

Behavioral/Cognitive

Stimulus Context and Reward Contingency Induce Behavioral Adaptation in a Rodent Tactile Detection Task

Christian Waiblinger, Caroline M. Wu, Michael F. Bolus, Peter Y. Borden, and Garrett B. Stanley

Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, Georgia 30332

Behavioral adaptation is a prerequisite for survival in a constantly changing sensory environment, but the underlying strategies and relevant variables driving adaptive behavior are not well understood. Many learning models and neural theories consider probabilistic computations as an efficient way to solve a variety of tasks, especially if uncertainty is involved. Although this suggests a possible role for probabilistic inference and expectation in adaptive behaviors, there is little if any evidence of this relationship experimentally. Here, we investigated adaptive behavior in the rat model by using a well controlled behavioral paradigm within a psychophysical framework to predict and quantify changes in performance of animals trained on a simple whisker-based detection task. The sensory environment of the task was changed by transforming the probabilistic distribution of whisker deflection amplitudes systematically while measuring the animal's detection performance and corresponding rate of accumulated reward. We show that the psychometric function deviates significantly and reversibly depending on the probabilistic distribution of stimuli. This change in performance relates to accumulating a constant reward count across trials, yet it is exempt from changes in reward volume. Our simple model of reward accumulation captures the observed change in psychometric sensitivity and predicts a strategy seeking to maintain reward expectation across trials in the face of the changing stimulus distribution. We conclude that rats are able maintain a constant payoff under changing sensory conditions by flexibly adjusting their behavioral strategy. Our findings suggest the existence of an internal probabilistic model that facilitates behavioral adaptation when sensory demands change.

Key words: adaptive; behavior; decision making; psychophysics; rodent; whisker

Significance Statement

The strategy animals use to deal with a complex and ever-changing world is a key to understanding natural behavior. This study provides evidence that rodent behavioral performance is highly flexible in the face of a changing stimulus distribution, consistent with a strategy to maintain a desired accumulation of reward.

Introduction

Our decisions in everyday situations are governed by sensory stimuli but also, and most importantly, by our experience and by a constantly changing sensory environment. How do we adapt to such contextual changes? What is our strategy to deal with a dynamically changing environment? In the last decade, there have been major advances toward answering this question by studying perceptual decision making across species and asking how and to what degree neuronal activity reflects behavioral choice and how sensory information is transformed into adaptive

action (Romo and Salinas, 2003; Gold and Shadlen, 2007; Nienborg and Cumming, 2009). An important finding is that behavioral performance in psychophysical experiments is often not determined by sensory processes alone but by a wide range of biasing factors, among them prior probabilities, reward payoff, changing stimulus–action associations, and trial history (Busse et al., 2011; Stüttgen et al., 2013; Jaramillo et al., 2014; Akrami et al., 2018; Waiblinger et al., 2018), potentially pointing to a complex, dynamic behavioral strategy reliant on a continuous interplay between these factors. Claims about neuronal coding schemes therefore depend crucially on behavioral context and a precise assessment of the observer's active task strategy. By considering these dynamic cognitive processes, much of the variability observed in behavior could be much more predictable compared with using simplified psychophysical approaches that do not take the contextual elements into account. Although there exist a variety of useful computational models to explain choice biases or contextual dependencies in human and animal psychophysical datasets (Nassar et al., 2010; Fründ et al., 2011, 2014; Wilson et al., 2013; Braun et al., 2018), a definitive understanding of behavior within a dynamic framework does not yet exist.

Here, we directly investigate behavioral strategies in the face of a dynamically changing sensory environment, using the rodent

Received Aug. 8, 2018; revised Oct. 30, 2018; accepted Nov. 21, 2018.

Author contributions: C.W. wrote the first draft of the paper; C.M.W., M.F.B., P.Y.B., and G.B.S. edited the paper; C.W. and G.B.S. designed research; C.W. and C.M.W. performed research; M.F.B. and P.Y.B. contributed unpublished reagents/analytic tools; C.W., M.F.B., and P.Y.B. analyzed data; C.W. wrote the paper.

C.W. was supported by a fellowship from the Deutsche Forschungsgemeinschaft (GZ: WA 3862/1-1) and the National Institutes of Health (Grant U01NS094302). P.Y.B. was supported by an NIH National Research Service Award (Predoctoral Fellowship F31NS09869). M.F.B. was supported by the National Science Foundation (Graduate Research Fellowship Grant DGE-1650044). C.M.W. was supported by an Undergraduate Research Scholarship from the Georgia Tech Petit Institute. G.B.S. was supported by the NIH (Grant R01NS085447 and Brain Grants R01NS104928 and U01NS094302).

The authors declare no competing financial interests.

Correspondence should be addressed to Garrett B. Stanley at garrett.stanley@bme.gatech.edu.

<https://doi.org/10.1523/JNEUROSCI.2032-18.2018>

Copyright © 2019 the authors 0270-6474/19/391088-12\$15.00/0

vibrissa pathway as a model system. Specifically, we challenge head-fixed rats in a Go/No-Go detection task with a variable whisker deflection amplitude drawn from a controlled but probabilistic distribution of amplitudes, which varies across sessions. In response to changing stimulus distribution, we observe a dynamic behavioral strategy reflected in a shift in psychometric sensitivity that is qualitatively consistent with the animal maintaining the expected accumulated reward in the face of a changing environment. We quantitatively challenge this hypothesis within a simple model framework of reward accumulation to predict and capture systematic changes in the performance of expert rats in a range of experiments with changing stimulus and reward contingencies. The observed change in psychometric sensitivity is largely predicted by a strategy seeking to maintain reward expectation in the face of a changing stimulus distribution and reversibly persists in the face of a range of transformations of the stimulus distribution. Finally, analysis of psychometric sensitivity within sessions reveals an asymmetric shift in sensitivity on even shorter time scales, suggestive of a dynamic strategy based on an internal model of reward accumulation. A direct manipulation of the reward feedback thereby reveals that the number of past rewarded trials not the volume corresponds with task engagement. Together, the results here provide a simple predictive framework for adaptive reward-based behaviors in a changing sensory environment.

Materials and Methods

Animals, surgery, and general procedures for behavioral testing

All experimental and surgical procedures were approved by the Georgia Institute of Technology Institutional Animal Care and Use Committee and were in agreement with guidelines established by the National Institutes of Health. Subjects were seven female Sprague Dawley rats aged 12–16 weeks at time of implantation. The basic procedures of head-cap surgery, habituation to head fixation, and behavioral training exactly followed the ones published in a technical review (Schwarz et al., 2010) and more recent studies (Waiblinger et al., 2015, 2018). In the following text, only procedures pertaining to the specific paradigm established here are described in detail.

Oral antibiotics (Baytril; Bayer injectable solution 2.27%, 20 ml) were provided for 1 d before surgery and 1 week postoperatively. The animals were anesthetized using isoflurane and a head cap for head fixation was implanted. The wound was treated with antibiotic ointment and sutured. Analgesia and warmth were provided after surgery. Rats were allowed to recover for at least 10 d before habituation training. Subjects were housed together with a maximum number of two in one group cage and kept under a 12/12 h inverted light/dark cycle. During successive days of behavioral testing, water intake was restricted to the experimental sessions in which animals were given the opportunity to earn water to satiety. Testing was paused and water was available *ad libitum* for 2 d a week. Body weight was monitored daily and was typically observed to increase during training. If the body weight dropped for more than ~10 g due to a higher task difficulty, supplementary water was delivered outside of training sessions to keep the animal's weight up. The first step of behavioral training was systematic habituation to head-fixation lasting for ~2 weeks. Once rats were trained on the behavioral task, one or two experiments were usually conducted per day, comprising 100–200 trials. During behavioral testing, a constant auditory white background noise (70 dB) was produced by an arbitrary waveform generator to mask any sound emission of the galvo-motor-based whisker actuator.

Whisker stimulation

For whisker stimulation, a galvo-motor (galvanometer optical scanner model 6210H; Cambridge Technology) as described in Chagas et al. (2013) was used. The rotating arm of the galvo-motor contacted a single whisker on the right of the rat's face at 5 mm (± 1 mm tolerance) distance from the skin and thus directly engaged the proximal whisker shaft, largely overriding bioelastic whisker properties. All of the remaining

whiskers were trimmed to prevent them from being touched by the rotating arm. Voltage commands for the actuator were programmed in MATLAB and Simulink (version 2015b; The MathWorks). A stimulus consisted of a single event: a sinusoidal pulse (half period of a 100 Hz sine wave, starting at one minimum and ending at the next maximum). The pulse amplitudes used ($A = [0, 0.25, 1, 4, 16]^\circ$ or maximal velocities, respectively: $V = [0, 78.49, 313.95, 1255.81, 5023.24]^\circ/\text{s}$) were well within the range reported for frictional slips observed in natural whisker movement (Ritt et al., 2008; Wolfe et al., 2008).

Behavioral paradigm and training

All seven rats were trained on a standard Go/No-Go detection task using a similar protocol as described before (Stüttgen et al., 2006; Ollerenshaw et al., 2012, 2014). In this task, the whisker is deflected at intervals of 4–10 s (flat probability distribution) with a single pulse (detection target). A trial was categorized a “hit” if the animal generated the “Go” indicator response, a lick at a water spout within 1000 ms of target onset. If no lick was emitted, the trial counted as a “miss.” In addition, catch trials were included in which no deflection of the whisker occurred ($A = 0^\circ$) and a trial was categorized as a correct rejection if licking was withheld (No-Go). However, a trial was categorized as a false alarm if random licks occurred within 1000 ms of catch onset. Premature licking in a 2 s period before the stimulus was mildly punished by resetting time (time-out) and starting a new intertrial interval of 4–10 s duration drawn at random from a flat probability distribution. Note that these trial types were excluded from the main psychometric data analysis because they can occur with a different likelihood as a particular stimulus or catch trial. However, we report these trial types separately and refer to them as “impulsive licks” in the rest of the manuscript.

