

Research Article

Improving the Reliability of Tinnitus Screening in Laboratory Animals

AIKEEN JONES¹ AND BRADFORD J. MAY¹

¹*Department of Otolaryngology-Head and Neck Surgery, Johns Hopkins University School of Medicine, Traylor Research Building, Room 521, 720 Rutland Avenue, Baltimore, MD 21205, USA*

Received: 26 July 2016; Accepted: 17 October 2016; Online publication: 2 November 2016

ABSTRACT

Behavioral screening remains a contentious issue for animal studies of tinnitus. Most paradigms base a positive tinnitus test on an animal's natural tendency to respond to the "sound" of tinnitus as if it were an actual sound. As a result, animals with tinnitus are expected to display sound-conditioned behaviors when no sound is present or to miss gaps in background sounds because tinnitus "fills in the gap." Reliable confirmation of the behavioral indications of tinnitus can be problematic because the reinforcement contingencies of conventional discrimination tasks break down an animal's tendency to group tinnitus with sound. When responses in silence are rewarded, animals respond in silence regardless of their tinnitus status. When responses in silence are punished, animals stop responding. This study introduces stimulus classification as an alternative approach to tinnitus screening. Classification procedures train animals to respond to the common perceptual features that define a group of sounds (e.g., high pitch or narrow bandwidth). Our procedure trains animals to drink when they hear tinnitus and to suppress drinking when they hear other sounds. Animals with tinnitus are revealed by their tendency to drink in the presence of unreinforced probe sounds that share the perceptual features of the tinnitus classification. The advantages of this approach are illustrated by taking laboratory rats through a testing sequence that includes classification training, the experimental induction of tinnitus, and postinduction screening. Behavioral indica-

tions of tinnitus are interpreted and then verified by simulating a known tinnitus percept with objective sounds.

Keywords: tinnitus, conditioned suppression, sound exposure, salicylate, perceptual grouping, stimulus generalization

INTRODUCTION

Investigators who study tinnitus in animals face a unique challenge (Jastreboff et al., 1988; Chao et al., 2014; Eggermont and Roberts, 2015). The perception of "phantom sound" must be experimentally induced and then verified. Reliable screening is essential because induction methods cannot be assumed always to produce tinnitus, and screening errors undermine the reliability of experimental outcomes. Unlike humans, animals cannot directly report their tinnitus status. There is no clear physiological or anatomical marker. Tinnitus screening must be based on quantifiable behavioral assessments (Zhang et al., 2016).

The past three decades have produced over 60 scientific publications that advocate a behavioral protocol for tinnitus screening in animals (for recent reviews, see Chao et al., 2014; Brozski and Bauer, 2016). The studies share the assumption that animals have a natural tendency to perceive tinnitus as an actual sound. This predisposition is usually observed in one of two behavioral contexts. In a conditioned suppression paradigm, tinnitus-positive animals begin to drink in silence after they have been trained to drink only in the presence of sound (Brennan and Jastreboff, 1991; Bauer et al., 1999; Heffner and

Correspondence to: Bradford J. May · Department of Otolaryngology-Head and Neck Surgery · Johns Hopkins University School of Medicine · Traylor Research Building, Room 521, 720 Rutland Avenue, Baltimore, MD 21205, USA. Telephone: 410-614-1878; email: bmay@jhu.edu

Harrington, 2002; Lobarinas et al., 2004). In a gap detection paradigm, tinnitus-positive animals show less gap prepulse inhibition of the acoustic startle reflex (GPIAS), presumably because tinnitus “fills in the gap” that signals the impending presentation of a startle-eliciting stimulus (Turner et al., 2006; Lobarinas et al., 2013; Longenecker et al., 2014).

The tendency to group tinnitus with objective sound does not mean that tinnitus is inseparable from other sounds. Tinnitus patients have normal gap detection thresholds (Campolo et al., 2013), as do laboratory rats (Radziwon et al., 2015). Performance does not deteriorate when the physical characteristics of the stimulus containing the gap are altered to correspond more closely to a self-reported tinnitus percept (Boyen et al., 2015). Tinnitus patients do not show a selective loss of startle inhibition when tested with the GPIAS paradigm (Fournier and Hebert, 2013). Results in laboratory mice are also inconsistent with the assumption that tinnitus fills in the gap (Hickox and Liberman, 2014). For each of these testing procedures, the stimulus differences between tinnitus and the background sound provide a rich source of information for gap detection. Consequently, responding to tinnitus as a special sound improves threshold performance in a gap-detection task or prepares the observer for the startle-eliciting stimulus in a GPIAS paradigm. The screening errors that occur when animals learn to separate tinnitus from other sounds are not unique for GPIAS-based screening procedures. Tinnitus discrimination undermines the reliability of any procedure that bases a positive test on the false perception of sound in silence.

This study presents a series of experiments that illustrate how tinnitus screening can be made more reliable by replacing conventional discrimination tasks with a sound classification paradigm. Classification procedures are an established method for isolating the natural perceptual categories of laboratory animals (May et al., 1988, 1989). Here, we use them to map the defining features of an unknown tinnitus percept. We describe preliminary training methods, induction of tinnitus with sound exposure or salicylate treatment, and the assessment of tinnitus behavior with a stimulus generalization paradigm. We demonstrate how generalization patterns reveal the perceptual characteristics of tinnitus, and how those interpretations can be verified by simulating tinnitus with objective sounds.

Our findings outline common pitfalls that should be considered when selecting any screening method that is promoted without independent verification. For qualified behaviorists who wish to advance the state of tinnitus screening, our classi-

fication procedures represent a rigorous protocol for the cross-validation of alternative approaches that may be better suited to the capabilities of the many non-behavioral laboratories now conducting tinnitus research.

METHODS

All experiments were performed on adult male Sprague-Dawley rats. Prior to training, the rats were habituated to the testing apparatus in 1-h sessions, where they were given free access to water. When rats began to drink in the testing apparatus, the water source in the home cage was removed. Restricted access to water promoted high rates of drinking behavior during training sessions. Water consumption and body weight were recorded daily to ensure that the rats received a sufficient ration and maintained normal growth. Water-infused gel packs were placed in the home cage on non-training days. The rats were housed in the laboratory’s quiet vivarium to avoid the uncontrolled sources of auditory stimulation that occur in institutional facilities (Lauer et al., 2009).

