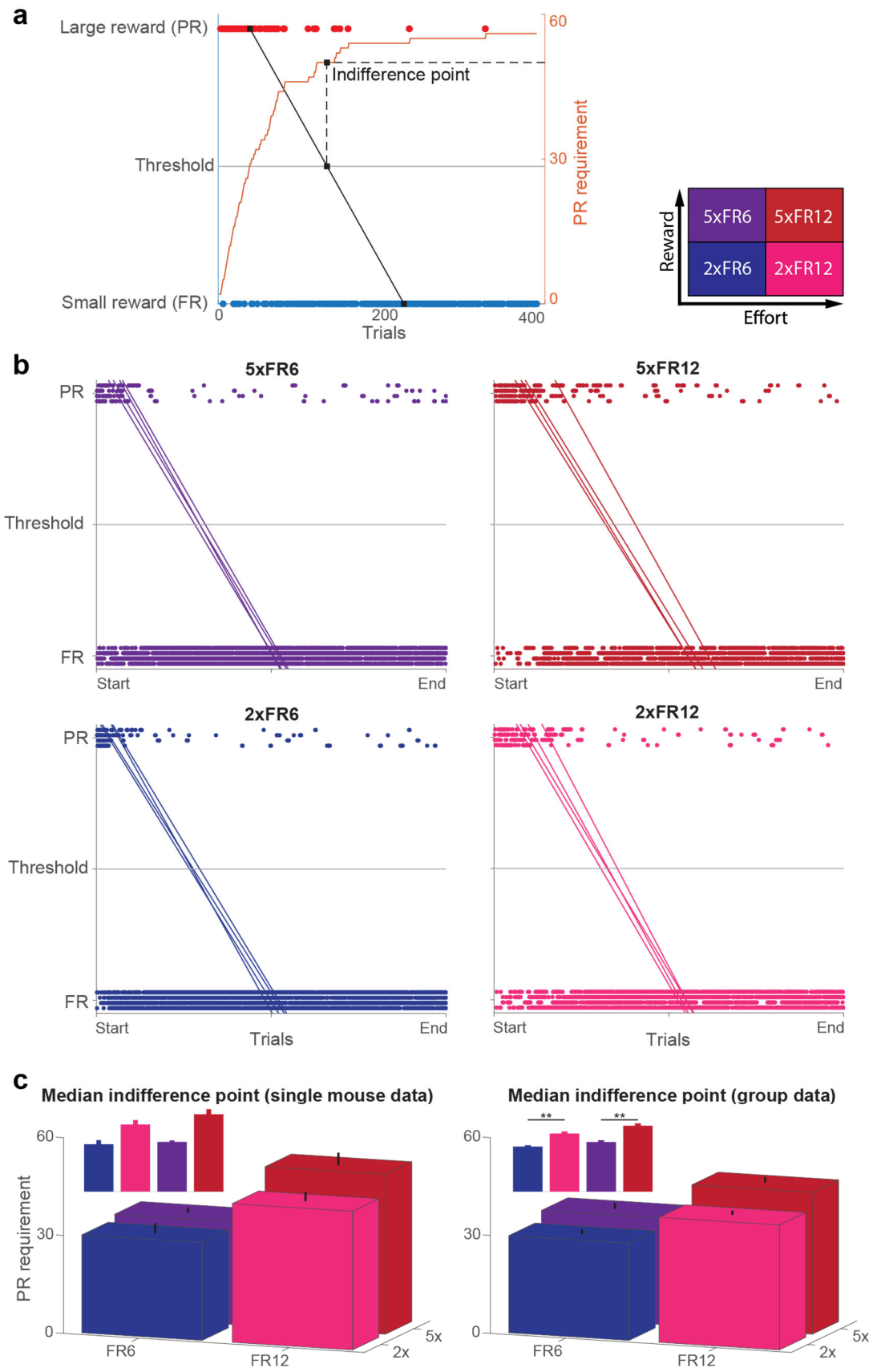
a
Large reward (PR)
Indifference point
Threshold
Small reward (FR)
60
30
0
PR requirement
0
200
400
Trials
Reward
Effort
5xFR6
5xFR12
2xFR6
2xFR12
b
5xFR6
5xFR12
2xFR6
2xFR12
PR
Threshold
FR
Start
End
Trials
c
Median indifference point (single mouse data)
Median indifference point (group data)
PR requirement
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 ($H_1 = 22.70$, $p = 2 \times 10^{-6}$). Insert shows the same data with significant pairs notated (post hoc Wilcoxon Rank-sum test with Bonferroni correction, $^{**}p < 0.01$).

