Novel two-alternative forced choice paradigm for bilateral vibrotactile whisker frequency discrimination in head-fixed mice and rats

Johannes M. Mayrhofer,* Vida Skreb,* Wolfger von der Behrens, Simon Musall, Bruno Weber, and Florent Haiss

Institute of Pharmacology and Toxicology, University of Zürich, Zürich, Switzerland

Submitted 8 June 2012; accepted in final form 5 October 2012

Mayrhofer JM, Skreb V, von der Behrens W, Musall S, Weber B, Haiss F. Novel two-alternative forced choice paradigm for bilateral vibrotactile whisker frequency discrimination in head-fixed mice and rats. J Neurophysiol 109: 273–284, 2013. First published October 10, 2012; doi:10.1152/jn.00488.2012.—Rats and mice receive a constant bilateral stream of tactile information with their large mystacial vibrissae when navigating in their environment. In a two-alternative forced choice paradigm (2-AFC), head-fixed rats and mice learned to discriminate vibrotactile frequencies applied simultaneously to individual whiskers on the left and right sides of the snout. Mice and rats discriminated 90-Hz pulsatile stimuli from pulsatile stimuli with lower repetition frequencies (10–80 Hz) but with identical kinematic properties in each pulse. Psychometric curves displayed an average perceptual threshold of 50.6-Hz and 53.0-Hz frequency difference corresponding to Weber fractions of 0.56 and 0.58 in mice and rats, respectively. Both species performed >400 trials a day (>200 trials per session, 2 sessions/day), with a peak performance of >90% correct responses. In general, rats and mice trained in the identical task showed comparable psychometric curves. Behavioral readouts, such as reaction times, learning rates, trial omissions, and impulsivity, were also very similar in the two species. Furthermore, whisking of the animals before stimulus presentation reduced task performance. This behavioral paradigm, combined with whisker position tracking, allows precise stimulus control in the 2-AFC task for head-fixed rodents. It is compatible with state-of-the-art neurophysiological recording techniques, such as electrophysiology and two-photon imaging, and therefore represents a valuable framework for neurophysiological investigations of perceptual decision-making.

head fixation; psychophysics; operant conditioning; whisker; rat; mouse; interhemispheric communication; somatosensory cortex

RATS AND MICE are nocturnal animals that use their array of movable whiskers to explore and navigate the environment. These tactile sensors enable rodents to acquire sensory information of the surrounding world. Several studies, mostly performed in rats, have reported that rodents are able to discriminate different textures (Carvell and Simons 1990; Morita et al. 2011), the size of apertures (Krupa et al. 2001), gap size (Jenkinson and Glickstein 2000), and the position of objects (Mehta et al. 2007). More recent work has shown that psychophysical measurements can be acquired during active palpation (Carvell and Simons 1995; Knutsen et al. 2006; Morita et al. 2011; Prigg et al. 2002) or during the passive deflections of individual stationary whiskers (Gerdjikov et al. 2010; Stüttgen and Schwarz 2008, 2010). Most of these studies employed paradigms in which the animals were able to move freely in the experimental arena. However, head-fixed paradigms are becoming increasingly popular as they allow precise stimulus presentation and monitoring of whisker motion (Hentschke et al. 2006). Furthermore, fixing the animal’s head is a requirement for several neuroscientific methods.

To investigate the neuronal code underlying perception and cognitive functions, it is necessary to combine behavioral measurements with recordings from the individual neuronal elements involved in these processes. Simultaneous acquisition of psychophysical and neural data is indeed one of the gold standards in cognitive neuroscience (Parker and Newsome 1998; Stüttgen 2010). The current knowledge in this field has been obtained mainly from the head-fixed nonhuman primate preparation (Wurtz 1968). Although recent developments in head-fixed rodent Go-NoGo paradigms allow single neurons to be recorded concomitantly with psychophysics, these complex head-fixed paradigms are still rare for rats and in particular for mice. On the other hand, the mouse is the major mammalian genetic model that allows targeting of specific cell types with genetically encoded proteins. These proteins can be used to manipulate and visualize neural activity in subsets of neurons. Combining such specific manipulations and measurements with behavioral paradigms will allow us to pinpoint the contribution of individual cell types to perception and decision-making processes (O’Connor et al. 2009). We therefore developed a two-alternative forced choice (2-AFC) task for mice and rats, allowing us to report psychophysical metrics for both species. We have chosen a 2-AFC paradigm because it has clear advantages over Go-NoGo paradigms since it allows a cleaner classification of behavior into correct rejections and omitted trials on a single-trial basis. Go-NoGo paradigms are more susceptible to biases due to fluctuations of motivation compared with 2-AFC tasks (for review, see Stüttgen et al. 2011).

We found that mice and rats learned to discriminate frequencies that were simultaneously applied to single whiskers on the left and right sides of the muzzle in the head-fixed situation. Furthermore, mice and rats showed similar psychometric curves. Other behavioral readouts, such as reaction times, learning rates, trial omissions, and impulsivity, were also very similar in the two species. Reaction times did not depend on the level of difficulty but were in general higher for error trials. Prior whisking of the animal before stimulus onset decreased the task performance significantly. To our knowledge, this is the first description of a 2-AFC task in head-fixed mice and rats. This paradigm allows precise behavioral control and readouts, which can be combined with techniques that take...
advantage of genetic tools available in the mouse and starting to become available in rats.