Training sessions were conducted inside a mesh cage at the center of a sound-attenuating chamber. Contact with a metal spout on the front wall of the cage produced water rewards or electrical shocks. A free-field speaker signaled the current consequence of contact with the spout. When the speaker was silent, contact produced water. When the speaker was presenting sound, contact produced an electrical shock. The shock was adjusted on an individual basis to the lowest level that suppressed drinking behavior. Rats quickly learned to suppress drinking during sound presentations and were rarely shocked during behavioral testing. All procedures were reviewed and approved by the Institutional Animal Care and Use Committee of The Johns Hopkins University.

Overview of the Screening Procedure

An overview of the stimulus conditions for each training stage is presented in Figure 1. For each training stage, rats indicated the detection of safe signals by drinking from the spout and the detection of warning signals by suppressing drinking behavior. The sound classification task trained rats to classify silence as a safe signal and sound as a warning signal. The tinnitus classification task established tinnitus as a new category of “safe sound.” The generalization task identified rats with tinnitus by confirming the presence of the safe sound category.

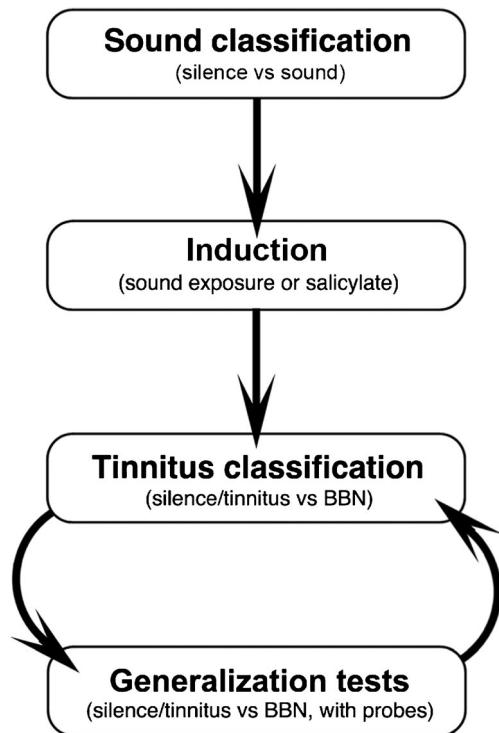


Fig. 1. Behavioral protocol for tinnitus screening. Stimulus conditions are listed for each stage of training. Generalization tests reveal tinnitus as stimulus-specific drinking behavior

Sound Classification Training

The sound classification task was based on the conditioned suppression of drinking behavior because the method has proven to be both robust and efficient in a variety of species (Heffner and Heffner, 1995; Heffner et al., 2013). Classical conditioning methods also have been used successfully for tinnitus screening (Jastreboff et al., 1988; Brennan and Jastreboff, 1991; Jastreboff and Sasaki, 1994). Those methods tend to be avoided in contemporary tinnitus research because training involves exposure to inescapable electric shocks. Classically conditioned responses show rapid extinction when the shock is eliminated, reducing the reliability of screening during prolonged testing.

Initial training was conducted in silence with unlimited access to water. When rats began to show high rates of drinking, periods of broadband noise (50-kHz bandwidth, 50 dB SPL) were introduced to indicate temporary inactivation of the water delivery system. Time spent in noise gradually increased until water was only available during 15-s silent intervals that occurred randomly throughout the session. Approximately 40 of these “safe” trials were presented in each session.

Drinking in the presence of sound was suppressed by adding “warning” trials to the behavioral procedure. Warning trials were equal in duration to safe trials but occurred with higher probability.

Approximately 100 warning trials were presented during a training session. When a rat attempted to drink during a warning trial, a mild electric shock was delivered through the spout. Warning trials were first signaled by broadband noise and therefore were indistinguishable from the background noise that was presented between trials. When broadband noise produced high suppression rates, the set of warning sounds was expanded to include pure tones (frequencies = 8, 12, 16, 22 kHz; 70 dB SPL) or ¼-octave noise bands (center frequencies = 8, 12, 16, 22 kHz; 70 dB SPL).

Our quantitative analysis of drinking behavior is summarized by the lick rate histograms in Figure 2. Each trial window was divided into contiguous 20-ms bins. Each bin was given a binary score of 1 or 0, depending on the presence or absence of lick responses in that time window. Bin scores estimate lick rates without requiring precise sampling of each contact with the spout. A peristimulus time histogram (PSTH) of drinking behavior was computed by aligning bin scores to the onset of water or shock

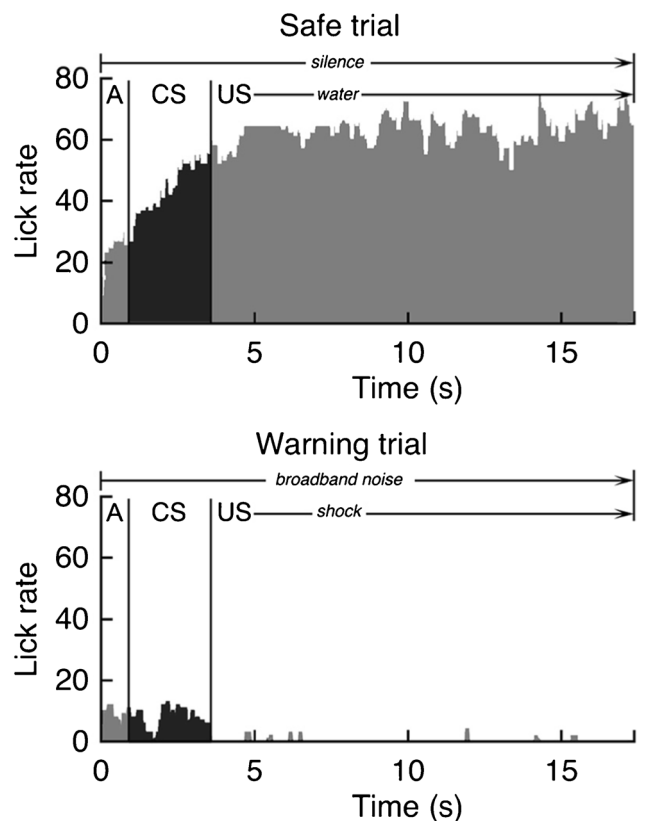


Fig. 2. Peristimulus time histograms of drinking behavior. Cumulative lick rates are shown for one session’s safe trials (upper panel) and warning trials (lower panel). Each histogram is divided into an approach interval (A), a conditioned stimulus interval (CS), and an unconditioned stimulus interval (US). Classification training is indicated by high lick rates in the CS interval of safe trials and low lick rates in the CS interval of warning trials

delivery and summing across all trials with the same stimulus condition.