During the first training phase, a single pulse with fixed amplitude was presented interspersed by catch trials ($P_{\text{stim}} = 0.8, P_{\text{catch}} = 0.2$). Immediately following stimulus offset, a droplet of water became available at the waterspout to condition the animal's lick response thereby shaping the stimulus–reward association. Once subjects showed stable and immediate consumption behavior (usually within a few sessions), water was only delivered after an indicator lick at the spout within 1000 ms, turning the task into an operant conditioning paradigm in which the response is only reinforced by reward if it is correctly emitted after the stimulus.

To assess differences in learning based on stimulus amplitude at this stage of training, we separated animals into two groups: Group 1 (rats 1–3) receiving only high-amplitude stimuli ($A = [0\ 16]^\circ$) and Group 2 (rats 4 and 5) receiving only small-amplitude stimuli ($A = [0\ 4]^\circ$). To assess learning, training was conducted without manual interference by the experimenter and with equal conditions across sessions from here on. The learning curve consisted of detection indices across training sessions each calculated by subtracting the false alarm rate from the hit rate for a given session ($D_i = P_{\text{HIT}} - P_{\text{FA}}$). A criterion of 0.75 was used to determine expert level. After expert level was achieved, the psychometric curve was measured using the method of constant stimuli, which entails the presentation of repeated stimulus blocks containing multiple stimulus amplitudes. The sequence of stimuli within a block was pseudo-random (i.e., each stimulus and catch was presented once in shuffled order and then repeated). The experimenter aborted a session when the animal stopped working on the task; that is, when it did not generate lick responses for an entire stimulus block, including the highest-amplitude stimulus.

The main experiments of this study were performed using different stimulus and reward conditions in a within-subject design. Each experimental condition consisted of 8–10 sessions performed over ~5 d. Each animal performed 1–2 sessions per day with a minimal break of 3 h in between. Stimulus or reward parameters changed once an animal had completed an experimental condition. The different animal groups and experimental conditions are described in detail in the following section (see also Fig. 1).

Experiment 1. Once our first group of rats (rats 1–3) had learned the task, psychometric tests were conducted with a manipulation of the stimulus distribution range. In a given session, the sequence of trials consisted of repeated stimulus blocks, each block containing multiple stimuli and catch trial once in shuffled order. Therefore, each stimulus occurred with

the same probability (i.e., a uniform distribution). In the “high-range” condition, three stimulus amplitudes plus catch trial were used ($A = [0, 1, 4, 16]^\circ$) and presented in multiple successive sessions (sessions 1–8). Following this (sessions 9–16), three new stimulus amplitudes were presented ($A = [0, 0.25, 1, 4]^\circ$), forming the low-range condition. Both stimulus distributions shared two of the three stimulus amplitudes; however, the largest stimulus amplitude of the high-range condition (16°) was not part of the low range and vice versa, the smallest amplitude of the low-range condition (0.25°) was not part of the high range. To test the reversibility of potential behavioral changes, another high-range condition was presented at the end (sessions 17–24). For the second group of rats (rats 4 and 5), the order of experiments was reversed; that is, the first psychometric curve was measured by presenting the low-range condition first ($A = [0, 0.25, 1, 4]^\circ$) before switching to the high-range condition ($A = [0, 1, 4, 16]^\circ$) and then again back to the low range. A third group of rats (rats 6 and 7) served as a control group and underwent experiments without any changes to the stimulus distribution range; that is, the psychometric curves were measured with exactly the same set of stimuli throughout (sessions 1–24).

Experiment 2. In the next experiment, the same animal groups were used as before, but this time the stimulus distribution range was fixed while the relative probabilities were changed (i.e., no longer uniform). Now, all four amplitudes plus a catch trial ($A = [0, 0.25, 1, 4, 16]^\circ$) were used for both the “big” and “small” conditions, but in the big condition, the two stimuli with large amplitudes were presented with a higher probability than the two small amplitudes ($P_{\text{big}} > P_{\text{small}}$; $P[4, 16]^\circ = 0.36$, $P[0.25, 1]^\circ = 0.09$). In the small condition, the order of stimulus probabilities was reversed; that is, the two stimuli with small amplitudes were presented with a higher probability than the two large amplitudes ($P_{\text{big}} < P_{\text{small}}$; $P[4, 16]^\circ = 0.09$, $P[0.25, 1]^\circ = 0.36$). The probability of a catch trial always remained the same for all conditions ($P_{\text{catch}} = 0.1$). The sequence of manipulations was the same as in Experiment 1: rats 1–3 started with the big condition (sessions 1–8) and then underwent shifts to small (sessions 9–17) and back to the big condition (sessions 17–24), whereas the order of manipulations was reversed for rats 4–5. In all of the above described conditions, the reward size was kept constant throughout; that is, each correctly detected stimulus (hit) resulted in a water reward of identical volume (volume per trial = 0.09 ml).

Experiment 3. In a subset of rats (rats 1–3), the distribution range and probability of stimuli was kept constant throughout all conditions ($A = [0, 1, 4, 16]^\circ$, $p = 0.25$), but the reward size was varied systematically; that is, the drop size was small (0.04 ml), medium (0.09 ml), or big (0.18 ml) in a given hit trial. The reward volume changed only between experimental conditions, not within sessions. The above-described high-range distribution served as baseline and was performed first because it already used the medium reward size. In the following sessions, the stimuli were kept constant but the drop size was decreased by half and then increased to double the volume. Note that not all animals of this group experienced all reward conditions (rat 1: half and medium only, rat 2–3: half, medium, and big).

Data analysis and statistics

Psychometric data were assessed as response probabilities averaged across sessions within a given stimulus or reward condition. This was done separately for each of the seven animals. Psychometric curves were fit using Psignifit (Wichmann and Hill, 2001a,b; Fründ et al., 2011). Briefly, a constrained maximum likelihood method was used to fit a modified logistic function with four parameters: α (the displacement of the curve), β (related to the inverse of slope of the curve), γ (the lower asymptote or guess rate), and λ (the higher asymptote or lapse rate) as follows:

$$P(\text{GO} | s_i) = \gamma + (1 - \gamma - \lambda) \frac{1}{1 + \exp(-z(s_i))}$$

$$z(s_i) = \frac{s_i - \alpha}{\beta}$$

Response thresholds were calculated from the average psychometric function for a given experimental condition using Psignifit. The term

“response threshold” refers to the inverse of the psychometric function at some particular performance level with respect to the stimulus dimension. Throughout this study, we use a performance level of 50% (probability of detection = 0.5). Statistical differences between psychophysical curves were assessed using bootstrapped estimates of 95% confidence limits for the response thresholds provided by the Psignifit toolbox.

Reward accumulation

Let the stimulus amplitude delivered on the i^{th} trial be denoted as s_i , the corresponding reward as r_i , and the accumulated reward for N trials as R_N . Over N trials, the expected accumulated reward is as follows:

$$E\{R_N\} = \sum_{i=1}^N P(s_i) P(\text{GO} | s_i) r_i$$

where $P(s_i)$ comes from the experimentally controlled stimulus distribution, $P(\text{GO} | s_i)$ is the probability of a positive response (or “Go”) for the given stimulus amplitude, and $E\{\}$ denotes statistical expectation.

We considered the null hypothesis of this behavioral paradigm to be that animals do not adapt their behavior in response to an experimentally forced change in stimulus distribution and thus operate from the same psychometric curve (represented as dotted curves in Figs. 2, 3, and 4). For example, in moving from the high-range to the low-range stimulus condition (see Fig. 2), this would result in a decrease in the total accumulated reward for the same number of trials.

As an alternative hypothesis, one possible strategy that the animal could take in response to a change in the stimulus distribution would be to adjust behavior to maintain the same amount of accumulated reward during a session. For example, in moving from high-range stimuli to low-range stimuli, the accumulated reward would be assumed to be fixed and we can determine a new set of probabilities $P(\text{GO} | s_i)$ that define an adapted psychometric function. Note that there is not a unique solution, but one simple possibility is that the original psychometric function maintains the same asymptotes (γ and λ) and false alarm rate, but is compressed, with a decrease in response threshold and an increase in slope to maintain the same total accumulated reward. We denote this situation as our hypothetical psychometric function, which is represented as dashed curves in Figure 2.

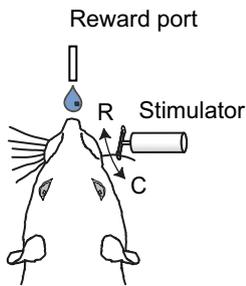
As outlined above, Experiment 1 involves the high- and low-range stimulus amplitude distributions, but both are uniform distributions, so $P(s_i)$ is the same for all s_i . As shown in Figure 3, we also conducted experiments (Experiment 2) with nonuniform distributions. Consider, for example, the distributions shown for the $P_{\text{big}} > P_{\text{small}}$ in the top panel of Figure 3A, where the probability of the big-amplitude stimuli of 4 and 16 degrees is higher than small amplitude stimuli of 0.25 and 1 degrees. Under the hypothesis that animals adapt their behavior to achieve the same amount of reward when moving from this distribution to the inverse distribution $P_{\text{big}} < P_{\text{small}}$, we can calculate a new psychometric function as done above. Again, there is not a unique solution, so we chose the simplest possibility by compressing the original psychometric function to the left so that it maintains the same asymptotes (γ and λ) and false alarm rate, but allows a change in response threshold and slope to maintain the same total accumulated reward. As before, we denote this situation as our hypothetical psychometric function, represented as dashed curves in Figure 3.

Finally, under the working hypothesis that the animal adjusts to maintain the total accumulated reward in response to changes in the stimulus distribution, then changing the amount of reward delivered upon a hit trial should produce predictable effects in the behavior. In a final set of experiments (Experiment 3; see Fig. 4), we thus changed the amount of reward delivered in a successful trial, which changes the value of r_i . This then generates predictable changes in the behavior, and thus the psychometric function, in comparing the small, medium, and big rewards of Figure 4A.

