

The Detection of Visual Contrast in the Behaving Mouse

Laura Busse,^{1,2} Asli Ayaz,¹ Neel T. Dhruv,¹ Steffen Katzner,¹ Aman B. Saleem,¹ Marieke L. Schölvinck,¹ Andrew D. Zaharia,¹ and Matteo Carandini¹

¹UCL Institute of Ophthalmology, University College London, London EC1V 9EL, United Kingdom, and ²Werner Reichardt Centre for Integrative Neuroscience, University of Tübingen, 72076 Tübingen, Germany

The mouse is becoming a key species for research on the neural circuits of the early visual system. To relate such circuits to perception, one must measure visually guided behavior and ask how it depends on fundamental stimulus attributes such as visual contrast. Using operant conditioning, we trained mice to detect visual contrast in a two-alternative forced-choice task. After 3–4 weeks of training, mice performed hundreds of trials in each session. Numerous sessions yielded high-quality psychometric curves from which we inferred measures of contrast sensitivity. In multiple sessions, however, choices were influenced not only by contrast, but also by estimates of reward value and by irrelevant factors such as recent failures and rewards. This behavior was captured by a generalized linear model involving not only the visual responses to the current stimulus but also a bias term and history terms depending on the outcome of the previous trial. We compared the behavioral performance of the mice to predictions of a simple decoder applied to neural responses measured in primary visual cortex of awake mice during passive viewing. The decoder performed better than the animal, suggesting that mice might not use optimally the information contained in the activity of visual cortex.

Introduction

In the quest to understand the relationship between neuronal activity and visual perception, an essential tool is the ability to measure visual performance in well controlled behavioral tasks. Such measurements are typically obtained through operant conditioning in monkeys, whose visual system most resembles that of humans, and are particularly fruitful when combined with recordings or perturbations of neuronal activity (Newsome et al., 1989; Britten et al., 1992; Nienborg and Cumming, 2010).

There is increasing interest in applying similar methods to the mouse. Mice have a simpler visual system than primates, with lower spatial acuity and simpler cortical micro-architecture (Chalupa and Williams, 2008). Nonetheless, mice are gaining popularity in visual neuroscience because of the readily available molecular and genetic tools. These tools allow cell-type-specific neurophysiology (Sohya et al., 2007; Kerlin et al., 2010; Runyan et al., 2010; Bock et al., 2011) and exquisite control of neuronal

activity (Huber et al., 2008; Cardin et al., 2009), and thus provide powerful approaches for resolving longstanding debates.

Mouse vision can be assessed behaviorally using practical and robust methods (Pinto and Enroth-Cugell, 2000). Some of these methods involve measurements of reflexive movements of eye or body induced by drifting gratings surrounding the animal (Prusky et al., 2004; van Alphen et al., 2009). Other methods involve training the mouse to swim in a water maze toward a submerged platform, indicated by a visual stimulus (Prusky et al., 2000; Prusky and Douglas, 2004).

These methods, however, have some disadvantages. The techniques based on reflexive movements are restricted to large moving stimuli and therefore cannot assess spatial vision; they do not probe cortical function (Douglas et al., 2005) and could hardly be used to investigate cognitive influences on visual processing. The swimming task, meanwhile, yields only a few dozen trials per day and cannot easily be paired with simultaneous recordings of brain activity.

We sought to overcome these limitations by measuring psychometric curves for two-alternative forced choice (2AFC) in freely moving mice. By adapting methods from various laboratories investigating sensory behavior in the rat (Kepecs et al., 2008; Yang et al., 2008; Meier et al., 2011), we trained mice to routinely perform hundreds of trials per session in a visual detection task. We analyzed the resulting psychometric curves using classical psychophysical methods to infer measures of contrast sensitivity.

While many sessions yielded high-quality psychometric curves, other sessions produced responses with large biases and high error rates. In such sessions, mice followed suboptimal strategies influenced by nonvisual factors such as past choices and expectations of reward. To account for such strategies, we developed a simple model by drawing from theories of value-based decision making (Corrado et al., 2005; Lau and Glimcher, 2005).

Received Dec. 21, 2010; revised June 15, 2011; accepted June 16, 2011.

Author contributions: L.B., S.K., and M.C. designed research; L.B., A.A., N.T.D., S.K., A.B.S., M.L.S., A.D.Z., and M.C. performed research; L.B., N.T.D., and A.B.S. analyzed data; L.B., and M.C. wrote the paper.

This work was supported by Fight for Sight (Project Grant 1779) and by an Advanced Investigator award from the European Research Council (project acronym: CORTEX). L.B. was supported partially by the German Academy of Sciences Leopoldina (BMBF-LPD9901/8-165) and by funds awarded to the Centre for Integrative Neuroscience (DFG Exec 307). N.D. was supported by a Newton International Fellowship from the Royal Society. M.C. holds the Glaxo-SmithKline/Fight for Sight Chair in Visual Neuroscience. We thank Pamela Reinagel for early encouragement and advice on the experimental procedures, and Paul Glimcher for advice on modeling techniques. Eyal Seidemann kindly pointed us toward his doctoral thesis, which had anticipated our model of nonsensory influences by over a decade. We thank Jessica Hill and Julie Scott for technical help, and the Smith-Kettlewell Eye Research Institute for providing space for the initial pilot experiments.

The authors declare no competing financial interests.

Correspondence should be addressed to Laura Busse, Centre for Integrative Neuroscience, University Tübingen, 72076 Tübingen, Germany. E-mail: laura.busse@cin.uni-tuebingen.de.

DOI:10.1523/JNEUROSCI.6689-10.2011

Copyright © 2011 the authors 0270-6474/11/3111351-11\$15.00/0

The resulting model captures the strategies and yields an estimated internal representation of stimulus contrast.

Finally, we compared visual behavior to predictions of a simple decoding model applied to responses of mouse primary visual cortex measured during passive viewing. The decoding model performed generally better than the mice, suggesting that the mice might make only suboptimal use of the information contained in the neural responses.

Materials and Methods

All experimental procedures were conducted according to the US National Institutes of Health Guidelines for the Care and Use of Animals for Experimental Procedures and to the UK Animals Scientific Procedures Act (1986). Experiments were performed at University College London under personal and project licenses released by the Home Office following appropriate ethics review.

Animals. Seven wild-type (C57BL/6J) and five transgenic (two HHrt-TAXK and three TRE/ASTBDN-1) mice performed a visual contrast-detection task for fluid reward. Two of these mice were males, the others were females. The transgenic animals were used as controls for a separate project aimed at studying retinal neovascularization and did not exhibit any retinal abnormalities (Wall et al., 2004). At the beginning of training, animals were between 2 and 8 months old; at the end of testing, they were between 15 and 16 months. The mice were kept on a 12 h light/dark cycle and tests were performed during the light cycle. An additional group of seven wild-type mice (C57BL/6J) were used for electrophysiological experiments.

Water control. Animals in the behavioral study had *ad libitum* access to water only during weekends (typically Friday afternoon to Sunday afternoon). During the rest of the week, they obtained water by performing the task. Signs of possible dehydration were monitored (reduced skin tension, sunken eyes, and marked variations in general behavior) and were absent in all animals. To ensure adequate hydration, we weighed each animal at the beginning and end of each experimental session and compared the weight to a standard weight updated weekly. If weight measured after the session was <90% of the standard weight, the animal would be temporarily taken out of the study and given *ad libitum* access to water until the weight recovered. This condition never occurred.

Apparatus. The choice box (Fig. 1A,B) was a translucent chamber facing an LCD screen (HX192D Hanns.G; mean luminance 110–140 cd/m², refresh rate 60 Hz). The box was divided by translucent walls into three connected areas, each centered on a port: a central port and two choice ports. Each port consisted of a hole large enough to accommodate the animal's snout (1.25 cm diameter). The presence of the snout was monitored by an infrared beam. In front of the hole, a small spout delivered water. Water pressure was obtained by placing the source ~50 cm above the chamber. Reward volume was controlled by solenoid valves (161T010; Neptune Research). Sound speakers attached to the stimulus monitor gave auditory feedback signals during key states of the experiment. Mouse behavior was monitored through a network camera (Linksys; Cisco). This setup was enclosed in a cabinet (Ikea) that provided some sound and light isolation from the surroundings. Software for experimental control and stimulus presentation was custom written in Matlab (MathWorks) with extensions from the Psychophysics toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007).