Each trial began with a 1-s approach interval (A) and ended with an unconditioned stimulus interval (US). Drinking during the approach interval was not used in the calculation of lick rate because the location of the rat was not controlled at the beginning of a trial. Drinking during the US interval was also excluded because lick rates reflected the reinforcement status of the spout. Rats licked continuously when water was being delivered and not at all when shocks were being delivered. Our analysis focused exclusively on the conditioned response to silence or sound that was sampled immediately before the spout was activated to deliver the unconditioned stimulus. This 2.5-s window is designated the conditioned stimulus interval (CS). The onset of the US interval varied randomly across trials (3.5–6.0 s) to prevent timed responses.

Representative learning curves are shown in Figure 3, which plots the average lick rates of two rats relative to session number. Initial sessions show the acquisition of suppression behavior for broadband noise. After approximately 1 week, pure tones (upper panel) or narrow bands of noise (lower panel) were added to the training set. A wide assortment of sounds made it more likely that a subset of the stimuli would share characteristics of the unknown tinnitus percept that was induced at the end of this training stage. In the present study, pure tones and noise bands were presented as discrete training sets to evaluate the relative merits of each stimulus type, but it is possible to create mixed training sets to increase efficiency. Training continued for approximately 3–4 weeks to enhance suppression and improve stability.

Tinnitus Induction

Tinnitus induction procedures were conducted when rats achieved stable performance on the sound classification task. Previous animal studies have relied almost exclusively on sound exposure or salicylate treatment to induce tinnitus. Sound exposure provides the best approximation of the chronic neurological changes that affect tinnitus patients but the induction process may take months to complete (Bauer et al., 1999; Longenecker et al., 2014). Acute salicylate treatment may be less representative of the clinical condition, but induction is rapid and more consistent (Rüttiger et al., 2003; Puel and Guitton, 2007; Stolzberg et al., 2012). Our study performed each method on separate groups of rats to illustrate the basic principles of the two approaches.

Unilateral Sound Exposure. Unanesthetized rats ($N = 9$) were sound exposed inside a slowly rotating mesh cage at the center of a sound-attenuating chamber.

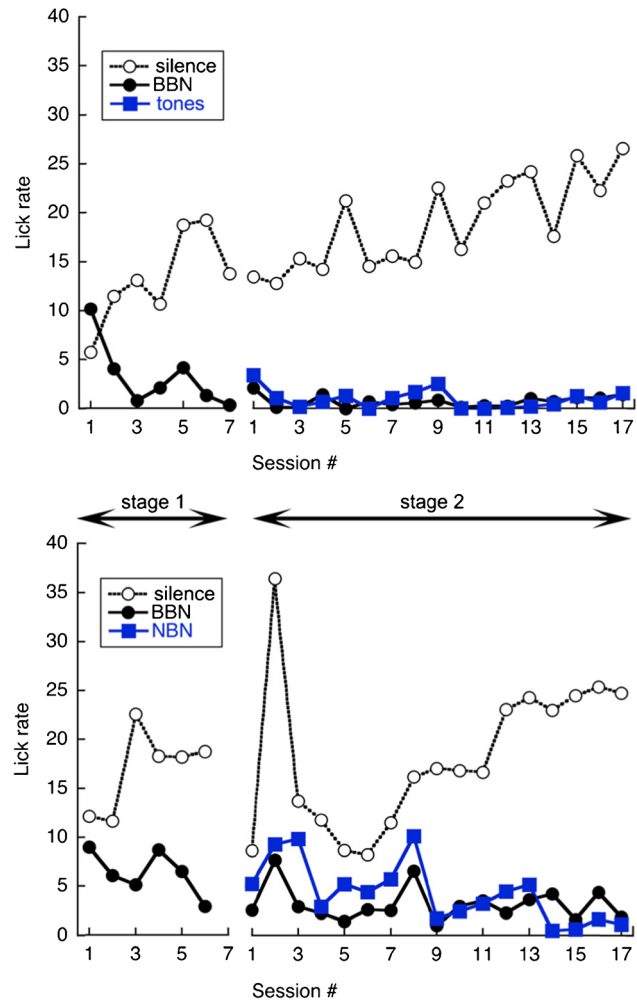


Fig. 3. Stages of sound classification training for a representative rat. Stage 1 contrasted silence (safe trials) with broadband noise (BBN, warning trials). Stage 2 expanded the set of warning sounds with tones (upper panel) or narrow bands of noise (NBN, lower panel)

Two high-output tweeters (Pyramid TW57) delivered a continuous 16-kHz tone. The level of the tone was gradually increased to 116 dB SPL during the initial 60 s of the exposure. Once attained, the peak level was maintained for 2 h.

One ear was fitted with an acoustic foam plug (EAR, 3M). The ear plug attenuated sound frequencies above 10 kHz by at least 55 dB. The maximum attenuation effect was determined in unexposed rats by placing a plug in the contralateral ear, positioning the head so that the ipsilateral ear was oriented toward the speaker, and measuring ABR thresholds with and without a plug in the ipsilateral ear. Protecting one ear ensured that sound-exposed rats maintained normal free-field thresholds during subsequent behavioral testing.

ABR audiograms were collected 1 week after sound exposure to assess hearing loss. Details of these

methods are described in previous publications (Ngan and May, 2001; May et al., 2011). Briefly, the rats were lightly anesthetized with ketamine and xylazine (40:10 mg/kg) and placed on a regulated heating pad to maintain a core temperature of 37 °C. Platinum subcutaneous electrodes were attached to the bulla of the test ear (active), vertex (reference), and ipsilateral leg (ground).

Recordings were made inside an electrically shielded sound-attenuating chamber with the head approximately 45 cm from a wide-range tweeter (Fostex FT28D). The test ear was directed toward the speaker. The opposite ear was plugged and oriented toward the foam-lined floor of the chamber. Based on our previous analysis of ear plugs in unexposed rats, these methods are expected to provide an accurate assessment of unilateral hearing loss when the threshold for the exposed ear is within 55 dB of the threshold for the protected ear. Larger threshold shifts may be underestimated because the more sensitive protected ear may produce an ABR at very high sound levels even when it is plugged.