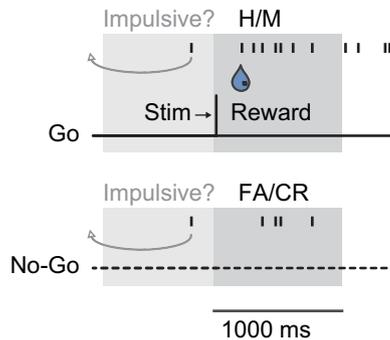
Results

We trained seven head-fixed rats using a tactile Go/No-Go detection task (Cook and Maunsell, 2002; Stüttgen et al., 2006;

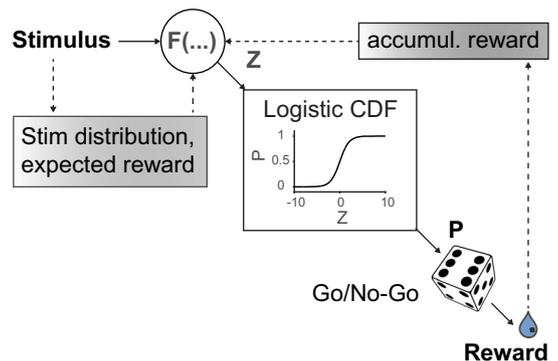
A Behavioral Setup



B Go/No-Go Detection



C Conceptual Framework



D Stimulus & Reward Conditions

	Stim-amp (deg)				Stim-prob (p)					Reward (ml/trial)					
E1	0	1	4	16	uniform (0.25)	0	0.09	0.09	0.09	0	0.09	0.09	0.09		
	0	0.25	1	4											
E2	0	0.25	1	4	16	0.1	0.09	0.09	0.36	0.36	0	0.09	0.09	0.09	0.09
						0.1	0.36	0.36	0.09	0.09					
E3	0	1	4	16	uniform (0.25)	0	0.04	0.04	0.04	0	0.09	0.09	0.09	0.09	
						0	0.09	0.09	0.09						
						0	0.18	0.18	0.18						

E Experimental Design

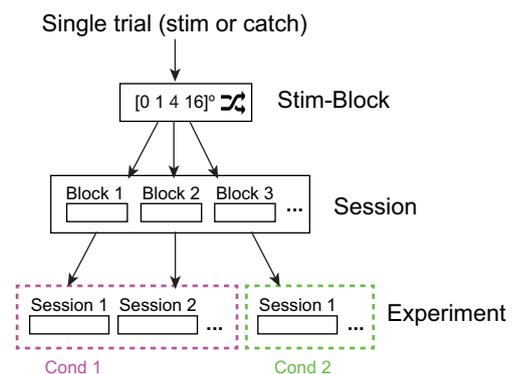


Figure 1. Experimental strategy. **A**, Behavioral setup with head-fixed rat, reward port, and whisker stimulator. A single whisker was deflected in the rostrocaudal plain (R, C). **B**, Go/No-Go detection task. A punctuate stimulus (10 ms) has to be detected by the rat with an indicator lick to receive reward. A Go trial is categorized as a hit (H) if the animal responds within 1 s upon stimulation and a miss (M) if no response is emitted. Catch trials in which stimulation is absent (No-Go) are categorized as a false alarm (FA) if random licking occurs or as a correct rejection (CR) if licks are withheld. Reward is only delivered in hit trials. Impulsive licks in a 2 s period before trial onset are mildly punished by a time-out (arrow, 4–10 s). **C**, Conceptual behavioral framework. The animal’s choice (Go or No-Go) is modeled as a function of the trial by trial response to a stimulus. The resulting variable z is passed through a logistic function that yields a probability for Go or No-Go responses. Two history-dependent variables are thought to influence the behavioral response: (1) reward expectation is formed by integrating the stimulus distribution (range or probability of a particular amplitude occurring) in a feedforward model (left gray box) and (2) the reward accumulation is integrated and leads to behavioral adjustments in a feedback model (right gray box). **D**, Overview of all experiments. Experiment 1 manipulated the range of stimulus amplitudes; that is, the upper and lower limits (top, magenta vs green). Experiment 2 manipulated the probability of a stimulus presentation (middle, magenta vs green). Experiment 3 manipulated the reward payoff determined by the water drop size in hit trials (bottom, different shades of blue). **E**, Experimental design. From top to bottom, A stimulus block consisted of three to four deflection amplitudes and one catch trial presented in pseudorandom order. Each session consisted of repeated stimulus blocks. Each experiment consisted of multiple sessions and was split into different conditions (colored boxes).

Ollerenshaw et al., 2012). In this task, the animal must detect pulse-shaped deflections of a single whisker and report the decision by either generating a lick on a waterspout (Go) or by withholding licking (No-Go) if no stimulus is present (Fig. 1A,B). A trial is categorized as a hit if the animal responds within 1 s upon stimulation and a miss if no response is emitted. Catch trials in which stimulation is absent are also included and a trial is categorized as a false alarm if random licking occurs or as a correct rejection if licks are withheld. However, reward is only delivered in hit trials. Premature licks 2 s before any trial were mildly punished by a time-out. We report these events as impulsive licks, but distinguish them from false alarms during catch trials for the main psychometric data analysis. All seven rats achieved expert level within 10 sessions of operant behavioral training (data not shown).

As in related studies (Busse et al., 2011), it is helpful to visualize the relationship between presented sensory stimuli and the resultant behavior in the form of a block diagram as in Figure 1C. At its core, the behavior is a function of the properties of the current sensory

stimulus through a classical psychometric relationship, where the presentation of a sensory stimulus maps directly to a probability of response (GO, lick) through a logistic function. To predict behavioral dynamics with regard to changing context, the behavioral sensitivity of the animal is potentially modulated by both “feedforward” representations of the stimulus distribution and corresponding reward expectation, requiring an internal model of stimulus history, and “feedback” representations of actual reward accumulation, requiring an internal model of reward history.

To investigate the influence of the contextual variables and disentangle potential sources of behavioral modulation, we designed psychophysical experiments in which probabilistic stimulus distributions and reward contingencies were systematically manipulated (Fig. 1D). Experiment 1 manipulated the distribution of whisker deflection amplitudes; that is, the upper and lower limits of amplitudes presented in a psychophysical test, but importantly involved amplitudes common to both high-range and low-range conditions. Experiment 2 held the stimulus range fixed, but manipulated the shape of the distribution by changing the

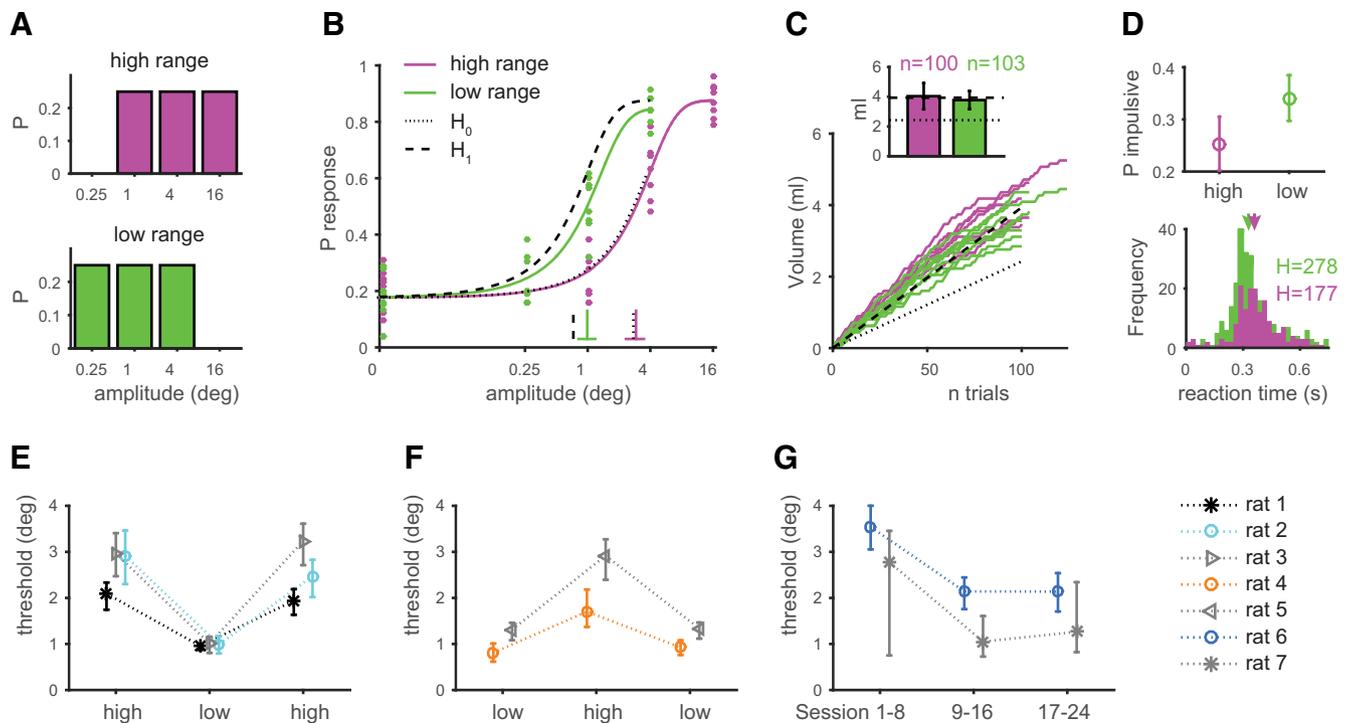


Figure 2. Experiment 1. **A**, Manipulation of stimulus range (magenta vs green). Every stimulus and catch trial (data not shown) was presented with equal probability ($p = 0.25$). **B**, Psychometric curves and response thresholds (vertical lines with CI95) for an example animal (rat 2). Each dot corresponds to response probabilities from a single session. Solid curves are logistic fits to the average data (seven to eight sessions). Dotted line is a hypothetical curve assuming no change in performance (H_0) from high-range to low-range stimuli. Dashed line is a hypothetical curve assuming a change in performance to maintain the same amount of accumulated reward (H_1) when switched from high-range to low-range stimuli. **C**, Water volume accumulated by the same animal under both conditions. Each line corresponds to one session. Inset, Average total reward volume per session for a given condition. The average number of trials worked by this animal is shown. Error bars represent SD; otherwise, the figure conventions are the same as in **B**. **D**, Metrics of task engagement. Top, Average probability of impulsive licks leading to time-outs. Error bars indicate CI95. Bottom, Histogram of RTs (poststimulus lick) for stimuli close to response threshold (1° and 4°). Median RTs are indicated by arrows. The number of hits for each condition is shown. **E**, Response thresholds in Experiment 1a with high-range stimuli presented first, followed by low-range stimuli and another high-range set ($n = 3$ rats). **F**, Response thresholds in Experiment 1b reversing the order of conditions ($n = 2$ rats). **G**, Control experiment with no change in condition ($n = 2$ rats). The data were split into three equal parts. Data points in **E–G** represent means across sessions within the same animal. Error bars indicate CI95.

