Two-dimensional reward evaluation in mice

Vladislav Nachev[1*§], Marion Rivalan[1,2¶], York Winter[1,2§]

[1] Institute of Biology, Humboldt University, Berlin, Germany [2] Charité University Medicine, Berlin, Germany

[*]**For correspondence:** vladislav.nachev@charite.de

Present Address: [§]Dept. of Biology, Humboldt University, Philippstr. 13, 10099 Berlin, Germany [¶]Exzellenzcluster NeuroCure, Charité University Medicine, Virchowweg 6, 10117 Berlin, Germany

Abstract

When choosing among multi-attribute options, integrating the full information may be computationally costly and time-consuming. So-called noncompensatory decision rules only rely on partial information, for example when a difference on a single attribute overrides all others. Such rules may be ecologically more advantageous, despite deviations from economical optimality. Here we present a study that investigates to what extent animals rely on integration versus noncompensatory rules when choosing where to forage. Groups of mice were trained to obtain water from dispensers varying along two reward dimensions: probability and volume. The choices of the mice over the course of the experiment suggested an initial reliance on integrative rules, later displaced by a sequential rule, in which volume was evaluated before probability. Our results also demonstrate that while the evaluation of probability differences may depend on the reward volumes, the evaluation of volume differences is seemingly unaffected by the reward probabilities.

Introduction

Animals confronted with options that differ on a single attribute generally make economically rational choices consistent with gain maximization (Monteiro, Vasconcelos, and Kacelnik 2013; Rivalan, Winter, and Nachev 2017). In multiattribute choice (Pitz and Sachs 1984; Jansen, Duijvenvoorde, and Huizenga 2012; Hunt, Dolan, and Behrens 2014) however, where reward attributes must be weighed against each other (price *vs.* quality, risk *vs.* payoff, etc.), consistent deviations from economical rationality have been described in humans (Tversky and Kahneman 1974; Rieskamp, Busemeyer, and Mellers 2006; Katsikopoulos and Gigerenzer 2008), non-human animals (Shafir, Waite, and Smith 2002; Bateson, Healy, and Hurly 2003; Schuck-Paim, Pompilio, and Kacelnik 2004; Scarpi 2011; Nachev and Winter 2012; Nachev et al. 2017; Constantinople, Piet, and Brody 2019). Some deviations from gain maximization can be accounted for by considering the ecological circumstances of an animal, which may confer fitness benefits to seemingly irrational choices (Kacelnik 2006; Houston, McNamara, and Steer 2007; Trimmer 2013; McNamara, Trimmer, and Houston 2014).

An animal foraging in its natural environment mostly encounters food items that differ on multiple attributes, but only some of those attributes affect the long-term gains. We refer to those attributes as reward dimensions. In multidimensional choice the decision task is considerably simplified if differences that are (nearly) equal are not evaluated but ignored (Tversky 1969; Pitz and Sachs 1984; Shafir 1994; Shafir and Yehonatan 2014). For example, an animal might only consider the one reward dimension (e.g. prey size) that most strongly affects the long-term gains. Such decision processes in which one reward dimension overrides the others have been described as noncompensatory (Pitz and Sachs 1984; Reid et al. 2015) and can potentially increase speed and decrease computation costs at the expense of accuracy. Attributes can be considered sequentially, for example ranked by salience, until a sufficient difference is detected on one attribute so that a decision can be reached (Brandstätter, Gigerenzer, and Hertwig 2006; Jansen, Duijvenvoorde, and Huizenga 2012). In compensatory decision-making (Pitz and Sachs 1984; Reid et al. 2015) on the other hand, choice is affected by multiple attributes that are integrated into a common decision currency (utility) (Levy and Glimcher 2012). A fully integrative approach that makes use of all the available information (also referred to as absolute reward evaluation Tversky (1969); Shafir (1994); Shafir and Yehonatan (2014)) is equivalent to gain maximization. For example, if options differ along the reward dimensions of amount and probability of obtaining this amount, maximizing the gain is ensured by selecting the option with the highest expected value, which is the product of the amount and probability. Even in two-dimensional reward evaluation, a range of strategies are possible, from sequential and other noncompensatory rules, up to full integration.

When studying animal decision-making, preferences are measured over many choices, especially when options differ in reward probability. Although a rational subject should exclusively select the most profitable option, animals can persist in choosing less profitable options even after long training, usually at some low frequency. The partial preference observed in choice experiments can be explained by profitability matching (Kacelnik 1984), which states that animals proportionally allocate their effort depending on the relative pay-off of the options.

Scalar Utility Theory (SUT; Kacelnik and Brito e Abreu (1998); Marsh and Kacelnik (2002)) is a framework that proposes a proximate mechanism that accounts for partial preferences in the context of reward amount and reward variability (Rosenström, Wiesner, and Houston 2016). Based on findings in psychophysics, SUT postulates that cognitive representations of stimuli exhibit a scalar property, i.e. they have error distributions that are normal with a mean equal to the magnitude of the stimulus and a standard deviation that is proportional to the mean. In other words, SUT states that the memory traces of perceived or expected outcomes of choices are subject to Weber's Law (Akre and Johnsen 2014) and that rewards are evaluated proportionally rather than linearly (Marsh and Kacelnik 2002; Rosenström, Wiesner, and Houston 2016). Therefore, according to SUT choice is modelled by sampling from the internal representations and selecting the most favorable sample. This allows for making quantitative predictions about the strength of preferences from the contrasts between options.

In previous experiments we have demonstrated that proportional processing can be used to predict the choice behavior of animals when options vary along a single dimension (Nachev, Stich, and Winter 2013; Rivalan, Winter, and Nachev 2017). In the present study we extend the application of proportional processing and SUT to two-dimensional choice tasks with the aim to test whether (contradictory) information from two reward dimensions generates choices more consistent with integrative or noncompensatory decision rules.

Results

We performed a series of four experiments (in chronological order) using mice in automated group cages (Haupt, Eccard, and Winter 2010; Rivalan, Winter, and Nachev 2017). Cages had four computer-controlled liquid dispensers that delivered drinking water as a reward. During each of the 18h-long drinking sessions each mouse had access to all dispensers, but received rewards at only two of them. The two rewarding dispensers differed on one or both reward dimensions, probability and volume (Rivalan, Winter, and Nachev 2017). An overview of the differences between choice options in the different experimental conditions is given in Table 1. All experiments were conducted with three different cohorts of eight mice each. Cohort 2 was housed in a different automated group cage than cohorts 1 and 3 (See Methods for differences between cages).

Experiment 1: Mice consistently preferred the more profitable option, even with a trade-off between reward probability and reward volume

In the baseline conditions rewards only differed on one dimension (the relevant dimension), but not on the other dimension (the background dimension). For example, in the BVP1 (baseline for volume at probability 1) condition, both options had the same probability of 0.2, but one option had a volume of 4 μL and the other, a volume of 20 μL (Table 1). Based on previous experiments (Rivalan, Winter, and Nachev 2017), we expected a baseline difference between 4 μL and 20 μL volumes to result in a similar discrimination performance (relative preference for the superior option) compared to a baseline difference between probabilities 0.2 and 0.5. In the C (congruent) condition one option was superior to the other on both dimensions. Finally, in the I (incongruent) condition each of the options was superior to the other on one of the reward dimensions, so that the option that had the higher volume had the lower probability and vice versa. Since the differences on both dimensions were chosen to be of comparable salience (Rivalan, Winter, and Nachev 2017), we expected

Table 1: Overview of the experimental conditions in all four experiments.

		option A			option B			EV_A/EV_B
experiment[a]	condition[b]	volume[c]	probability	EV[d]	volume[c]	probability	EV[d]	relative value
1	BPV1	4	0.2	0.8	4	0.5	2.0	0.40
1	BPV2	20	0.2	4.0	20	0.5	10.0	0.40
1	BVP1[e]	4	0.2	0.8	20	0.2	4.0	0.20
1	BVP2	4	0.5	2.0	20	0.5	10.0	0.20
1	C	4	0.2	0.8	20	0.5	10.0	0.08
1	I	4	0.5	2.0	20	0.2	4.0	0.50
2	BPV1	4	0.2	0.8	4	1.0	4.0	0.20
2	BPV2	20	0.2	4.0	20	1.0	20.0	0.20
2	BVP2	4	1.0	4.0	20	1.0	20.0	0.20
2	C	4	0.2	0.8	20	1.0	20.0	0.04
2	I	4	1.0	4.0	20	0.2	4.0	1.00
3	PV1	4	0.2	0.8	4	0.5	2.0	0.40
3	PV2	10	0.2	2.0	10	0.5	5.0	0.40
3	PV3	15	0.2	3.0	15	0.5	7.5	0.40
3	PV4	20	0.2	4.0	20	0.5	10.0	0.40
3	VP1	4	0.2	0.8	10	0.2	2.0	0.40
3	VP2	4	0.5	2.0	10	0.5	5.0	0.40
3	VP3	4	0.7	2.8	10	0.7	7.0	0.40
3	VP4	4	0.8	3.2	10	0.8	8.0	0.40

[a] conditions in experiment 1 and 4 were identical; only conditions for experiment 1 are shown here for brevity;

[b] condition sequences were randomized for each mouse;

[c] volumes (in microliters) shown are for cohorts 1 and 3. In cohort 2 the volumes were 4.7 instead of 4, 9.4 instead of 10, 14.0 instead of 15, and 20.3 instead of 20 microliters;

[d] EV: expected value;

[e] condition BVP1 in experiment 1 was not repeated in experiment 2, but instead the results from experiment 1 were reused in further analyses

the mean discrimination performance in the incongruent condition to be at chance level (0.5), despite the difference in expected value (Table 1).

In experiment 1 and in all subsequent experiments, each mouse had its individual pseudo-random sequence of conditions. However, each condition was experienced by each mouse in two consecutive drinking sessions (first exposure and reversal), with a spatial reversal of the two reward conditions between the two sessions. In order to investigate how the two reward dimensions contributed towards choice, we looked at the contrasts between the baselines (when only one dimension was relevant) to the conditions when the two dimensions were congruent or incongruent to each other. We used equivalence tests (Lakens 2017) with an *a priori* smallest effect size of interest (sesoi) of 0.1, chosen based on variance observed in a previous study (see Fig.4 in Rivalan, Winter, Nachev 2017). When using equivalence tests, if the 90% confidence interval (CI) of 105 the result estimate falls within the equivalence bounds (+sesoi, -sesoi) the effect is statistically smaller than any effect deemed worthwhile (Methods). If the 90% CI is not fully bounded by the sesoi, but the 95% CI includes the effect size of zero, the results are deemed inconclusive. Therefore, we only considered absolute differences of at least 0.1 percentage points to be of biological relevance. Smaller differences, regardless of their statistical significance using other tests, were considered to be trivial.

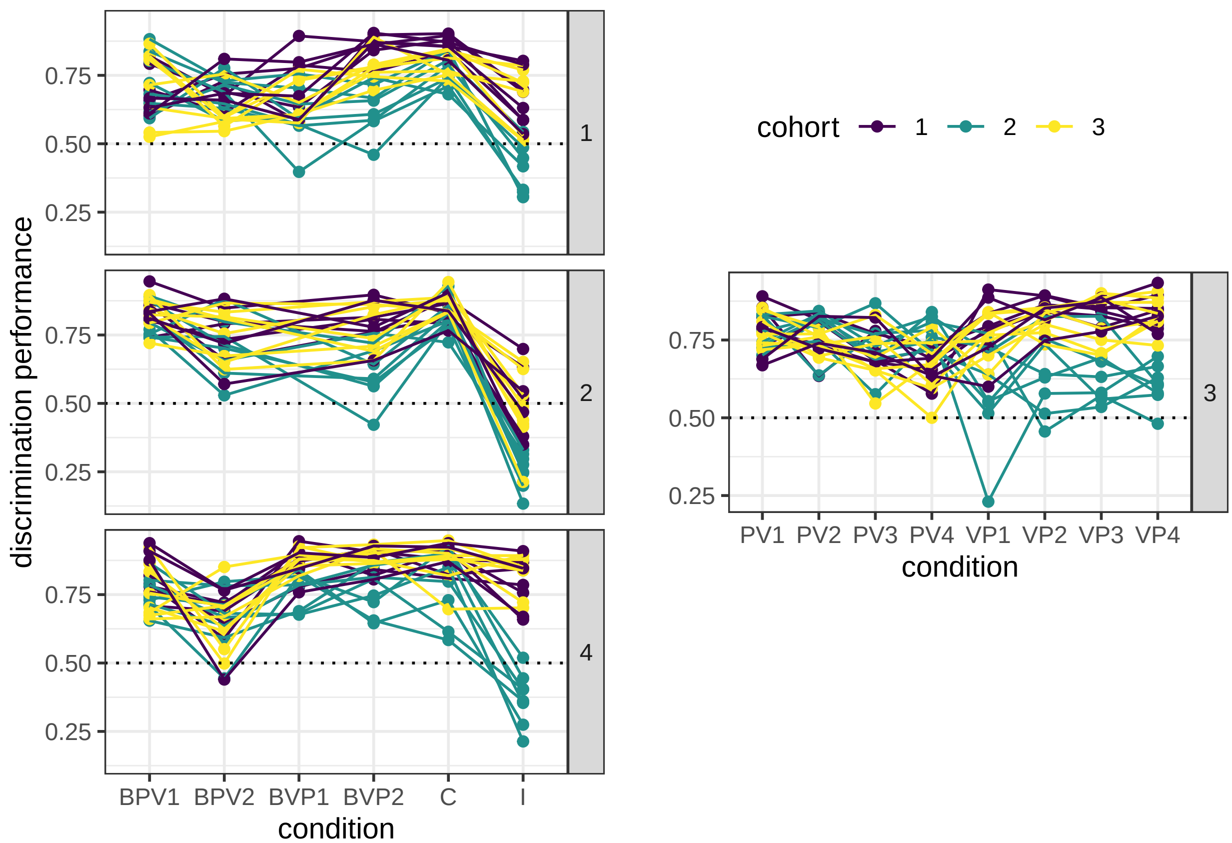

Figure 1: **Overview of discrimination performance for all mice in all experiments.** Experiments 1 through 4 are shown in different panels (1-4). Each colored dot is the mean discrimination performance of an individual mouse over two presentations of the same condition (first exposure and reversal). The experimental conditions are described in detail in Table 1. The discrimination performance gives the relative visitation rate of the more profitable option, or, in the incongruent condition, the option with the higher volume. Dotted line gives the chance level of 0.5. Data are shown in different colors for three different cohorts of eight mice each (total $n = 24$). Data from the same individuals are connected with lines. Cohort 2 (green) was tested in a different cage set-up than cohorts 1 and 3 (see Methods for details).