The ABR audiograms of sound-exposed rats are shown in Figure 4. Threshold is defined as the sound level that produced an ABR peak-to-peak magnitude 2 SDs above background activity. The threshold level was interpolated from an input-output function with 10-dB resolution. The threshold at each frequency is plotted as hearing loss relative to the average threshold of unexposed Sprague-Dawley rats ($N = 174$). Exposed ears exhibited variable patterns of acoustic damage (upper panel) but always featured a sharp transition from normal to impaired thresholds. The magnitude and location of threshold shifts did not show consistent differences between rats that ultimately produced positive (black lines) or negative tinnitus tests (gray lines). Thresholds for the protected ear remained within 20 dB of normal hearing levels (lower panel).

Acute Salicylate Injections. A second group of rats ($N = 7$) was given acute salicylate injections (250 mg/kg, IP). Salicylate was administered on four consecutive days, 2 h before the start of the behavioral session. On days 1–3, the rats were trained with the tinnitus classification task to introduce the safe sound of salicylate-induced tinnitus. On day 4, the rats were probed for tinnitus with the generalization task. These behavioral procedures are described below.

The dose and timing of salicylate injections varies widely across prior studies (Jastreboff et al., 1988; Rüttiger et al., 2003; Yang et al., 2007; Radziwon et al., 2015). Our protocol was derived from preliminary assessments of commonly reported values. In our paradigm, lower doses did not produce reliable tinnitus behavior. Higher doses disrupted behavioral performance, presumably because of acute toxicity. At

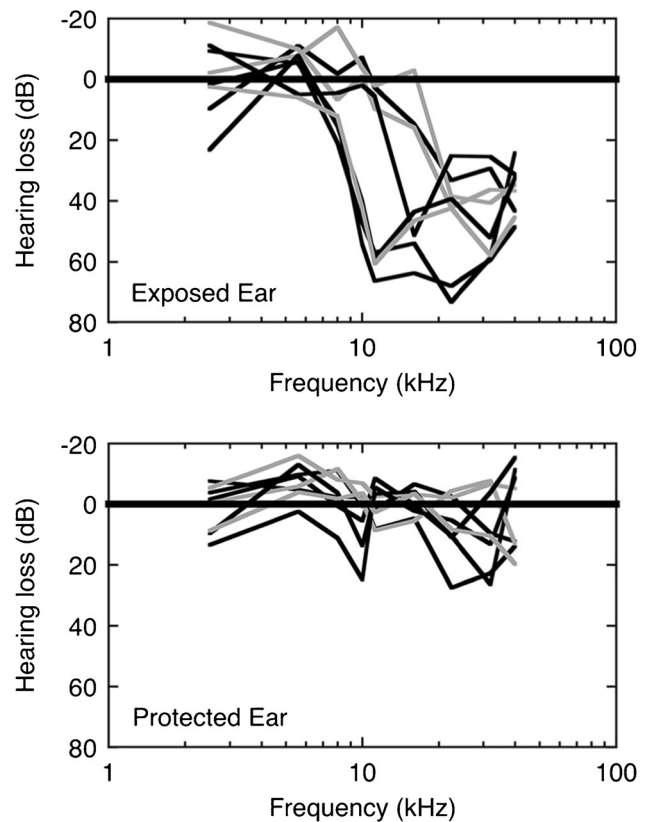


Fig. 4. Auditory brainstem response (ABR) audiograms of nine unilaterally sound-exposed rats. The ABR threshold at each frequency is plotted relative to the average threshold of 174 unexposed ears. Exposed and protected ears are compared in the upper and lower panels, respectively. The rats were assigned to the tinnitus-positive group (black) or tinnitus-negative group (gray) during subsequent generalization tests

the 250 mg/kg dosage, the rats required a 2-h post-injection recovery period before they showed reliable performance in the classification task.

Tinnitus Classification Training

The tinnitus classification task established tinnitus as a special safe sound. Rather than eliminate potentially problematic discrimination effects, tinnitus classification training relied on the subject's ability to discriminate tinnitus from objective sound as a means to establish a behavioral response that was found only in tinnitus-positive animals.

Tinnitus classification training was identical to sound classification training, except stimulus conditions were limited to silent safe trials and broadband noise warning trials. Rats without tinnitus maintained performance with no change in previously learned suppression behavior. Rats with tinnitus, on the other hand, learned to associate the recently acquired “sound” of tinnitus with access to water. Under these conditions, the optimal strategy for maximizing water rewards and minimizing shocks re-

quired the formation of a new safe sound category. That is, the rats should drink only when presented with the sound of tinnitus.

Salicylate-treated rats received three consecutive days of tinnitus classification training. Each training session started 2 h after the daily drug injection. On the fourth day, the rats received salicylate treatment and were screened for tinnitus with the generalization task.

Generalization Testing

The goal of the generalization task was to confirm safe sound behavior without altering previously learned classification groups (May et al., 1988, 1989; Wilson and Pearce, 1989). This goal was achieved by adding a third trial condition. Generalization trials reintroduced pure tones or noise bands as unreinforced probe sounds. Classification behaviors were maintained by interspersing the generalization probes with intervals of silence/tinnitus or broadband noise that continued to serve as safe and warning stimuli. Tinnitus-positive rats were expected to classify probes as safe sounds when they matched salient features of the tinnitus percept. Tinnitus-negative rats were expected to classify all probes as warning sounds.

Generalization tests were repeated on sound-exposed rats no more than 1 day a week and usually at intervals of 2–5 weeks. Accurate tinnitus classification was maintained by conducting training sessions with intervals of broadband noise and silence/tinnitus on other days. Screening was stopped when the rats produced stable behavioral indications of tinnitus, usually within 4–6 weeks. Training was continued for a minimum of 8 weeks in cases where rats failed to exhibit behavioral indications of tinnitus. If rats did not show stable tinnitus behaviors within 8 weeks, they were assigned to the tinnitus-negative group. Occasionally, rats were tested for as long as 4 months without additional training. Even longer periods of observation may be possible. Sound classification training can be repeated if subjects begin to show a general break down of suppression behavior.