METHODS

Animals. All experimental and surgical procedures were approved by the local veterinary authorities, conforming to the guidelines of the Swiss Animal Protection Law, Veterinary Office, Canton Zürich (Act of Animal Protection 16 December 2005 and Animal Protection Ordinance 23 April 2008). In total, six female adult Sprague-Dawley rats (r1, r2, r3, r4, r5, r6; 250–350 g) and one male and two female adult C57BL/6J mice (m1, m2, m3; 20–26 g) were trained in a 2-AFC task. The animals had to discriminate between two simultaneously presented vibrotactile stimuli in a head-fixed situation. The age of the animals on the day of headpost implantation was 10–15 wk. The rats were housed in groups of two with food ad libitum, whereas the mice were housed individually with food ad libitum. The animals were subjected to water deprivation for 5 days/wk during the behavioral training. Body weight was monitored prior to each of the two daily training sessions, during which water acted exclusively as reward. When body weight dropped below 90% of the initial weight, additional water was given after the training session. The animals were housed under an inverted 12:12-h light-dark regime such that they were trained during their dark cycle, when they are active.

Surgical procedure. The animals were anesthetized with isoflurane (1–3%; Abbott, North Chicago, IL). The depth of the anesthesia was monitored by both hind paw withdrawal and corneal reflexes. The animal’s temperature was monitored with a rectal temperature probe and maintained at 37°C by a feedback-controlled heating pad (Hartford Apparatus, Holliston, MA). The head was fixed in a stereotaxic apparatus, and a local anesthetic (Bucain, 5 mg/ml; Actavis, Steinhausen, Switzerland) was given subcutaneously preceding the scalp incision. Eyes were protected with ointment (vitamin A eye cream; Bausch & Lomb, Zug, Switzerland). For rats, nine titanium cortical screws (Modus 1.5, 3-mm length; Medartis, Basel, Switzerland) were inserted into the cleared skull, acting as anchors for the headcap. The exact positions of these screws have previously been described (Schwarz et al. 2010). Cortical screws were not used in mice. A bonding agent (Gluuma Comfort Bond; Heraeus Kulzer, Hanau, Germany) was applied to the cleaned skull and was polymerized with a handheld blue light source (600 mW/cm2; Demetron LC; Biogö, Switzerland). The headcap was formed by layers of transparent light-curing dental cement (Tetric EvoFlow; Ivoclar Vivadent, Schaan, Liechtenstein) applied on top of the bonding layer. A screw (M5x15) turned head-down was cemented into the headcap at the midline (2 mm caudal to lambda) and subsequently acted as the post for rat head fixation. In mice, a custom-made aluminum headpost was built for this purpose (Fig. 1A). The headrotor mechanics consist of a brake (Fig. 1A, inset) mounted on the waterspouts (LDT0-028K; Measurement Specialties, Hampton, VA). Two solenoid valves (Bürkert, Ingelgrin- gen, Germany) were used to control the water delivery to the drinking spouts. All of the components of the behavioral apparatus were controlled and monitored with millisecond temporal precision by a custom-written LabVIEW program (National Instruments) running on personal computers using multifunctional data acquisition cards (PCI-6259; National Instruments).

Behavioral paradigm. Handling of the animals started after 2 wk of recovery from the headcap implantation (Fig. 2). The animals were water deprived 24 h before the head fixation training started. Animals were slowly familiarized with the experimenter in two or three daily sessions lasting roughly 10 min for 1–2 wk. In the following 2–3 wk

Innovative Methodology
the animals were trained to tolerate head fixation. Water was given manually from a syringe in the head fixation training. As soon as the animal tolerated the head-fixed situation for up to 10 min without any signs of stress, the animal was placed into the behavior setup. A drinking spout was placed in front of the animal’s snout. In the first three sessions water was given every 5–6 s. Later, the animal had to first lick once and then a couple of water drops were delivered through the drinking spout (referred to as the “reward on lick” mode). In the next step, the procedural training started (Schwarz et al. 2010). The rodents were subjected to a prompting procedure (6–15 sessions) in order to establish an association of the vibrotactile stimulus location with the reward location. The 90-Hz target stimulus was presented without a distractor stimulus, and the animal’s head was gently turned by the motorized system to the stimulus side where water was delivered via the spout. After about two sessions the “reward on lick” mode was turned on as soon as the animal started licking without waiting to sense water coming from the correct spout. Next, actual behavior measurements started and the animal’s performance was recorded. The animal had to turn its head to reach the reward location. First, rats and mice had to detect the side of the 90-Hz target stimulus. When they reached a stable performance (rats: 11 sessions; mice: 11 sessions), gradually more difficult distractor stimuli (10- to 80-Hz pulse trains in steps of 10 Hz) were introduced and they had to discriminate the higher frequency (90-Hz pulse train target stimulus) from the lower frequency. The distractor and the rewarded target stimuli were presented simultaneously on both sides (Fig. 1 D). To preserve the animal’s motivation, more difficult distractor stimuli (40–80 Hz) were always intermingled with less difficult ones (10–30 Hz). In the initial training phase, the head rotation brake was released.

Fig. 1. Setup, paradigm, head rotation mechanics, and stimulus. A: schematic representation of a rat in the behavioral apparatus that is connected to the rotation mechanics via an implanted headpost (1). A single whisker is inserted into a glass tube (2) that is glued to a piezo actuator (3a, 3b) used to deliver simultaneously the vibration stimuli, e.g., 90 Hz (3a) as target and 40 Hz (3b) as distractor. The water reward is delivered with 2 waterspouts (4), which are mounted on a semicircular rail in front of the animal and equipped with piezo lick sensors (5). Inset: mouse apparatus with an opened head rotation brake. The head rotation mechanics consist of a brake (blue part), which is moved by a DC motor with the help of a knee lever (green part). The headpost holder (red part) is pivoted by a ball bearing. For the prompting procedure the headpost holder was directly connected with the motor through an additional lever (lever not shown; note: knee lever was detached in this case). This allowed direct setting of the rotation angle. C and D: schematic representation of a single stimulus pulse (C) and the time sequence of an individual trial (D).