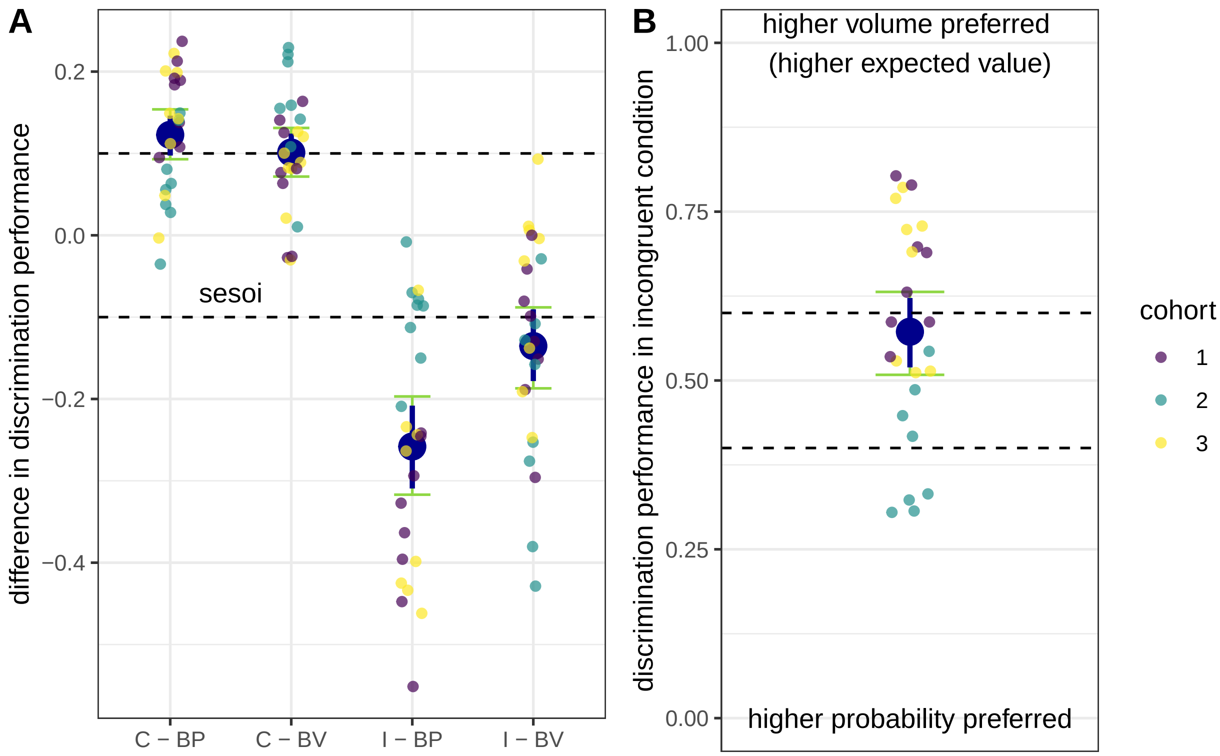

Figure 2: **Discrimination performance in experiment 1.** (**A**) Difference between discrimination performance in the baseline conditions and in the congruent and incongruent conditions. Colored dots show the individual differences in discrimination performance for the given conditions of each individual mouse (total $n = 24$), with different cohorts ($n = 8$) shown in different colors. Positive differences indicate an increase in performance and negative differences - a decrease in performance, compared to the baseline. Large blue circles give the means and the blue vertical lines the 90% confidence intervals from non-parametric bootstraps. The smallest effect size of interest (sesoi) is represented by the dashed lines. Green whiskers give the 95% CI from non-parametric bootstraps. When the blue confidence intervals lie completely within the sesoi interval there is statistical support for equivalence (Lakens 2017). When the green confidence intervals do not cross the zero line, and the blue confidence intervals are not bounded by the sesoi, there is statistical support for difference. Discrimination performance in the baseline conditions was calculated from the mean values from the two different baseline conditions for each reward dimension (volume and probability), i.e. BP was the mean of BPV1 and BPV2, and BV was the mean of BVP1 and BVP2 (Table 1, Fig. 1). The discrimination performance in the incongruent condition was calculated as the relative preference for the higher probability dispenser when contrasted with the probability baseline (I - BP) and for the higher volume dispenser when contrasted with the volume baseline (I - BV). (**B**) Discrimination performance in the incongruent condition. Dashed lines give the sesoi around chance level performance. Remaining notation is the same as in (A). In this experiment the option with the higher volume was also the more profitable option.

An overview of all experimental results is seen in Fig. 1. Compared to the baselines, mice showed an increase in discrimination performance in the congruent condition and a decrease in performance in the incongruent condition (Fig. 2A). Contrary to our expectations based on previous work, the trade-off between volume and probability chosen for this experiment did not abolish preference in the incongruent condition, with a discrimination performance significantly higher than the chance level of 0.5 (0.57, 95% CI = [0.5, 0.63], Fig. 2B). Thus, in the incongruent condition mice preferred the more profitable option and the subjective contrast in probability was not stronger than the subjective contrast in volume.

Experiment 2: Some evidence for equal weighing of reward probability and reward volume

In previous experiments (Rivalan, Winter, and Nachev 2017), we had shown that the relative stimulus intensity (i), i.e. the absolute difference between two options divided by their mean (difference/mean ratio), was a good predictor of discrimination performance for both volume and probability differences. Another finding from these experiments was that, at least initially, mice responded less strongly to differences in volume than to differences in probability, despite equivalence in expected values (Rivalan, Winter, and Nachev 2017). We aimed to correct for this effect in experiment 1 by selecting options with a higher relative intensity for volume (4 μL vs. 20 μL, $i = 1.33$) than for probability (0.2 vs. 0.5, $i = 0.857$). However, the results from experiment 1 were not consistent with a subjective equality between the chosen volume and probability differences. In order to test whether we had over-corrected for decreased sensitivity to volume in experiment 1, we replaced the 0.5 probability with a probability of 1 in each experimental condition of experiment 2 (Table 1). With the two choice options having the same relative intensities ($i = 1.33$) for both reward dimensions and the same expected values, we hypothesized that the discrimination performance in the incongruent condition would be at chance level if both dimensions were equally weighed and equally perceived. On the other hand, if mice were less sensitive for volume than for probability differences as in our previous experiments, then the discrimination performance in the incongruent condition should be skewed towards probability (< 0.5).

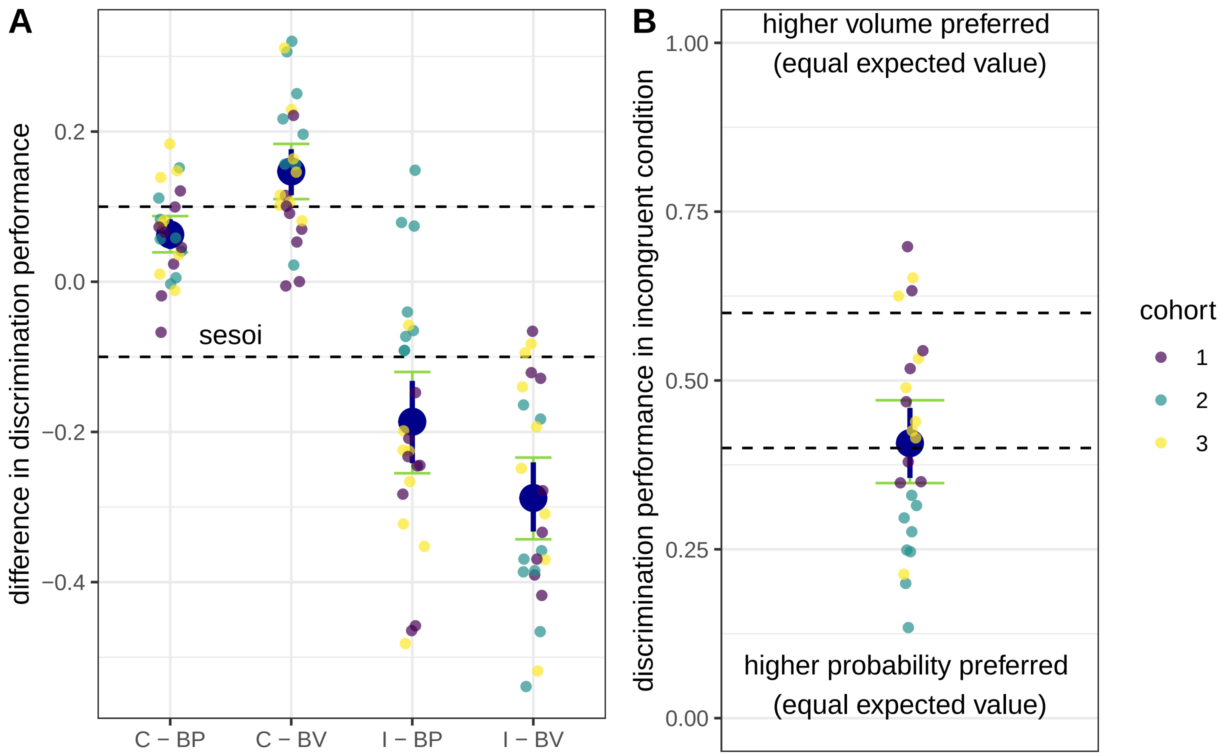

Figure 3: **Discrimination performance in experiment 2.** Same notation as in Fig. 2. (**A**) Difference between discrimination performance in the baseline conditions and in the congruent and incongruent conditions. Discrimination performances in the baseline conditions was calculated from the mean values from the two different baseline conditions for each reward dimension (volume and probability), i.e. BP was the mean of BPV1 and BPV2, and BV was the mean of BVP1 and BVP2, where the values for condition BVP1 were taken from experiment 1 (Table 1). The discrimination performance in the incongruent condition was calculated as the relative preference for the higher probability dispenser when contrasted with the probability baseline (I - BP) and for the higher volume dispenser when contrasted with the volume baseline (I - BV). (**B**) Discrimination performance in the incongruent condition. In this experiment both options were equally profitable and had the same expected value.

In contrast to experiment 1, in experiment 2 mice showed an increase in discrimination performance in the congruent condition only when compared to the volume baseline, but not when compared to the probability baseline (Fig. 3A). As in experiment 1, the discrimination performance in the incongruent condition was lower than in either of the two baselines (Fig. 3A). Although the discrimination performance in the incongruent condition was again different from 0.5 (0.41, 95% CI = [0.35, 0.47]), it was lower than chance, thus skewed towards probability (Fig. 3B). However, when the data from cohort 2 were excluded, the discrimination performance became equivalent to 0.5 (0.48, 95% CI = [0.42, 0.54]). We return to the differences between cohorts in the discussion.

Experiment 3: Probability discrimination decreased with an increase in reward volume, but volume discrimination was not affected by changes in reward probability

In the previous experiments we used two different baseline conditions for each dimension (BPV1, BPV2, BVP1, and BVP2, Table 1), in order to exhaust all combinations of reward stimuli and balance the experimental design. However, we also wanted to test whether the level of the background dimension despite being the same across choice options nevertheless affected the discrimination performance on the relevant dimension. If mice use a noncompensatory decision rule, we can predict that regardless of the level of the background dimension, the discrimination performance on the relevant dimension should remain constant. Alternatively, with absolute reward evaluation the subjective difference between the options is said to decrease as the background (irrelevant) dimension increases and therefore the discrimination performance is also expected to decrease (Shafir and Yehonatan 2014). This prediction is derived from the concave shape of the utility function, which is generally assumed to increase at a decreasing rate with the increase in any reward dimension (Kahneman and Tversky 1979; Kenrick et al. 2009; but see also Kacelnik and Brito e Abreu 1998). The same prediction can be made if we assume that motivation decreases with satiety, i.e. the strength of preference decreases under rich environmental conditions (Schuck-Paim, Pompilio, and Kacelnik 2004), for example at high reward volume or probability. In order to test whether the two reward dimensions (volume and probability) interact with each other even when one of them is irrelevant (as background dimension that is the same across choice options), we performed experiment 3.

The conditions in experiment 3 were chosen to be similar to the baseline conditions in the previous experiments, by having one background and one relevant dimension (Table 1). The relevant dimension always differed between the two options. For the probability dimension, we selected the same values of 0.2 and 0.5 ($i = 0.86$), as in the previous experiments. For the volume dimension we selected the values of 4 μL and 10 μL (4.8 μL and 9.6 μL in cohort 2, Table 1), because the combination of a higher volume with a probability of 0.8 was expected to result in an insufficient number of visits for analysis. Cohort 2 had different reward volumes due to differences in the pumping process between the two cages used (Methods), which also resulted in a lower relative intensity for volume (0.67 instead of 0.86; we will return to this point in the discussion). There were four different levels for each background dimension (volume and probability, Table 1). Each mouse had its own pseudo-random sequence of the eight possible conditions. In order to test whether the background dimension affected discrimination performance we fitted linear regression models for each mouse and each dimension, with discrimination performance as the dependent variable and background level as the independent variable. The background level was the proportion of the actual value to the maximum of the four values tested, e.g. the background levels for volumes 4, 10, 15, 20 were 0.2, 0.5, 0.75, 1, respectively. We defined *a priori* a smallest effect size of interest (sesoi), as 0.125, which is the slope that would result from a difference of 0.1 in discrimination performance between the smallest and the largest background levels (PV1 and PV4, 0.2 and 1, respectively). A slope estimate (whether positive or negative) within the sesoi interval was considered equivalent to zero and demonstrating a lack of an effect of background dimension.

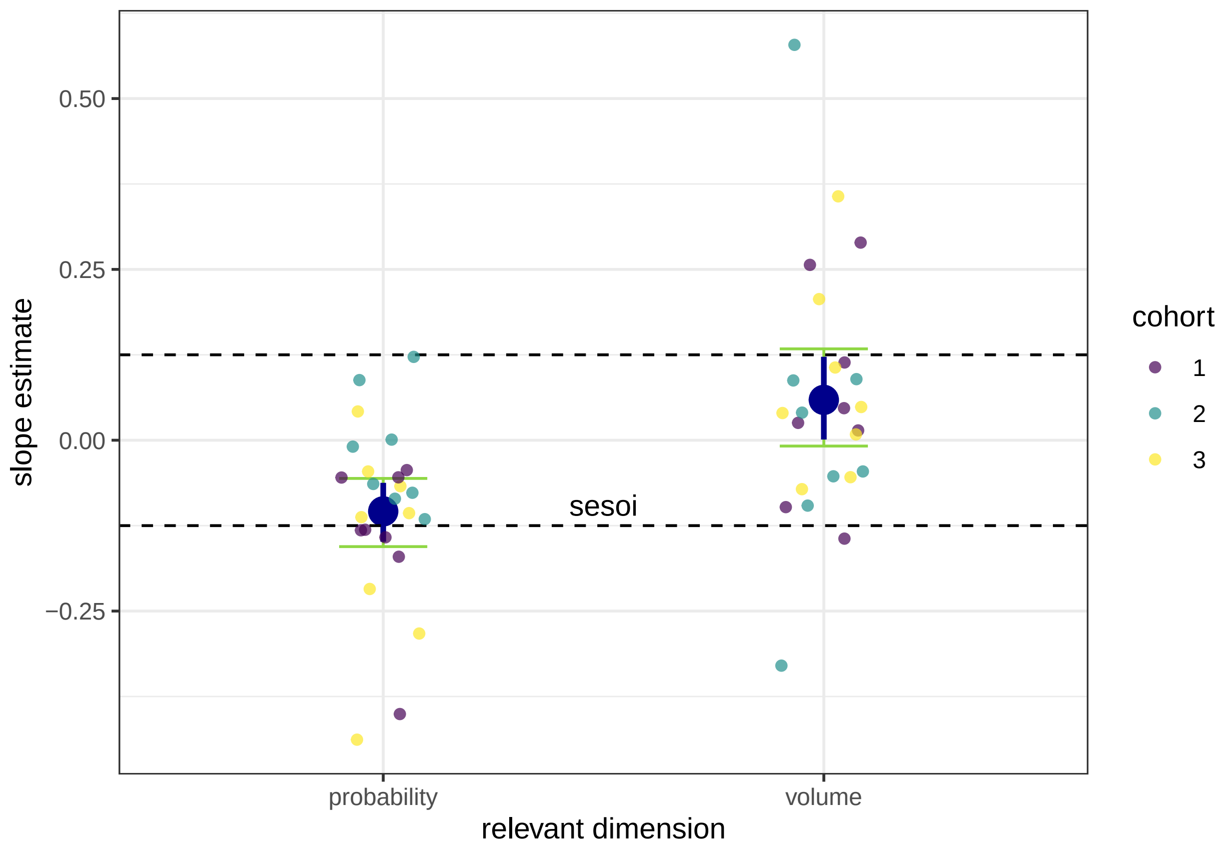

Figure 4: **Slope estimates for the effect of the background dimension on the discrimination performance in the relevant dimension**. The two choice options always differed along the relevant dimension (either probability or volume, given on the abscissa) at a fixed relative intensity. The discrimination performance for each mouse was measured at four different levels of the background dimension, which was set at the same values on both choice options during a single drinking session, but differed from condition to condition (Table 1). Each dot is the individual slope estimate over the four different background dimensions, color-coded for cohort number. The smallest effect size of interest (sesoi, dashed lines) was determined to be the slope (0.125) that would have resulted in a difference in discrimination performance of 0.1, from the lowest to the highest level of the background dimension. Large blue circles give the means and the blue vertical lines the 90%-confidence intervals from non-parametric bootstraps. Green whiskers give the 95% CI from non-parametric bootstraps.