Task. We used a two-alternative forced-choice (2AFC) task to measure contrast thresholds. The mouse started each trial by inserting its snout into the central port. Upon breaking the infrared beam, it received ~3 μ l of water and a circular window containing a vertical grating was presented 40° to the right or to the left. The grating drifted until the mouse retracted its snout from the central port, and remained stationary thereafter. The grating's spatial and temporal frequency were set to the optimum values found in studies of reflexive optomotor behavior: 0.13 cycles/deg and 1.5 Hz (Umino et al., 2008). Grating contrast varied from trial to trial between 10% and 100%. Grating position (left or right) changed randomly from trial to trial. To avoid learning of spurious patterns in behavior, we reduced the probability of more than three successive stimulus repetitions on the same side or three successive stimulus alternations between sides. The task of the mouse was to nose-poke the

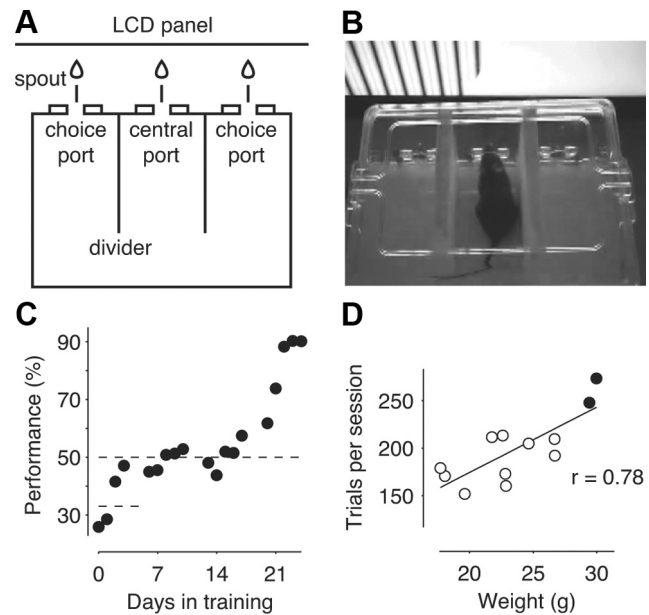


Figure 1. Apparatus and training. **A**, Scheme of the behavioral choice box. **B**, Picture showing the mouse triggering stimulus presentation by placing the snout in the central port. **C**, Mice learn the visual contrast detection within 3–4 weeks of training. Mice were trained for 1 week in stage 2 and 3 weeks in stage 3. Dashed lines indicate chance performance. On days 1–5, chance is 33% for a naive mouse, 50% for a mouse that understands the basic task. $N = 3$. **D**, Heavier mice performed more trials per session. Each data point represents a single animal ($N = 12$). Relation between average number of trials per session and average body weight ($r = 0.78$, $p < 0.01$). Open and filled circles indicate females ($N = 10$) and males ($N = 2$), respectively. Line is given by linear regression.

choice port corresponding to the stimulus location. Upon a correct choice, it received a reward of 6–8 μ l of water. Upon an incorrect choice, it received a 6 s timeout accompanied by a Gaussian noise sound, during which a new trial could not be initiated. This timeout condition also occurred when the mouse did not initiate a trial within 60 s or did not poke a choice port within 30 s of leaving the central port. On each experimental day, mice performed one to three sessions of 15–40 min each. Average response times in correct trials tended to decrease with increasing stimulus contrast, resulting in significant negative correlation coefficients for 11 of 12 animals (Spearman rank correlation, average across animals $\rho = -0.29 \pm 0.16$ SD, $p < 0.001$).

Training. Training was performed in three stages. In the first stage (1–2 d), the mouse learned to explore the choice box and obtain fluid from the ports, and formed a positive association with the reward sound. A stationary grating was presented at each port and the mouse was rewarded with a large amount of water (16 μ l) for inserting its snout into any of the three ports. The stimuli then disappeared and a new trial initiated automatically soon afterward. In the second stage (<1 week), the mouse was exposed to the different trial outcomes (reward vs timeout), formed negative associations with the timeout sound, and learned that it is essential to change ports to collect rewards. In each trial, the grating was presented at only one of the ports and the mouse was rewarded only for choosing that port. The grating was repeated at the same location until the mouse collected the reward. After a success, the grating was always presented at one of the other ports. After incorrect behavior, there was a timeout of 1.8 s. The animal was promoted to the next training stage when it frequently chose different ports. In the third stage (2–3 weeks), the mouse learned the remaining skills: to initiate a trial in the central port, to associate visual stimuli with reward, and to indicate choice by poking the appropriate port. All aspects of the behavioral task were now the same as in the final task, except that (1) the mouse received the same amount of water in the central port (for initiating the trial) as in the choice port (for giving the correct answer) and (2) the timeout following incorrect choices was short (2.5 s). Gradually, these two variables were brought toward their final value. This training phase was over when the

animal’s performance had risen from 50% (chance value) to 80–90%; when this happened, stimuli of lower contrast were introduced.

Data processing. Sessions with <10 trials were excluded. If a session contained three or more consecutive trials in which the mouse did not initiate the task or did not give a response, we deemed the mouse to have lost motivation and we ignored the remaining trials. We concatenated all data obtained within a day and used it for further analyses if it contained at least 50 trials.

Psychometric analysis. For the psychometric analysis, we considered all trials except those that ended in a timeout. We calculated the percentage of rightward choices as a function of signed contrast c , where negative or positive contrast denotes trials in which the stimulus was presented on the left or right side. We fitted these data with the psychometric function

$$\psi(c) = \lambda + (1 - 2\lambda)F(c; \mu, \sigma), \tag{1}$$

where $F(x)$ is a cumulative Gaussian. The parameters μ and σ represent the mean and standard deviation of the underlying Gaussian, respectively, and determine the left–right bias and slope of the psychometric function. The parameter λ represents the lapse rate. This function was fitted via maximum likelihood estimation using `psignifit` (Wichmann and Hill, 2001a). We defined threshold as the standard deviation of the Gaussian. Confidence intervals of the parameters were found by the bootstrap method based on 2000 simulations (Wichmann and Hill, 2001b).

Likelihood analysis. To obtain an overall estimate of contrast threshold for each animal, we calculated the likelihood of each contrast representing the true contrast threshold. For each session i , we computed the estimated mean c_i and standard error e_i of the contrast threshold via bootstrapping (Wichmann and Hill, 2001b). For simplicity, we assumed that the resulting distribution was Gaussian, $G[c_i, e_i]$. We then computed, for each contrast, the likelihood L of the true contrast threshold being c by multiplying the probabilities of observing c across sessions:

$$L(c) = G[c_1, e_1](c) \times G[c_2, e_2](c) \times \dots \times G[c_n, e_n](c), \tag{2}$$

The value of contrast threshold c that we report for each animal is the one that maximized this likelihood.

Probabilistic choice model. To account for nonsensory factors that contribute to the behavioral responses, we used a probabilistic model, a generalized linear model. In the model, the observer makes a decision by tossing a coin (i.e., by sampling from a Bernoulli distribution) with probability p of heads (going right) and $1 - p$ of tails (going left). The log odds of the fairness of the coin is determined by a decision quantity z that is positive if the observer is inclined to go right, and negative if it is inclined to go left. The relation between p and z is given by the sigmoidal (logistic) function

$$p = \frac{1}{1 + e^{-z}} \tag{3}$$

In each trial t , the decision quantity z depends on the signed contrast in the present trial, $c(t)$, and on the success, $s(t - 1)$, and failure, $f(t - 1)$, in the preceding trial:

$$z(t) = v[c(t)] + b_s s(t - 1) + b_f f(t - 1) + b_0, \tag{4}$$

where v weighs the stimulus contrast in the present trial, b_s and b_f weigh the successes and failures in the preceding trial, and b_0 indicates overall bias. The visual term $v(c)$ is negative for stimuli on the left and positive for stimuli on the right; it is an odd-symmetric function of visual contrast, $v(-c) = -v(c)$. Successes s and failures f are sequences of -1 (for left port) and 1 (for right port). They are complementary in that if one is nonzero the other must be zero. However, they can both be zero if the trial was aborted. The bias term b_0 is negative for leftward biases and positive for rightward biases.

Model fitting and simulations. Model fits and simulations were performed on a session-by-session basis. We fitted the model using the Matlab function `glmfit` applied to a matrix defined as follows. The first row was a constant (to estimate b_0). The second row was the sequence $s(t)$

shifted by one trial (to estimate b_s). The third row was the sequence $f(t)$ shifted by one trial (to estimate b_f). The remaining rows (one for each absolute contrast c_1, \dots, c_n used) constituted an indicator matrix with $c_{it} = 1$ if $c(t) = c_i$ (the contrast c_i was shown on the right), $c_{it} = -1$ if $c(t) = -c_i$ (the contrast c_i was shown on the left), and zero otherwise. These latter rows allowed us to estimate the values of the visual weights $v(c_i)$ for each of the possible absolute contrasts c_i . Extending the model to trials further in the past did not significantly improve the goodness of fit as assessed by the model deviance. Trial-by-trial simulations of behavior were performed by drawing randomly from a binomial distribution with p given by Equation 3, using at each time interval the true history of successes and failures experienced by the animal in the preceding trial.

Choice tendency. To evaluate the dependency of behavior on recent history, we measured the tendency to make a choice (left or right) given the previous choice (left or right) and the resulting outcome (success or failure). For instance, to measure the tendency to choose right following a successful left choice, we took all trials in which the previous choice was left and was a success, and counted in how many of them the animal went right, N_{choice} , and in how many of them the stimulus was indeed on the right, N_{stimulus} . We defined tendency as $(N_{\text{choice}}/N_{\text{stimulus}}) - 1$, and computed it for the actual choices by the animal, for the choices predicted by the full model, and for the choices predicted by a vision-and-bias-only model in which the strategy terms b_s and b_f were set to zero. To evaluate the predictions of these models, we correlated the predicted tendencies with the observed ones to obtain a model quality index ranging from -1 (perfectly wrong predictions) through 0 (random predictions) to $+1$ (perfect predictions).