relative probabilities within the distribution. Experiment 3 held the stimulus distribution fixed, but manipulated the reward payoff determined by the water drop size in hit trials. The experiments were designed such that, on a single trial, one of four different stimuli or a catch trial was presented after a variable time interval, each with equal or unequal probability (Fig. 1E). A stimulus block consisted of a trial sequence comprising three to four stimuli and a catch trial in pseudorandom order (e.g., each trial type once per block). A behavioral session consisted of repeated stimulus blocks until the animal disengaged from the task. Therefore, the chosen stimuli occurred repetitively but randomly within a session. Multiple behavioral sessions (maximum two per day) were then performed to measure the psychometric performance for a given condition. An experiment consisted of multiple sessions comprising two or three conditions, where each condition is defined by a specific stimulus distribution and reward contingency. To assess long-term behavioral effects, a specific condition was always kept constant within and across multiple behavioral sessions (seven to eight per condition) before the task was changed.

Influence of a shift in stimulus distribution

For the first group of rats (rats 1–3), the stimulus distribution consisted of three different stimulus amplitudes and a catch trial ($A = [0, 1, 4, 16]^\circ$; Fig. 2A, magenta), each occurring with the same probability throughout the first part of Experiment 1 (i.e., a uniform distribution). We define this as the high-range condition. In the second part, the stimulus distribution consisted of

three new stimulus amplitudes and a catch trial ($A = [0, 0.25, 1, 4]^\circ$; Fig. 2A, green), each occurring with the same probability, which we define as the low-range condition. Importantly, both conditions share two of the three stimulus amplitudes. However, the largest amplitude of the high range (16°) is not part of the low range and, vice versa, the smallest amplitude of the low range (0.25°) is not part of the high range. Figure 2B depicts a typical psychometric curve (magenta) of an individual animal performing the task under the high-range condition. In response to a shift in stimulus distribution (Experiment 1), we consider two extreme hypotheses. The null hypothesis (H_0) asserts that the animal does not adjust its behavior and thus operates from the same psychometric function (black dotted curve on top of magenta curve). In moving from the high-range to the low-range condition, this would result in a decreased reward rate for the same number of trials. We consider an alternative hypothesis (H_1), that the animal adapts its behavior to maintain accumulated reward in the face of a changing stimulus distribution. In this case, we use a simple reward expectation model to predict a hypothetical psychometric performance that would maintain reward intake when the stimulus distribution changes from the high stimulus range to the low stimulus range (see Materials and Methods). Note that this prediction represents one of many possible ways to maintain the same reward in the face of the changing stimulus probabilities. The black dashed curve in Figure 2B denotes the hypothetical psychometric function with the same lapse and guess rate as the original curve from the rat, but allowing it to shift

to the left such that the expected reward per trial remains constant. The experimentally measured psychometric function in the low-range condition (green) comes quite close to the hypothetical performance level based on the assumption of maintained reward expectation. Consistent with the model's prediction, the observed shift results in a significant decrease in response threshold ($T_{\text{high}} = 2.92$, $T_{\text{low}} = 0.99$, $\text{CI}_{95\text{high}} = [2.34\ 3.51]$, $\text{CI}_{95\text{low}} = [0.81\ 1.81]$, all in units of degrees; data are reported as mean and 95% confidence interval, CI_{95}) and an increase in slope (note the difference in slope is obscured due to the logarithmic plot; the increase in probability of response per degree near threshold is significantly higher for the green compared with magenta curve). Figure 2C depicts the actual trial-by-trial accumulation of water volume in each session with different conditions color coded. Note that overlaid are results for $n = 8$ sessions with the high-range distribution (magenta) and $n = 8$ sessions with the low-range distribution (green). The slope of reward accumulation in the low-range condition almost matches that of the high-range condition and the slope for low case (green) is close to the prediction from the maintenance of accumulated reward hypothesis, H_1 , while being clearly separable from the slope representing the null hypothesis (dotted line). The total reward volume acquired on average per session in both conditions further confirms this ($V_{\text{high}} = 4.04 \pm 0.88$ ml, $V_{\text{low}} = 3.78 \pm 0.61$ ml; Fig. 2C, inset, dashed line is H_1 hypothesis, dotted line is H_0 hypothesis). An alternative strategy to maintain the total reward in the low-range condition could also be to work more trials, which we could not confirm in this particular dataset (see numbers in Fig. 2C, inset).

Can the behavioral adaptation be explained by a change in the animal's task engagement; for example, by overall changes in arousal, vigilance, or fluctuations in general attention and motivation that may tightly depend on the experimentally imposed task structure? Although it is challenging to differentiate between these variables within the Go/No-Go paradigm, we provide two measures to infer task engagement, the subject's impulsivity ($I =$ probability of impulsive lick) leading to time-outs between trials (Fig. 2D, top) and reaction times (RTs) in correctly detected Go trials (Fig. 2D, bottom).

When the task is switched from the high-range to the low-range condition, the animal generates more impulsive licks on average ($I_{\text{high}} = 0.25$, $I_{\text{low}} = 0.34$, $\text{CI}_{95\text{high}} = [0.2\ 0.31]$, $\text{CI}_{95\text{low}} = [0.3\ 0.38]$). In addition, the same rat shows a trend of slightly shorter lick RTs upon near threshold stimuli (median $\text{RT}_{\text{high}} = 360$ ms, $\text{RT}_{\text{low}} = 329$ ms). Both the animal's increase in impulsivity and slightly shorter RT suggest an increase in arousal. However, even though performance improves systematically across all animals undergoing task changes, this effect is clearly inconsistent (e.g., rat 3; $I_{\text{high}} = 0.21$, $I_{\text{low}} = 0.21$, $\text{RT}_{\text{high}} = 244$ ms, $\text{RT}_{\text{low}} = 259$ ms, data not shown). To further rule out whether impulsivity has an effect on performance, we repeated our psychometric analysis by including all spontaneous licks; that is, from a pool of impulsive trials (time-outs) and catch trials, we sampled randomly the same number as stimulus presentations. Across animals and sessions, an increase in hits does not correspond to an increase in impulsivity and false alarms, indicating that changes in performance are not due to random guessing (data not shown).

It is possible that the observed transformation of the psychometric function is due to a steady increase in learning, which could be irreversible and not tied to the changing stimulus distribution itself. To address this question, we performed an additional series of tests with reciprocal changes, as well as a control experiment with fixed conditions. Figure 2E shows response thresholds from the first group of rats (rat 1–3) with high-range

stimulus distribution presented first, followed by low-range stimulus distribution and another high-range stimulus distribution. Changing the stimulus distribution range reciprocally from high to low and back to high leads to a dynamic threshold progression, which closely follows the direction of task manipulation, showing a decrease at the first transition and a rebound after the second transition. We also performed the inverse task sequence with a different group of rats (Fig. 2F); that is, we trained naive subjects using low-range stimulus distributions first (rats 4 and 5) and then changed the task to the high-range and back to the low-range condition. Again, all response thresholds follow the direction of task manipulation, indicating that the animals are able to apply a reciprocal task strategy modulated by the shift in stimulus distribution. Interestingly, the second group of rats started with a higher performance level, which is likely due to their training history (early training with small amplitude $A = 4^\circ$ and first psychophysical test with low-range stimulus distribution $A = [0, 0.25, 1, 4]^\circ$).

Although both the reversibility of the shift in psychometric sensitivity and the persistence of the effect in response to a reversal of switching (e.g., high to low and low to high) suggest that the observed phenomenon cannot just be attributed to a learning effect, learning obviously plays some role in these trained behaviors. To better quantify this effect, we conducted an additional set of experiments and analyses. Figure 2G depicts response thresholds of two animals that never experienced any change in stimulus distributions (rats 6 and 7). In addition to the high variability of the individual animal's response threshold from one session to another, a permanent decrease in thresholds over many sessions was observed (averages of $n = 8 \times 3$ sessions). Because the context was never changed experimentally, we consider this steady improvement in detection performance to be shaped by progressing training over the course of many experiments. Such long-term perceptual learning effects have been shown before in primate psychophysical datasets (Gold et al., 2008).