The results of experiment 3 show that the discrimination performance for probability decreased with increasing volumes, although the effect size was small (PV1-PV4, Fig. 1, Fig. 4). In contrast, the discrimination performance for volume was independent from probability as the background dimension, since the slope was smaller than the sesoi (VP1-VP4, Fig. 1, Fig. 4). These results partially support the hypothesis that decision-makers may ignore a reward dimension along which options do not vary.

Experiment 4: Mice improved their volume discrimination over time

For laboratory mice that have unrestricted access to a water bottle, the volume of a water reward is not usually a stimulus that predicts reward profitability. In previous experiments (Rivalan, Winter, and Nachev 2017), mice had shown an improved discrimination performance for volume over time. This suggests that with experience mice become more attuned to the relevant reward dimension. In order to test whether the discrimination performance for one or both dimensions improved over time, we performed experiment 4, which had the same conditions as experiment 1 (Table 1), but with a new pseudo-random order. The same mice participated in all experiments (1 to 4), with about seven weeks between experiment 1 and experiment 4. As in the previous experiments, we also used equivalence tests on the contrasts between the baselines and the congruent and incongruent conditions.

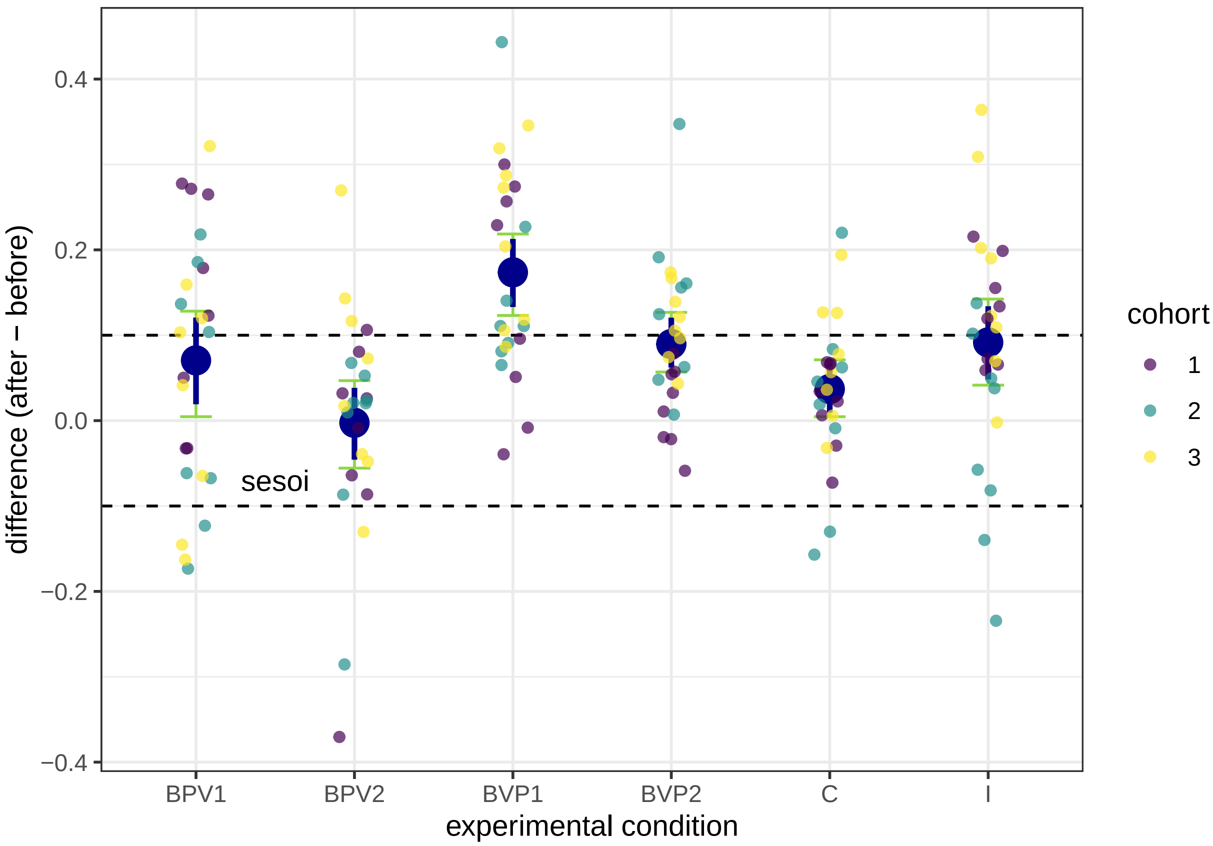

Figure 5: **Difference in discrimination performance between identical conditions in experiment 1 and experiment 4**. Same notation as in in Fig. 2. The sequence of conditions was pseudo-random in each experiment and different for each individual. Positive differences indicate an increase in discrimination performance with time. Mice were seven weeks old at the beginning of experiment 1 and 13-14 weeks old at the beginning of experiment 4. The discrimination performance in the incongruent condition was calculated as the relative preference for the higher volume dispenser.

In the comparison between experiment 1 and experiment 4, mice showed an improved discrimination performance in both volume baselines, as well as in the incongruent and BPV1 conditions (Fig. 5). There was no change in the C condition. When we applied a familywise error control procedure, only the BPV1 result changed from an increase to inconclusive. Thus, consistent with our prior findings, mice improved their volume discrimination over time. The discrimination performance in the congruent condition was better than in the probability baseline, but the same as in the volume baseline (Fig. 6A). The discrimination in the incongruent condition was lower than in any of the two baselines, but the difference to the volume baseline was smaller (Fig. 6A). Finally, compared to experiment 1 the influence of the volume dimension on choice was even more pronounced (Fig. 6B).

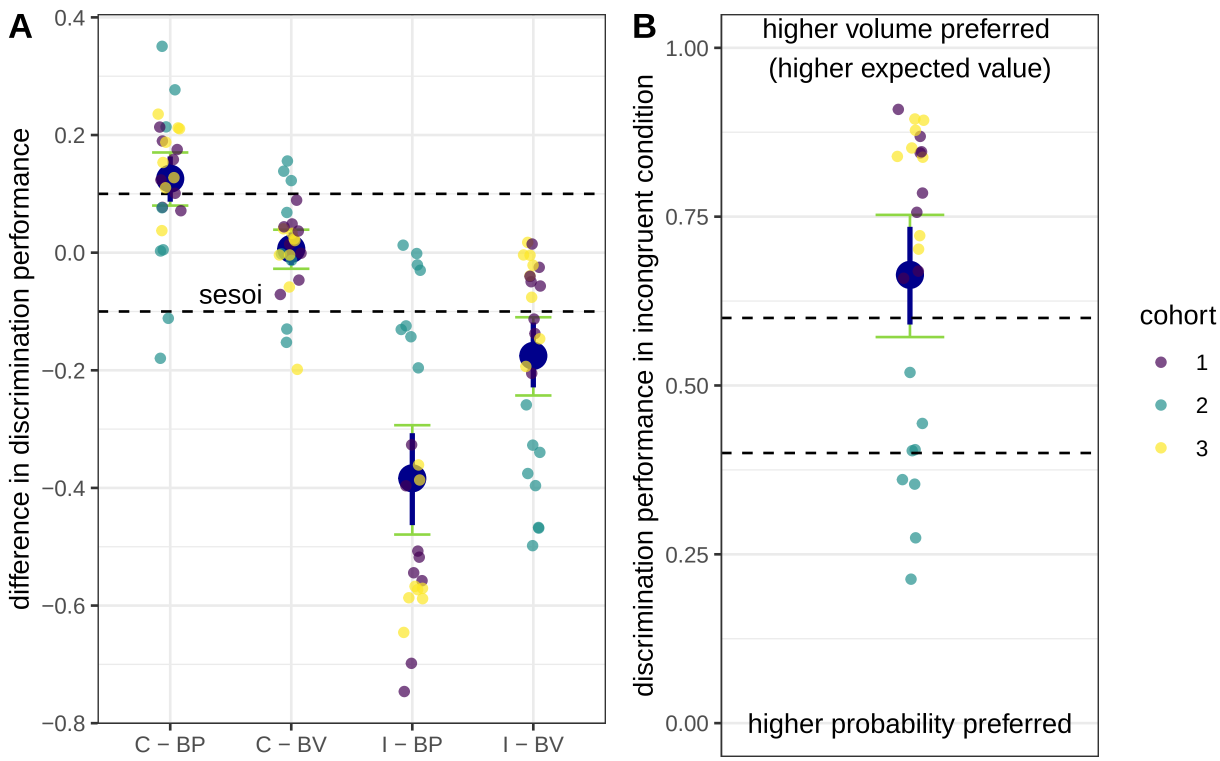

Figure 6: **Discrimination performance in experiment 4, with identical conditions to experiment 1.** Same notation as in Fig. 2. (**A**) Difference between discrimination performance in the baseline conditions and in the congruent and incongruent conditions. The discrimination performance in the incongruent condition was calculated as the relative preference for the higher probability dispenser when comparing to the probability baseline and for the higher volume dispenser when comparing to the volume baseline. (**B**) Discrimination performance in the incongruent condition. In experiments 1 and 4 the option with the higher volume was also the more profitable option. Compare to Fig. 2.

Decision models of two-dimensional choice suggest that mice initially relied on both reward volume and reward probability, but then developed a bias for reward volume

We based our decision models on the Scalar Utility Theory (SUT, Kacelnik and Brito e Abreu (1998); Rosenström, Wiesner, and Houston (2016)), which models memory traces for reward amounts (or volumes) as normal distributions rather than point estimates. The scalar property is implemented by setting the standard deviations of these distributions to be proportional to their means. Choice between two options with different volumes can be simulated by taking a single sample from each memory trace distribution and selecting the option with the larger sample.
As previously explained, the discrimination performance for reward probabilities can be reasonably predicted by the relative intensity of the two options (Rivalan, Winter, and Nachev 2017). This suggests that the memory traces of reward probabilities also exhibit the scalar property, so that discrimination of small probabilities (e.g. 0.2 vs. 0.5, i = 0.86) is easier than discrimination of large probabilities (e.g. 0.5 vs. 0.8, i = 0.46). Consequently, discrimination (of either volumes or probabilities) when options vary along a single dimension can be predicted by SUT.
In order to extend the basic model for multidimensional choice situations, we implemented six variations that differed in the use of information from the volume and probability dimensions (Table 2), including integrative

and noncompensatory models. The information from the different reward dimensions was used to obtain for each choice option a *remembered value* (utility), which exhibited the scalar property. Choice was simulated by single sampling from the *remembered value* distributions with means equal to the *remembered values* and standard deviations proportional to the *remembered values.* The *remembered value* in the scalar expected value and two-scalar models relied on the full integration of both the volume and the probability dimensions, but differed in the implementation of the scalar property, which either affected only the volume dimension (scalar expected value) or both dimensions (two-scalar; Table 2). In the randomly noncompensatory model, the *remembered value* for each choice was determined by only one of the reward dimensions, selected at random. In the winner-takes-all model, choice was exclusively driven by the more salient of the two dimensions. In the last two models the saliences of the reward dimensions were considered sequentially, either probability first, or volume first, and a decision was reached if the salience surpassed a given threshold, estimated in previous discrimination experiments (Rivalan, Winter, and Nachev 2017).

If we assume that mice do not change strategies over time, the best model should predict their choices in all experiments. We thus used the probability baselines (BPV1 and BPV2) in experiments 1 and 4 to estimate the free parameters of the models and then used simulations to predict choices in all remaining experiments. For each model we generated 100 choices by 100 virtual mice for each experimental condition in each of the four experiments. We then quantified the out-of-sample model fits to the empirical data by calculating root-mean-square-errors (RMSE) and ranked the models by their RMSE scores.

Table 2: Decision-making models.

abbreviation	model	remembered value	criterion	γ
sev	scalar expected value	$\pi\mathcal{N}(v,\gamma v)$	-	1.05
2scal	two-scalar	$\mathcal{N}(\pi,\gamma\pi)\times\mathcal{N}(v,\gamma v)$	-	0.65
rnonc	randomly noncompensatory	$\mathcal{N}(r,\gamma r)$	$\theta_v = 0.5$	0.05
wta	winner-takes-all	$\mathcal{N}(r,\gamma r)$	$\theta = 1$	0.7
pfirst	probability first	$\mathcal{N}(r,\gamma r)$	if $s(\pi) > 0.8$ then $r = \pi$, if $s(v) > 0.8$ then $r = v$, otherwise $\theta = 0.5$	0.95
vfirst	volume first	$\mathcal{N}(r,\gamma r)$	if $s(v) > 0.8$ then $r = v$, if $s(\pi) > 0.8$ then $r = \pi$, otherwise $\theta = 0.5$	0.5

Note: π - probability estimate; v - volume estimate; γ - coefficient of variation; r - either v or π depending on the *criterion*; θ_v - probability of selecting the volume dimension; θ - probability of selecting the dimension with the higher salience; $s(r)$ - salience of dimension r, calculated as $\frac{max(r)-min(r)}{\overline{r}}$, where $\overline{r}$ is the arithmetic mean of r over all options.

There was no single model that could best explain the choice of the mice in all four experiments, but the scalar expected value, two-scalar, and winner-takes-all models were in the top-three performing models most frequently (Tables 2, 3, see also Appendix 1 Figures A5, A6, A7, and A8). However, due to the unexpected differences in performance between cohort 2 and the other cohorts (e.g. Appendix 1 Figure A8), we also ranked the models separately for the different mouse groups, depending on which cage they performed the experiments in (cohorts 1 and 3 in cage 1 and cohort 2 in cage 2). Indeed, two different patterns emerged for the different cages. For the two cohorts in cage 1, scalar expected value and two-scalar were the best supported models, followed by the winner-takes-all and volume first models (Table 4. Notably, the volume first model was the best performing model in the later experiments 3 and 4, but the worst model in the

Table 3: Best performing models ranked by root-mean-square-errors (RMSE).

	experiment			
rank	1	2	3	4
1	sev	sev	vfirst	2scal
2	2scal	2scal	sev	wta
3	wta	wta	2scal	sev
4	rnonc	pfirst	wta	vfirst
5	pfirst	rnonc	pfirst	rnonc
6	vfirst	vfirst	rnonc	pfirst

earlier experiments 1 and 2. In contrast, the probability first model was the best supported model for cohort 2, followed by the randomly noncompensatory, scalar expected value, and two-scalar models (Table 5.