Electrophysiology in awake mice. Seven adult wild-type mice (C57BL/6J, 20–35 g) were implanted with a custom-designed headpost and recording chamber over the left visual cortex using dental acrylic (Superbond C&B; Prestige Dental) under isoflurane anesthesia (5% induction, 1.5–2% maintenance). A $\sim 1 \text{ mm}^2$ craniotomy was performed centered at coordinates 3 mm lateral from the midline and 0.5 mm anterior from lambda. The chamber was sealed with Kwik-Cast (WPI). The animal was administered an analgesic (Rimadyl, 0.01 ml/g s.c.) during surgery and for 3 d after surgery (Rimadyl, oral route). Following recovery (1 week), the animals were handled and acclimatized to the recording room. Mice were head-fixed and placed on an air-suspended Styrofoam ball (Hölscher et al., 2005; Dombeck et al., 2007). Extracellular neural activity was recorded by advancing a 16-channel electrode into visual cortex (model A1X16-3mm-50-413; NeuroNexus Tech) by $\sim 900 \mu\text{m}$. The tissue was allowed to settle for 20 min before the start of each 2–3 h recording session. After each session, the recording chamber was resealed and the animal returned to the home cage until the subsequent recording session (mean 3, maximum 6 sessions per animal). Visual stimuli were presented using custom-written software across three calibrated LCD monitors (HA191, Hanns.G, mean luminance 50 cd/m^2), covering an angle of 140° horizontal and 36° vertical. We recorded spike responses (multiunit activity) to contrast-reversing sinusoidal gratings with a spatial frequency of 0.05 cycles/deg and a temporal frequency of 2 Hz, presented for 1 s over the receptive field. We used eight contrast levels (0, 10, 25, 40, 55, 70, 85, and 100%), each shown 40 times in random order.

Contrast responses. To compare the contrast responses inferred from behavior $v(c)$ to the contrast responses obtained by extracellular recordings of neuronal activity $r(c)$, we fitted a hyperbolic ratio function of contrast (Albrecht and Hamilton, 1982):

$$f(c) = R_0 + R_{\text{max}} \frac{c^n}{c_{50} + c^n}. \tag{5}$$

Here, the parameters R_{max} , c_{50} , and n determine the overall responsiveness, the semisaturation contrast, and the exponent of an accelerating nonlinearity related to spike threshold, respectively. The parameter R_0 allows for a positive baseline.

Calculating neurometric performance. We obtained the population activity r in response to the k th presentation of a stimulus with contrast c by

summing the average firing rates measured at different sites i across the multielectrode array and across experiments:

$$r(c,k) = \sum_i r_i(c,k). \quad (6)$$

Following the approach of Britten et al. (1992), we considered each site to appear twice, once in viewing the side with the stimulus ($c > 0$) and once in viewing the side without the stimulus ($c = 0$). Based on this assumption, we calculated the pair of response distributions, $r(c, \cdot)$ and $r(0, \cdot)$, for each stimulus condition. The fraction of correct choices for the ideal observer is calculated as the area under the receiver operating characteristic (ROC) curve for this pair of distributions (Green and Swets, 1966; Tolhurst et al., 1983; Britten et al., 1992). By summing activity across sites and experiments (Eq. 6), we are ignoring correlations across neurons. This might not be a major limitation when modeling contrast sensitivity due to the broad similarity of contrast response curves across neurons (Montani et al., 2007).

Results

We present our results in four sections. First, we describe the visual detection task and detail how we trained mice to perform it. Second, we analyze the resulting psychometric curves using classical methods and estimate contrast sensitivity. Third, we describe multiple occasions in which the mice were influenced by nonvisual factors, and present a probabilistic choice model that accounts for these factors. Fourth, we compare the visual performance of the animal to that of neurons in V1 of awake, passively viewing mice.

Behavioral training

We trained mice to perform a two-alternative forced-choice (2AFC) contrast-detection task using operant conditioning (Fig. 1A,B). We adapted to mouse vision a three-port nose-poking protocol used previously in rats to probe olfaction (Uchida and Mainen, 2003), hearing (Yang et al., 2008), and vision (Meier et al., 2011). The mouse puts its snout in a central port to trigger the presentation of a visual stimulus on the left or right side. The mouse then indicates the stimulus position by poking its snout into the corresponding choice port. Correct trials are rewarded by water. Stimuli were drifting sinusoidal gratings whose contrast was 100% during training, and randomly chosen from a set of contrasts afterward.

The training consisted of three stages and typically lasted ~4 weeks. First, the mice learned to obtain fluid from the ports, and formed a positive association with the reward sound. Second, the mice were exposed to the different trial outcomes (reward vs timeout), formed negative associations with the timeout sound, and learned that it is essential to change ports to collect rewards. Finally, the mice learned to initiate a trial in the central port, to associate visual stimuli with reward, and to indicate choice by poking the appropriate port. Training ended when the animal's performance rose from 50% (chance value) to 80–90% (Fig. 1C). We found it advantageous to work with heavier mice (and hence with males) because they performed markedly more trials per session (Fig. 1D). Small animals are likely to reach fluid satiation quicker than large animals.

In pilot studies, we identified several aspects of these methods that are necessary for the animals to learn the task. A pilot study determined that the box dividers (Fig. 1A,B) were essential, as they increased the costs of mistakes and therefore discouraged a guessing strategy that ignored visual stimuli. Two additional pilot studies determined that it was essential to give reward in the central port for initiation of a trial and to maintain the stimulus present (albeit stationary) after the mouse exited that port.

A key pilot study established the necessity for a regime of complete water control, in which mice received water only by performing the task in the choice box. A group of five naive mice were trained in a less stringent regime: they received *ad libitum* water every day for 1 h at the end of each training day. After 20 d of phase three training, these animals had still not learned the task (performance <55% correct). They did learn the task, however, once they were subjected to the regime of complete water control.

Analysis of psychometric curves

To estimate contrast sensitivity, we analyzed psychometric curves relating stimulus contrast to the proportion of times the animal chose a port (Fig. 2). To plot contrast on a single axis, we give it a sign: positive for stimuli on the right (R) and negative for stimuli on the left (L). To illustrate the span of our results, we select nine example sessions in five animals (Fig. 2). As expected, the proportion of R choices systematically increases as a function of signed contrast. We fitted these data with a standard psychometric function (Eq. 1), which is determined by three parameters: lapse rate, bias, and contrast threshold.

The lapse rate is the proportion of times the animal made a mistake even though the stimulus was clearly visible. It determines the degree to which the asymptotes of the curve differ from 0% and 100%. Lapse rate was high in all of our data: it was 14%, 11%, and 10% in the first three example sessions (Fig. 2A–C), and it was similarly high across sessions in all animals ($8.5 \pm 1.6\%$, SD; $N = 12$). These values are significantly larger than zero ($p < 0.001$).

The bias reflects a preference of the animal for one or the other choice. It describes the degree to which the curve is displaced along the contrast axis: if bias is absent, the curve crosses the ordinate of 50% at zero contrast. Bias was small in the first three example curves (1.3%, –4.3%, and 0.4%) (Fig. 2A–C), but in other cases it could be substantially higher. For instance, a given mouse could show a strong positive bias of 34.5% on day 1 (Fig. 2D), a strong negative bias of –57.7% 36 d later (Fig. 2E), and a negligible bias of 1.3% 48 d later (Fig. 2F). These large biases for one side or the other tended to vary slowly over sessions and were not random because they were often consistent across animals (Fig. 3). They were probably due to common factors such as minor imbalances (on the order of 1 μ l) in the size of reward. Indeed, bias was substantially reduced once we adopted a policy of more accurate valve calibration.