Each salicylate-treated rat received one series of injections and one generalization test. Treatment was limited to 4 days and was not replicated to avoid the health risks associated with salicylate toxicity. Tinnitus simulations with objective sounds confirmed that tinnitus classifications could be established within this relatively short time period (see below).

RESULTS

Generalization Profiles for Sound-Induced Tinnitus

The generalization profile of a representative tinnitus-positive rat is shown in Figure 5. The upper panel of the figure presents the average lick rates for the 5 days

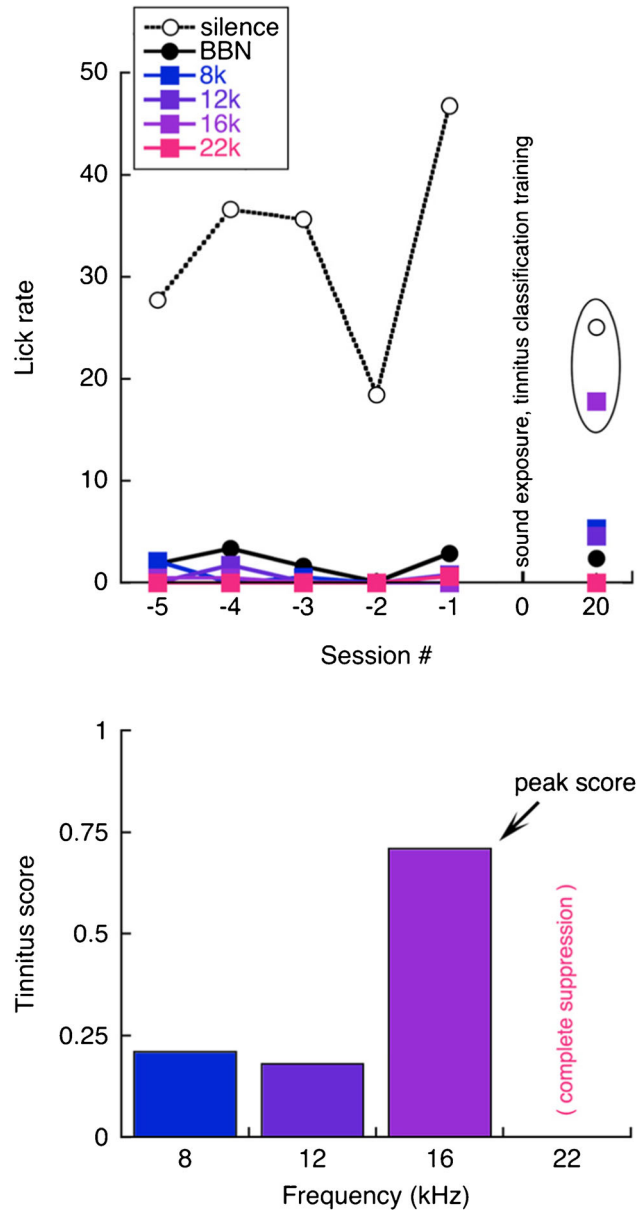


Fig. 5. Generalization profile of a tinnitus-positive rat. Lick rates are shown for five sessions immediately before sound exposure and one generalization test 20 days later (upper panel). The silent safe signal and the 16-kHz generalization probe elicited similar rates (ellipse). Tinnitus scores derived from the probe rates suggest a tuned tinnitus (lower panel)

of sound classification training that immediately preceded sound exposure. During this time, the rat maintained high lick rates in silence and showed strong suppression during sound presentations. After sound exposure, the rat completed tinnitus classification training and was probed with unreinforced pure tones on post-exposure day 20. It is assumed that 16-kHz probes elicited high lick rates because they approximated the sound of this subject's tinnitus. Responses to other probes fell near the actively

suppressed lick rates of the broadband noise warning trials.

Responses to generalization probes were converted to “tinnitus scores” using Eq. 1, where R_{probe} is the lick rate for probe trials, $R_{warning}$ is the suppressed lick rate associated with the presentation of warning trials (BBN), and R_{safe} is the positively reinforced lick rate associated with the presentation of safe trials (silence/tinnitus). Normalization reduces the effects of intersubject variations in drinking behavior. Tinnitus scores approach 1 when probes are treated as safe sounds (positive test). The scores reduce to 0 when probes are treated as warning sounds (negative test).

$$\text{Tinnitus score} = \frac{R_{probe} - R_{warning}}{R_{safe} - R_{warning}}$$

The tinnitus scores in the lower panel of Figure 5 were derived from the lick rates in the upper panel. The magnitude of safe sound behavior indicates a positive tinnitus test. Frequency specificity suggests that the rat’s tinnitus approximated a tonal pitch near 16 kHz. As seen here, the generalization of a pitch cue becomes weaker at frequencies remote to the matching stimulus. The contingencies for reward and punishment can be arranged to sharpen this generalization gradient; for example, an 8-kHz tone could be used as the warning sound. A more selective screening method increases the risk of false-negative screening errors when evaluating an unknown tinnitus percept.

Responses to pure tone probes are summarized for all sound-exposed rats in Figure 6. Each point represents the tinnitus score of one rat at one of the four probe frequencies. The scores have been separated into categories based on the shape of the generalization profile. Each profile has one peak score (circles) and three baseline scores (Xs). A line separates the scores in the upper quartile of the distribution from the scores in the lower three quartiles ($Q3 = 0.385$).

Our study used a categorical screening procedure that was based on the statistical distribution of tinnitus scores. Rats with peak scores in the upper quartile (above $Q3$) were assigned to the tinnitus-positive group (filled circles). These are rats with highly peaked generalization profiles, similar to the example in Figure 5. Rats with peak scores in the lower three quartiles were assigned to the tinnitus-negative group (open circles).