Table 4: Best performing models ranked by root-mean-square-errors (RMSE) for cohorts 1 and 3.

	experiment			
rank	1	2	3	4
1	sev	2scal	vfirst	vfirst
2	2scal	sev	sev	2scal
3	wta	wta	2scal	wta
4	rnonc	rnonc	wta	sev
5	pfirst	pfirst	rnonc	rnonc
6	vfirst	vfirst	pfirst	pfirst

Table 5: Best performing models ranked by root-mean-square-errors (RMSE) for cohort 2.

	experiment			
rank	1	2	3	4
1	pfirst	pfirst	pfirst	pfirst
2	rnonc	rnonc	wta	rnonc
3	sev	sev	2scal	wta
4	wta	2scal	sev	2scal
5	2scal	wta	rnonc	sev
6	vfirst	vfirst	vfirst	vfirst

Discussion

The foraging choices of the mice in this study provide evidence both for and against full integration of reward volume and probability. In the first two experiments, mice differed in discrimination performance in the conditions in which both reward dimensions were relevant (congruent and incongruent conditions) compared to the baselines, in which only one of the two dimensions was relevant (Figs. 2, 3). Consequently, the best supported models for these two experiments (cohort 2 excluded, see discussion about differences between cohorts below) were the models that made use of the full information from both reward dimensions (sev, 2scal), or from the dimension that was subjectively more salient (wta, Table. 4). Although these models were good predictors of choices in experiments 3 and 4 as well, the best-performing model was the one that considered the probability dimension only if differences on the volume dimension were insufficient to reach a decision (Table 4). Thus, it appears that mice initially used information from all reward dimensions without bias and with experience started to rely more on one reward dimension and disregarded the other when both dimensions differed between choice options. Interestingly, in human development the use of integrative decision rules has also been shown to decrease with age (Jansen, Duijvenvoorde, and Huizenga 2012).

In similar and more complex choice situations when options vary on several dimensions, an animal has no immediate method of distinguishing the relevant from the background dimensions. Instead it must rely on its experience over many visits before it can obtain information about the long-term profitability associated with the different reward dimensions. Under such circumstances a decision rule that considers all or the most salient reward dimensions initially and prioritizes dimensions based on gathered experience can be profitable without being too computationally demanding. Indeed, with the particular experimental design in this study, a mouse using a "volume first" priority heuristic would have preferentially visited the more profitable option (whenever there was one) in every single experimental condition, including the incongruent conditions.

Scalar property considerations

An alternative explanation of our main results is that the mice used the "volume first" heuristic from the beginning of the experiment, but only became better at discriminating volumes (their coefficient of variation γ decreased) in the last two experiments. This interpretation is supported by the comparison between experiments 1 and 4 (Fig. 5), as well as from previous experiments (Rivalan, Winter, and Nachev 2017), in which mice improved their volume discrimination over time. However, it is not possible with these data to distinguish whether the effect was caused by training or age. Perhaps an increase in mouth capacity (Vora, Camci, and Cox 2016) or, potentially, in the number of acid-sensing taste receptors (Zocchi, Wennemuth, and Oka 2017) due to growth and aging could allow adult mice to better discriminate water volumes. We assumed that mice consumed all water without spilling, but perhaps less experienced mice spill some water. Comparing the discrimination performance of older naive and younger trained mice would help clarify this confound.

The increase in discrimination performance for volume between experiments 1 and 4 (Fig. 5) suggests that the scalar property only approximately holds, and that the γ for volume is not truly constant over a long period of time. This can be seen as evidence against the scalar expected value model, which assumes that the same coefficient of variation affects performance along each reward dimension. Instead, the improving volume discrimination supports a version of the two-scalar model, in which there are two different scalars ($\gamma_\pi \neq \gamma_v$). Alternatively, there might be only one scalar, associated with dynamic relative weights of the two dimensions (which can be implemented as a changing θ_v in the randomly noncompensatory model, Fig. A2). Yet another model extension that can account for the improving volume discrimination would be to introduce an explicit sampling (exploration-exploitation balance) method (Sih and Del Giudice 2012; Nachev and Winter 2019). In natural conditions reward dimensions rarely remain stable over time and foragers can benefit from making sampling choices to gather information about the current state of the environment. Thus, not all choices need to be based on expected values and individuals may differ in their sampling rates (Sih and Del Giudice 2012; Rivalan, Winter, and Nachev 2017; Nachev and Winter 2019). With such an implementation it is not the scalar but the frequency of sampling visits that changes over time, causing differences in discrimination performance. The biggest challenge is that when it comes to volumes and probabilities, no direct method of interrogating an animal's estimate and coefficient of variation exist, so that researchers have to infer these values from choice behavior, which is also affected by motivation and sampling frequency. In contrast, when it comes to time intervals, the peak procedure gives us a more direct measurement of the time estimation of animal subjects (Kacelnik and Brito e Abreu 1998).

Interaction between dimensions and noncompensatory decision-making

Although mice were roughly equally good at discriminating volume rewards at each different probability, the discrimination of probabilities decreased at higher volumes (Fig. 4; the estimated effect size was a decrease of 0.12 between a volume background at 4 μL and at 20 μL). This suggests that the two dimensions interact with each other. Absolute reward evaluation (Shafir 1994; Shafir and Yehonatan 2014) and state-dependent evaluation (Schuck-Paim, Pompilio, and Kacelnik 2004) are both consistent with this decrease in discrimination performance, but not with the lack of effect in the conditions in which the probability was the background dimension. With comparable expected values (Table 1) between the two series of conditions, these hypotheses make the same predictions regardless of which dimension is relevant and which is background. An alternative explanation is that arriving at a good estimate of probability requires a large number of visits and when the rewards are richer (of higher volume), mice satiate earlier and make a smaller total number of visits, resulting in poor estimates of the probabilities and poorer discrimination performance. Consistent with this explanation, mice made on average ($\pm$ SD) 474 $\pm$ 199 nose pokes at the relevant dispensers at 4 μL, but only 306 $\pm$ 64 nose pokes at 20 μL (Fig. A9, PV1 and PV4, respectively).

As mentioned earlier, researchers have proposed that with absolute reward evaluation the difference/mean ratio in an experimental series like our experiment 3 should decrease with the increase of the background dimension, leading to a decrease in the proportional preference for the high-profitability alternative (i.e. discrimination performance) (Shafir and Yehonatan 2014). However, this is only the case if the difference is calculated from the relevant dimension and divided by the mean utility. We suggest that both the difference and the mean should be calculated from the same entity, either utility or one of the reward dimensions. When, as in our sev and 2scal models 1, we calculate utility by multiplying the estimates for each dimension together, the difference/mean ratio of the utility does not change with the change in the background dimension between treatments. In fact, none of our models in experiment 3 exhibited an effect of the background dimension on the discrimination performance, with all slopes equivalent to zero (Fig. A10). Thus, our results also show that absolute reward evaluation does not necessarily predict an effect of background dimension on discrimination performance.

Difference between cohorts

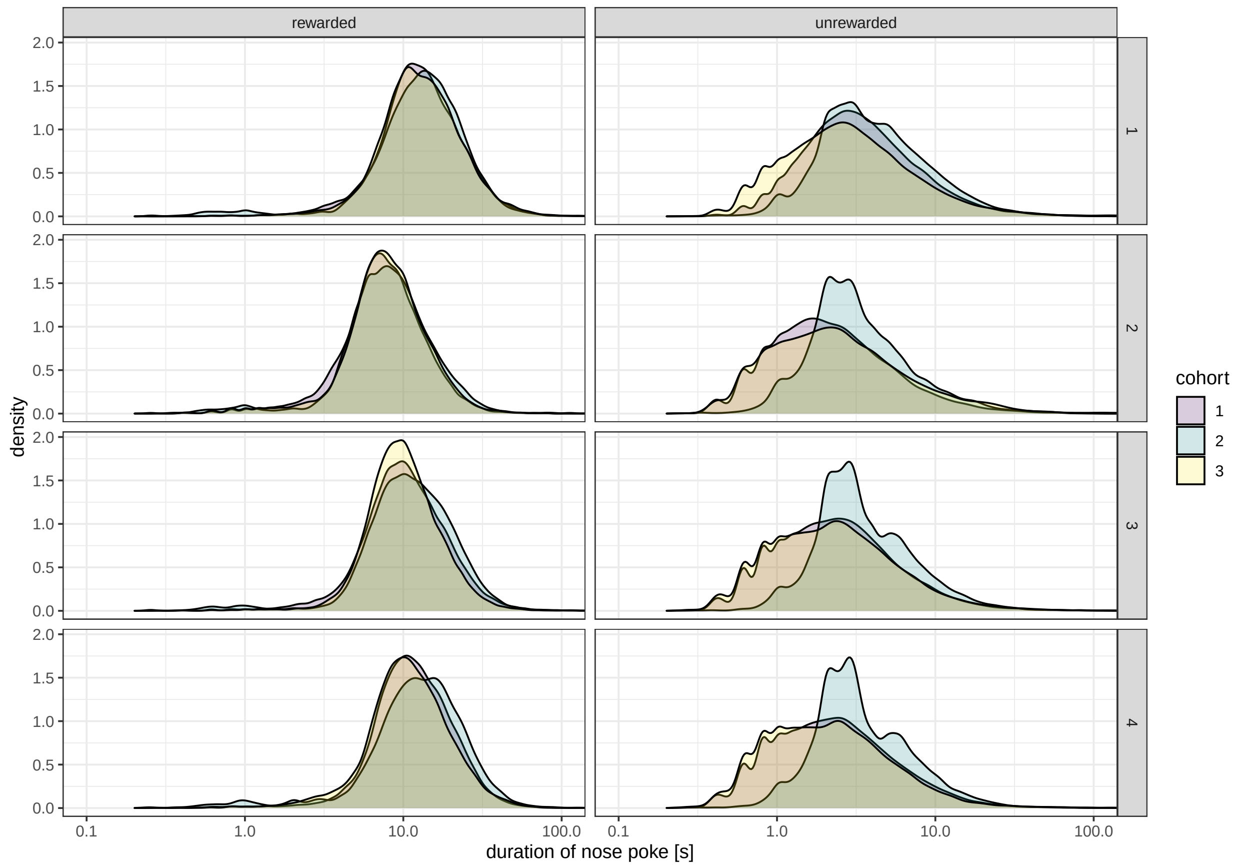

Figure 7: **Visit durations during rewarded and unrewarded nose pokes for the three cohorts in all experiments.** Columns give the status of the nose poke (rewarded or unrewarded) and rows, the experiment number (1-4). Data from the three cohorts are represented by differently color-filled density curves from the observed individual nose poke durations. Note the logarithmic scale on the abscissa.

Our results revealed some striking differences in behavior between cohort 2 and cohorts 1 and 3 (most obvious in Fig. 6). The most likely explanation for this is an effect of the specific experimental apparatus. As explained in Methods, the precision of the reward volumes was lower in cage 2, which housed cohort 2. However, it is unlikely that such a small magnitude of the difference ($0.33 \pm 0.03\ \mu Lstep^{-1}$ in cage 1 vs. $1.56 \pm 0.24\ \mu Lstep^{-1}$ in cage 2) could influence volume discrimination to the observed extent. Future experiments can address this issue by specifically manipulating the reliability of the volume dimension using the higher-precision pump. Instead, we suspect that the difference between cohorts might have been caused by the acoustic noise and vibrations produced by the stepping motors of the pumps. The pump in cage 1 was much louder, whereas the one in cage 2 was barely audible (to a human experimenter). This could have made it harder for mice in cage 2 to discern whether a reward was forthcoming, which could have influenced their choices (Ojeda, Murphy, and Kacelnik 2018). As a result, mice in cage 2 waited longer before leaving the dispenser during unrewarded nose pokes (Fig. 7). This potentially costly delay might have increased the relative importance of the probability dimension (decreased θ_v), resulting in the observed discrimination performance in cohort 2. Furthermore, the same line of reasoning can also explain the improving volume discrimination: from the first to the fourth experiment there was a shift towards shorter unrewarded nose poke durations in the loud cage (cohorts 1 and 3, Fig. 7), suggesting that mice had learned over time to abort the unrewarded visits. This could have decreased the relative importance of the probability dimension (increased θ_v), resulting in better volume discrimination. In an unrelated experiment we tested two cohorts of mice in both cages simultaneously and then translocated them to the other cage. The results demonstrated that differences in discrimination performance were primarily influenced by cage and not by cohort (Nachev, in prep.). Thus, the sound cue associated with reward delivery may be an important confounding factor in probability discrimination in mice, as it provides a signal for the reward outcome (Ojeda, Murphy, and Kacelnik 2018).

Conclusion

In summary, our results show that mice could integrate reward volume and reward probability, which allowed them to select the more profitable option when the two reward dimensions varied independently. The resulting partial preference was consistent with Scalar Utility Theory. However, we also found that, with time mice improved their performance in volume (but not in probability) discrimination tasks and their choices became more consistent with a noncompensatory decision rule, in which volume is evaluated before probability. Finally, we found that mice could discriminate the same pair of probabilities better when reward volumes were smaller, but changes in the reward probability did not affect their volume discrimination performance.

Animals, Methods, and Materials

Animals

The experiments were conducted with C57BL/6NCrl female mice (Charles River, Sulzfeld, Germany, total n = 30). Mice were five weeks old on arrival. The mice from each cohort were housed together, before and during the experiments. They were marked with unique Radiofrequency Identification tags (RFID: 12 × 2.1 mm, 125 kHz, Sokymat, Rastede, Germany) under the skin in the scruff of the neck and also earmarked at age six weeks. At age seven weeks mice were transferred to the automated group home cage for the main experiment. Pellet chow (V1535, maintenance food, ssniff, Soest, Germany) was always accessible from a trough in the cage lid. Water was available from the operant modules of the automated group cage, depending on individual reward schedules. Light conditions in the experiments were 12:12 LD and climatic conditions were 23 ± 2 °C and 50–70% humidity.

Ethics statement

The experimental procedures were aimed at maximizing animal welfare. During experiments, mice remained undisturbed in their home cage. Data collection was automated, with animals voluntarily visiting water dispensers to drink. The water intake and health of the mice was monitored daily. Due to the observational nature of the study, animals were free from damage, pain, and suffering. The animals were not sacrificed at the end of the study, which was performed under the supervision and with the approval of the animal welfare officer (*Tierschutzbeauftragter*) heading the animal welfare committee at Humboldt University. Experiments followed national regulations in accordance with the European Communities Council Directive 10/63/EU.

Cage and dispenser system

We used automated cages (612 × 435 × 216 mm, P2000, Tecniplast, Buggugiate, Italy) with woodchip bedding (AB 6, AsBe-wood, Gransee, Germany), and enriched with two grey PVC tubes and paper towels as nesting material. The cage was outfitted with four computer-controlled liquid dispensers. The experimental set-up is described in detail in Rivalan, Winter, and Nachev (2017). Briefly, mice were detected at the dispensers via infrared beam-break sensors and RFID-sensors. Water delivery at each dispenser could be controlled, so that it could be restricted or dispensed at different amounts on an individual basis. Mice were therefore rewarded with droplets of water from the dispenser spout that they could remove by licking. We changed cage bedding and weighed all animals on a weekly basis, always during the light phase and at least an hour before the start of the testing session. Data were recorded and stored automatically on a laptop computer running a custom-written software in C#, based on the .NET framework. Time-stamped nose poke events and amounts of water delivered were recorded for each dispenser, with the corresponding mouse identity.