The contrast threshold yields an assessment of contrast sensitivity. It is inversely related to the slope of the psychometric curve: steeper curves correspond to lower thresholds, i.e., higher sensitivity. We define contrast threshold as the increment in R contrast needed to go from 50% R choices to 68% of the upper bound in R choices (where the upper bound is 100% minus the lapse rate). In the first three example datasets, contrast thresholds were rather similar at 10%, 10%, and 11% (Fig. 2A–C), and could be estimated with high confidence (95% confidence intervals were 5–19%, 5–16%, and 5–18%). These values were typical of sessions with particularly reliable estimates of the psychometric curve. These sessions tended to occur commonly toward the end of the week, presumably due to increased levels of motivation (the animals received *ad libitum* access to water during the weekend). The estimates of contrast threshold varied from session to session, with the most reliable estimates corresponding to lower thresholds. Consider, for example, three sessions from the same mouse (Fig. 2G–I). The first yielded a low threshold (13%) and a rather tight confidence interval (7–23%). The second yielded a higher threshold (20%) and an extremely wide confidence interval (5–80%). The third yielded an even higher threshold (41%) and a similarly wide confidence interval (8–70%). These noisy

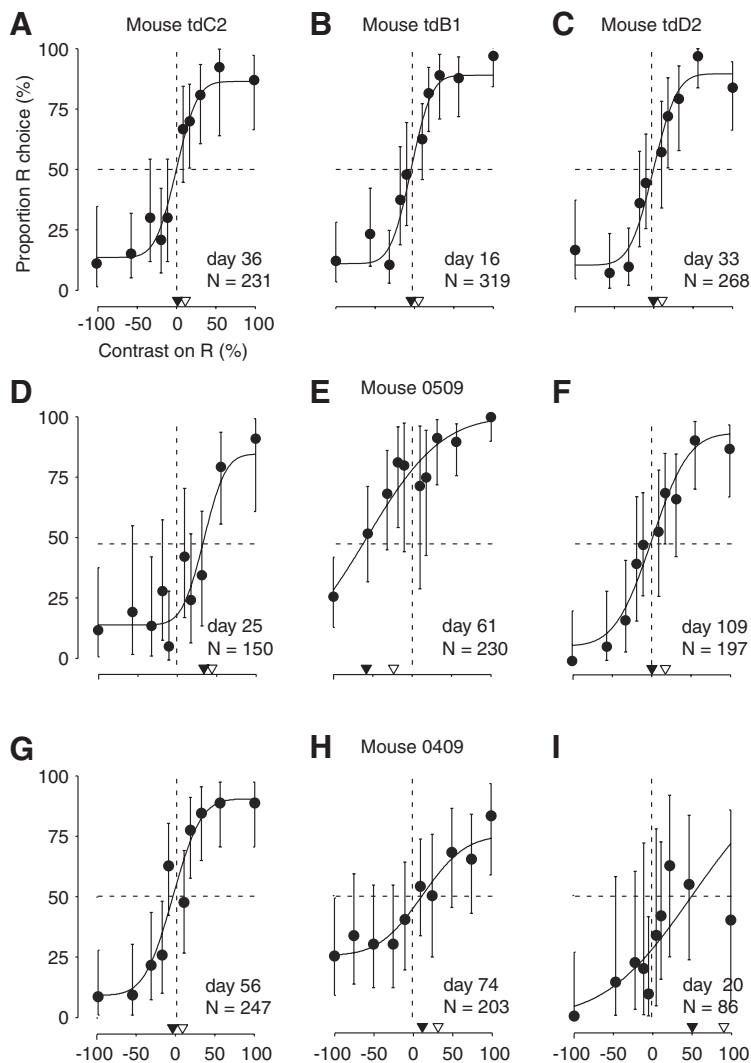


Figure 2. Psychometric functions for contrast detection in mice. Each panel represents data acquired in a single 15–40 min session. Proportion of right port (R) choices is plotted as a function of signed stimulus contrast, with negative contrasts representing stimuli shown on the left. *N*, Number of trials per session. Filled triangles indicate bias—the contrast at which the animal went 50% of the times to the left and 50% of the times to the right. The difference in contrast between the filled triangles and the open triangles indicates contrast threshold. Error bars are binomial confidence intervals. **A–C**, Three sessions with low bias, yielding precise estimates of contrast threshold. Sessions 100331, 100310, 100326. **D–F**, Three sessions for a single mouse at different days of testing, showing a strong bias for L choices (34.5%, **D**), a strong bias for R choices (−57.7%, **E**) and a negligible bias (1.3%, **F**). Sessions 090910, 091016, 091203. **G–I**, Three sessions for a single mouse at different days, showing decreasing degrees of reliability of the estimated threshold, with confidence intervals 6.7–23.1% in **G**, 4.7–80.3% in **H**, and 8.0–69.7% in **I**. Sessions 091204, 100427, and 090914.

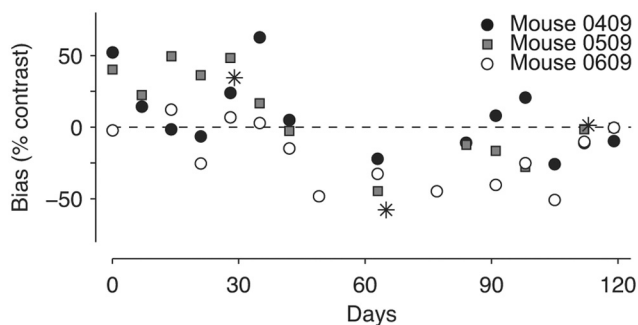


Figure 3. Small imbalances in fluid reward cause large biases that change gradually with time and are shared across animals. Each data point represents the average bias across all training sessions within a week. Different symbols represent data from three different animals. Asterisks indicate bias for the sessions in Figure 2, **D–F**. The animals favored the left port in the first few weeks and the right port in the subsequent weeks, most likely because of tiny imbalances in fluid reward, in the order of 1 μ l.

sessions are simply not usable to estimate contrast threshold using classical psychophysics.

We confirmed that noisier data tended to yield higher estimates of threshold by rank-ordering the sessions based on estimated threshold (Fig. 4A); clearly, sessions with high contrast thresholds tended to be associated with large confidence intervals. The noisier sessions were often the ones showing large biases, whether positive or negative (Fig. 4B), regardless of number of trials (Fig. 4C). These biases undermined our ability to obtain high-confidence estimates of threshold.

Taking into account the confidence in each session’s estimate of threshold, we obtained an overall estimate of threshold for each animal (Fig. 4D). For each animal, we computed the likelihood of each contrast being the true contrast threshold, and defined the contrast threshold as the contrast where this likelihood peaks. For example, the entirety of the session-by-session measurements for an individual mouse (Fig. 4A) is most consistent with a single contrast threshold of 19.0% (Fig. 4D). This value was typical of our sample; the threshold averaged across all animals was 21.3% (± 3.9 , SD; $N = 12$). This estimate represents a consensus value that summarizes the responses obtained in all sessions and all animals. Such values are high compared to those of primates and humans, whose contrast threshold is in the order of ~ 0.1 – 0.3% (De Valois et al., 1974; Kiorpes and Kiper, 1996).

Probabilistic choice model

As shown by the analysis above, our measurements can be analyzed using classical psychophysics to yield estimates of contrast threshold in mouse. Sessions with little bias and lapse rate yield psychometric curves of high quality (Fig. 2), allowing a precise measure of the contrast that gives a specified level of performance. This classical approach is satisfactory for most purposes, but it suffers from at least two shortcomings.

First, the classical approach does not estimate how the inferred neural responses depend on contrast. The contrast responses of neurons in various stages of the visual system are nonlinear (Sclar et al., 1990), so one cannot extrapolate their shape from a single measurement of slope made at threshold. We would like the psychophysical measurements to estimate the neural response elicited by each contrast.

Second, the many sessions with high lapse rate or strong bias constitute a challenge for classical psychophysics. The classical approach does not explain the common tendency of the mice to guess even when confronted with a patently visible stimulus. For example, a lapse rate of 25% means that even when presented with a clearly visible, high-contrast visual stimulus, the mouse acted randomly in half of the trials (Fig. 2H). Why would an observer make these apparently random choices?

To address these questions, we developed a probabilistic model of choice that takes into account both an internal neural response to contrast and systematic influences of nonvisual factors on the choices of the animal. To obtain this model, we adopted a framework that had been developed to study value-based decision making (Corrado et al., 2005; Lau and Glimcher, 2005) and endowed this framework with a sensory term (Seidemann, 1998; Gold et al., 2008). The result is a powerful model that accounts for both sensory and nonsensory determinants of choice.

In the probabilistic model, the observer makes a decision by flipping a coin whose fairness depends on the sum of a sensory term, two strategy terms, and an overall bias (Fig. 5). The sensory term $v(c)$ is applied to the contrast c of the stimulus in the present trial. The fitted values of $v(c_i)$ are the estimated neural response to each contrast c_i in the session. The strategy terms weigh the outcome—success or failure—of the previous trial. Their weights b_s and b_f can be positive (indicating a tendency to return to the previous choice) or negative (indicating the opposite tendency). For instance, if b_s is positive and b_f is negative, then the overall strategy is to repeat the previous choice if it was a success and to change port if it was a failure. A bias term b_0 determines an overall preference for left or right due, for example, to different estimates of average reward. The sum of all these terms is the decision variable z , which is positive if the animal is inclined to choose right and negative if it is inclined to choose left. The decision variable z determines through a logistic function the probability p of choosing the right port in the coin flip (in other words, z equals the log odds of the coin flip).

We fitted the probabilistic model to the data of each session, obtaining weights for both visual and strategy terms. Fits were obtained through logistic regression. As we show next, the model was able to predict behavior, capturing the magnitude of both sensory and nonsensory influences on choice. To quantify the performance of the model and to compare it to a simpler model with only the sensory term and bias, we measured its deviance. We found that the full model performed significantly better than the simpler model in 64% of all sessions (643 of 998 sessions, $p < 0.05$). Even when we concentrated on the 20% of sessions with the lowest thresholds for each animal, where the influences of nonsensory terms are presumably weakest, the full model still had more predictive power in 55% of the sessions (113 of 204 sessions, $p < 0.05$).