Influence of stimulus redistribution

In Experiment 1, the change in behavioral performance was in response to a relatively large manipulation of the stimulus distribution in shifting the range. Is the animal's strategy sensitive to more subtle changes in the stimulus distribution? In Experiment 2, we performed a separate set of experiments with the same animal group, where the stimulus distribution range was fixed, but the relative probabilities were changed (i.e., no longer uniform). Specifically, all four amplitudes plus a catch trial ($A = [0, 0.25, 1, 4, 16]^\circ$) were used for both the big and small conditions, but in the big condition, the two stimuli with large amplitudes were presented with a higher probability than the two small amplitudes ($P_{\text{big}} > P_{\text{small}}$; $P [4, 16]^\circ = 0.36$, $P [0.25, 1]^\circ = 0.09$; Fig. 3A, magenta). In the small condition, the order of stimulus probabilities was reversed; that is, the two stimuli with small amplitudes were presented with a higher probability than the two large amplitudes ($P_{\text{big}} < P_{\text{small}}$; $P [4, 16]^\circ = 0.09$, $P [0.25, 1]^\circ = 0.36$; Fig. 3A, green). The probability of a catch trial remained the same ($P [0^\circ] = 0.1$, data not shown). Figure 3B depicts typical measured psychometric curves for an individual animal performing the task under both conditions (magenta vs green). Based on the changes in stimulus distributions, we again used the simple reward–expectation model to predict possible changes in psychometric sensitivity. Again, the null hypothesis H_0 and the maintenance of reward expectation hypothesis H_1 are shown as dotted and dashed lines, respectively. Similar effects become apparent as seen above with the modulated stimulus distribution range: if the distribution of stimulus presentations tilts to the left;

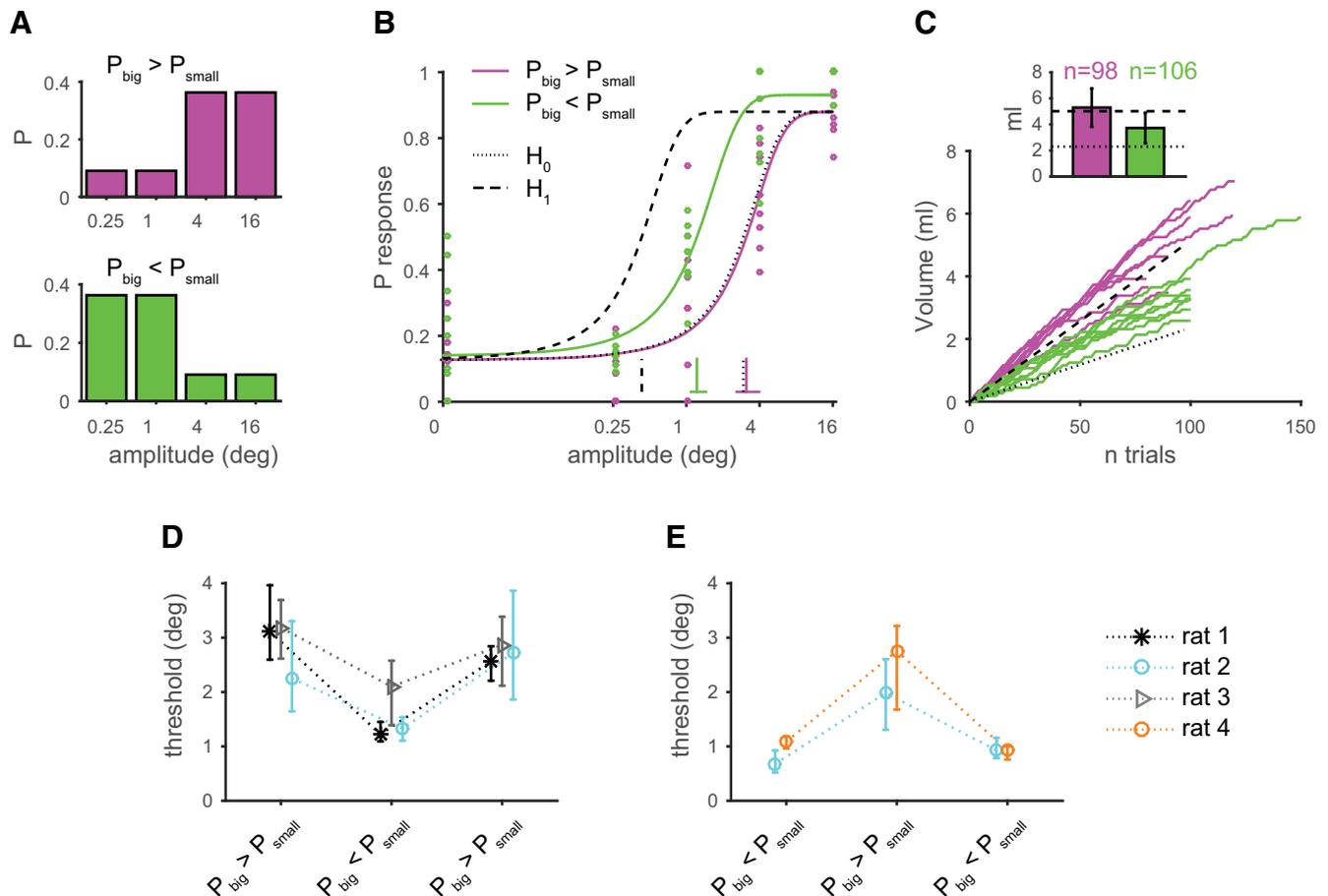


Figure 3. Experiment 2. **A**, Manipulation of stimulus statistics. Stimuli were presented with unequal probabilities, either big amplitudes more often than small amplitudes (magenta, $P_{big} = 0.36$, $P_{small} = 0.09$) or vice versa (green, $P_{big} = 0.09$, $P_{small} = 0.36$). Catch trials occurred at an intermediate probability ($P_{catch} = 0.1$, data not shown). **B**, Psychometric curves and response thresholds (vertical lines with CI95) for an example animal (rat 1). Each dot corresponds to response probabilities from a single session. Solid curves are logistic fits to the average data (seven to eight sessions). Dotted line is a hypothetical curve assuming no change in performance (H_0). Dashed line is a hypothetical curve assuming a change in performance to maintain the same amount of accumulated reward (H_1) when switched from easy ($P_{big} > P_{small}$) to difficult ($P_{big} < P_{small}$). **C**, Water volume accumulated by the same rat under both conditions. Each line corresponds to one session. Inset, Average total reward volume per session for a given condition. The average number of trials worked by this animal is shown. Error bars indicate SDs; otherwise, the figure conventions are the same as in **B**. **D**, Response thresholds in Experiment 2a with the easy condition presented first ($P_{big} > P_{small}$), followed by the difficult condition ($P_{big} < P_{small}$) and another easy one ($n = 3$ rats). **E**, Response thresholds in Experiment 2b reversing the order of conditions ($n = 2$ rats). Note that the first and third conditions were always the same. Data points in **D** and **E** represent means across sessions within the same animal. Error bars indicate CI95.

that is, if stimuli with smaller amplitudes occur with a higher probability, then the predicted psychometric curve also shifts toward the left with an increase in slope. Again, note that this is plotted logarithmically and that the small (green) psychometric curve has a significantly larger increase in probability of response per degree near threshold compared with the big (magenta) psychometric curve. Even though the experimentally measured psychometric function does not reach this hypothetical performance level, it shows a significant decrease in response threshold ($T_{high} = 3.12$, $T_{low} = 1.22$, $CI95_{high} = [2.63\ 3.96]$, $CI95_{low} = [1.08\ 1.43]$; all numbers in degrees, reported are means and CI95) consistent with this kind of prediction. The reward volume acquired by the rat is not the same between the two conditions ($V_{high} = 5.29 \pm 1.46$ ml, $V_{low} = 3.73 \pm 1.17$ ml), but it is clearly separable from the theoretical accumulation predicted by the null hypothesis (i.e., where performance does not change in response to the change in distribution; Fig. 3C). Interestingly, this animal seemed to use the strategy of working more trials in the difficult condition; however, across all animals, we did not find a significant effect of trials worked under the different task conditions. Again, we performed this experiment for multiple animals with the dif-

ferent conditions in reversing order and further confirmed the finding described above (Fig. 3D,E).

These results are consistent with the possibility that reward accumulation plays an important role in adaptively shaping the animal's behavior, where the reward-focused strategy incorporates an inferred model of the stimulus statistics that enables detecting the same exact stimulus with different accuracy in different contexts. We further find that changes in performance are reversible: the animal's response threshold decreases significantly with a change from an easy to a difficult condition and vice versa.

Influence of reward

The results described above suggest that a change in sensory context can cause an adaptation of the animal's performance, presumably to maintain a certain level of reward. However, it is possible that such a scenario exists either through an internal model of reward accumulation driven by the properties of the sensory stimulus or through direct feedback to the animal in the form of the actual reward. In Experiments 1 and 2, it is the case that the stimulus distribution influences the reward accumulation, so these two possibilities are conflated. To further investi-