In Figure 6a, the tinnitus scores are plotted relative to probe frequency. In five out of six cases, tinnitus-positive rats produced a peak score at 16 kHz. The relatively high secondary peaks of these rats produced baseline scores that exceeded $Q3$ at 22-kHz. This

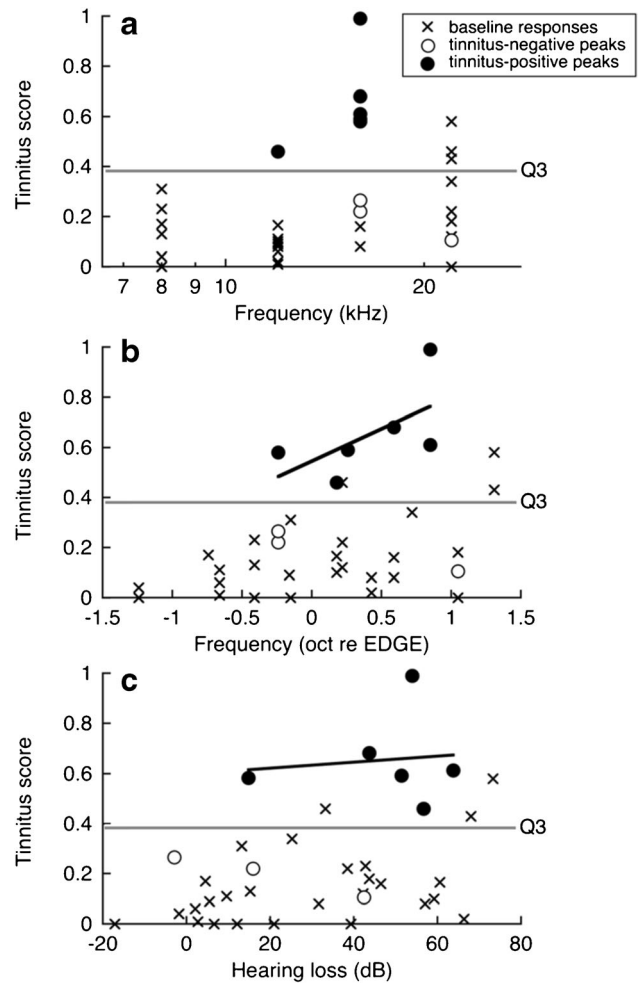


Fig. 6. Distribution of tinnitus scores by frequency in sound-exposed rats. Each rat produced one peak score, which was defined as the highest tinnitus score, and three baseline scores. Rats that produced peak scores in the upper quartile (above $Q3$) were assigned to the tinnitus-positive group. The peak scores of tinnitus-positive rats clustered near 16 kHz (a) increased with EDGE frequency (b) and were not predicted by the frequency of maximum hearing loss (c). Lines were fit to the tinnitus-positive peaks in b and c

frequency grouping suggests that our method of sound exposure consistently induced tinnitus with a pitch close to 16 kHz. In two out of three cases, tinnitus-negative rats produced smaller magnitude peak scores at the same frequency. Therefore, when using categorical criteria, a negative test may indicate the weak expression of tinnitus rather than normal behavior. Studies of the neuropathology of tinnitus may avoid these ambiguities by relating continuous physiological metrics, such as neural spontaneous rate, to the actual magnitude of the tinnitus score (Kaltenbach et al., 2004).

In Figure 6b, the same tinnitus scores are plotted relative to the EDGE frequency of hearing loss. EDGE

frequency is derived from each rat's exposed-ear audiogram. It is the frequency at the midpoint of the sharp drop in hearing sensitivity. EDGE frequencies ranged from 8.9 to 18.9 kHz. The peak scores of tinnitus-positive rats tended to increase with EDGE frequency, although the effect was not statistically significant in our relatively small sample of tinnitus-positive rats, $R(4) = 0.61$, $P = 0.10$. These are rats that showed the transition to hearing loss at lower frequencies (Fig. 4, upper panel). A low EDGE frequency was often associated with large, broadly distributed hearing loss.

In Figure 6c, the tinnitus scores are plotted relative to the ABR threshold shift at the probe frequency. Most tinnitus-positive peaks occupied regions of significant hearing loss, but the magnitude of hearing loss was not correlated with tinnitus score: $R(4) = 0.11$, $P = 0.42$. The distribution of tinnitus-positive peaks also shows considerable overlap with tinnitus-negative peaks and baseline responses. Previous studies have noted that significant hearing loss is usually associated with successful induction, but it is not always necessary or entirely sufficient (Bauer et al., 1999).

The generalization profiles of tinnitus-positive and tinnitus-negative rats are compared in Figure 7. The profiles in the upper panel were collected with pure tones. Tinnitus-positive rats produced a sharply tuned profile with a prominent peak score at 16 kHz. Tinnitus-negative rats produced a featureless profile.

The tuned generalization profile of tinnitus-positive rats contrasts with the broadly distributed patterns of hearing loss that were produced by our exposure paradigm (Fig. 4) and therefore confirm that tinnitus behavior did not reflect the subject's inability to hear probe stimuli. As an additional safeguard, the probes were presented at 70 dB SPL to ensure their audibility in subjects with one normally functioning ear. This sound level is higher than the expected loudness of a tinnitus percept. Tinnitus-positive rats showed excellent generalization despite these probable loudness differences because classification procedures are based on perceptual grouping and not the discrimination of any detectable stimulus difference. Adding probe levels to the generalization paradigm would provide a direct measure of tinnitus loudness, at the expense of test efficiency.

The generalization profiles in the lower panel of Figure 7 were collected with $\frac{1}{4}$ -octave noise bands. Although these probes were expected to share the pitch quality of pure tones, the resulting profiles did not show equivalent frequency tuning. On average, tinnitus-positive and tinnitus-negative rats emitted drinking behavior regardless of noise band center frequency. Individual tinnitus scores varied greatly at each frequency.