Fig. 2. Training procedure. The different training phases are shown, starting with the headpost surgery through to the acquisition of the psychometric curves. Gray shadings indicate that the duration of the individual phases was adapted for each animal.
0–0.2 s after stimulus onset. After an additional waiting time of 0–0.2 s, the animal was given a 2-s-long window of opportunity (Fig. 1D) to respond with a lick at one of the two waterspouts. The animal was able to turn its head to the left and right sides, reaching the drinking spout with its tongue to report the perceived target stimulus side. Correct choices led to water delivery (rats: 20–50 μl/trial; mice: 5–20 μl/trial). During the aforementioned window of opportunity, a lick emitted to the incorrect waterspout (distractor side; not rewarded) or no-lick response for 2 s after stimulus onset led to a closure of the head rotation mechanics, bringing the rodent back to the central position. Initiation of the next trial was preceded by a randomly varied intertrial interval (ITI, 1.5 ± 0.3 s; Fig. 1D). Lick events in a 2.5-s-long prestimulus time window (no-lick period; Fig. 1D) were punished with a temporal delay (1–2.5 s) of the stimulus presentation. A temporal jitter of maximal ±30% was added to the no-lick period in order to avoid prediction of the stimulus onset. In the learning phase, a bias correction algorithm was used to avoid stereotypical response biases (Knutsen et al. 2006), thus preventing the animals from developing a preference to one side and acting against stereotypic licking such as left-right alternation. When the animals reached a stable performance and nonimpulsive behavior, both the head rotation angle and the angle of the waterspouts were gradually narrowed until the brake remained closed (time span over 1–2 wk). The rats performed up to 400 trials per session and the mice up to 380 trials per session. Two sessions per day were conducted, one in the morning and one in the evening.

**Data analysis.** The data set consisted of behavioral recordings from six rats (r1, r2, r3, r4, r5, r6) and three mice (m1, m2, m3). There was no significant difference in performance between the morning and evening sessions (data not shown). A trial was counted as a correct response when the animal licked at the rewarded spout first. An error was counted when the animal licked at the nonrewarded spout first. A no-lick response within the 2-s time window after the start of the decision period was classified as a missed trial. To compute the performance, only trials with a behavioral response were taken into account (correct response and error); correct responses (correct responses + errors). To analyze whisker position traces, whisker velocity was computed by taking the first derivative of the low-pass filtered (Butterworth 4th order, cutoff frequency 200 Hz) whisker movement recording. For each trial, the root mean square (RMS) velocity was computed in the prestimulus period from −200 ms to stimulus onset. If prestimulus RMS velocity exceeded a threshold of 0.02 m/s, the trial was classified as a movement trial; otherwise it was classified as a nonmovement trial (as in Hentschke et al. 2006). To test whether behavioral performance was different in movement versus nonmovement trials, a fourfold χ2-test was done. All data analysis was performed with MATLAB (MathWorks, Natick, MA).

**Psychophysics.** A psychometric curve was obtained by computing the performance for each stimulus pair and plotting it against the difference between distractor and rewarded frequency. The confidence intervals were computed on the basis of a binomial distribution with a confidence level of 95%. Psychometric functions were fitted with the software package psignifit (version 2.5.6; see [http://bootstrap-software.org/psignifit/](http://bootstrap-software.org/psignifit/), which implements the maximum-likelihood method described by Wichmann and Hill (2001). A logistic function

\[
\Psi(x; \alpha, \beta, \gamma, \lambda) = \gamma + (1 - \gamma - \lambda)F(x; \alpha, \beta)
\]

where

\[
F(x; \alpha, \beta) = \frac{1}{1 + \exp \left[ (\alpha - x) / \beta \right]}
\]

was fitted (fit parameters: \(\alpha, \beta, \gamma, \lambda; \gamma \) fixed to 0.5 and \(\lambda\) constrained between [0 0.2]) to the data points and used to obtain the discrimination threshold and slopes. The discrimination threshold (DTh) was defined as the frequency difference between distractor and target frequency at which the logistic fit reached 50% of its cumulative value. This value was used to compute an additional measure of sensitivity, the Weber fraction: DTh/90 Hz. For the psychometric curves, only the last 270–300 trials (20–40 sessions) in each category were taken into account. For \(r5, r6, r1, r3, m1,\) and \(m2,\) the frequency range from 10 to 80 Hz in steps of 10 Hz was sampled. A reduced set of distractor frequencies was presented to \(r2, r4,\) and \(m3 (r2 \text{ and } r4: 0 \text{ Hz}, 10 \text{ Hz}, 40 \text{ Hz}, 60 \text{ Hz}, \text{ and } 80 \text{ Hz}; m3: 0 \text{ Hz}, 10 \text{ Hz}, 20 \text{ Hz}, 40 \text{ Hz}, \text{ and } 60 \text{ Hz}).\) Peak performance of each animal was determined at 10-Hz distractor frequency and with no distractor present. The learning curves were generated and quantified by analyzing the performance over sessions in trials with no distractor frequency. A cumulative Weibull function was fitted to the mean performance over sessions (first 60 sessions, \(r1, r2, r3, r4, m1, m2\)) except for \(m3,\) as this mouse reached stable performance within the first session after the prompting procedure. The dynamic phase of learning was defined as the range between the first and ninth decile ordinate values of the fitted cumulative Weibull function. The performance stability was computed by taking the standard deviation of the session performance during detection (0-Hz distractor frequency) over the last 60 sessions. Reaction times were defined as the time span between stimulus onset and first lick response and were estimated by the median of the first lick latencies of the trials (same data set as for the psychophysics). Latencies were measured in 10-ms bins and used to construct cumulative plots on a probit scale with a reciprocal time axis (Carpenter and Williams 1995)—the median first lick latency was the interception of the 50% ordinate value with the cumulative lick count distribution. The relative reaction time difference between error trial (timeerror) and correct trial (timecorrect) was computed per session (\(t = 1, 2, 3, \ldots N\)) and averaged over sessions:

\[
\text{reaction time} = 2/N \sum_i (\text{timeerror}(i) - \text{timecorrect}(i))/\text{(timeerror}(i) + \text{timecorrect}(i))
\]