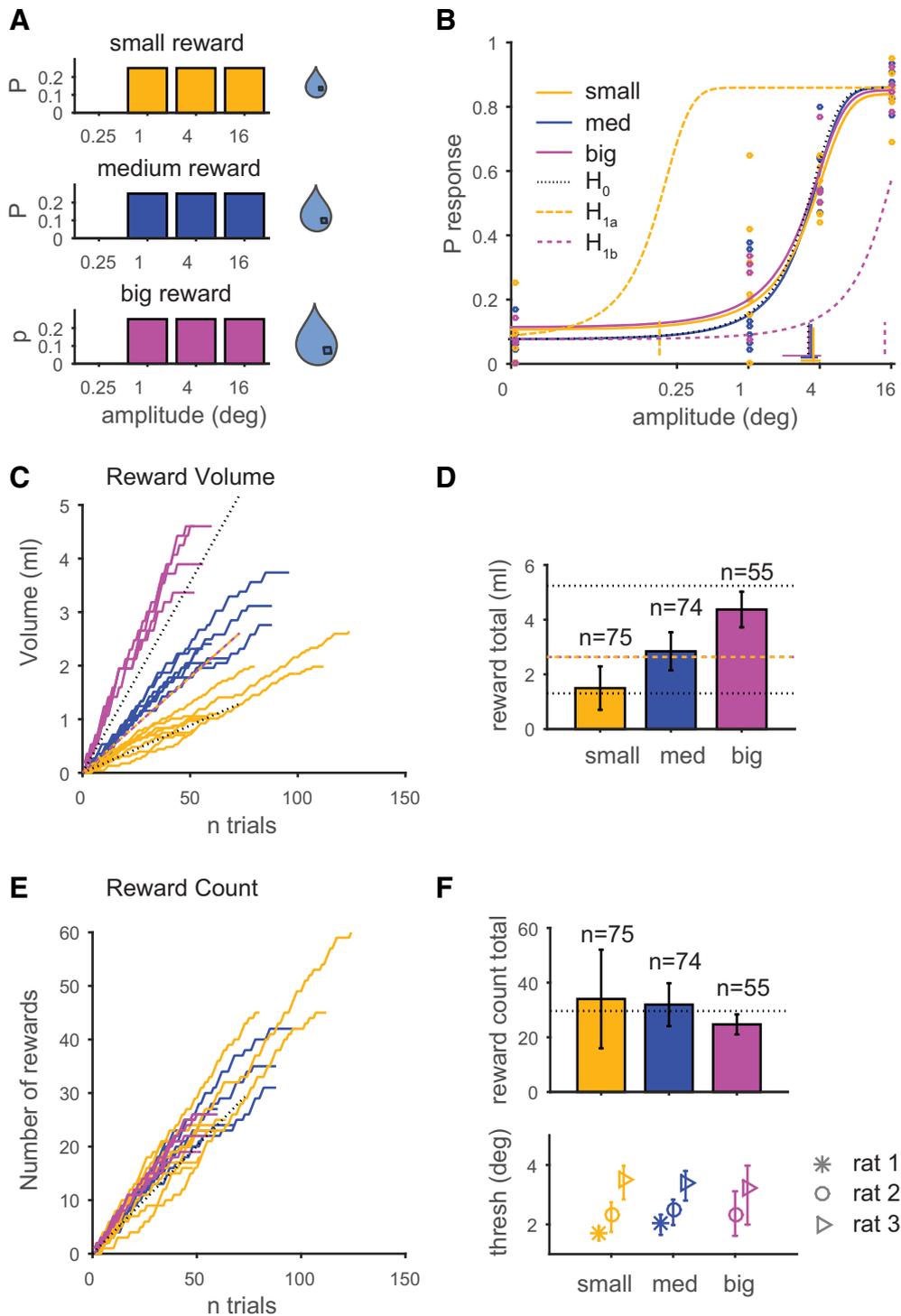


Figure 4. Experiment 3. **A**, Manipulation of reward. The water drop size in hit trials was changed from the original drop size (blue, medium = 0.09 ml/trial) to half the size (yellow, small = 0.04 ml/trial) or to double the size (magenta, big = 0.18 ml/trial). **B**, Psychometric curves and response thresholds (vertical lines with C195) for an example animal (rat 3) under the three reward conditions. Same conventions as in Figure 2 and 3. Dotted line is a hypothetical curve assuming no change in performance (H_0). Yellow dashed line is a hypothetical curve assuming a change in performance to maintain the same volume of reward per trial (H_{1a}) when switched from a medium drop size to a small drop size. Magenta dashed line is a hypothetical curve assuming a change in performance to maintain the volume of reward (H_{1b}) when switched from a medium drop size to a big drop size. **C**, Reward volume accumulation of the same rat under all three conditions. **D**, Average total reward volume acquired per session for a given condition. The average number of trials worked by this rat is depicted on top. Error bars indicate SDs. **E**, Same data as in **C** but plotted as reward count. **F**, Top, Average total reward count for a given condition. Bottom, Response thresholds of three rats under different conditions (rat 1 experienced only medium and small drops). Data points represent means across sessions within the same animal. Error bars indicate C195.

gate this issue, we performed another set of experiments with the same animal group by keeping the stimulus distribution constant throughout ($A = [0, 1, 4, 16]^\circ$), but systematically varying the volume of deterministic reward delivery (Fig. 4A). Specifically,

the drop size in a hit trial was either small (0.04 ml, yellow), medium (0.09 ml, blue), or big (0.18 ml, magenta) in separate experimental conditions. Figure 4B depicts typical experimentally measured psychometric curves for an individual animal per-

forming the task under all three conditions, starting with the medium drop size (blue curve). The null hypothesis assumes that the animal's performance is independent of the delivered reward volume and thus operates from the same psychometric function across all conditions (black dotted curve on top of blue curve). In contrast, the feedback model makes several predictions for an experimentally forced change in volume. One possible strategy could be to adjust behavior and compensate to maintain a constant accumulated reward volume. In this case, the model predicts the hypothetical performance levels needed to maintain reward volume when moving from the medium drop size to the smaller (H_{1a}) or the bigger (H_{1b}) drop size. The dashed curves in Figure 4B represent hypothetical psychometric functions (yellow for a change to small and magenta for a change to big rewards) with the same lapse and guess rate as the original experimental curve, but allowing it to shift left or right such that the expected reward remains constant. Note that one of the model's predictions (yellow dashed curve) is located outside the animal's perceptual range because the chosen reduction from medium to small drop size was rather extreme. Interestingly, we did not find any differences in performance; the psychometric curves and response thresholds for all conditions are almost identical ($RT_{\text{small}} = 3.54$, $RT_{\text{med}} = 3.38$, $RT_{\text{big}} = 3.24$), suggesting that the reward volume itself has no effect on the animal's task strategy. Indeed, when plotting the animal's reward accumulation under the different conditions (each hit trial multiplied with the respective drop volume), a drastic difference in slope and total acquired reward volume per session becomes obvious (Fig. 4C,D). Even though the animal established a tendency to work for more trials when smaller drops were available, the acquired volume at the end of an experimental session never reached previous levels. Note that an experiment was always aborted when the animal stopped working on the task by not licking for an entire stimulus block. Figure 4, E and F, shows the same data, but now plotted as a function of reward count (i.e., the cumulative number of times rewarded), thereby ignoring volume (hit trials not multiplied). In this case, the accumulated reward counts per session are very similar, but the sessions had markedly different numbers of trials. Across all three animals that underwent this experiment, the effects were robust because we did not see any significant changes in response threshold (Fig. 4F, bottom).

This finding suggests that a constant sensory environment with a systematic change in reward volume is not sufficient to modulate the animal's task strategy. However, we do not exclude the possibility that the number of past rewarded trials could be important for the animal's task engagement.

Performance adaptation on a finer timescale

The preceding results demonstrate clear, robust effects of dynamically changing stimulus context on behavioral performance. Thus far, we have only examined behavioral changes on a session-by-session or experiment-by-experiment basis. However, it is known from various studies that a subject's task strategy can change within sessions (Boneau and Cole, 1967) and even from one trial to another (Nienborg and Cumming, 2009; Busse et al., 2011; Fründ et al., 2014; Waiblinger et al., 2018). This suggests that the adaptive behaviors that we observe could also be the result of a dynamic process on a finer time scale. To estimate within-session fluctuations in performance, we parsed each experimental session into three parts of equal trial numbers (e.g., $n = 3 \times 50$; Fig. 5A) and calculated the response probabilities separately. Figure 5, B and C, shows separate psychometric functions of the first and the last part. Consistent with the classic

findings, the curves shift over to the right (arrows) during the run time of a session. Interestingly, we find that the shift is clearly dependent on the stimulus context because it is qualitatively larger in the easy high-range condition (Fig. 5B, magenta) compared with the difficult low-range condition (Fig. 5C, green). To rule out the possibility that this effect was due to poor fitting of the psychometric curve, we inspected the goodness-of-fit metric of deviance (D) as well as estimates of where the goodness-of-fit lay in bootstrapped cumulative probability distributions of this error metric (CPE) using the psignifit toolbox (Wichmann and Hill, 2001a). Our deviance statistics revealed that the data was fit equally well across conditions with the exception of two outliers that likely occurred due to subsampling of trials with highly variable performance (rat 1, low range, third part, $D = 5.33$, and rat 2, high range, third part; $D = 9.51$, both with a CPE > 95, indicating a goodness-of-fit outside the upper confidence limit).

When further comparing the derived response thresholds for first, second, and third part of the session (Fig. 5D), a significant increase is obvious toward the end of the high-range condition (magenta, third part), whereas very little to almost no change occurs in the low-range condition (green, third part). Therefore, the difference in behavioral performance as measured by the psychometric curves between the high- and low-range stimulus distributions in Experiment 1 and between the big and small stimulus distributions in Experiment 2 does not appear to result from a static, fixed property of the behavior in these different stimulus contexts. Instead, it appears to result from a dynamic process whereby the animal relaxes vigilance gradually through the session for the "easy" (high-range and big) stimulus distributions. In contrast, the behavioral performance was relatively invariant for the more "difficult" stimulus distributions.

To rule out satiety effects, we repeated this analysis by parsing each experimental session into three parts of accumulated reward volume (e.g., 0–1 ml, 1–2 ml, 2–3 ml), thereby ignoring trial numbers (data not shown). Because animals work different numbers of trials and acquire different amounts of water in each session, the upper volume limit was determined for an animal's typical session (rat 1: 4.5 ml, rat 2: 3 ml, rat 3: 3.5 ml) and all trials beyond this limit were discarded. Consistent with the previous results (Fig. 5), all three rats show performance changes that are again more dramatic for the high-range condition and this effect persists even though the same amount of water is consumed across conditions. This indicates that the behavioral adaptation described here is stimulus dependent and cannot solely be attributed to satiety effects.

The results here validate our predictions of behavioral modulation due to both feedforward representations of the stimulus distribution and corresponding reward expectation and feedback representations of actual reward accumulation (Fig. 1C). It is in principle difficult to disentangle these two sources of modulation, but the results here suggest that, indeed, the statistical properties of the stimulus–reward relationship strongly influence the behavior and that the effects are clearly invariant to the absolute amount of reward, consistent with a feedforward influence. However, an alternate possibility is that the relevant feedback signal is not the accumulated volume of reward, but instead an accumulated running count of rewards received.