For some sessions the visual weights and the bias term were all that was needed (Fig. 6A–C). For instance, in our first example session, the visual weights were fairly large (a value of 2 for z yields $p = 0.88$, i.e., it enforces a choice in 88% of the cases) and the strategy weights were negligible (Fig. 6A). The resulting psychometric responses were strongly driven by contrast, both for the actual measurements (Fig. 6B, dots) and for the responses predicted by the model (Fig. 6B, gray area). The tendency of the animal to choose L or R (Fig. 6C, left) differed due to a bias for left choices, but it depended little on whether the preceding choice was on L or R or whether it was a success or a failure. The full model captured this tendency (Fig. 6C,

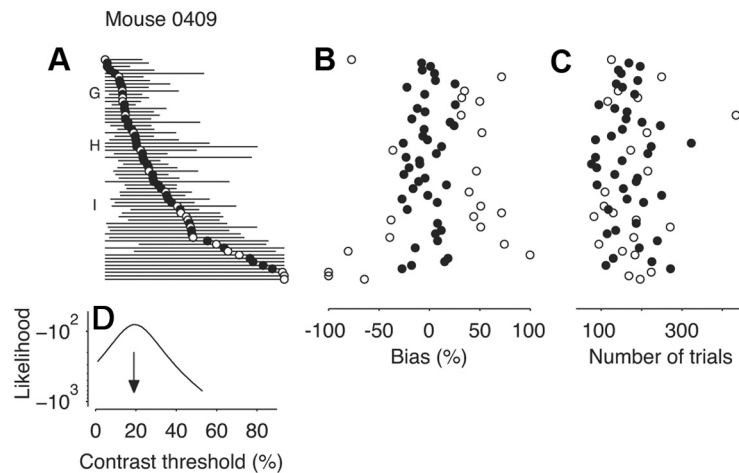


Figure 4. The distribution of contrast thresholds estimated across days in a single mouse is consistent with a single estimate of $\sim 20\%$. **A**, Contrast thresholds for individual sessions, sorted in increasing order. Symbols indicate estimated threshold and lines indicate 95% confidence intervals obtained by bootstrapping. Open symbols indicate sessions in which bias exceeded $\pm 30\%$ contrast. Letters G–I indicate sessions used as examples in Figure 2, G–I, respectively. **B**, Corresponding biases. **C**, Number of trials. **D**, Likelihood of the distribution of contrast thresholds. The maximum corresponds to the single estimate that is consistent with all the observations and is located at 19% contrast (arrow).

middle), but so did a simpler model involving only the visual terms and the bias term (Fig. 6C, right). In accordance with these observations, there is no difference in deviance between the two models ($p = 0.99$).

In other sessions, however, the strategy terms played a fundamental role (Fig. 6D–H). For instance, in our second example session, the sensory weights were smaller and both strategy weights were negative (Fig. 6D). The strategy weights approached the magnitude of visual weight for stimuli of substantial contrast, indicating strong influence of nonvisual factors. Indeed, the corresponding psychometric data (Fig. 6E) were less related to stimulus contrast, with a higher lapse rate and higher threshold than in the previous example. The effects of history (i.e., a tendency to switch after a success) were captured by the full model but not by the model without strategy terms ($p < 0.001$) (Fig. 6F). In our third example session (Fig. 6G–I), contrast sensitivity was rather high (the contrast threshold was 19%) and visual weights were substantial, yet the strategy terms were also required to predict accurately the dependence of behavior on history. As confirmed by the difference in deviance, the full model makes significantly better predictions than the model with only sensory terms and bias ($p < 0.001$).

The strategy terms were important to explain the data in a majority of sessions (Fig. 7). We computed a model quality index that measures the similarity in the history-dependence of behavior predicted by the model and exhibited by the observer (Fig. 6C). Model quality index was much higher for the full model (0.90 ± 0.16 , SD) (Fig. 7A) than for the simpler model that ignored the history terms (0.59 ± 0.34 , $p < 0.001$) (Fig. 7B). Finally, in many sessions, both strategy weights were negative (Fig. 7C), indicating that the strategy was often more directly related to choice than to reward or failure: the animals tended to alternate between ports regardless of outcome ($58.3 \pm 15.0\%$, SD, of sessions; $N = 12$ mice). Thus, the distribution of signs for the two strategy weights is biased toward negative (χ^2 , $p < 0.001$).

The probabilistic choice model gave estimates of perceptual sensitivity that were on a par or superior to those obtained directly from the psychometric curves (Fig. 8). First, the two approaches yielded highly consistent estimates of overall bias (Fig.

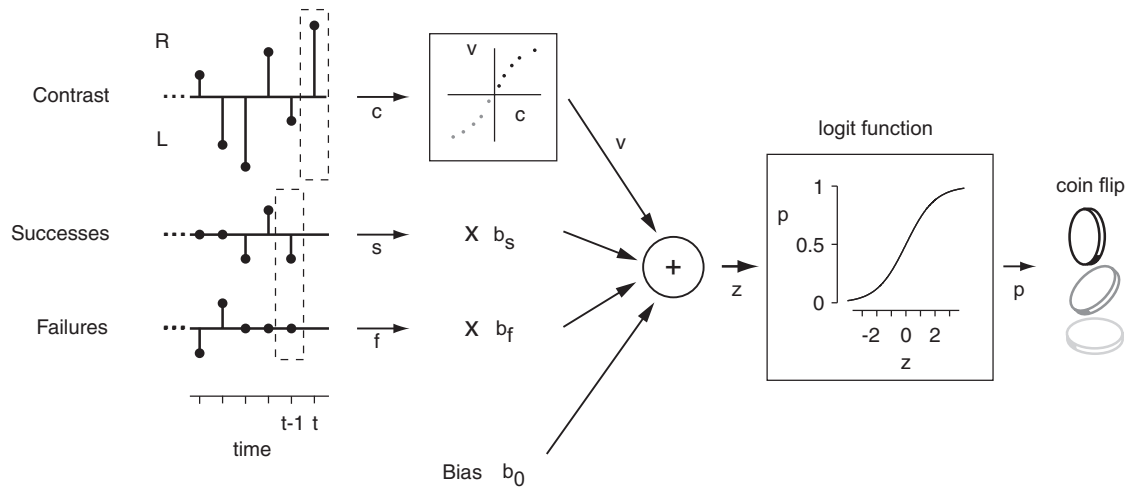


Figure 5. Schematic illustration of the probabilistic choice model. The animal’s choice is modeled as a weighted sum of a sensory term (depending on contrast), two strategy terms (weighing past-trial success and failure), and a bias term. Predictors for an example trial are indicated by the dashed boxes, fitted weights are given by $v(c)$ (contrast), b_s (success), b_f (failure), and b_0 (bias). The resulting variable z is passed through a logistic function that yields a probability for right-port choices (p). Trial-by-trial choices are simulated by randomly drawing from a Bernoulli distribution with parameter p . This step corresponds to flipping a coin, with p being the fairness of the coin.

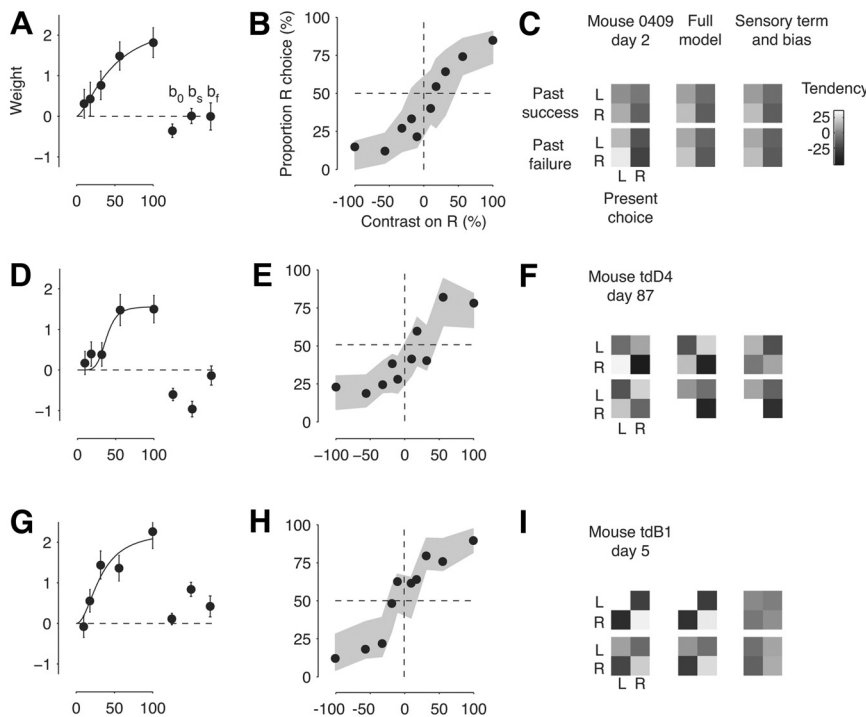


Figure 6. The probabilistic choice model fitted to data. **A**, Parameters fitted by logistic regression to data from one session (Mouse 0409, session 090813). They determine the sensory term $v(c)$, the bias b_0 , and the strategy weights b_s and b_f for success and failure. The solid line is the fit of a hyperbolic ratio function. **B**, The model predicts the corresponding psychometric data. Shaded area indicates the 95% confidence interval of responses simulated by the fitted model. **C**, The model predicts the history-dependence of behavior. Gray level indicates the tendency to choose L or R in the present trial given the possible outcomes of the previous trial: success on the L, success on the R, failure on the L, failure on the R. Columns show the mouse behavior (left), the prediction of the probabilistic model (middle), and the prediction of the model with strategy weights set to zero (right). In this example, strategy weights are low, and both models perform equally well. **D–F**, Another example session (animal tdD4, session 100630). The model predicts both the leftwards bias and a strong negative weight for past success. Only the full model can capture the tendency to switch after a success. **G–I**, A third example (animal tdB1, session 100212). Again, the full model is needed to capture the history effects.

analysis (Fig. 8B). In the choice model, threshold was defined as the lowest contrast at which the visual weights departed significantly from zero. While the estimates of threshold based on the parameter σ of the psychometric function and the visual weights of the choice model were broadly correlated, the thresholds obtained by the choice model rarely exceeded 50%. Finally, across animals, the median threshold given by the choice model was closely related to the threshold obtained from the psychometric function (Fig. 8C).