This relative value was computed in order to compensate for session-to-session variability over weeks. In addition, only trials with first lick emission 200 ms after stimulus onset were considered. The influence of session number on the reaction time was tested with the Kruskal-Wallis test. The significance of a behavioral response was tested by the Wilcoxon signed-rank test. The average number of trials in which an early lick was detected served as a measure for impulsivity and represented the average number of trials in which a lick was emitted within the no-lick period. This value was taken as zero when there were no early licks before the stimulus onset. A value of 10% denoted that the animal did on average an early lick every tenth trial. For each frequency the percentage of missed trials was computed and statistically tested with the Kruskal-Wallis test. To facilitate the comparison between different animals, the curves were normalized to mean value over all stimulus categories. A control session for each of the animals was performed to rule out the use of auditory cues in the discrimination task. In the first half of the session, the animal had to discriminate between 90 Hz and no distractor frequency or an 80-Hz frequency difference between the distractor and the target frequency. In the second half of the session the whiskers on both sides were unplugged from the whisker stimulator and the animal had to continue without mechanical stimulation. It was ensured that the stimulators were as close to the head as in the normal stimulus situation but without touching any whisker. The influence of the bias correction was measured by comparing the percentage correct of the left and right sides with each other, % correct left side − % correct right side/(% correct left side + % correct right side), for each session and testing the first 15 sessions against the last 15 sessions with bias correction switched on.

**Statistics.** Statistical errors are SE of the mean unless noted otherwise. For binary values a binomial distribution was used to compute the 95% confidence intervals. The Kruskal-Wallis test, the Wilcoxon signed-rank test, and the χ2-test (MATLAB implementations) were used for statistical analyses.
RESULTS

Training procedure. All animals entering procedural training reached the frequency discrimination training after 7–10 wk of training with two daily sessions (Fig. 2). Slow adaptation to the head fixation and regular daily trainings (5 days/wk) were important for a good performance (data not shown; see Schwarz et al. 2010). The criterion to enter discriminative training was that the animals accepted drinking water from a syringe in the head-fixed condition. The animals that required more handling sessions to reach this criterion did not perform worse than the animals that adapted more rapidly (e.g., m1 9.7 wk: 96% and 222 trials/session compared with m2 7.5 wk: 93% and 228 trials/session). The mice adapted to the head fixation and behavioral apparatus in a time frame similar to that for the rats [rats (r1, r2, r3, r4): 3 wk; mice (m1, m2, m3): 2.5 wk].

Discriminative training. The discriminative training started with a prompting procedure in which the animal’s head was gently turned to the side where the stimulus was presented (passive training; Fig. 2). The motorized head rotation mechanics allowed well-controlled direction of the animal's head to the rewarded waterspout during presentation of the vibrotactile stimulus. This helps the animal to establish an association between stimulus and reward location but may not be necessary for the animal to learn the task. The rewarded frequency (90-Hz pulse train, 1-s duration; Fig. 1C) without a distractor was presented (detection paradigm). The rats only needed six sessions (~500 trials) to establish an association between stimulus and reward location. A lick response to the side of the target stimulus presentation was taken as a criterion for the association. The mice needed 11 sessions for this step (~1,600 trials). After the passive training, the animals had to indicate their decision by turning the head to the reward location in order to collect the water. In the first sessions, water was given automatically without an initial lick to the spout. During the early phase of training, the animals performed a detection task without any distractor stimulus. After animals reached a stable performance in the detection paradigm (rats: ~18 sessions, mice: ~11 sessions of active training), task complexity was increased by introducing distractor frequencies in the paradigm. The animals were trained to discriminate between the rewarded 90-Hz stimulus and the simultaneously presented nonrewarded distractors ranging from 10 to 80 Hz (in 10-Hz steps) frequency difference. The animals performed several hundred trials per session, twice a day. The average median number of trials per session was 202 ± 18 (rats) and 231 ± 6 (mice).