Discussion

In this study, we have investigated adaptive behavior in a rodent tactile detection task. Our findings provide evidence that a chang-

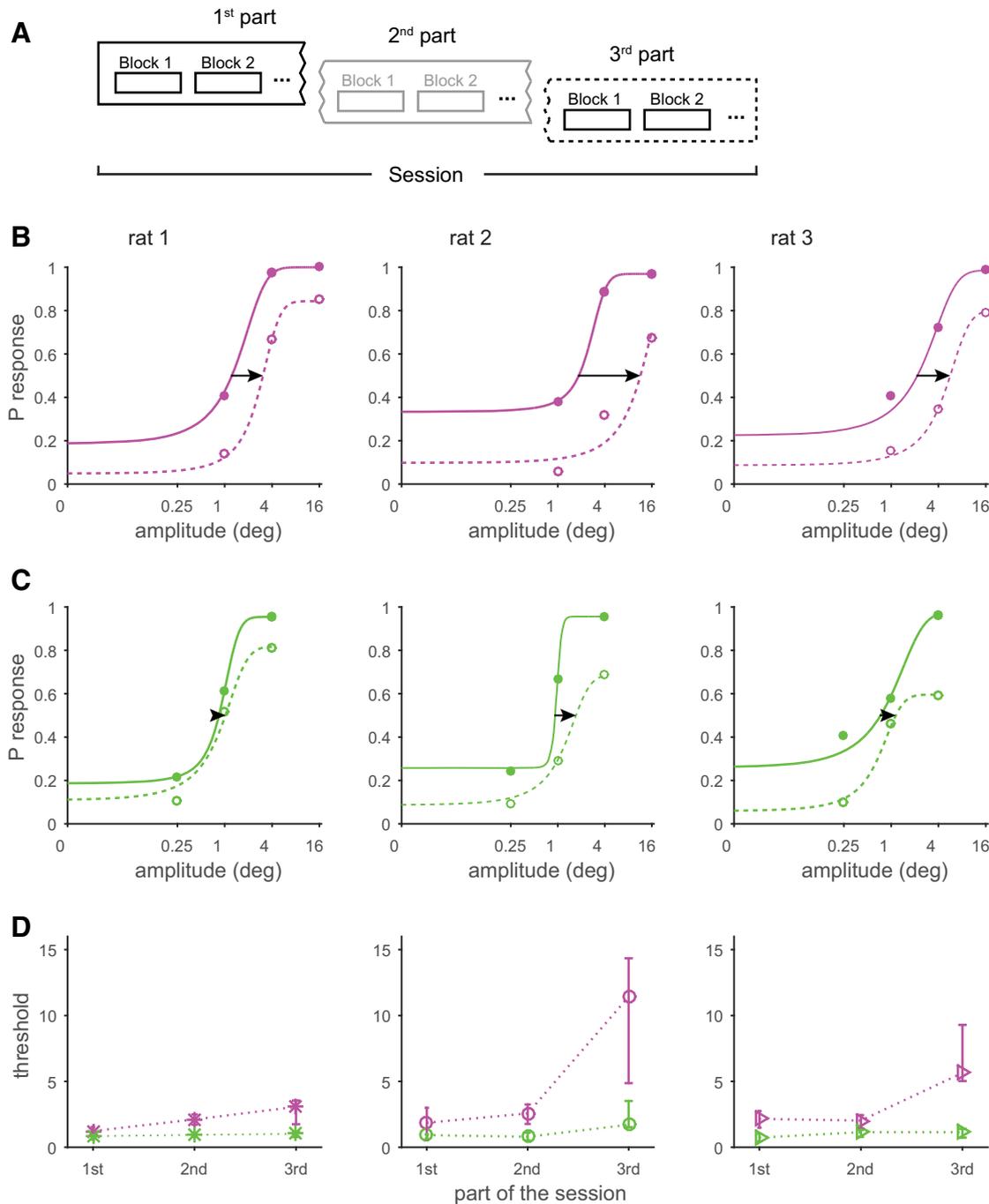


Figure 5. Behavioral adaptation within sessions. **A**, Each session was subdivided into three equal parts to assess detection performance across the run time of a session. **B**, Psychometric curves for the first and last part of a session in the high-range experiment. **C**, Psychometric curves for the first and last part of a session in the low-range experiment. Circles represent average response probabilities (filled circles are the first part, empty circles are the last part). Curves are logistic fits. Arrows indicate the shift in response threshold. **D**, Evolving response thresholds for first, second, and third part of a session separately plotted for the high-range condition (magenta) and low-range condition (green). All data points represent means across corresponding parts of multiple sessions ($n = 7–8$). Error bars indicate C195. Data are shown separately for three different animals.

ing sensory environment and associated reward expectation have a substantial impact on the animal’s behavior. We present the following novel aspects. First, we show that metrics of performance deviate significantly and reversibly depending on the probabilistic distribution of stimulus amplitudes. Second, we show that this change in performance relates to accumulating a constant reward count across trials. Third, the behavioral adaptation determines task engagement within a behavioral session.

Metrics of behavioral adaptation

It seems reasonable to assume that the behavioral adaptation described here can be explained by cognitive aspects such as different levels of arousal, vigilance, or fluctuations in general attention and motivation that may tightly depend on the experimentally imposed task structure. Furthermore, learning and satiety effects could dominate our results because they play a major role in Go/No-Go detection behavior. For instance, over the

course of a session, the animal may respond progressively less often because of decreased motivation to obtain reward, which could lead to the false conclusion that the ability to detect stimuli has diminished.

On the experimental side, we address these issues in multiple ways. We control each animal's impulsivity by using time-outs upon early guesses, therefore focusing the animal's attention on the presence of a stimulus. We control satiety by flexibly adjusting the number of trials until responding stops or by disentangling the probabilistic distribution of sensory inputs and accumulation of reward volume in separate tests (Experiments 1 and 2 vs Experiment 3). Finally, we control learning effects by either reversing all changes (e.g., high to low, back to high range) or by keeping all task-related parameters constant throughout experiments.

On the analytical side, we infer behavioral performance from estimates of the psychometric function (Wichmann and Hill, 2001a). In our case, the function is obtained by fitting a logistic curve to the measured data points that represent the animal's response probabilities given a distribution of stimuli. We consider the response threshold at $p = 0.5$ as an optimal metric of detection performance because it refers to the critical horizontal shift of the psychometric function along the stimulus amplitude axis. In this context, it is important to note that theories that assume a hard threshold cannot explain decision making in psychophysical tasks (Swets, 1961). Because our effects of changing response thresholds are highly significant, persistent across experiments (Experiment 1 and 2), and further reversible, we conclude that changes in performance are clearly stimulus dependent and cannot be explained by learning effects.

In addition to psychometrics, we provide several measures of task engagement: (1) the subject's spontaneous guessing or impulsivity leading to time-outs between trials, (2) RTs, and (3) we reanalyzed the data given different satiety levels.

When the task is switched from the easy high-range condition to the difficult low-range condition, some individuals indeed show an increase in impulsivity, resulting in more time-outs and slightly shorter RTs, suggesting that performance changes due to increased levels of arousal. However, the inconsistencies in these results across subjects and experiments lead us to conclude that behavioral adaptation is not solely due to a change in the animal's arousal or motivational state.

To rule out satiety effects, we split sessions into equal subsets and repeated the psychometric analysis (parsed by number of trials or reward volume). Performance changes are more dramatic for the easier stimulus distribution and this effect persists even though the same amount of water is consumed across conditions. This result shows that satiety can have a general and substantial influence on behavior (increase in lapse rate), but it cannot account for the changes in performance under different task conditions. This notion is further supported by data from Experiment 3 (Fig. 4) showing that psychometric curves are highly persistent toward dramatic manipulations in reward volume. Surprisingly, the same data suggest that the number of past rewarded trials seems to determine the animal's task engagement, thus showing the ability to integrate a distribution of accomplishments across trials.

The experiments here were designed to directly probe reward expectation in the case when the relationship between performance and reward is fixed (i.e., a hit was always rewarded). However, the findings also predict that a manipulation of reward probability upon hit trials, an experiment that has not been performed here, would have an impact on task performance if the adaptive behavior were indeed tied to the actual reward as op-

posed to the stimulus properties that dictate rewards. Indeed, many studies using paradigms with asymmetric reward contingencies show that animals are highly sensitive to changes in the frequency or probability of reward (Herrnstein, 1961; Reynolds, 1961; Nevin et al., 1975; Balcı et al., 2009; Teichert and Ferrera, 2010; Stüttgen et al., 2011, 2013). This would open up a set of additional questions in the context of this study that would be important to explore.

These findings suggest that the behavioral adaptation described here cannot be explained by a single parameter such as changes in arousal, satiation, or learning. Instead, we propose that the change in performance results from a complex integration of the above.

A recent study identifying relevant features for Go/No-Go behavioral variability on timescales from a few trials to an entire session further supports this notion (Waiblinger et al., 2018).

Decision theoretical aspects

Our focus on adaptive behavior in a dynamic sensory environment is motivated by the fact that human and animal behavior is often consistent with probabilistic computations (Bayes, 1958; De Finetti et al., 1993; Van Horn, 2003). Especially in tasks involving uncertainty, it is efficient to represent knowledge with probability distributions and to acquire new knowledge by following the rules of probabilistic reasoning. Recent theories have evolved to investigate probabilistic computations in the sensory, motor, and cognitive domains at the level of neural circuits (for review, see Pouget et al., 2013). An important aspect of these theories is that they circumscribe a wide range of tasks from sensory processing to high-level cognition. However, insights into the neural basis of perceptual decisions have come mainly from primates and computational learning models have been characterized mostly for complex human psychophysical datasets (Nassar et al., 2010). New questions are arising that might be ideally answered in the rodent, especially with more recent advances in tools for measurement and manipulation (Knöpfel et al., 2006; Jin et al., 2012; Borden et al., 2017).

By using principles of signal detection theory (Green and Swets, 1966), we consider two hypothetical scenarios how the change in performance reported here could be explained within neuronal circuits. Behavioral adaptation can either be due to internal changes in sensitivity (discriminability, d'), decision criterion (bias), or both (Luo and Maunsell, 2018). A decision maker may improve sensitivity by reducing the overlap between signal and noise distributions. Alternatively, the decision maker may value hits and false alarms differently by altering the criterion. These two changes can be distinguished by the decision makers' false alarm rate. An improvement in hit rate brought about by a decrease in criterion is associated with more false alarms, whereas the same increase in hit rate brought about by an increase in sensitivity is associated with fewer false alarms.