A second automated group cage (cage 2) was made for the purposes of this study, nearly identical to the one described above (cage 1). The crucial modification was that the stepping-motor syringe pump was replaced with a model that used disposable plastic 25-mL syringes instead of gas-tight Hamilton glass syringes (Series 1025). Thus, the pumping systems in the two cages differed in the smallest reward that could be delivered and in the precision of reward delivery (mean ± SD: 0.33 ± 0.03 $\mu Lstep^{-1}$ in cage 1 vs. 1.56 ± 0.24 $\mu Lstep^{-1}$ in cage 2). The precision of each pump was estimated by manually triggering reward visits at different preset pump steps (17 and 42 in cage 1, 3 and 12 in cage 2) and collecting the expelled liquid in a graduated glass pipette placed horizontally next to the cage. Each dispenser was measured at least 20 times for each pump step value.

Experimental schedule

The general experimental procedure was the same as in Rivalan, Winter, and Nachev (2017). The water dispensers were only active during a 18h-long drinking session each day, which began with the onset of the dark phase and ended six hours after the end of the dark phase. The reward properties (volume and probability) were dependent on the experimental condition. Rewards were drawn from fixed pseudo-random repeating sequences. These sequences were: 11101111101101111110 for 80%, 11011101110101101110 for 70%, 10110101101001001010 for 50%, 10010100100001001000 for 30%, and 10001000010001000000 for 20%, where 1 is a rewarded nose poke and 0 is an unrewarded nose poke.

Although individual mice shared the same dispensers inside the same cage, they were not necessarily in the same experimental phase or experimental condition. The three cohorts (1-3 in chronological order) were tested consecutively, with cohort 2 housed in cage 2 and the other cohorts housed in cage 1. If after any drinking session during any experimental phase a mouse drank less than 1000 μL of water, we placed two water bottles in the automated cage, gently awakened all mice and allowed them to drink freely until they voluntarily stopped.

Exploratory phase

At the beginning of this phase there were ten mice in each cohort, except for cohort 2, in which one mouse was excluded due to the loss of the RFID tag after implantation (the mouse was in good health condition). The mice were transferred to the automated cages 1-2 hours before the first drinking session of the exploratory phase. The purpose of this phase was to let mice accustom to the cage and learn to use the dispensers to obtain water. Therefore, each nose poke at any dispenser was rewarded with a constant volume of 20 μL. The criterion for advancing to the following training phase was consuming more than 1000 μL in a single drinking session. Mice that did not reach the criterion remained in the exploratory phase until they either advanced to the following phase or were excluded from the experiment ($n = 1$ mouse in cohort 2).

Training phase

In this phase the reward volume was reduced to 10 μL and the reward probability was reduced to 0.3 at all dispensers. These reward values ensured that mice remained motivated to make several hundred visits per drinking session. The training phase was repeated for two to three days until at least eight mice fulfilled the criterion of consuming more than 1000 L of water in one drinking session. The purpose of the training phase was to introduce mice to the reward dimensions (volume and probability) that would be used in the following discrimination experiments. In cohorts 1 and 2, mice were excluded from the experiment if they did not reach the criterion in two days, or, alternatively, if more than eight mice had reached the criterion, mice were excluded at random to ensure a balanced number of mice per dispenser. These mice were returned to regular housing and available for reuse in other experiments.

Autoshaping phase

We introduced an autoshaping phase for the mice in cohort 3, because after two days only six of them had advanced to the training phase. The unusually low number of visits made by mice that did not pass the exploratory phase suggested that the noise produced by the pumping systems might scare naive, shy mice away from the dispensers. In order to ensure that all mice were successfully trained, we designed the autoshaping phase so that rewards at all four dispensers were delivered at regular intervals (7 μL every minute), regardless of the behavior of the mice. After two days, all mice had made at least 200 nose pokes and the cohort was then moved to the previous phase before autoshaping, either exploratory or training. Two days later all mice successfully completed the training phase and two mice were randomly selected out of the experiment, bringing the number of mice to eight. We therefore updated our training procedure to always begin with the autoshaping phase, followed by the exploratory phase and the training phase.

General procedure in the main experiments

After eight mice had successfully passed the training phase, they proceeded with experiment 1 from the main experiments (1-4). In all of the main experiments mice had a choice between four dispensers, where two were not rewarding and the other two gave rewards with volumes and probabilities that depended on the experimental condition (Table 1). In most conditions one of the rewarding dispensers (high-profitability dispenser) was more profitable than the other (low-profitability dispenser). The sequence of conditions was randomized for each individual, so that any given mouse was usually experiencing a different experimental condition than all other mice. On any given day two of the dispensers were rewarding for four mice and the other two were rewarding for the other four mice. Within each group of four, each pair of mice shared the same high and low-profitability dispensers, which were spatially inverted between pairs of mice. This pairing was done to increase the throughput of the experiments, while controlling for potential social learning effects and distributing mice evenly over the dispensers to minimize crowding effects.
As a control for positional biases, each condition was followed by a reversal on the next day, so that the high and low-profitability dispensers were spatially inverted for all mice, whereas the two non-rewarding dispensers remained unchanged. Reversal was followed by the next experimental condition, with random distribution of the dispensers among the pairs of mice following the constraints described above. Over the 50 total days in the main experiment (twice the number of conditions shown in Table 1, because of reversals, plus experiment 4), each mouse experienced each dispenser as a high-profitability dispenser between 11 and 14 times. In the event of an electrical or mechanical malfunction, data from the failed condition and its reversal were discarded and the failed condition was repeated at the end of the experiment. Such a failure occurred once in cohort 1, four times in cohort 2 and did not occur in cohort 3. After experiments 1 and 2, mice were given another training phase (rewards with 10 μL and 0.3 probability) for a single day, before they proceeded with the next experiment. After experiment 3 mice were given water *ad libitum* from a standard water bottle for four days, followed by one day in the training phase, before proceeding with experiment 4. At the end of experiment 4 mice were returned to the animal facility.

Data analysis

On average (mean $\pm$ SD), mice made 477 $\pm$ 163 nose pokes per drinking session (Fig. A9), with an average proportion of $\pm$ 0.1 nose pokes at the rewarding dispensers. In order to focus on post-acquisition performance (Rivalan, Winter, and Nachev 2017), we excluded the first 150 nose pokes at the rewarding dispensers. We then calculated the *discrimination performance* for each mouse and each condition of each experiment. Since each condition was repeated twice (first exposure and reversal), we calculated the discrimination performance as the total number of nose pokes at the high-profitability dispenser divided by the sum of the total number of nose pokes at the high- and at the low-profitability dispensers. Nose pokes at the non-rewarding dispensers were ignored. In the I condition of experiment 2 in which the profitability was equal (relative value = 1, Table 1), the dispenser with the higher reward volume was treated as the "high-profitability" dispenser.
When comparing discrimination performances, we used the two one-sided procedure (TOST) for equivalence testing (Lauzon and Caffo 2009; Lakens 2017). First, we picked a smallest effect size of interest (sesoi) *a priori* as the difference in discrimination performance of 0.1 units in either direction. (The sesoi can be graphically represented as the [-0.1, 0.1] interval around the difference of zero, or as [-0.6, 0.6] around the chance performance of 0.5.) Then, we estimated the mean differences and their confidence intervals (CIs) from 1000 non-parametric bootstraps using the `smean.cl.boot` function in the package `Hmisc` (Harrell and Dupont 2019). For a single equivalence test the 90% CI is usually constructed, i.e. $1 - 2\alpha$ with $\alpha = 0.05$, because both the upper and the lower confidence bounds are tested against the sesoi (Lauzon and Caffo

2009; Lakens 2017). Thus, equivalence was statistically supported if the 90% CI was completely bounded by the sesoi interval around the effect size of zero (the null hypothesis). A difference was considered to be statistically supported if the 95% CI did not contain zero and the 90% CI was not completely bounded by the sesoi interval. If the 95% CI contained zero, but the 90% CI was not completely bounded by the sesoi, then results were inconclusive. Researchers have shown that in order to correct for multiple comparisons in equivalence tests, it suffices to only apply a familywise correction of the α for the problematic cases where the type I error is most likely (Davidson and Cribbie 2019), i.e. when equivalence is supported, but the mean difference is close to the sesoi bound. The families of tests, for which multiple comparisons occur in our study, are the four contrasts in each of experiments 1, 2, and 4 (three families), the tests on the two slopes in experiment 3, and the six before-after contrasts between experiment 1 and 4. For each of these five families the α was divided by $k^2/4$, where k was the number of problematic cases in each family (Caffo, Lauzon, and Röhmel 2013). However, the number of problematic cases did not exceed two in any of the test families, which resulted in the corrected *alpha* equal to the original value of 0.05. Furthermore, even with k equal to four, two, and six (the total number of tests in each test family), only a single result changed from non-equivalent to inconclusive. We therefore report the uncorrected 90% and 95% CIs.

Data analysis and simulations were done using R (Team 2019). All data and code is available in the Zenodo repository: https://doi.org/10.5281/zenodo.3726829.

Simulations

Environment

Each of the experimental conditions was recreated in the simulations as a binary choice task between the high-profitability and the low-profitability options. We did not simulate the two non-rewarding options. Upon a visit by a virtual mouse, a choice option would deliver a reward with its corresponding volume and probability (Table 1). The virtual environment was not spatially and temporally explicit. Thus, no reversal conditions were simulated and the test of each experimental condition consisted in a sequence of 100 choices. All experimental conditions in all four experiments were tested.

Virtual mice

For simplicity and in order to simulate post-acquisition discrimination performance, we assumed that each mouse had a precise estimate of each of the two reward dimensions for both choice options. The virtual mice thus began each experimental condition in a learned state and (further) learning was not simulated.

From its memory traces a virtual mouse generated one *remembered value* distribution for each choice option, according to one of six different rules (models, Table 2). Action selection was then implemented by taking a single sample from each distribution and selecting the option with the larger sample.

Remembered value models

All six models implemented the *scalar property* from the Scalar Utility Theory (SUT, Kacelnik and Brito e Abreu (1998); Rosenström, Wiesner, and Houston (2016)), because the remembered value was modelled as a normal distribution with a standard deviation proportional to its mean. However, the models differed in the way information from the two reward dimensions was used (either through integration of the full information or by one dimension overriding the other).

These models were:

1. *Scalar expected value model.* There is a single memory trace for each option and it consists in the simple product of the estimate for the volume and the estimate for the probability (expected value). The scalar property is implemented as $\pi\mathcal{N}(v, \gamma v)$, where π is the probability estimate. $\mathcal{N}(\mu, \sigma)$ is a normal distribution with mean μ and standard deviation σ, v is the volume estimate, and γ is a free parameter, the coefficient of variation. This model thus utilizes information from all dimensions for every decision.
2. *Two-scalar model.* There are traces for each dimension for every option, where each trace exhibits the scalar property independently and the value is obtained by simple multiplication of the traces for each dimension: $\mathcal{N}(\pi, \gamma\pi) \times \mathcal{N}(v, \gamma v)$. This model also utilizes information from all dimensions for every

decision. Although it allows each dimension to have its own scalar factor, e.g. $\gamma_\pi \neq \gamma_v$, for the sake of simplicity we assume that they are both equal.

The memory traces in the remaining models are identical to the traces in the two-scalar model, but these models usually consider only a single dimension.

3. *randomly noncompensatory model.* Each decision is based on a single dimension, selected with probability $\theta_v = 0.5$.

4. *Winner-takes-all model.* Each decision is based only on the dimension with the highest salience. The salience for a vector of estimates from memory traces (mean values) along one dimension, e.g. volume $v = (v_1, v_2, ..., v_n)$, is calculated as $\frac{max(v)-min(v)}{\overline{v}}$, where n is the number of options. In the case of $n = 2$, the salience is equivalent to the previously described relative intensity measure. For dimensions of equal salience the model reverts to random choice.

The last two models are examples of a lexicographic rule, in which the dimensions are checked in a specific order. If the salience of a dimension is higher than a given threshold, then a decision is made based only on this dimension. Otherwise the next-order dimension is checked. If all dimensions have saliences below the threshold, the model reverts to random choice. The value of the threshold was set at 0.8, the psychometric function threshold for probability (Rivalan, Winter, and Nachev 2017), but we also performed sensitivity analyses on the threshold values (Fig. @**??**fig:senspfirst), Fig. @**??**fig:sensvfirst)).

5. *Probability first model.* Probability is checked first, then volume.

6. *Volume first model.* Volume is checked first, then probability.

Model fits

All models described above share the same free parameter, the scalar factor γ. In order to obtain baseline estimates for γ for each of the models (Table @ref(tab:conds_tab)), we focused on the probability baseline discrimination performances of all mice in experiments 1 and 4 (conditions BPV1 and BPV2). We performed a grid search sensitivity analysis by varying γ with steps of 0.05 in the range of (0.05, 2). We generated 100 decisions by 100 mice for each cell in this grid and then used locally weighted scatterplot smoothing (loess) to fit a model for each condition. The free parameter values that resulted in the smallest RMSEs compared to the observed baseline data were selected for the comparison of the six models (Table 2). We also performed a sensitivity analysis for different values of the free parameters θ_v in the randomly noncompensatory model and of the thresholds for volume and probability in the volume first and probability first models, in the range of (0, 1), with a step of 0.05. The resulting free parameter estimates (across animals) were then used in out-of-sample tests of the six models. For each of the experimental conditions in the four experiments (Table @ref(tab:conds_tab)) and for each of the six models we simulated 100 choices by 100 (identically parametrized) mice. Over the 100 choices we calculated the discrimination performance for each mouse and then used the median of the individual discrimination performances as the model prediction. We then quantified the model fits to the empirical data by calculating root-mean-square-errors (RMSE), excluding the BPV1 and BPV2 conditions in experiments 1 and 4. Finally, we ranked the models by their RMSE scores.

Appendix

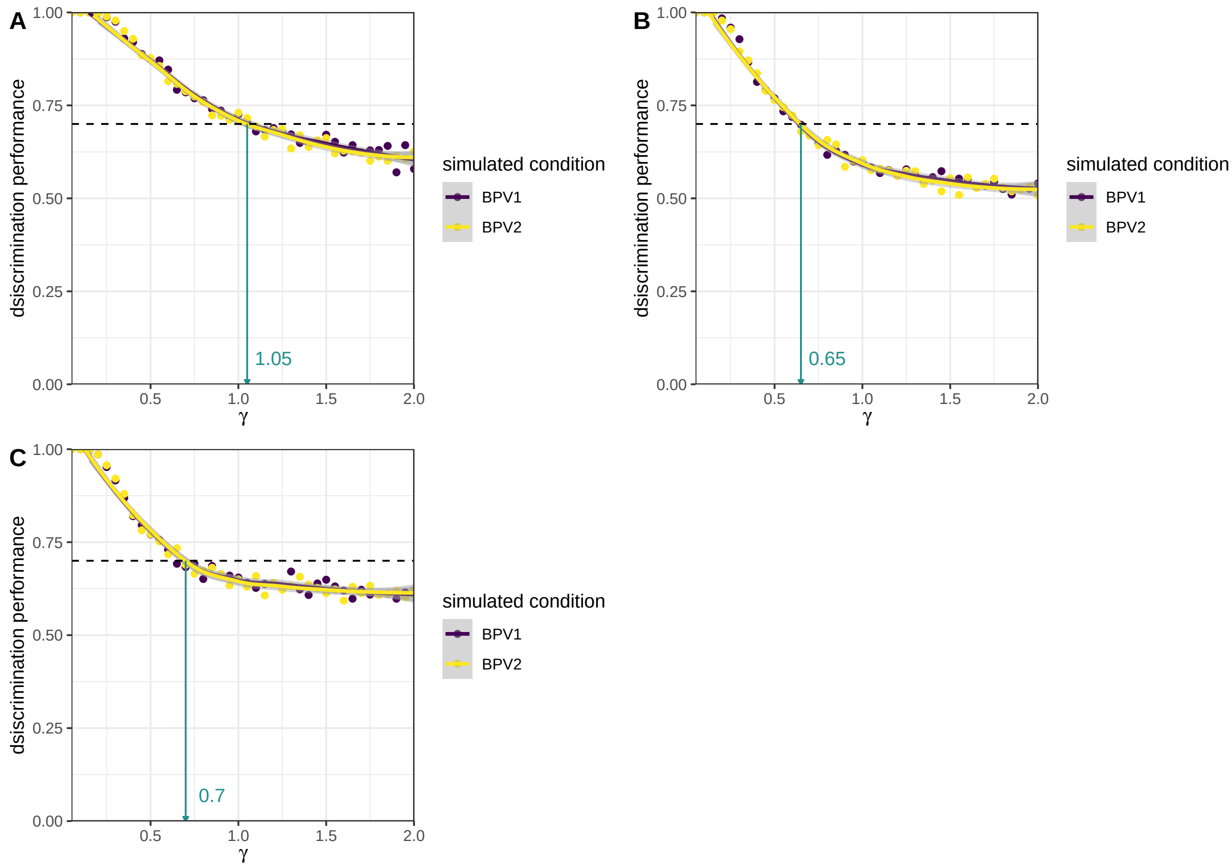

Figure A1: **Sensitivity tests for the models that only had γ as a free parameter.** Dots give the discrimination performances calculated from 1000 choices for each value of γ tested [0.05 , 2] and for each of the baseline conditions "BPV1" (purple) and "BPV2" (yellow). Lines give the corresponding fits based on locally weighted scatterplot smoothing (loess). The dashed line gives the empirical mean discrimination performance from the baseline conditions "BPV1" and "BPV2" and the green arrows point to the value of gamma that resulted in the smallest root-mean-square-errors (RMSEs). These values were then used in the main simulations (Table 2). The different panels give the results for the scalar expected value (**A**), two-scalar (**B**), and winner-takes-all (**C**) models.