Psychophysics. Figure 3 shows the psychometric curves computed from 20–40 sessions for each of the nine animals. For each stimulus category, only the last 270–300 presented trials were taken into account, thus excluding the initial learning phase (Fig. 3, A–E). For each animal a logistic function was fitted to the stimulus pair performance values to obtain the thresholds and slopes in the transition region of the psychometric curve (see METHODS). The best-performing rat (r6, not shown separately) had a discriminative threshold of 42.0-Hz frequency difference [95% confidence interval (CI95): 38.8–45.1 Hz] and a slope of 2.32 ± 0.50 Hz. The highest threshold for the rats (r4) was 71.9-Hz frequency difference (CI95: 69.5–76.7 Hz). The population mean of the thresholds of all six rats was 53.0 ± 4.4 Hz frequency difference. Peak performance in one animal (rat r3; Fig. 3B) for the least difficult discriminative category (80-Hz distractor to target frequency difference) went up to 93.0% (CI95: 90.3–96.0%) (group mean: 84.1 ± 2.5%) and for the detection up to 97.3% (CI95: 95.3–99.0%) (group mean: 93.0 ± 1.2%). In the three mice the psychometric curves showed similar characteristics. Peak performance for discriminating a distractor to target frequency difference of 80 Hz reached 93.0% (CI95: 90.0–95.7%) for the first mouse (m1; Fig. 3C) and 85.0% (CI95: 81.0–89.0%) for the second mouse (m2; Fig. 3D). The mean for three mice was 85.0 ± 5.6%. In the detection task, a performance of 96.0% (CI95: 98.0–93.7%) for the first mouse (m1) and 87.7% (CI95: 84.0–91.4%) for the second mouse (m2) was reached (group mean: 89.7 ± 3.2%). The discriminative threshold was 49.4-Hz frequency difference (CI95: 48.1–51.9 Hz) for m1 and 48.5-Hz frequency difference (CI95: 42.5–56.2 Hz) for m2 (group mean: 50.6 ± 1.7 Hz frequency difference). In summary, rats and mice displayed very similar psychometric measures. The discrimination threshold averaged over three mice was 50.6 ± 1.7 Hz frequency difference compared with a mean of 53.0 ± 4.4 Hz frequency difference for all six rats. These values correspond to Weber fractions of 0.56 ± 0.02 and 0.58 ± 0.04, respectively. The peak performance for the detection task was 93.0 ± 1.2% for the rats and 89.7 ± 3.2% for the mice. The slope of the psychometric curve in the transition regime, which captures the sensitivity fluctuation of the sensory process, was 2.00 ± 0.25 Hz/1 Hz and 1.82 ± 0.36 Hz/1 Hz for rats and mice, respectively. Note that an ideal sensory process would have a step function as a psychometric function and therefore a high slope value at the threshold, e.g., 5% Hz and above. For each animal at least one control session was recorded to ensure that the animal was exclusively using the vibrotactile whisker stimulus for making the decision (Fig. 3F). In the first half of the control session, the animal was in the normal situation with a single whisker stimulated on each side. In the second half of the session, the whiskers were unplugged from the stimulator. The stimulator was placed as close as possible to the animal’s head without touching the whiskers. All of the animals performed above chance level in the first half of the control session [binomial test of performance against chance level, P value < 0.01 for all rats (n = 6) and mice (n = 3)]. As soon as the stimulators were detached from the whiskers, the performance dropped to chance level (binomial test of performance against chance level, P value > 0.05 for all animals).

Learning and behavioral performance stability. To characterize the learning phase, a cumulative Weibull function was fitted to the mean performance of the detection paradigm over sessions (Fig. 4). The dynamic range where most of the performance change was observed was characterized as the range between the first and the ninth decile of a fitted cumulative Weibull function (Fig. 4, insets). The first session after the dynamic learning phase was chosen as a criterion for reaching stable performance. Rats r1, r2, r3, and r4 were able to perform the detection task (90-Hz target frequency and no distractor frequency) within 18.2 ± 9.0 sessions (2,256 ± 1,373 trials, 1.8 wk; Fig. 4, A and B), while the mice needed 16.5 ± 10.8 sessions (1,312 ± 835 trials, 1.7 wk; Fig. 4, C and D) to reach the aforementioned criterion. The dynamic learning phase varied across the animals and was 30.9 sessions (mouse m1), 1.2 sessions (mouse m2), 12.2 sessions (rat r1), and 11.0
sessions (rat r3) long. The animals showed session-to-session variability but never dropped to chance level after the dynamic learning phase. To quantify the performance stability over days, the standard deviation was computed over the last 60 sessions of the data set shown. The mean variability was 5.6% for the rats (r1, r2, r3, r4) and 3.9% for the mice (m1, m2, m3), demonstrating that rats and mice had comparable levels of performance variability in the 2-AFC discrimination task.

Reaction times. In Figure 5 the distribution and analysis of the first lick latencies for behavioral (error and correct) and different stimulus (easy and difficult) categories are shown. The time span between the stimulus onset and the first lick detected was used to compute the reaction time of the animals. These measurements were performed when the animals were not allowed to rotate their heads. The head rotation mechanism brake was kept closed during the whole session, and the rodents solely indicated their decision by licking one of the two

---

**Fig. 3. Psychophysical performance**—performance over difference between distractor and target frequency (90 Hz). A–D: psychometric curves of single animals (rats r1, r3; mice m1, m2) with logistic fits (blue) and thresholds (green). The detection performance is shown as a black square symbol at 90-Hz difference and was not used for the fitting procedure. **Inset** in A shows the stimulus pair close to the threshold regime for the 50-Hz frequency difference data point (90-Hz target and 40-Hz distractor). **E**: psychometric curve fits from all animals. **F**: control session for all 9 animals. The first half of the training session was performed with normal stimulation (whisker attached to the piezo stimulator), where all animals performed above chance (P < 0.05). In the second half of the training session, the stimulators were retracted from the whiskers (no mechanical contact). The performance of correct trials dropped to the chance level for all animals when the stimulators were retracted (P < 0.05). Error bars are plotted as 95% confidence intervals.
waterspouts that were within reachable distance. To depict the stochastic distribution of the reaction times (first lick event), the cumulative probability is plotted on a probit scale as a function of reciprocal latency (reciprobit lick histogram in Fig. 5A) (Carpenter and Williams 1995). The median first lick latency is the interception of the 50% ordinate value with the cumulative lick count distribution. The shortest median reaction time was 232 ms for the rats and 337 ms for the mice. The average of all rats was 334 ms and for all mice was 486 ms. A straight line in this type of diagram denotes that the underlying distribution has a Gaussian-like shape. All the curves are composed of two almost linear components, one with a shallow inclination and the other with a steep inclination. The reaction time curves shown in Fig. 5A have a form similar to saccadic eye latency distributions where subjects had to do a saccade from a central fixation point to a left- or right-appearing target (Carpenter and Williams 1995). Note Fig. 5, inset, where similar numbers of counts in the correct and error conditions occurred in the early phase and the form of correct and error distributions in the late phase look similar, suggesting two underlying processes for the observed distribution. Nevertheless, all the animals except one showed a significantly longer reaction time for error trials compared with correct trials (Wilcoxon signed-rank test $P < 0.05$ for 6 rats and 2 mice; Fig. 5B). The reaction time difference between the error and correct trials was computed for each session and averaged over sessions (Fig. 5B). This step was necessary in order to compensate for session-to-session variability over weeks (Kruskal-Wallis test, $P < 0.001$, indicating session dependence). Figure 5B, inset, shows example raw reaction times for the error and correct conditions and the resulting relative reaction time over sessions. With the exception of one animal, we did not observe a significant difference in reaction times depending on the level of difficulty (i.e., distractor frequency).