The generalization profiles of tinnitus-positive rats were evaluated using a two-factor ANOVA with

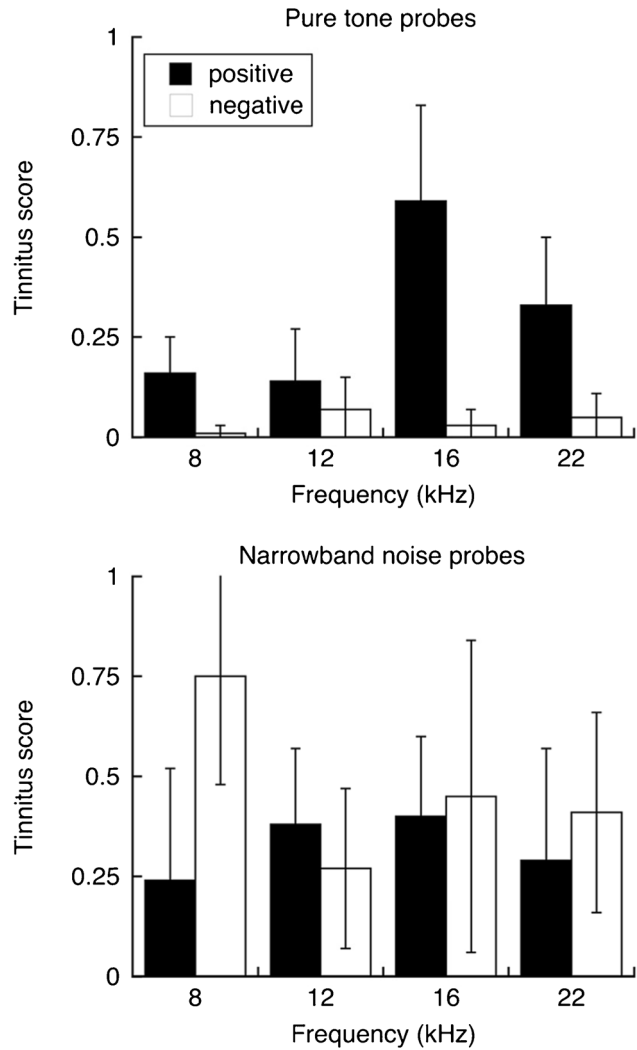


Fig. 7. Mean tinnitus scores of sound-exposed rats. Responses to pure tones and narrowband noise are shown in the upper and lower panels, respectively. Pure tones produced tuned, less variable profiles in tinnitus-positive rats. Error bars indicate 95% confidence intervals

repeated measures. The interaction between probe type (i.e., tone vs. noise band) and frequency was statistically significant ($F(3, 15) = 4.175$, $p = 0.025$), confirming that pure-tone probes produced a more restricted distribution of high tinnitus scores. The main effects of probe type and frequency were not significant.

Generalization Profiles for Salicylate-Induced Tinnitus

The generalization profiles of salicylate-treated rats are presented in Figure 8. Once again, $\frac{1}{4}$ -octave noise bands produced highly variable, poorly tuned profiles. Pure tones also failed to reveal frequency-specific drinking behavior. The highest tinnitus scores were obtained with noise band probes.

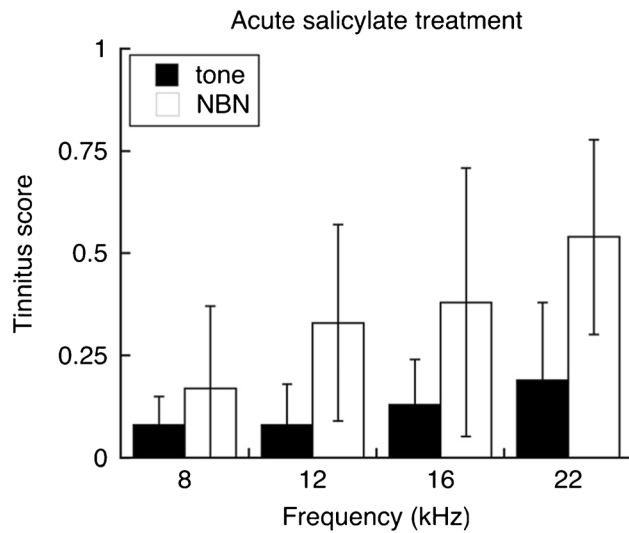


Fig. 8. Generalization profiles of seven salicylate-treated rats. Mean tinnitus scores are compared for tests with pure tones and narrowband noise (NBN). Narrowband noise produced higher tinnitus scores than did pure tones but did not show a clear frequency bias. As noted with sound-exposed rats, responses to narrowband noise were more variable than responses to tones. Error bars indicate 95 % confidence intervals

Statistical analysis confirmed that noise bands produced significantly higher generalization scores than did pure tones ($F(1, 6) = 10.949, p = 0.016$). The effects of probe frequency and the interaction between probe type and frequency were not significant. In combination, these results suggest that our method of salicylate treatment induced a noisy, atonal tinnitus (Cazals, 2000).

Generalization Profiles for Tinnitus Simulations

Tinnitus simulations were performed to relate the inferred characteristics of tinnitus to the physical properties of objective sounds. Rats with simulated tinnitus completed the same sound classification training as sound-exposed and salicylate-treated rats but entered tinnitus classification training without experiencing either method of induction. Instead, a constant objective sound was added to the background of the training task to simulate the sound of tinnitus. Like actual tinnitus, simulated tinnitus was partially masked by broadband noise except when the noise was turned off during safe trials. The availability of water during safe trials was expected to modify previously learned suppression behavior for the pitch and bandwidth characteristics of the simulated tinnitus.

Tinnitus classification training was identical to the procedure used with sound-exposed and salicylate-treated rats. Broadband noise signaled warning trials. The sound of (simulated) tinnitus was present during

safe trials. Pure tones or noise bands served as unreinforced generalization probes. As with acute salicylate treatments, rats received 3 days of simulated tinnitus training before generalization testing.

Generalization profiles for tinnitus simulations are presented in Figure 9. The tinnitus scores in the upper panel were obtained when tinnitus was simulated with a 16-kHz tone (70 dB SPL). Pure tone probes produced a tuned profile with a prominent peak at 16 kHz. Narrowband noise elicited drinking behavior but produced variable, poorly tuned responses. In both cases, the generalization patterns closely resembled the tinnitus-positive profiles of sound-exposed rats. Tinnitus scores were somewhat