Comparison with contrast responses in primary visual cortex

Finally, we asked whether the behavioral contrast sensitivity of the mice could be mediated by the responses of neurons in primary visual cortex (V1). We measured the contrast response functions of V1 neurons in awake, passively viewing mice and asked how they compare with the sensory term inferred by the choice model that describes the behavioral data. This sensory term is composed of visual weights whose dependence of contrast constitutes an inferred response of visual neurons. How does this inferred response compare with the response of V1 neurons?

To obtain contrast response functions, we recorded multiunit activity elicited by stimuli of different contrast in V1 of awake mice (Fig. 9A–C). The contrast responses were well fitted by a hyperbolic ratio function (Albrecht and Hamilton,

1982) with a mean explained variance of 77% ($N = 85$), allowing us to summarize the responses using the fitted curves. As expected, neural responses to the different levels of contrast in-

1982) with a mean explained variance of 77% ($N = 85$), allowing us to summarize the responses using the fitted curves. As expected, neural responses to the different levels of contrast in-

creased and eventually saturated. For example, a typical recording gave a semisaturation contrast, c_{50} , of 42% and an exponent, n , of 1.8 (Fig. 9A). This behavior was representative of the population of recordings (Fig. 9B) where the median semisaturation contrast was 34% and exponent was 2.3. These measures, however, varied widely across sites (Fig. 9C). For example, the 10% most sensitive and the 10% least sensitive contrast responses exhibited a $c_{50} < 11\%$ and $> 50\%$ (Fig. 9C, gray triangles). Similarly, the recording sites also varied broadly in the slope of the contrast response functions, captured by the exponent n of the hyperbolic ratio function.

These neural responses resemble the estimated internal responses inferred by the choice model that describes the behavioral data (Fig. 9D). The distributions of semisaturation contrasts for the internal representation of contrast inferred by the model had a median c_{50} of 39%, while the 10% most and least sensitive functions exhibited a c_{50} of 14% and 61%.

Finally, to obtain a more direct comparison between V1 responses and behavioral performance, we returned to the psychometric data, concentrating on the very best sessions, which provided the lowest thresholds for contrast detection. We asked whether the choices made by the animals in those sessions could be mediated by activity in V1. If so, we would expect to be able to decode V1 activity and obtain performance that is equal or better than that of the animal. In fact, considering that the animals often used suboptimal strategies based on prior choices, we might find that the performance predicted by V1 activity is substantially superior to the performance of the animal.

To decode V1 activity, we applied a simple model to the neural responses (Fig. 10). Following the approach of Britten et al. (1992), we computed the response of a population of neurons to multiple presentations of a given stimulus contrast ($c > 0$) and asked how well an ideal observer could distinguish this response distribution from the distribution of responses to the gray screen ($c = 0$). Separately for each contrast, we calculated neurometric performance from the area under the ROC curve of these distributions (Green and Swets, 1966; Tolhurst et al., 1983; Britten et al., 1992). In the model, we varied the size of the population contributing to the response distributions between 10 and 85 recording sites. We found that with increasing population size, neurometric performance was increasingly steeper and had its asymptote at progressively wider values (Fig. 10A).

Only the smallest population yielded thresholds similar to the best thresholds of the animals in the behavioral study (Fig. 10B). To quantify the effects of population size on predicted contrast threshold, we fitted the neurometric performance data with the same psychometric function that we used to fit the psychophysical data (Eq. 1). We found that contrast thresholds of the ideal observer reading out populations of 10 and 15 sites were 10.2% and 9.0%, respectively, which is broadly consistent with the best thresholds of $\sim 10\%$ obtained in our animals. Applying the decoder to larger populations resulted in progressively lower thresholds. Furthermore, the decoder predicts small lapse rates of $\sim 4\%$ even for the smallest sample size, while the animals had an average lapse rate of 8.5%.

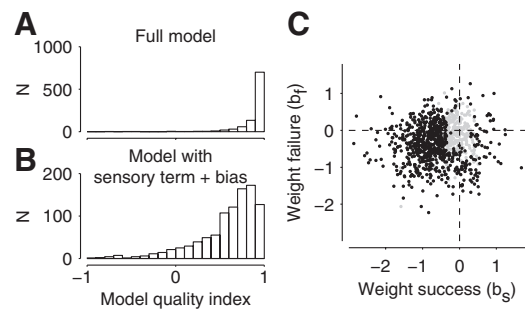


Figure 7. The probabilistic choice model makes accurate predictions of behavior when strategy terms are included. **A**, Model quality index for the full model, including weights for past successes and failures. The model quality index ranges from -1 (bad predictions) to $+1$ (perfect predictions). **B**, Model quality index for the model with only the sensory term and overall bias. **C**, Distribution of strategies across animals and sessions. Black points correspond to sessions in which at least one of the weights is significantly different from zero.

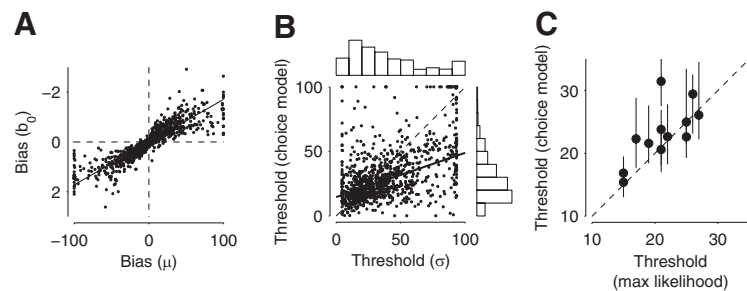


Figure 8. Comparison of key parameters of the psychometric analysis and the probabilistic choice model. **A**, Comparison of bias. The parameter b_0 of the choice model is plotted as a function of the parameter μ of the psychometric function for all sessions and animals. Note that a positive bias in the psychometric analysis denotes a preference for the left port, which corresponds to a negative bias in the choice model. The solid line is given by robust regression (offset = 0.006, $p = 0.5$; slope = -0.02 , $p < 0.001$). **B**, Comparison of contrast threshold. The contrast threshold as given by the choice model is plotted for all sessions as a function of parameter σ of the psychometric function. The solid line is given by robust regression (offset = 14.36, $p < 0.001$; slope = 0.34, $p < 0.001$). **C**, For each animal, the median threshold across sessions as obtained by the choice model is plotted as a function of the overall threshold given by the maximum likelihood analysis (offset = 6.0, $p = 0.24$; slope = 0.78, $p < 0.01$).

While the behavioral responses are broadly consistent with neural responses in visual cortex, this analysis suggests that mice do not use signals in area V1 optimally. This analysis, however, can only constitute a coarse comparison between neural and behavioral signatures of contrast sensitivity, because the neural activity was not been acquired during the same trials as the behavioral data. A more detailed and informative comparison will have to be performed by recording from neuronal populations as the same time as the animal is performing a task.