The comparison of difficult (50- to 80-Hz distractor frequency) to less difficult (0 to 40-Hz distractor frequency) stimulus pairs in the first part of the reciprocit lick histogram (20–100 ms; Fig. 5C) for correct trials revealed no significant differences in the distribution in all but one animal (Kolmogorov-Smirnov test, $P < 0.01$ for 1 rat). This finding suggests that the underlying process for the lick responses in the 20–100 ms time window is independent of the level of difficulty and may reflect impulsive behavior. Analyzing only those trials in which the animal responded in the 20–100 ms time window after stimulus onset showed that, except for two animals, the performance of correct trials never exceeded chance level (binomial test, $P < 0.05$ for 1 rat and 1 mouse).

**Impulsive behavior and motivation.** To address the questions of how much impulsive behavior and motivation vary over different animals or species, the percentage of early licks and misses are plotted in Fig. 6. One of the major advantages of a 2-AFC paradigm is that correct rejections can be distinguished from omitted trials (e.g., lack of motivation) on a single-trial basis in contrast to a Go-NoGo paradigm. However, it is essential that the animals do not develop a response bias in the 2-AFC task (e.g., positional bias). This can be avoided by training the animals with a bias correction algorithm in the

---

**Fig. 4.** Learning and stability—performance in detection paradigm over initial learning phase and period of stable behavioral performance. **A–D:** performance over time for the detection stimuli for 4 of the animals (rats r1, r3; mice m1, m2). **Insets** show the initial learning phase, which was fitted with a Weibull function. Gray-shaded areas show the dynamic learning phase.
BILATERAL FREQUENCY DISCRIMINATION IN HEAD-FIXED RODENTS

Innovative Methodology

learning phase (see METHODS). Comparing the early phase with the late phase of the training with bias correction switched on, four of nine animals had significantly less response bias to one side (see METHODS, Kruskal-Wallis test, \( P < 0.05 \)). The response bias was quantized as the absolute value of % correct on the left side minus % correct on the right side divided by the sum of both (\( r_1 \) from 0.144 ± 0.044 to 0.052 ± 0.008, \( r_2 \) from 0.138 ± 0.028 to 0.052 ± 0.014, \( r_3 \) from 0.226 ± 0.069 to 0.076 ± 0.011, \( r_4 \) from 0.087 ± 0.017 to 0.083 ± 0.015, \( r_5 \) from 0.035 ± 0.006 to 0.048 ± 0.011, \( r_6 \) from 0.048 ± 0.010 to 0.036 ± 0.007, \( m_1 \) from 0.131 ± 0.023 to 0.065 ± 0.0151, \( m_2 \) from 0.214 ± 0.052 to 0.050 ± 0.01, \( m_3 \) from 0.053 ± 0.011 to 0.049 ± 0.009). To characterize the impulsive behavior, the licks before stimulus presentations were measured and classified as early licks in the situation where the animals were not allowed to rotate their heads and solely indicated their decision by licking one of the two waterspouts. This impulsive licking behavior can be completely abolished if the head rotation brake is released and the animal can only reach the waterspouts by rotating its head. Nevertheless, impulsive behavior could be quite different across species. Thus licking behavior in the condition in which the head rotation brake was closed was measured. The early lick count represents the average number of trials in which a lick occurred before the actual stimulus presentation (Fig. 6A). Note that in the case of an early lick, the stimulus presentation was shifted backwards so that each stimulus was preceded by a “lick-free” period (Fig. 1D). The average value number of trials with early licks over the last sessions (same data set as for the psychophysics) was 19.3 ± 2.5% for rats and 22.1 ± 7.7% for mice. These almost identical early lick counts for rats and mice indicate that the species are highly comparable in terms of impulsivity. It is worth mentioning that the task design allowed chance performance only if the animal licked irrespective of the stimulus side. As shown in the probitFITs in Fig. 5A, only a small fraction of licks occurred at very short latencies (<100 ms) and these lick time points appear to originate from a different behavioral process reflected in a separate distribution compared with the majority of the responses (linear fits in Fig. 5C).

Fig. 5. Reaction times—distribution and analysis of the first lick latencies for behavioral (error and correct) and different stimulus categories (easy and difficult) for all animals. A: cumulative first lick latency count for the different animals (rats r5, r6, r1, r2, r3, r4 in green; mice m1, m2, m3 in black) with a probit scale as a function of reciprocal latency. Red dots depict the median first lick latency, corresponding to the interceptions of the 50% ordinate value with the cumulative lick count distributions. Inset shows the first lick histogram for m2 for correct and error trials. The second part (>180 ms) was normalized to the peak of the histograms to highlight the similar form of the distributions. B: % of the median reaction time of error trials relative to correct trials. *Significance level \( P < 0.05 \) (\( n = 9 \)). Inset shows raw median reaction times for error and correct trials and the resulting relative reaction times over sessions. C: same plot as in A but split into easy (0–40 Hz distractor) and more difficult (50–80 Hz distractor) stimulus pairs for mouse m2. Inset shows an overview for all animals (mice in black and rats in green). Solid lines correspond to the difficult category and dashed lines to the easy category.