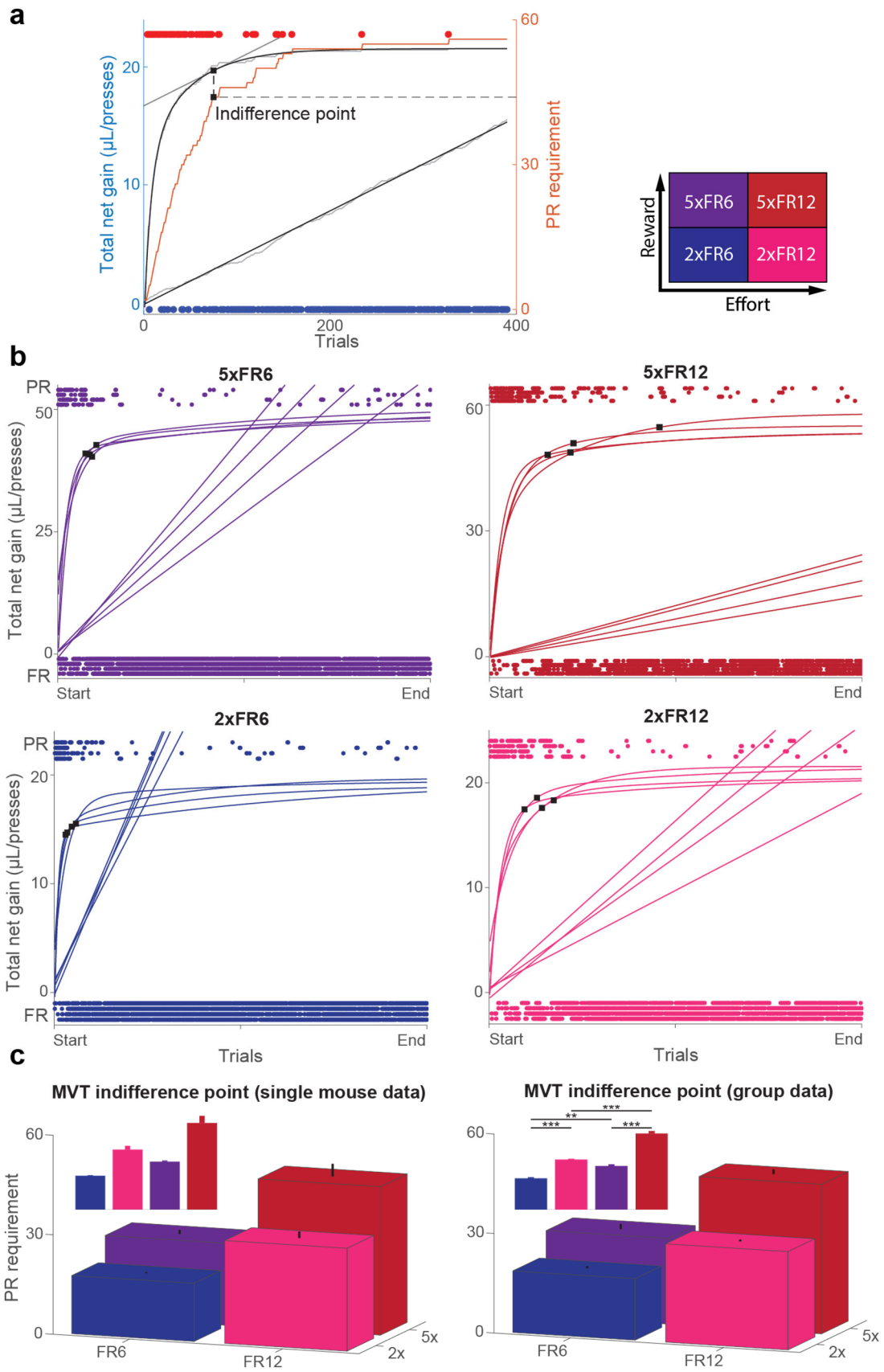
a
Total net gain (µL/presses)
PR requirement
Indifference point
Trials
Reward
Effort
5xFR6
5xFR12
2xFR6
2xFR12
b
5xFR6
5xFR12
2xFR6
2xFR12
PR
FR
Start
End
Trials
c
MVT indifference point (single mouse data)
MVT indifference point (group data)
PR requirement
FR6
FR12
5x
2x

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, $^{**}p < 0.01$, $^{***}p < 0.001$).