Discussion

Using a two-alternative forced choice (2AFC) method in freely moving mice, we measured psychometric functions for contrast detection and showed that mice can be trained to perform several hundred trials per session, allowing the estimate of high-quality psychometric functions. We also found substantial nonsensory influences on mouse behavior and proposed a simple generalized linear model to capture these influences. Finally, we measured contrast response functions in mouse V1 and used an ideal observer analysis to predict neurometric curves for contrast. A comparison of predicted and measured behavior revealed that the model generally outperforms the animals, suggesting that the mice do not optimally use the information about stimulus contrast contained in the population response. Together, this work brings to the mouse, the key species in biomedical research, tech-

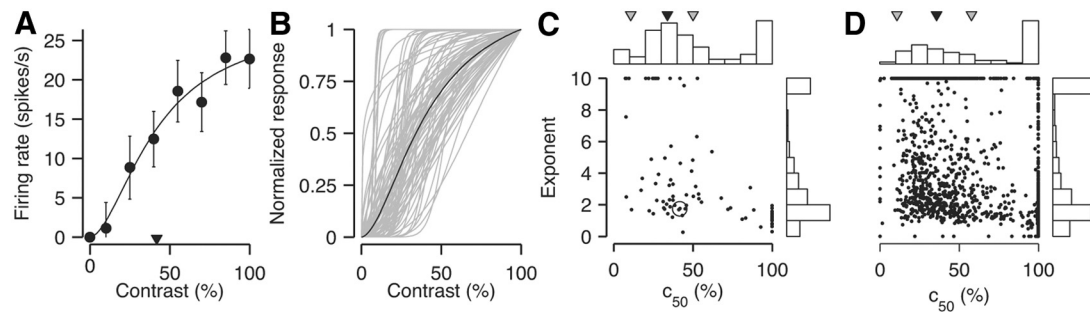


Figure 9. Contrast response functions in visual cortex of awake mice and comparison with inferred internal representation of contrast as revealed by the probabilistic model. **A**, Representative example of a contrast response function measured at a single recording site. The solid line is a fit by the hyperbolic ratio function. Triangle indicates semisaturation contrast. **B**, Contrast responses of all sites in the sample ($N = 85$). Black line corresponds to the example in **A**. **C**, The relationship of semisaturation contrast and exponent across the population of recorded sites. Circle corresponds to the example in **A**. Triangles indicate the semisaturation contrasts for the median responses across the neuronal population (black triangle, $c_{50} = 34\%$) and for the 10% most sensitive and least sensitive contrast responses (gray triangles, $c_{50} < 11\%$ and $> 50\%$). **D**, The relationship of semisaturation contrast and exponent for the inferred internal representation of contrast by the probabilistic model across all animals and sessions.

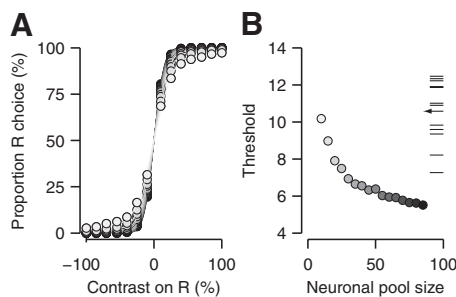


Figure 10. Comparison of performance of a simple decoding model applied to V1 contrast responses with the best behavioral performance of the animals in the behavioral study. **A**, Psychometric curves predicting the proportion of rightward choices as a function of stimulus contrast. The fill color of the data points (white to black) indicates the size of the population available to the decoder (10–85 sites). Lines correspond to fitted psychometric functions (Eq. 1). **B**, Contrast threshold as a function of the size of the neuronal pool. For comparison, small horizontal lines at the left side indicate examples for the best contrast thresholds for each animal in the behavioral study. The arrow points at the mean of those measurements.

niques that have been hitherto used with enormous success to probe visual perception in humans and nonhuman primates.

The three-port choice method that we have adapted to the study of mouse vision has some disadvantages over the most common tests probing mouse vision, the optomotor drum (Prusky et al., 2004) and the Y-shaped water maze (Prusky et al., 2000; Prusky and Douglas, 2004): it takes 3–4 weeks of training compared with minutes or days for the other tasks, and it yields estimates of varying quality across sessions. By comparison, the existing methods are more robust, easier, and quicker.

However, the three-port choice method also has major advantages over these existing methods. First, our task allows investigating spatial vision and top-down influences on the processing of visual information. This would not be possible with the optomotor drum, which requires large stimuli drifting sideways and which probes responses of subcortical structures (Douglas et al., 2005). Second, our task yields hundreds of trials per day and could be readily combined with chronic microelectrode recordings (Kepecs et al., 2008; Yang et al., 2008) or photostimulation (Huber et al., 2008). Little of this would be possible with the swimming task.

Our methods, just like those based on swimming (Prusky et al., 2000), probe visual perception in mice using a 2AFC design. This design has multiple advantages over simpler Go-NoGo tasks. In Go-NoGo tasks, animals can have large biases for indi-

cating presence or absence, which might change during the course of the session due to changing levels of impulsivity and motivation. By contrast, the 2AFC design forces the observer to place its internal criterion at the neutral point, because in every trial it has to indicate the presence of a stimulus in one location and the absence in another. Furthermore, observers usually perform better in 2AFC than Go-NoGo tasks, as they can sample simultaneously from the noise and the noise-plus-signal distribution (Gescheider, 1997; Macmillan and Creelman, 2005).

In the future, the three-port choice method can be extended and improved. The method could be easily extended to discrimination and to more complex visual stimuli (Morton et al., 2006). There is the intriguing possibility of obtaining these measurements in a high-throughput version of the task performed by the animals in the home cage (Meier et al., 2011). Reliability of performance could potentially be enhanced by further increasing the costs for incorrect responses, e.g., by delivering an air-puff after a mistake (Andermann et al., 2010; O'Connor et al., 2010) or by providing a reward that is more appetitive than plain water. In combination, such measures might reduce the high lapse rates that we encountered in our mice. To avoid biases for one or the other choice port, calibration and maintenance of the solenoid valves is crucial since, as we have shown, minor imbalances of reward in the order of 1 μ l can have a dramatic impact on behavior. Finally, to facilitate neurophysiological measurements, a head-fixed version of the task could be developed by training mice to perform a similar task on a spherical treadmill in a virtual reality environment (Hölscher et al., 2005; Harvey et al., 2009). This extension would achieve the same level of stimulus control and ease of brain access as recently achieved in a visual Go-NoGo design (Andermann et al., 2010).

Our estimates of contrast threshold of $\sim 10\%$ falls in the middle of a wide distribution of published thresholds for mice, which range from $< 1\%$ (van Alphen et al., 2009) to 24% (Schmucker and Schaeffel, 2006). Studies relying on optomotor eye and head reflexes typically yield lower thresholds than the swimming task (Prusky and Douglas, 2004). Furthermore, different studies have used different definitions of threshold. Some studies define contrast threshold as the minimum contrast at which a response can be elicited (Prusky et al., 2000; Schmucker and Schaeffel, 2006; van Alphen et al., 2009). Others define it as the contrast that yields 70% correct responses (Prusky and Douglas, 2004; Umino et al., 2008). Yet, other studies extrapolate their measurements to zero response to obtain a contrast threshold (Porciatti et al., 1999). In

our own measurements, we found average thresholds of $\sim 20\%$ across animals and sessions and best thresholds of $\sim 10\%$. These wide ranges of results show that care needs to be taken when comparing contrast thresholds across studies.

Though our data could clearly be analyzed using classical psychophysics, we found that numerous aspects of the responses called for a novel modeling approach. Indeed, choice behavior in the mouse depended partly on nonsensory factors such as past history of failures and rewards. In the absence of a sensory stimulus, it is well known that behavior can be influenced by implicitly learned reinforcement contingencies (Herrnstein, 1961) or past choices (Corrado et al., 2005; Lau and Glimcher, 2005). We extended this approach to the case of sensory uncertainty and obtained a model that quantifies both the observer's sensory responses and behavioral strategies.

Our model of nonsensory influences can be applied in psychophysical experiments to obtain estimates of internal variables contributing to the behavioral response. This approach might be particularly useful in experiments with children or patients who cannot be extensively trained or might be more susceptible to nonsensory decision factors. Beyond this potential application, we have shown that the model yielded—compared with classical psychophysics—better predictions in more than half of the data-sets even when we concentrated on the best psychometric data for each animal. Likewise, Seidemann (1998) found that a similar model could summarize well the responses of highly trained monkeys. Finally, our model could be used to refine the training of observers by online fitting and instantaneous adjustments in the task to counterbalance and minimize behavioral strategies. Thus, our proposed approach is likely to be useful to a wide community, including those who work with humans.

Finally, we compared behavioral performance of the animals to the predictions of an ideal observer applied to the population responses in area V1 of awake, passively viewing animals. We found that predicted contrast thresholds based on neurometric performance were similar to the best contrast thresholds of the animals in the behavioral study, but only when considering small neural pool sizes for the prediction; when increasing the neural pool, predicted thresholds were much lower than those observed behaviorally. Furthermore, the decoder failed to capture the rather high lapse rates of the animals. This suggests that animals did not optimally use for task performance the information contained in the pooled activity of primary visual cortex and is consistent with measurements made in behaving primates (Chen et al., 2006, 2008). However, since neural activity used for decoding has not been acquired during the same trials as the behavioral data, this analysis can only constitute a coarse comparison between neural and behavioral signatures of contrast sensitivity.

Our comparison of neural responses and perceptual behavior is only the first step to a much more rigorous examination: future studies of mouse vision should pair electrophysiological recordings with visual behavior to investigate, on a trial-by-trial basis, how neuronal responses relate to behavior. Because of the possible importance of neuronal correlations (Averbeck et al., 2006), it would be beneficial to record large numbers of individual neurons at the same time (Cohen and Maunsell, 2009). Training head-fixed mice to perform visual tasks on a spherical treadmill in a virtual reality environment (Hölscher et al., 2005; Harvey et al., 2009) would be an ideal method to test thoroughly the relationship between responses of cortical neurons and the detection of visual contrast by behaving mice.