Fig. 6. Impulsivity and missed trials—% of early lick trials and missed trials are similar for rats and mice. A: average number of early lick trials expressed in % for all animals (\( n = 9 \)). B: % of missed trials over frequencies normalized to the mean over all frequencies (rats in green, mice in black). Inset shows the mean values for the corresponding animals (\( n = 9 \)).
Does the level of difficulty influence the number of trials without a response? The average numbers of omitted trials over all distractor frequencies are shown in Fig. 6B. To compare between animals, the values were normalized to the mean over all frequencies. Of each session, only the first 30 omitted trials were taken, thus avoiding a dominance of omitted trials usually occurring at the end of a session when the animal is not engaged in the task anymore. The mean values for each animal are shown in Fig. 6B, inset. There was no significant effect of the distractor frequency on the number of omitted trials in eight animals (Kruskal-Wallis test P values: m1 0.080, m2 0.223, m3 0.429, r5 0.358, r6 0.063, r1 0.287, r2 0.455, r3 0.114, r4 0.016).

**Whisker movements and behavioral performance.** To clarify whether whisker motion has an impact on the psychometrics measured during the bilateral frequency discrimination paradigm, the position of one whisker (C2) was tracked with a linear CCD sensor (Fig. 7A) in four rats. The measured RMS velocity of this whisker was used for classifying trials in nonmovement and movement trials (Fig. 7B). About 9% of all

---

**Fig. 7.** Whisker motion and behavioral performance—whisking prior stimulus presentation decreased task performance. **A:** example trials demonstrating the presence or absence of whisker movements in the prestimulus phase of sensory stimulation (~200 ms, gray shading). Top 2 traces show whisker velocity for a nonmovement and a movement trial. Red and green dots indicate left and right licks, respectively. *Bottom* 2 traces show sensory stimuli as measured by a strain gauge sensor on each stimulator. Sensory stimulation was the same for both trials with the target being presented on the right side. **B:** histogram shows the distribution of prestimulus root mean square (RMS) velocity over all trials. Vertical black line indicates the RMS velocity threshold for the classification of movement trials. This was the case for ~9% of all trials (2,268 of 25,287 trials). To visualize the distribution of behavioral performance for different RMS velocities, trials were binned into 0.1-quantiles (gray shading) and behavioral performance was computed for each bin. Error bars represent 95% confidence intervals and horizontal lines the 95% confidence interval for all trials. **C:** behavioral performance in trials without whisker movements is better than trials with whisker movements. Analyzing only detection trials led to a stronger and significant drop (χ²-test, P < 0.01) compared with trials in the discrimination case (at the threshold, 50-Hz difference). **Significance level P < 0.01.
Innovative Methodology

BILATERAL FREQUENCY DISCRIMINATION IN HEAD-FIXED RODENTS

trials ($r_1$: 15.44% of 3,095 trials; $r_2$: 6.76% of 6,095 trials; $r_3$: 10.68% of 7,807 trials; $r_4$: 5.15% of 6,022 trials) were classified as movement trials. Comparing these two categories showed that trials with motion had a significantly reduced discrimination performance ($\chi^2$-test, $P < 0.01$). The mean performance for all four animals was 70.01% (CI95: 67.88–72.09%) in movement trials and 73.93% (CI95: 73.28–74.59%) in nonmovement trials. When only detection trials were considered, a larger effect was observed (movement: 71.79%, CI95: 61.53–82.05%; nonmovement: 89.62%, CI95: 87.15–92.09%), which is in line with the study of Ollerenshaw et al. (2012). At the discrimination threshold (50-Hz frequency difference) a similar trend was present, but no significance was reached (movement: 48.00%, CI95: 28.00–68.00%; nonmovement: 59.29%, CI95: 52.65–65.93%; Fig. 7C).

DISCUSSION

In the present study, we describe a frequency discrimination paradigm for head-fixed rats and mice. Two vibrotactile frequencies were presented simultaneously to individual whiskers on the left and right sides of the animal’s snout. The task was to report the presentation side of the target frequency by licking from one of two possible waterspouts. To our knowledge, this is the first bilateral frequency discrimination paradigm and the first description of a 2-AFC task in head-fixed rats and mice. We found that both the perceptual thresholds and the psychometric curves were very similar for rats and mice. Hence, our results show that mice can achieve complex discriminative behavior in a paradigm initially designed for rats in our laboratory. The average perceptual threshold was 50.6-Hz (Weber fraction of 0.56) and 53.0-Hz (Weber fraction of 0.58) frequency difference in mice and rats, respectively. The slopes of the psychometric curves of mice and rats were 1.82 and 2.00, respectively. Mice and rats performed on average >400 trials a day (>200 trials/session, 2 sessions/day), with a peak performance of >90% for the detection task in both species. The average reaction time was 334 ms for rats and 486 ms for mice and was longer for error trials compared with correct trials. Reaction time distributions found in this study were similar to previously reported tasks (Carpenter and Williams 1995). Motivation and impulsivity varied over animals but were similar for the two species. Finally, we could demonstrate that trained animals rarely moved their whisker prior to stimulus presentation (9% of trials). In trials where whisker motion occurred, we observed a significant drop of discrimination performance. Response biases could successfully be reduced by using a previously published bias correction (Knutsen et al. 2006). Although the head rotation mechanics was closed for most of the data presented in this study and it is not essential for the entire training procedure, it is a valuable tool for prompting the animals and avoiding early lick events. Furthermore, it allows a simple implementation of a working memory task. Animals can be forced to wait after stimulus offset by keeping the brake system closed for a variable delay period before a response is allowed by opening the brake. The fact that the animal can turn its head to retrieve a reward from two different locations would also allow one to study neuronal correlates head orientation movements (Erlich et al. 2011; Taube et al. 1990) while keeping the advantages of the head-fixed preparation.