Our data clearly show an increase of hits for a particular stimulus within the low-range distribution; however, animals did not always exhibit full adaptation as predicted by our reward model (Fig. 3B). Because there is no consistent change of false alarms or impulsivity in our data, we rule against the interpretation of changes in criterion or sensitivity alone and propose that rats indeed adopt a mixed strategy that is further compromised by some amount of cognitive effort. Again, this implies that decisions do not occur in isolation, but rather depend on accomplishments or failures at different points in time. This hypothesis is supported by a large body of literature suggesting that behavioral actions are not simply based on current sensory observations, but are often based on a statistically optimal integration of sensory observations and the subjects predic-

tions or prior knowledge (Shadmehr et al., 2010). Priors, history biases, and changing stimulus–action associations can partly affect neuronal computations at primary sensory and higher-order cortical levels (Busse et al., 2011; Jaramillo et al., 2014; Akrami et al., 2018; Waiblinger et al., 2018).

Our current study supports the notion of statistical integration and probabilistic computations in the rodent brain and presents the novel aspect of behavioral adaptation in a well controlled dynamic framework. By systematically changing the sensory environment, we are able to modulate the rats' behavioral strategy consistent with the probabilistic distribution of sensory inputs. Our simple model of behavioral adaptation captures the observed change in psychometric sensitivity and predicts a strategy seeking to maintain reward counts in the face of the changing stimulus distribution. Therefore, we propose that rats rely on an internal model integrating the distribution of sensory inputs across trials and altering their responses in a probabilistic manner to maintain the desired payoff with minimal effort.

References

- Akrami A, Kopec CD, Diamond ME, Brody CD (2018) Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature* 554:368–372. [CrossRef Medline](#)
- Balci F, Freestone D, Gallistel CR (2009) Risk assessment in man and mouse. *Proc Natl Acad Sci U S A* 106:2459–2463. [CrossRef Medline](#)
- Bayes T (1958) Essay towards solving a problem in the doctrine of chances. *Biometrika* 45:293–315. [CrossRef](#)
- Boneau CA, Cole JL (1967) Decision theory, the pigeon, and the psychophysical function 1. *Psychol Rev* 74:123–135. [CrossRef Medline](#)
- Borden PY, Ortiz AD, Waiblinger C, Sederberg AJ, Morrisette AE, Forest CR, Jaeger D, Stanley GB (2017) Genetically expressed voltage sensor ArcLight for imaging large scale cortical activity in the anesthetized and awake mouse (erratum). *Neurophotonics* 4:039801. [CrossRef Medline](#)
- Braun A, Urai AE, Donner TH (2018) Adaptive history biases result from confidence-weighted accumulation of past choices. *J Neurosci*. Advance online publication. Retrieved January 25, 2018. doi: 10.1523/JNEUROSCI.2189-17.2017
- Busse L, Ayaz A, Dhruv NT, Katzner S, Saleem AB, Scho ML, Zaharia AD, Carandini M (2011) The detection of visual contrast in the behaving mouse *J Neurosci* 31:11351–11361. [CrossRef](#)
- Chagas AM, Theis L, Sengupta B, Stüttgen MC, Bethge M, Schwarz C (2013) Functional analysis of ultra high information rates conveyed by rat vibrissal primary afferents. *Front Neural Circuits* 7:190. [CrossRef Medline](#)
- Cook EP, Maunsell JH (2002) Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat Neurosci* 5:985–994. [CrossRef Medline](#)
- De Finetti B, Machi A, Smith, A (1993) *Theory of Probability: a Critical Introductory Treatment*. New York, NY: Wiley.
- Fründ I, Haenel NV, Wichmann FA (2011) Inference for psychometric functions in the presence of nonstationary behavior. *J Vis* 11:16. [CrossRef Medline](#)
- Fründ I, Wichmann FA, Macke JH (2014) Quantifying the effect of intertrial dependence on perceptual decisions. *J Vis* 14:9. [CrossRef Medline](#)
- Gold JJ, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574. [CrossRef Medline](#)
- Gold JJ, Law CT, Connolly P, Bennur S (2008) The relative influences of priors and sensory evidence on an oculomotor decision variable during perceptual learning. *J Neurophysiol* 100:2653–2668. [CrossRef Medline](#)
- Green DM, Swets JA (1966) *Signal detection theory and psychophysics*. Society. New York: Wiley.
- Herrnstein RJ (1961) Relative and absolute strength of response as a function of frequency of reinforcement, 12. *J Exp Anal Behav* 4:267–272. [CrossRef Medline](#)
- Jaramillo S, Borges K, Zador AM (2014) Auditory thalamus and auditory cortex are equally modulated by context during flexible categorization of sounds. *J Neurosci* 34:5291–5301. [CrossRef Medline](#)
- Jin L, Han Z, Platasa J, Wooltorton JR a, Cohen LB, Pieribone VA (2012) Single action potentials and subthreshold electrical events imaged in neurons with a fluorescent protein voltage probe. *Neuron* 75:779–785. [CrossRef Medline](#)
- Knöpfel T, Díez-García J, Akemann W (2006) Optical probing of neuronal circuit dynamics: genetically encoded versus classical fluorescent sensors. *Trends Neurosci* 29:160–166. [CrossRef Medline](#)
- Luo TZ, Maunsell JHR (2018) Attentional changes in either criterion or sensitivity are associated with robust modulations in lateral prefrontal cortex. *Neuron* 97:1382–1393.e7. [CrossRef Medline](#)
- Nassar MR, Wilson RC, Heasly B, Gold JI (2010) An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J Neurosci* 30:12366–12378. [CrossRef Medline](#)
- Nevin JA, Olson K, Mandell C, Yarensky P (1975) Differential reinforcement and signal detection. *J Exp Anal Behav* 24:355–367. [CrossRef Medline](#)
- Nienborg H, Cumming BG (2009) Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* 459:89–92. [CrossRef Medline](#)
- Ollerenshaw DR, Bari BA, Millard DC, Orr LE, Wang Q, Stanley GB (2012) Detection of tactile inputs in the rat vibrissa pathway. *J Neurophysiol* 108:479–490. [CrossRef Medline](#)
- Ollerenshaw DR, Zheng HJV, Millard DC, Wang Q, Stanley GB (2014) The adaptive trade-off between detection and discrimination in cortical representations and behavior. *Neuron* 81:1152–1164. [CrossRef Medline](#)
- Pouget A, Beck JM, Ma WJ, Latham PE (2013) Probabilistic brains: knowns and unknowns. *Nat Neurosci* 16:1170–1178. [CrossRef Medline](#)
- Reynolds GS (1961) Behavioral contrast. *J Exp Anal Behav* 4:57–71. [CrossRef Medline](#)
- Ritt JT, Andermann ML, Moore CI (2008) Embodied information processing: vibrissa mechanics and texture features shape micromotions in actively sensing rats. *Neuron* 57:599–613. [CrossRef Medline](#)
- Romo R, Salinas E (2003) Flutter discrimination: neural codes, perception, memory and decision making. *Nat Rev Neurosci* 4:203–218. [CrossRef Medline](#)
- Schwarz C, Hentschke H, Butovas S, Haiss F, Stüttgen MC, Gerdjikov TV, Bergner CG, Waiblinger C (2010) The head-fixed behaving rat: procedures and pitfalls. *Somatosens Mot Res* 27:131–148. [CrossRef Medline](#)
- Shadmehr R, Smith MA, Krakauer JW (2010) Error correction, sensory prediction, and adaptation in motor control. *Annu Rev Neurosci* 33:89–108. [CrossRef Medline](#)
- Stüttgen MC, Rüter J, Schwarz C (2006) Two psychophysical channels of whisker deflection in rats align with two neuronal classes of primary afferents. *J Neurosci* 26:7933–7941. [CrossRef Medline](#)
- Stüttgen MC, Yildiz A, Güntürkün O (2011) Adaptive Criterion Setting in Perceptual Decision Making. *J Exp Anal Behav* 96:155–176. [CrossRef Medline](#)
- Stüttgen MC, Kasties N, Lengersdorf D, Starosta S, Güntürkün O, Jäkel F (2013) Suboptimal criterion setting in a perceptual choice task with asymmetric reinforcement. *Behav Processes* 96:59–70. [CrossRef Medline](#)
- Swets JA (1961) Is there a sensory threshold? *Science* 134:168–177. [CrossRef Medline](#)
- Teichert T, Ferrera VP (2010) Suboptimal integration of reward magnitude and prior reward likelihood in categorical decisions by monkeys. *Front Neurosci* 4:1–13. [CrossRef Medline](#)
- Van Horn KS (2003) Constructing a logic of plausible inference: a guide to Cox's theorem. *Int J Approx Reason* 34:3–24. [CrossRef](#)
- Waiblinger C, Brugger D, Whitmire CJ, Stanley GB, Schwarz C (2015) Support for the slip hypothesis from whisker-related tactile perception of rats in a noisy environment. *Front Integr Neurosci* 9:53. [CrossRef Medline](#)
- Waiblinger C, Whitmire CJ, Sederberg A, Stanley GB, Schwarz C (2018) Primary tactile thalamus spiking reflects cognitive signals. *J Neurosci* 38:4870–4885. [CrossRef Medline](#)
- Wichmann FA, Hill NJ (2001a) The psychometric function: I. fitting, sampling, and goodness of fit. *Percept Psychophys* 63:1293–1313. [CrossRef Medline](#)
- Wichmann FA, Hill NJ (2001b) The psychometric function: II. bootstrap-based confidence intervals and sampling. *Percept Psychophys* 63:1314–1329. [CrossRef Medline](#)
- Wilson RC, Nassar MR, Gold JI (2013) A mixture of delta-rules approximation to bayesian inference in change-point problems. *PLoS Comput Biol* 9:e1003150. [CrossRef](#)
- Wolfe J, Hill DN, Pahlavan S, Drew PJ, Kleinfeld D, Feldman DE (2008) Texture coding in the rat whisker system: slip-stick versus differential resonance. *PLoS Biol* 6:e215. [CrossRef Medline](#)