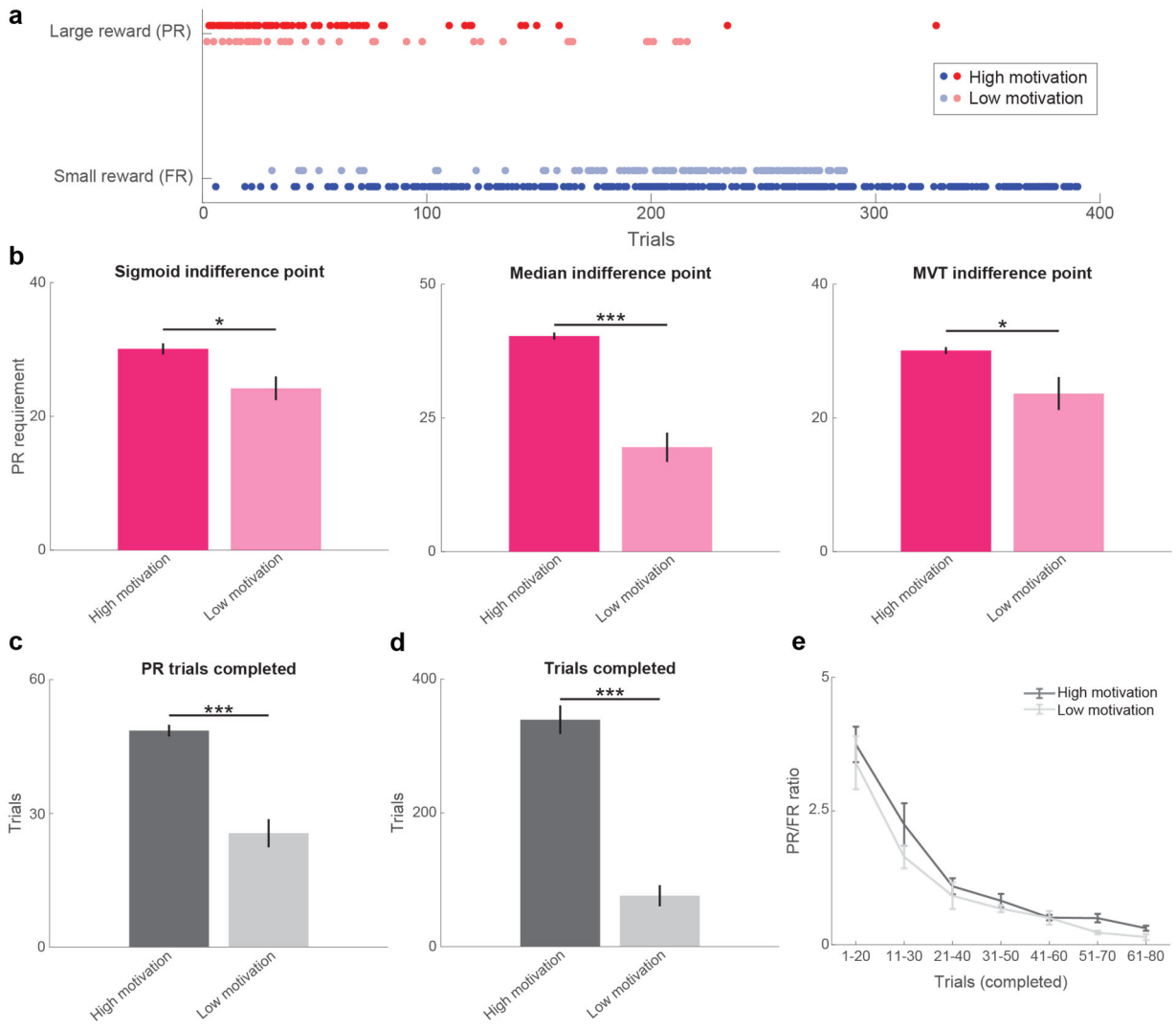

Figure 6. Motivation level effects choices in optimal switching task. a, Example sessions of a high and low motivational state from the same mouse. **b**, Mice switch from PR to FR earlier in a low motivational state (Wilcoxon Rank-sum test, *p < 0.05, ***p < 0.001, n = 7 mice). Error bars reflect standard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

	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