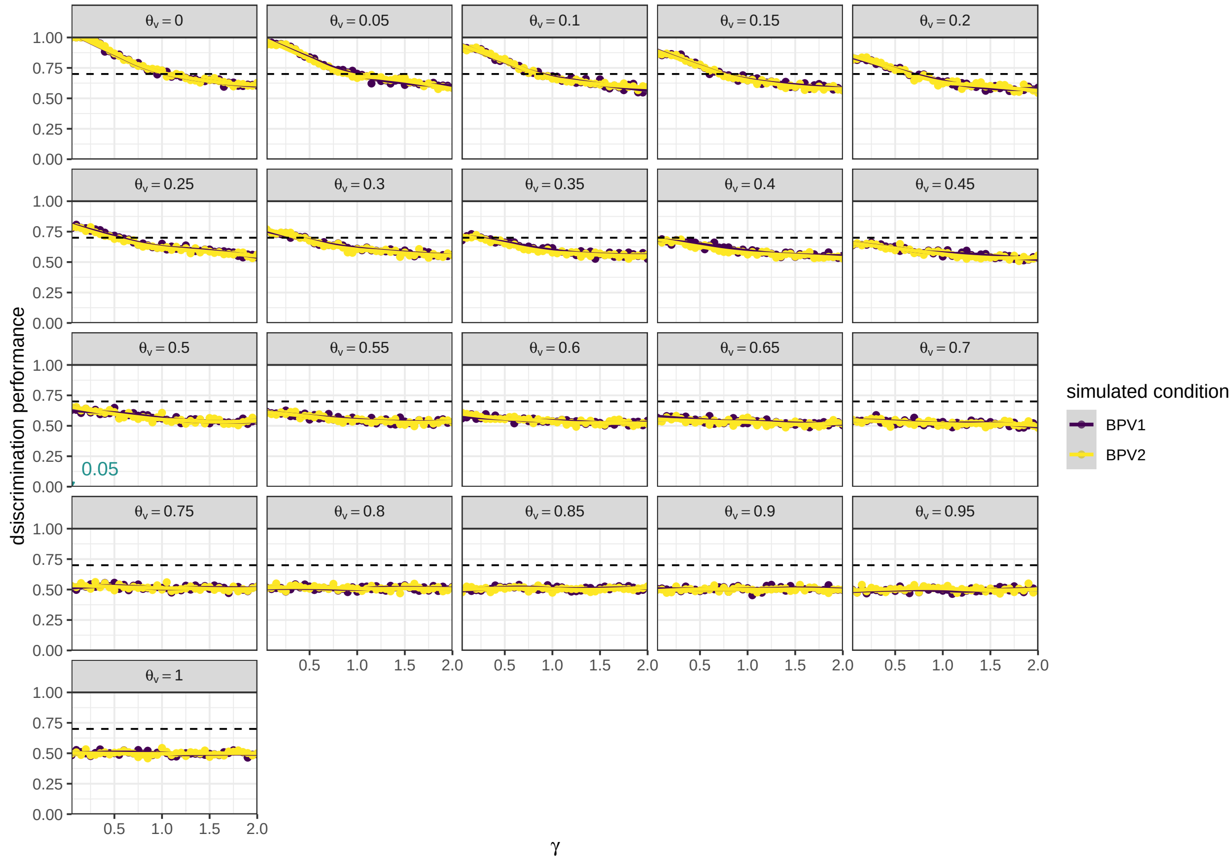

Figure A2: **Sensitivity tests for the randomly noncompensatory model.** Same notation as in Fig. A1. The different panels give the different values of the probability with which the volume dimension was chosen (θ_v). For a non-biased randomly noncompensatory model we set $\theta_v = 0.5$.

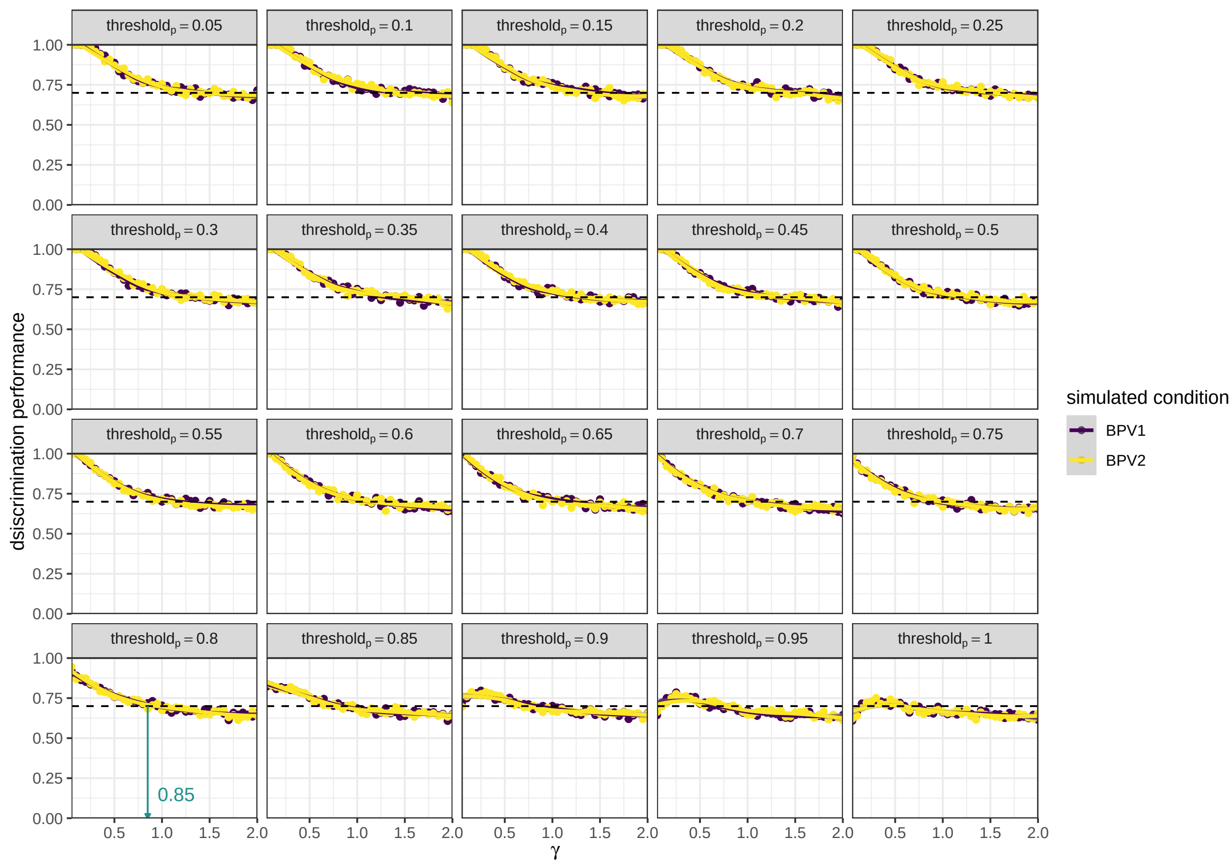

Figure A3: **Sensitivity tests for the probability first model.** Same notation as in Fig. A1. The different panels give the different values of the salience threshold that needed to be reached for one option to be preferred over the other. We set the value of the threshold for both the volume and probability dimensions to 0.8, based on the psychometric function threshold for probability (Rivalan, Winter, and Nachev 2017).

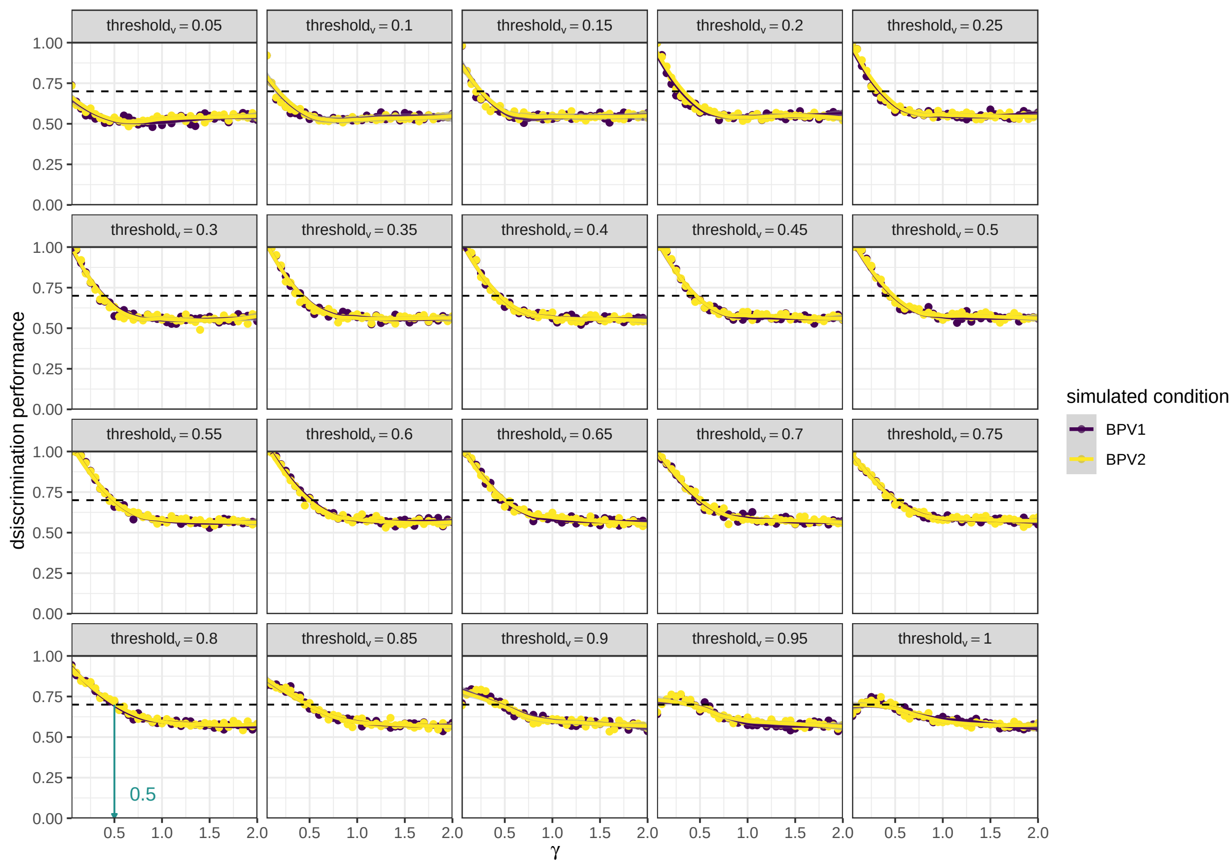

Figure A4: **Sensitivity tests for the volume first model.** Same notation as in Fig. A1. The different panels give the different values of the salience threshold that needed to be reached for one option to be preferred over the other. We set the value of the threshold for both the volume and probability dimensions to 0.8, based on the psychometric function threshold for probability (Rivalan, Winter, and Nachev 2017).

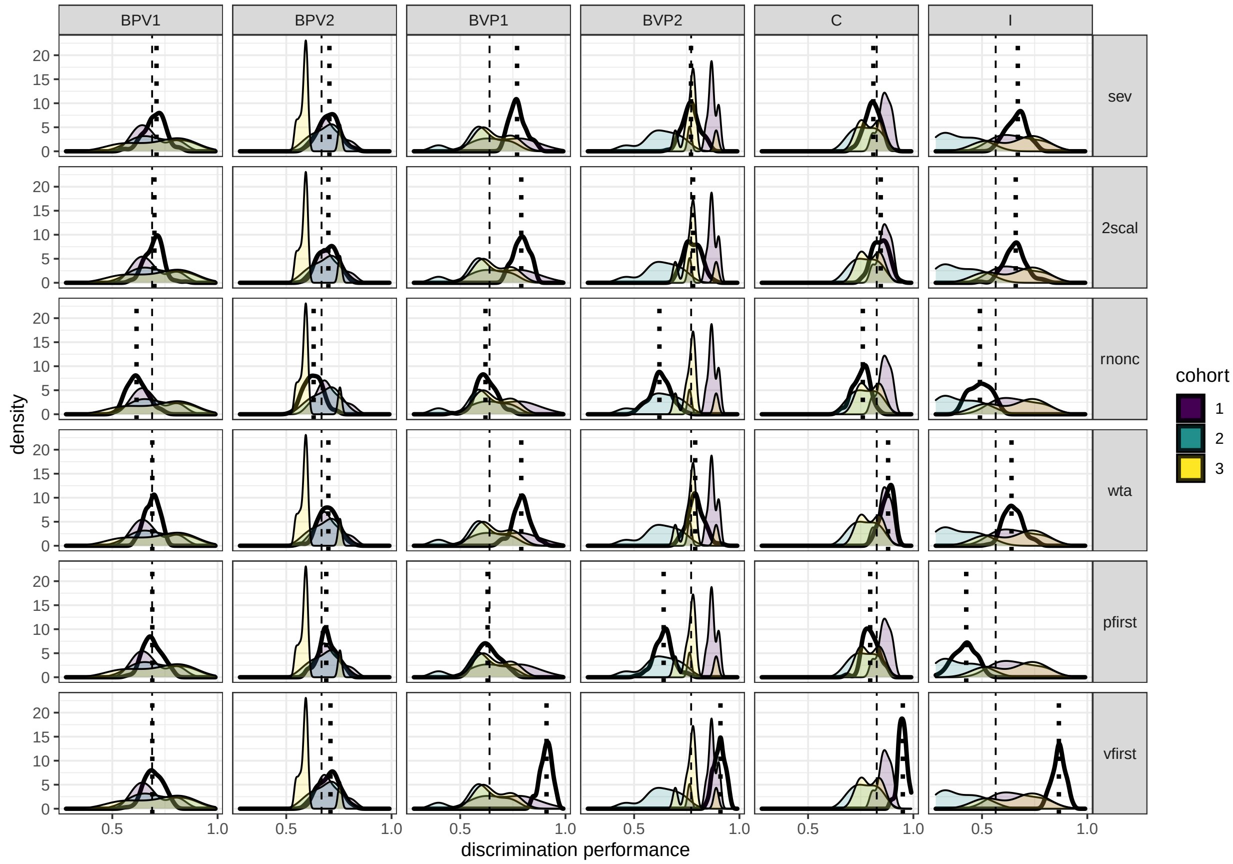

Figure A5: **Comparison of discrimination performance in all six simulation models and in the three mouse cohorts in Experiment 1**. Columns give the condition names (Table 1) and rows, the model number (Table 2). Empirical data from the three cohorts are represented by differently color-filled density curves from the observed discrimination performances. Simulation data are represented by an empty thick-lined density curve. The dashed line gives the median of the empirical data and the dotted line - the median of the simulated data. The discrimination performance gives the relative visitation rate of the more profitable option, or, in the incongruent condition, the option with the higher volume.