Frequency discrimination in different species. Frequency discrimination paradigms have been performed previously in rats, monkeys, and humans (Gerdjikov et al. 2010; LaMotte and Mountcastle 1975; Mountcastle et al. 1990). The perceptual threshold for frequency differences, however, depends on the location and the timing of the stimuli. The majority of tasks published to date involved a working memory component. In these paradigms, frequencies were presented consecutively to the sensory organ and thus the subject had to compare them in sequence (2-interval forced choice; see, e.g., Hernández et al. 1997). Our paradigm is different in that the simultaneously applied single-whisker stimuli had to be compared between the left and the right whisker pads while the stimuli were presented. The interhemispheric communication and the concomitant perception of two stimuli required to solve this task may involve neuronal processes that are distinct from those involved in a working memory task. Frequency thresholds and the resulting Weber fractions were higher as reported in working memory tasks in rats (Gerdjikov et al. 2010) and monkeys and humans (LaMotte and Mountcastle 1975; Mountcastle et al. 1990). Additionally, it has been shown in human subjects that discrimination performance was more accurate when frequencies were applied sequentially to the same finger. Discrimination performance dropped when the frequencies were presented to corresponding fingers on both hands (Harris et al. 2001). However, vibrotactile input arriving simultaneously from both sides might be behaviorally relevant when rodents navigate in darkness in narrow environments. Whisker vibrations may be generated by the texture of walls and ground on both sides of the animal and lead to simultaneous inputs that can be used for orienting in these situations. Passive whisker vibrations (i.e., when the animal is not actively whisking) may be a common setting that rodents encounter in their natural environment. Previous work showed that freely moving rats were able to accurately discriminate small variations of aperture sizes without the need of active whisker movements (Krupa et al. 2001). Furthermore, we observed lower discrimination performance when the animals moved their whiskers just before stimulus presentation. Active whisker movements are often observed when animals have to lean over a gap to discriminate different textures (Carvell and Simons 1990; Cybulska-Klosowicz and Kossut 2000; Von Heimendahl et al. 2007; Morita et al. 2011) and therefore hamper the use of body movements to contact textures. However, whisker tracking data from freely moving animals discriminating wall properties in very close proximity are still lacking.

Mice and psychophysics. The question as to why there is a need to transfer a psychophysical task that has been developed for rats to mice requires addressing. The mechanical difficulties of downsizing the behavioral apparatus components and the potential problems of tracking the whisker motion of an animal about a tenth of the weight of a rat argue against such an endeavor. There are, however, several benefits, especially with respect to the combination with state-of-the-art imaging and electrophysiology techniques that demand well-controlled behavioral paradigms for head-fixed mice. The mouse is currently the best developed mammalian genetic model organism, allowing cell type-specific readouts and manipulation of neuronal activity (Luo et al. 2008). The parallel acquisition of psychometric and neurometric curves is one of the gold standards in cognitive and systems neuroscience for understanding
the role of individual neurons in behavior (Parker and Newcombe 1998). The psychophysics described here will therefore allow the expansion of these types of investigations to a cell type-specific level and shed light on the contribution of individual neuronal elements to perceptual decision-making processes. Here we demonstrate that mice are able to perform tasks that have long been a standard in head-fixed monkey behavior and neurophysiology (Wurtz 1968). The animals were highly motivated at all levels of difficulty and performed several hundreds of trials per day with high maximum performance levels. In conclusion, we can report that the two rodent species (rats and mice) have very similar capacities for frequency discriminations in the 2-AFC paradigm described here.

2-AFC paradigm vs. Go-NoGo paradigm. Several laboratories are currently using Go-NoGo paradigms to measure psychophysics in head-fixed rodents (Andermann et al. 2010; Gerdjikov et al. 2010; O’Connor et al. 2010). These paradigms are well suited to obtain psychometric curves from head-fixed rodents performing sensory discrimination and detection tasks. In all of these tasks animals are subjected to water deprivation for motivation. Deprived animals, however, are prone to impulsive behavior as the measured responses are tightly linked to the pressure of retrieving a reward. In a Go-NoGo paradigm it is difficult to implement that the animal receives a reward in the Go condition and in the NoGo condition. This is more straightforward in 2-AFC tasks as the animal can retrieve a reward in each trial. Another obstacle present in the Go-NoGo paradigm is that animals can contaminate the measured threshold by developing a strong response bias toward Go or NoGo stimuli (for review of different task strategies see Stüttgen et al. 2011). Although animals did not show strong stereotypical response behavior (positional bias: left or right side preference, left/right switching behavior, etc.), this can be a problem in a 2-AFC task. Nevertheless, this kind of behavior can easily be inhibited by implementing a bias correction algorithm. In the Go-NoGo task it is difficult to distinguish a lack of motivation or lapses of attention from false rejections or correct rejections. The 2-AFC paradigm is therefore ideal for investigations of behavioral state-dependent neuronal processing. The uncoupling of reward retrieval per se (i.e.,licks) from the stimulus report action (i.e., the choice of 1 of 2 reward locations) allows a clear distinction of correct, incorrect, and missed trials. Furthermore, the task can easily be modified to accommodate a yes-no paradigm by providing only a single stimulus per trial.

Outlook. The presented apparatus and behavioral paradigm have been designed to combine two-photon imaging and electrophysiology with simultaneous psychophysical measurements. The dura mater of mice is transparent and therefore allows the use of transgenic animals such as Cre driver lines permits the expression of functional indicators or markers in specific subsets of neurons (O’Connor et al. 2009; Zariwala et al. 2011). The head-fixed 2-AFC paradigm presented here therefore promises to be a valuable framework in which to conduct two-photon imaging in the mouse while measuring its psychophysics.

ACKNOWLEDGMENTS

We thank Stefan Weber, Markus Küpfer, and Vasiliki Panagiotopoulou for technical assistance. We also thank Peter Roth for the illustration in Fig. 1A and Renaud Jolivet, Björn Kampa, David Margolis, and Maik Stüttgen for comments on an earlier version of the manuscript. We also thank Medartis for providing the cortical screws.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS


REFERENCES