Notes

Supplemental material for this article is available at www.carandini.com. This video shows a mouse engaged in contrast detection task in the laboratory. This material has not been peer reviewed.

References

- Albrecht DG, Hamilton DB (1982) Striate cortex of monkey and cat: contrast response function. *J Neurophysiol* 48:217–237.
- Andermann ML, Kerlin AM, Reid RC (2010) Chronic cellular imaging of mouse visual cortex during operant behavior and passive viewing. *Front Cell Neurosci* 4:3.
- Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. *Nat Rev Neurosci* 7:358–366.
- Bock DD, Lee WC, Kerlin AM, Andermann ML, Hood G, Wetzell AW, Yurgenson S, Soucy ER, Kim HS, Reid RC (2011) Network anatomy and in vivo physiology of visual cortical neurons. *Nature* 471:177–182.
- Brainard DH (1997) The Psychophysics Toolbox. *Spatial Vision* 10:433–436.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12:4745–4765.
- Cardin JA, Carlén M, Meletis K, Knoblich U, Zhang F, Deisseroth K, Tsai LH, Moore CI (2009) Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature* 459:663–667.
- Chalupa LM, Williams RW, eds (2008) *Eye, retina, and visual system of the mouse*. Cambridge, MA: MIT.
- Chen Y, Geisler WS, Seidemann E (2006) Optimal decoding of correlated neural population responses in the primate visual cortex. *Nat Neurosci* 9:1412–1420.
- Chen Y, Geisler WS, Seidemann E (2008) Optimal temporal decoding of neural population responses in a reaction-time visual detection task. *J Neurophysiol* 9:1366–1379.
- Cohen MR, Maunsell JH (2009) Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci* 12:1594–1600.
- Corrado GS, Sugrue LP, Seung HS, Newsome WT (2005) Linear-nonlinear-Poisson models of primate choice dynamics. *J Exp Anal Behav* 84:581–617.
- De Valois RL, Morgan H, Snodderly DM (1974) Psychophysical studies of monkey vision. 3. Spatial luminance contrast sensitivity tests of macaque and human observers. *Vision Res* 14:75–81.
- Dombeck DA, Khabbaz AN, Collman F, Adelman TL, Tank DW (2007) Imaging large-scale neural activity with cellular resolution in awake, mobile mice. *Neuron* 56:43–57.
- Douglas RM, Alam NM, Silver BD, McGill TJ, Tschetter WW, Prusky GT (2005) Independent visual threshold measurements in the two eyes of freely moving rats and mice using a virtual-reality optokinetic system. *Vis Neurosci* 22:677–684.
- Gescheider GA (1997) *Psychophysics: the fundamentals*, 3rd edition. Hillsdale, NJ: Erlbaum.
- Gold JJ, Law CT, Connolly P, Benucci S (2008) The relative influences of priors and sensory evidence on an oculomotor decision variable during perceptual learning. *J Neurophysiol* 100:2653–2668.
- Green DM, Swets JA (1966) *Signal detection theory and psychophysics*. New York: Wiley.
- Harvey CD, Collman F, Dombeck DA, Tank DW (2009) Intracellular dynamics of hippocampal place cells during virtual navigation. *Nature* 461:941–946.
- Herrnstein RJ (1961) Relative and absolute strength of response as a function of frequency of reinforcement. *J Exp Anal Behav* 4:267–272.
- Hölscher C, Schnee A, Dahmen H, Setia L, Mallot HA (2005) Rats are able to navigate in virtual environments. *J Exp Biol* 208:561–569.
- Huber D, Petreanu L, Ghitani N, Ranade S, Hromádka T, Mainen Z, Svoboda K (2008) Sparse optical microstimulation in barrel cortex drives learned behaviour in freely moving mice. *Nature* 451:61–64.
- Kepecs A, Uchida N, Zariwala HA, Mainen ZF (2008) Neural correlates, computation and behavioural impact of decision confidence. *Nature* 455:227–231.
- Kerlin AM, Andermann ML, Berezovskii VK, Reid RC (2010) Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron* 67:858–871.
- Kiorpes L, Kiper DC (1996) Development of contrast sensitivity across the

- visual field in macaque monkeys (*Macaca nemestrina*). *Vision Res* 36:239–247.
- Kleiner M, Brainard D, Pelli D (2007) What's new in Psychtoolbox-3? *Perception* 36:ECVP Abstract Supplement.
- Lau B, Glimcher PW (2005) Dynamic response-by-response models of matching behavior in rhesus monkeys. *J Exp Anal Behav* 84:555–579.
- Macmillan NA, Creelman CD (2005) *Detection theory: a user's guide*, 2nd edition. Mahwah, NJ: Lawrence Erlbaum.
- Meier P, Flister E, Reinagel P (2011) Collinear features impair visual detection by rats. *J Vis* 11:p11.22.
- Montani F, Kohn A, Smith MA, Schultz SR (2007) The role of correlations in direction and contrast coding in the primary visual cortex. *J Neurosci* 27:2338–2348.
- Morton AJ, Skillings E, Bussey TJ, Saksida LM (2006) Measuring cognitive deficits in disabled mice using an automated interactive touchscreen system. *Nat Methods* 3:767.
- Newsome WT, Britten KH, Movshon JA (1989) Neuronal correlates of a perceptual decision. *Nature* 341:52–54.
- Nienborg H, Cumming B (2010) Correlations between the activity of sensory neurons and behavior: how much do they tell us about a neuron's causality? *Curr Opin Neurobiol* 20:376–381.
- O'Connor DH, Daniel H, Peron SP, Huber D, Svoboda K (2010) Neural activity in barrel cortex underlying vibrissa-based object localization in mice. *Neuron* 67:1048–1061.
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision* 10:437–442.
- Pinto LH, Enroth-Cugell C (2000) Tests of the mouse visual system. *Mammalian Genome* 11:531–536.
- Porciatti V, Pizzorusso T, Maffei L (1999) The visual physiology of the wild type mouse determined with pattern VEPs. *Vision Res* 39:3071–3081.
- Prusky GT, Douglas RM (2004) Characterization of mouse cortical spatial vision. *Vision Res* 44:3411–3418.
- Prusky GT, West PW, Douglas RM (2000) Behavioral assessment of visual acuity in mice and rats. *Vision Res* 40:2201–2209.
- Prusky GT, Alam NM, Beekman S, Douglas RM (2004) Rapid quantification of adult and developing mouse spatial vision using a virtual optomotor system. *Invest Ophthalmol Vis Sci* 45:4611–4616.
- Runyan CA, Schummers J, Van Wart A, Kuhlman SJ, Wilson NR, Huang ZJ, Sur M (2010) Response features of parvalbumin-expressing interneurons suggest precise roles for subtypes of inhibition in visual cortex. *Neuron* 67:847–857.
- Schmucker C, Schaeffel F (2006) Contrast sensitivity of wildtype mice wearing diffusers or spectacle lenses, and the effect of atropine. *Vision Res* 46:678–687.
- Sclar G, Maunsell JH, Lennie P (1990) Coding of image contrast in central visual pathways of the macaque monkey. *Vision Res* 30:1–10.
- Seidemann E (1998) Neuronal mechanisms mediating conversion of visual signals into perceptual decisions in a direction discrimination task. Ph.D. dissertation, Stanford University.
- Sohya K, Kameyama K, Yanagawa Y, Obata K, Tsumoto T (2007) GABAergic Neurons are less selective to stimulus orientation than excitatory neurons in layer II/III of visual cortex, as revealed by in vivo functional Ca2+ imaging in transgenic mice. *J Neurosci* 27:2145–2149.
- Tolhurst DJ, Movshon JA, Dean AF (1983) The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res* 23:775–785.
- Uchida N, Mainen ZF (2003) Speed and accuracy of olfactory discrimination in the rat. *Nat Neurosci* 6:1224–1229.
- Umino Y, Solessio E, Barlow RB (2008) Speed, spatial, and temporal tuning of rod and cone vision in mouse. *J Neurosci* 28:189–198.
- van Alphen B, Winkelman BH, Frens MA (2009) Age- and sex-related differences in contrast sensitivity in C57Bl/6 mice. *Invest Ophthalmol Vis Sci* 50:2451–2458.
- Wall DS, Gendron RL, Good WV, Miskiewicz E, Woodland M, Leblanc K, Paradis H (2004) Conditional knockdown of tubedown-1 in endothelial cells leads to neovascular retinopathy. *Invest Ophthalmol Vis Sci* 45:3704–3712.
- Wichmann FA, Hill NJ (2001a) The psychometric function. I. Fitting, sampling, and goodness of fit. *Percept Psychophys* 63:1293–1313.
- Wichmann FA, Hill NJ (2001b) The psychometric function. II. Bootstrap-based confidence intervals and sampling. *Percept Psychophys* 63:1314–1329.
- Yang Y, DeWeese MR, Otazu GH, Zador AM (2008) Millisecond-scale differences in neural activity in auditory cortex can drive decisions. *Nat Neurosci* 11:1262–1263.