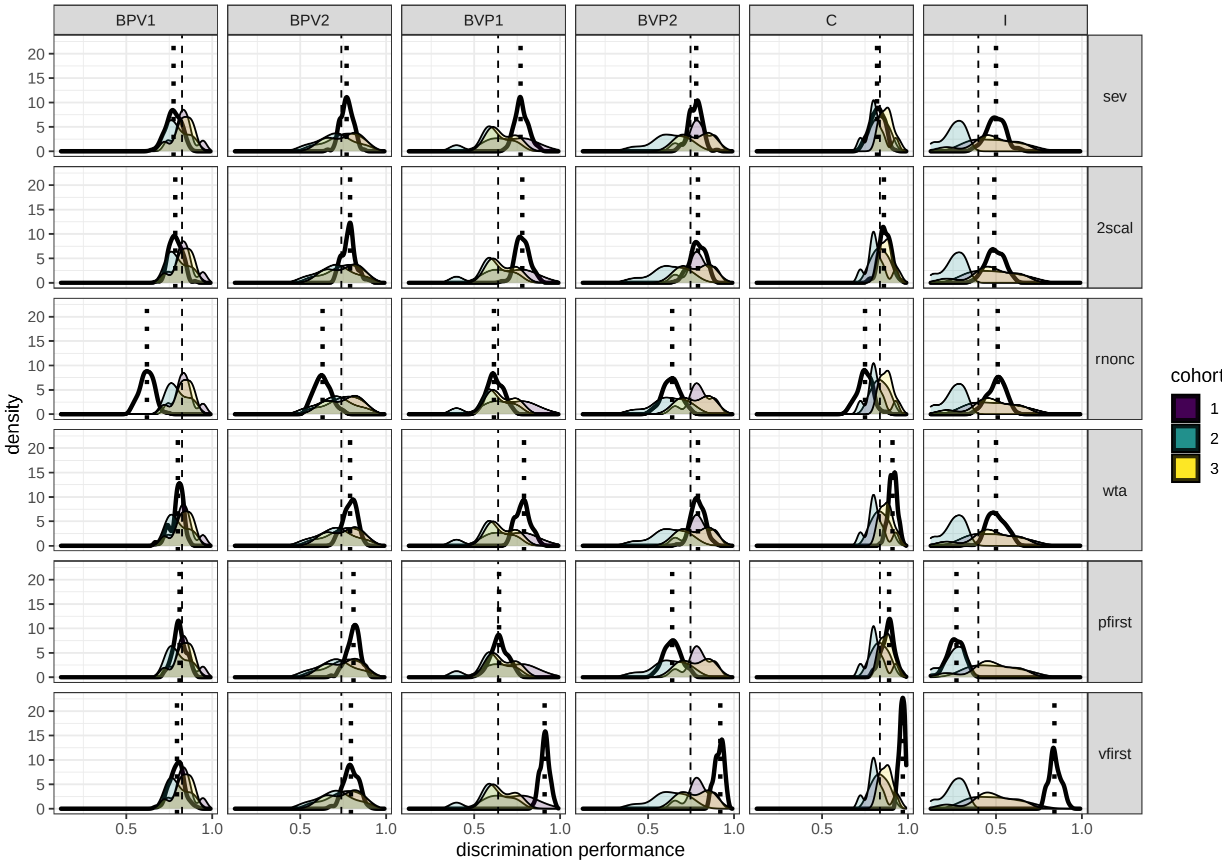

Figure A6: **Comparison of discrimination performance in all six simulation models and in the three mouse cohorts in Experiment 2**. Same notation as in Fig. A5.

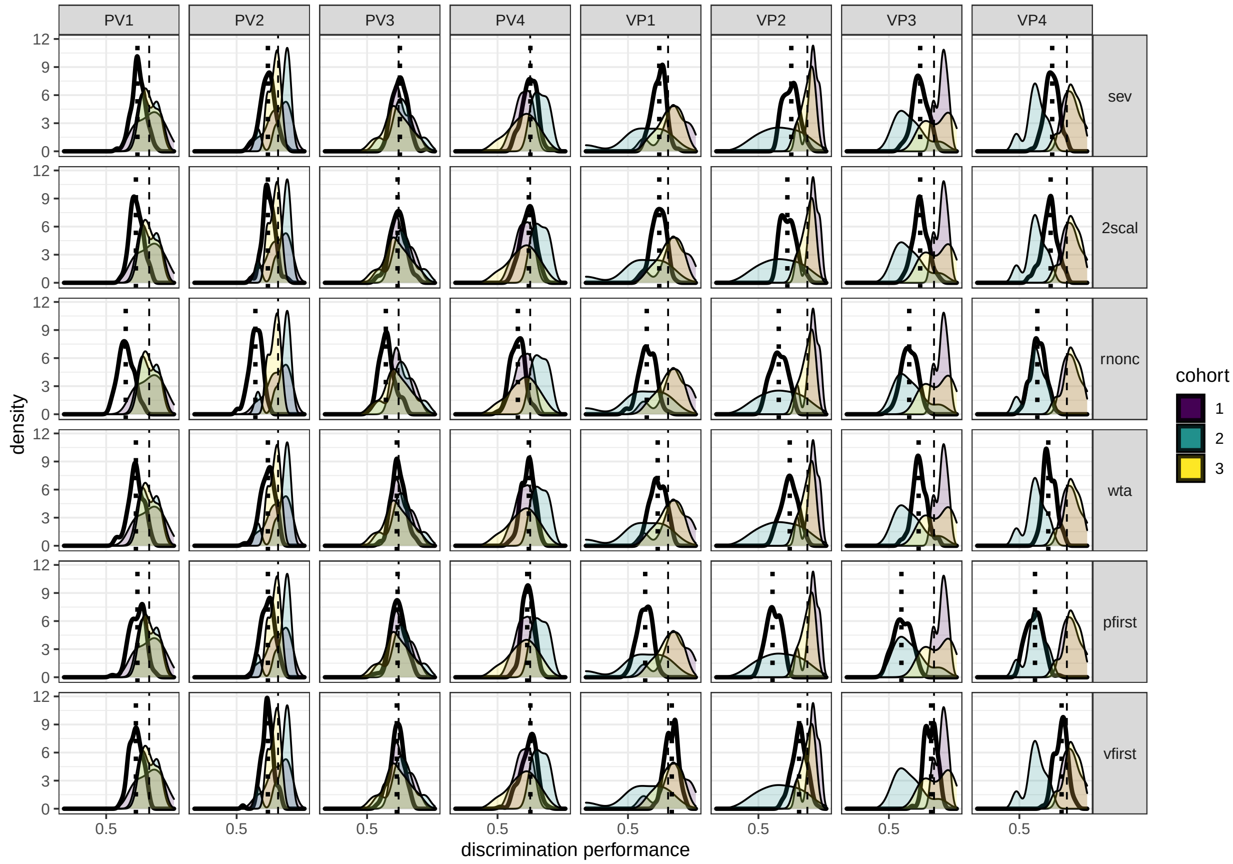

Figure A7: **Comparison of discrimination performance in all six simulation models and in the three mouse cohorts in Experiment 3**. Same notation as in Fig. A5.

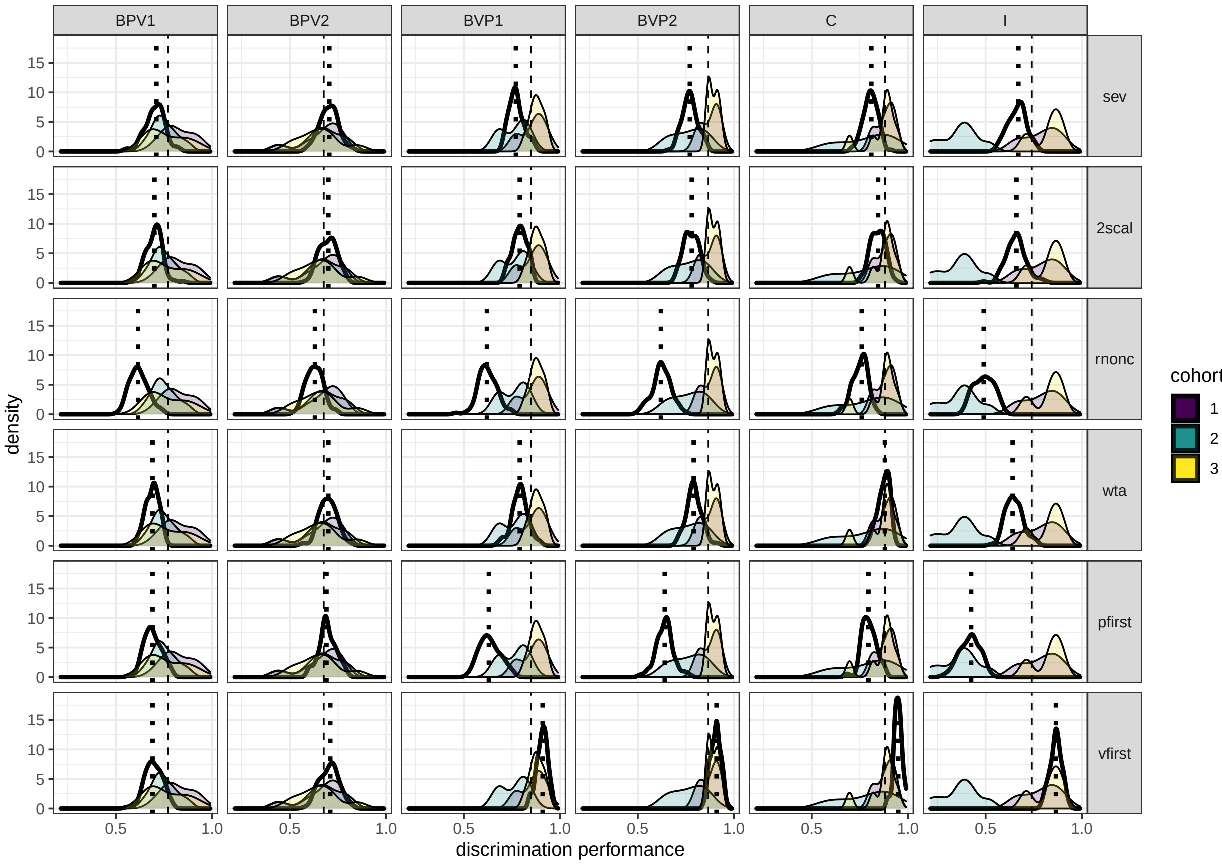

Figure A8: **Comparison of discrimination performance in all six simulation models and in the three mouse cohorts in Experiment 4**. Same notation as in Fig. A5.

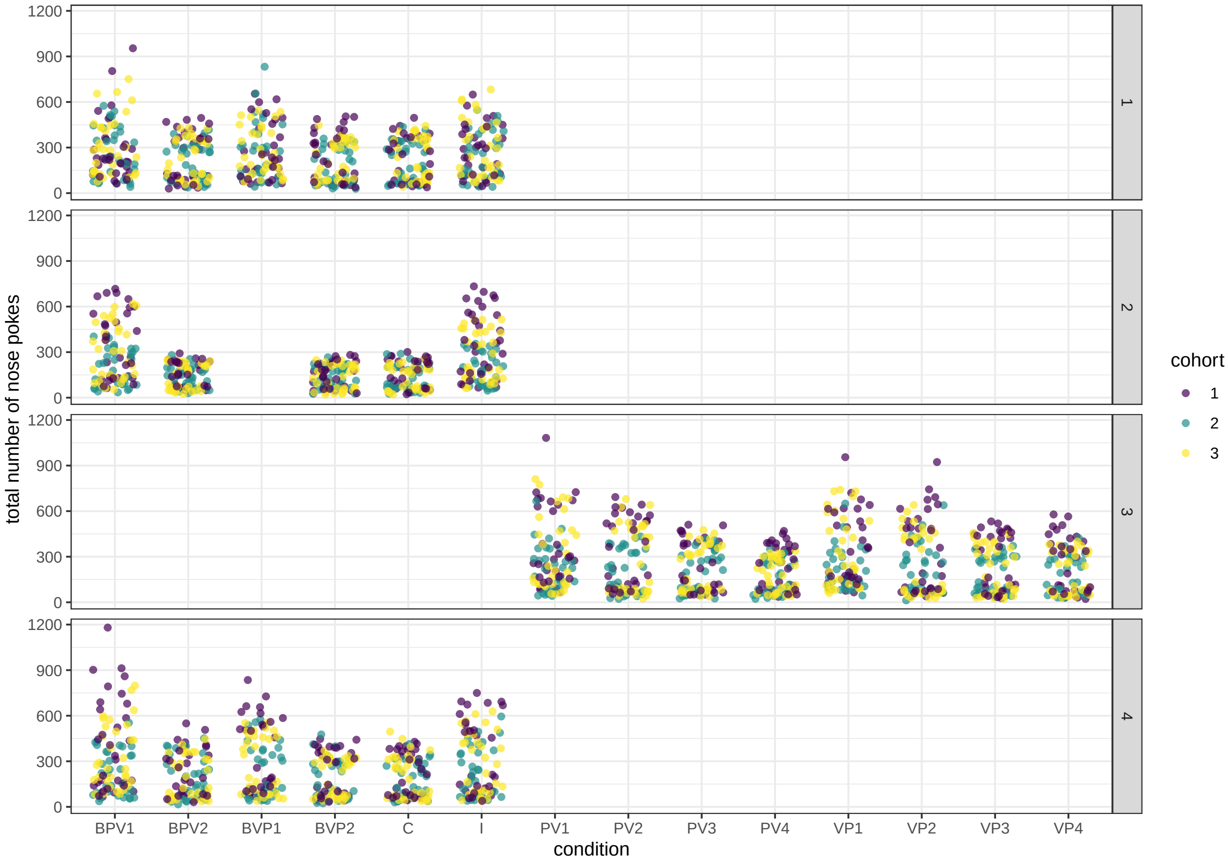

Figure A9: **Total number of nose pokes for each experimental condition in the three cohorts in all experiments.** Rows show different experiments (1-4). Each symbol represents the total number of nose pokes for a single mouse over one of the two experimental days of the given condition.

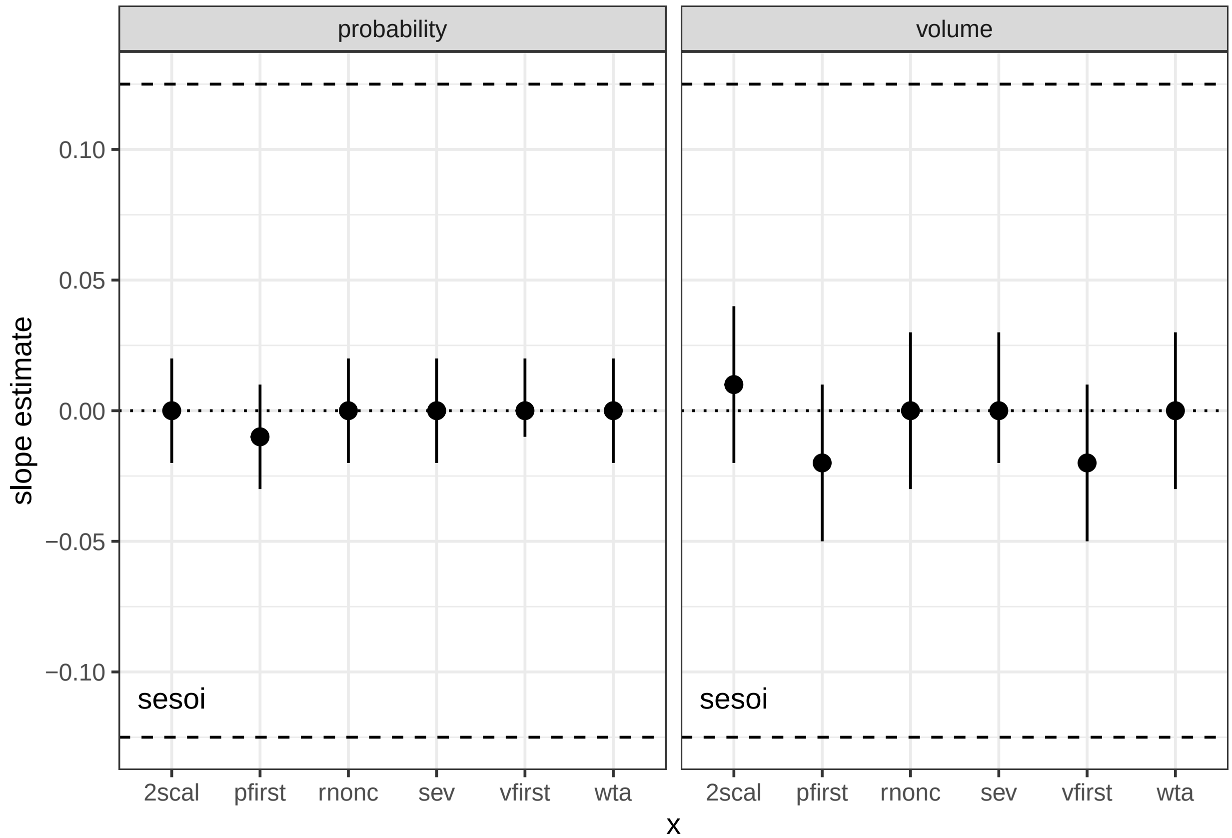

Figure A10: **Slope estimates for the effect of the background dimension on the discrimination performance in the relevant dimension for different decision models**. The two choice options always differed along the relevant dimension (either probability or volume) at a fixed relative intensity. The discrimination performance for 100 virtual mice making 100 decisions each was measured at four different levels of the background dimension. Symbols and whiskers give means and 98% confidence intervals estimated from bootstraps. The smallest effect size of interest (dashed lines) was determined to be the slope that would have resulted in a difference in discrimination performance of 0.1, from the lowest to the highest level of the background dimension. Compare to Fig. 4.

Acknowledgments

We thank Miléna Brunet, Alexia Hyde, and Sabine Wintergerst for data acquisition, Katja Frei for assistance with the mice, Alexej Schatz for programming of the control software, and our colleagues of the Winter lab for a fruitful discussion.

Authorship and contribution

V.N. Conceptualization, Methodology, Software, Formal Analysis, Data curation, Writing—original draft, Writing—review and editing, Visualization, Supervision, Project Administration.
M.R. Methodology, Writing—review and editing, Supervision.
Y.W. Resources, Methodology, Writing—review and editing, Supervision.

Competing interests

The authors declare that they have no competing interests.

References

Akre, Karin L., and Sönke Johnsen. 2014. "Psychophysics and the Evolution of Behavior." *Trends in Ecology & Evolution* 29 (5): 291–300. https://doi.org/10.1016/j.tree.2014.03.007.

Bateson, Melissa, Susan D Healy, and T Andrew Hurly. 2003. "Context-Dependent Foraging Decisions in Rufous Hummingbirds." *Proceedings of the Royal Society B: Biological Sciences* 270 (1521): 1271–6. https://doi.org/10.1098/rspb.2003.2365.

Brandstätter, Eduard, Gerd Gigerenzer, and Ralph Hertwig. 2006. "The Priority Heuristic: Making Choices Without Trade-Offs." *Psychological Review* 113 (2): 409–32. https://doi.org/10.1037/0033-295X.113.2.409.

Caffo, Brian, Carolyn Lauzon, and Joachim Röhmel. 2013. "Correction to 'Easy Multiplicity Control in Equivalence Testing Using Two One-Sided Tests'." *The American Statistician* 67 (2): 115–16. https://doi.org/10.1080/00031305.2012.760487.

Constantinople, Christine M., Alex T. Piet, and Carlos D. Brody. 2019. "An Analysis of Decision Under Risk in Rats." *Current Biology* 29 (12): 2066–2074.e5. https://doi.org/10.1016/j.cub.2019.05.013.

Davidson, Heather, and Robert A. Cribbie. 2019. "A More Powerful Familywise Error Control Procedure for Evaluating Mean Equivalence." *Communications in Statistics - Simulation and Computation* 0 (0): 1–16. https://doi.org/10.1080/03610918.2018.1530783.

Harrell, Frank E., and Charles Dupont. 2019. "R Package Hmisc: Harrell Miscellaneous." https://CRAN.R-project.org/package=Hmisc.

Haupt, Moritz, Jana A. Eccard, and York Winter. 2010. "Does Spatial Learning Ability of Common Voles (*Microtus Arvalis*) and Bank Voles (*Myodes Glareolus*) Constrain Foraging Efficiency?" *Animal Cognition* 13 (6): 783–91. https://doi.org/10.1007/s10071-010-0327-8.

Houston, Alasdair I., John M. McNamara, and Mark D. Steer. 2007. "Violations of Transitivity Under Fitness Maximization." *Biology Letters* 3 (4): 365–67. https://doi.org/10.1098/rsbl.2007.0111.

Hunt, Laurence T., Raymond J. Dolan, and Timothy E. J. Behrens. 2014. "Hierarchical Competitions Subserving Multi-Attribute Choice." *Nature Neuroscience* 17 (11): 1613–22. https://doi.org/10.1038/nn.3836.

Jansen, Brenda R. J., Anna C. K. van Duijvenvoorde, and Hilde M. Huizenga. 2012. "Development of Decision Making: Sequential Versus Integrative Rules." *Journal of Experimental Child Psychology* 111 (1): 87–100. https://doi.org/10.1016/j.jecp.2011.07.006.

Kacelnik, Alejandro. 1984. "Central Place Foraging in Starlings (*Sturnus Vulgaris*). I. Patch Residence Time." *Journal of Animal Ecology* 53 (1): 283–99. https://doi.org/10.2307/4357.

Kacelnik, Alex. 2006. "Meanings of Rationality." In *Rational Animals?*, edited by Susan L. Hurley and Matthew Nudds. Oxford University Press.

Kacelnik, Alex, and Fausto Brito e Abreu. 1998. "Risky Choice and Weber's Law." *Journal of Theoretical Biology* 194 (2): 289–98. https://doi.org/10.1006/jtbi.1998.0763.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica* 47 (2): 263–91. https://doi.org/10.2307/1914185.

Katsikopoulos, Konstantinos V., and Gerd Gigerenzer. 2008. "One-Reason Decision-Making: Modeling Violations of Expected Utility Theory." *Journal of Risk and Uncertainty* 37 (1): 35–56. https://doi.org/10.1007/s11166-008-9042-0.

Kenrick, Douglas T., Vladas Griskevicius, Jill M. Sundie, Norman P. Li, Yexin Jessica Li, and Steven L. Neuberg. 2009. "Deep Rationality: The Evolutionary Economics of Decision Making." *Social Cognition* 27 (5): 764–85. https://doi.org/10.1521/soco.2009.27.5.764.

Lakens, Daniël. 2017. "Equivalence Tests." *Social Psychological and Personality Science* 8 (4): 355–62. https://doi.org/10.1177/1948550617697177.

Lauzon, Carolyn, and Brian Caffo. 2009. "Easy Multiplicity Control in Equivalence Testing Using Two One-Sided Tests." *The American Statistician* 63 (2): 147–54. https://doi.org/10.1198/tast.2009.0029.

Levy, Dino J, and Paul W Glimcher. 2012. "The Root of All Value: A Neural Common Currency for Choice." *Current Opinion in Neurobiology*, Decision making, 22 (6): 1027–38. https://doi.org/10.1016/j.conb.2012.06.001.

Marsh, Barnaby, and Alex Kacelnik. 2002. "Framing Effects and Risky Decisions in Starlings." *Proceedings of the National Academy of Sciences* 99 (5): 3352–5. https://doi.org/10.1073/pnas.042491999.

McNamara, J. M., P. C. Trimmer, and A. I. Houston. 2014. "Natural Selection Can Favour 'Irrational' Behaviour." *Biology Letters* 10 (1): 20130935. https://doi.org/10.1098/rsbl.2013.0935.

Monteiro, Tiago, Marco Vasconcelos, and Alex Kacelnik. 2013. "Starlings Uphold Principles of Economic Rationality for Delay and Probability of Reward." *Proceedings of the Royal Society B: Biological Sciences* 280 (1756): 20122386. https://doi.org/10.1098/rspb.2012.2386.

Nachev, Vladislav, Kai Petra Stich, Clemens Winter, Alan Bond, Alan Kamil, and York Winter. 2017. "Cognition-Mediated Evolution of Low-Quality Floral Nectars." *Science* 355 (6320): 75–78. https://doi.org/10.1126/science.aah4219.

Nachev, Vladislav, Kai Petra Stich, and York Winter. 2013. "Weber's Law, the Magnitude Effect and Discrimination of Sugar Concentrations in Nectar-Feeding Animals." Edited by Jacob Engelmann. *PLoS ONE* 8 (9): e74144. https://doi.org/10.1371/journal.pone.0074144.

Nachev, Vladislav, and York Winter. 2012. "The Psychophysics of Uneconomical Choice: Non-Linear Reward Evaluation by a Nectar Feeder." *Animal Cognition* 15: 393–400. https://doi.org/10.1007/s10071-011-0465-7.

———. 2019. "Behavioral Repeatability and Choice Performance in Wild Free-Flying Nectarivorous Bats (*Glossophaga Commissarisi*)." *Behavioral Ecology and Sociobiology* 73 (2): 24. https://doi.org/10.1007/s00265-019-2637-4.

Ojeda, Andrés, Robin A. Murphy, and Alex Kacelnik. 2018. "Paradoxical Choice in Rats: Subjective Valuation and Mechanism of Choice." *Behavioural Processes*, Quantitative and Comparative Analyses of Behavior, 152 (July): 73–80. https://doi.org/10.1016/j.beproc.2018.03.024.

Pitz, Gordon F., and Natalie J. Sachs. 1984. "Judgment and Decision: Theory and Application." *Annual Review of Psychology* 35 (1): 139–64. https://doi.org/10.1146/annurev.ps.35.020184.001035.

Reid, Chris R., Simon Garnier, Madeleine Beekman, and Tanya Latty. 2015. "Information Integration and Multiattribute Decision Making in Non-Neuronal Organisms." *Animal Behaviour* 100 (February): 44–50. https://doi.org/10.1016/j.anbehav.2014.11.010.

Rieskamp, Jörg, Jerome R. Busemeyer, and Barbara A. Mellers. 2006. "Extending the Bounds of Rationality: Evidence and Theories of Preferential Choice." *Journal of Economic Literature* 44 (3): 631–61. https://doi.org/10.1257/jel.44.3.631.

Rivalan, Marion, York Winter, and Vladislav Nachev. 2017. "Principles of Economic Rationality in Mice." *Scientific Reports* 7 (1): 17441. https://doi.org/10.1038/s41598-017-17747-7.

Rosenström, Tom, Karoline Wiesner, and Alasdair I. Houston. 2016. "Scalar Utility Theory and Proportional Processing: What Does It Actually Imply?" *Journal of Theoretical Biology* 404 (September): 222–35. https://doi.org/10.1016/j.jtbi.2016.06.003.

Scarpi, Daniele. 2011. "The Impact of Phantom Decoys on Choices in Cats." *Animal Cognition* 14 (1): 127–36. https://doi.org/10.1007/s10071-010-0350-9.

Schuck-Paim, Cynthia, Lorena Pompilio, and Alex Kacelnik. 2004. "State-Dependent Decisions Cause Apparent Violations of Rationality in Animal Choice." *PLOS Biology* 2 (12): e402. https://doi.org/10.1371/journal.pbio.0020402.

Shafir, Sharoni. 1994. "Intransitivity of Preferences in Honey Bees: Support for 'Comparative' Evaluation of Foraging Options." *Animal Behaviour* 48 (1): 55–67. https://doi.org/10.1006/anbe.1994.1211.

Shafir, Sharoni, Tom Waite, and Brian Smith. 2002. "Context-Dependent Violations of Rational Choice in Honeybees (*Apis Mellifera*) and Gray Jays (*Perisoreus Canadensis*)." *Behavioral Ecology and Sociobiology* 51 (2): 180–87. https://doi.org/10.1007/s00265-001-0420-8.

Shafir, Sharoni, and Lia Yehonatan. 2014. "Comparative Evaluations of Reward Dimensions in Honey Bees: Evidence from Two-Alternative Forced Choice Proboscis-Extension Conditioning." *Animal Cognition* 17 (3): 633–44. https://doi.org/10.1007/s10071-013-0694-z.

Sih, Andrew, and Marco Del Giudice. 2012. "Linking Behavioural Syndromes and Cognition: A Behavioural Ecology Perspective." *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1603): 2762–72. https://doi.org/10.1098/rstb.2012.0216.

Team, R Development Core. 2019. "R: A Language and Environment for Statistical Computing." Vienna: R Foundation for Statistical Computing. http://www.R-project.org.

Trimmer, Pete C. 2013. "Optimal Behaviour Can Violate the Principle of Regularity." *Proceedings of the Royal Society of London B: Biological Sciences* 280 (1763): 20130858. https://doi.org/10.1098/rspb.2013.0858.

Tversky, Amos. 1969. "Intransitivity of Preferences." *Psychological Review* 76 (1): 31–48. https://doi.org/10.1037/h0026750.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment Under Uncertainty: Heuristics and Biases." *Science* 185 (4157): 1124–31. https://doi.org/10.1126/science.185.4157.1124.

Vora, Siddharth R., Esra D. Camci, and Timothy C. Cox. 2016. "Postnatal Ontogeny of the Cranial Base and Craniofacial Skeleton in Male C57BL/6J Mice: A Reference Standard for Quantitative Analysis." *Frontiers in Physiology* 6. https://doi.org/10.3389/fphys.2015.00417.

Zocchi, Dhruv, Gunther Wennemuth, and Yuki Oka. 2017. "The Cellular Mechanism for Water Detection in the Mammalian Taste System." *Nature Neuroscience* 20 (7): 927–33. https://doi.org/10.1038/nn.4575.