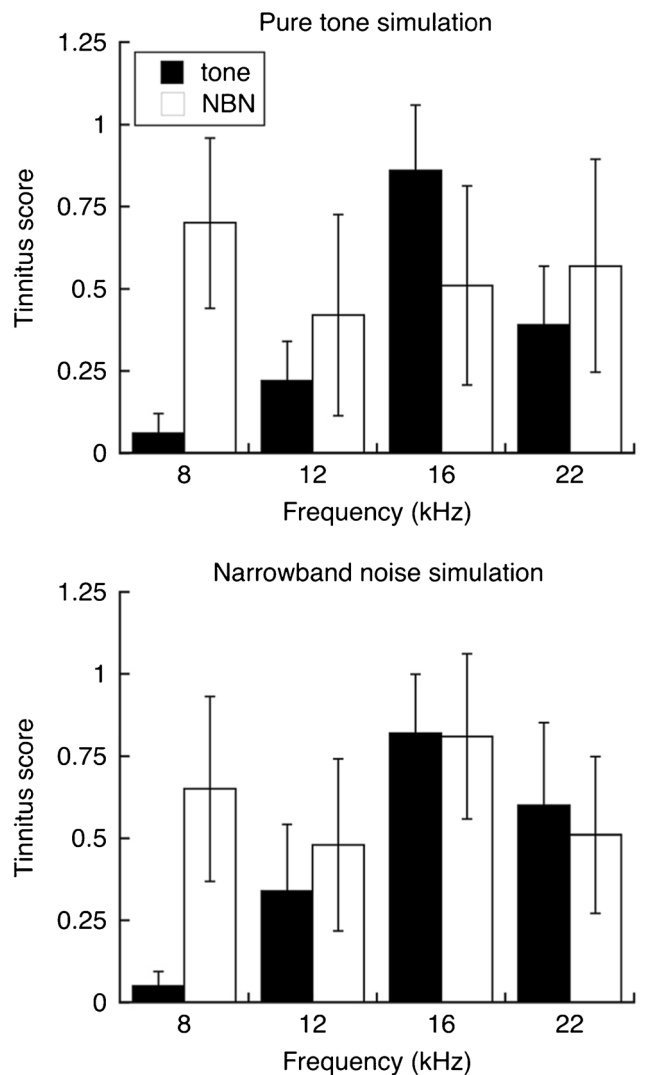


Fig. 9. Generalization profiles for tinnitus simulations. Mean tinnitus scores are shown for tests with pure tones (filled bars) and narrowband noise (NBN, open bars). Tinnitus was simulated with a 16-kHz pure tone (upper panel, $N = 8$) or a 16-kHz narrowband noise (lower panel, $N = 7$). Both simulations produced frequency-tuned profiles when pure tones were used as generalization probes. Error bars indicate 95 % confidence intervals

higher overall, perhaps because the simulated tinnitus perfectly matched the loudness and directional properties of the generalization probes. Those additional cues were not investigated.

The generalization profiles in the lower panel of Figure 9 were obtained when tinnitus was simulated with a $\frac{1}{4}$ -octave noise band (70 dB SPL, center frequency = 16 kHz). An interesting aspect of the data is the lack of frequency tuning for responses to narrowband probes. Although the noise band with a 16-kHz center frequency was identical to the simulated tinnitus and generated a very high tinnitus score, all rats generalized the safe sound classification to noise bands with other center frequencies. The classification appears to be dominated by the common bandwidth feature that separated all probes from the broadband warning stimulus. Additional probe manipulations are needed to confirm the selectivity of the bandwidth cue.

The availability of pitch information became evident when the narrowband simulation was probed with pure tones. In fact, the magnitude and tuning of the generalization profile were equivalent to results obtained with the pure-tone simulation. These results make it necessary to elaborate on our previous interpretation of stimulus generalization in sound-exposed rats (Fig. 7). A frequency-tuned profile indicates tinnitus with a dominant but not necessarily tonal pitch.

When designing a sensitive screening procedure, it is important to bear in mind that the shape of a generalization profile is dictated not only by the properties of tinnitus but also by the selection of probes. In our study, pure-tone probes established a pitch bias that sharpened generalization along the frequency dimension. Noise probes of fixed bandwidth produced a less informative frequency profile, presumably because probe grouping was not based on pitch. If tinnitus has a definite bandwidth quality, generalization behavior is expected to be selective along the bandwidth dimension. In future studies, a more complete description of the tinnitus percept could be gained by varying the bandwidth of noise probes that are centered on the pitch frequency, as determined by initial testing with pure tones.

Statistical analysis of the pure-tone simulation indicated that the effect of probe type was statistically significant ($F(1, 6) = 6.622, p = 0.042$), as was the effect of frequency ($F(3, 18) = 3.882, p = 0.0266$). There was also a statistically significant interaction between the effects of probe type and frequency ($F(3, 18) = 9.340, p < 0.001$). Analysis of the noise band simulation indicated significant effects of frequency ($F(3, 12) = 11.364, p < 0.001$) and the interaction between probe type and frequency, ($F(3, 12) = 16.784,$

$p < 0.001$). For both simulations, the high statistical significance of the interaction reflects the powerful influence of probe type on the frequency distribution of the highest tinnitus scores.

DISCUSSION

Sound Classification Training

Animals achieve their best performance in a conventional conditioned suppression paradigm by learning to respond to any detectable stimulus difference (Moody, 1995). It may take months of training before an animal produces a psychophysical threshold. The method is excessively time consuming when the ultimate goal of the experiment is to characterize some non-behavioral correlate of tinnitus. Moreover, accurate discrimination decreases the reliability of a testing procedure that requires perceptual grouping. With training, animals learn to separate tinnitus from objective sound and stop showing the expected behavioral indications of the disorder.

The learning curves in Figure 3 illustrate typical acquisition rates for the sound classification task. Rats learned to suppress drinking in the presence of broadband noise in 1 week and generalized the behavior to other training sounds almost immediately. Performance improved gradually over the following 2–3 weeks. Testing was prolonged to improve behavioral stability, but most of our subjects were ready for induction after 1 month of sound classification training. While the conditioned suppression method does involve significant investment in each subject, preparation is more efficient than is conventional threshold testing because subjects do not need to be pushed to the limits of their sensory abilities.

The GPIAS paradigm has become increasingly popular as a means to avoid the time and training investments that are required by a conditioned suppression procedure (Turner et al., 2006). Although the approach would seem ideal for the high-throughput screening of laboratory animals, GPIAS testing has proven to be unreliable in tinnitus patients (Boyen et al., 2015). It is now well established that tinnitus does not fill in the gap, violating the basic assumption of the testing procedure (Campolo et al., 2013). It also must be recognized that prolonged testing, even when it involves a reflexive startle response, will promote tinnitus discrimination (Zou et al., 2007). For example, a silent gap always precedes the startle-eliciting stimulus during GPIAS testing. Tinnitus-positive animals will show normal prepulse inhibition once they learn to prepare for the aversive stimulus by attending to the sound of tinnitus.

Tinnitus Classification Training

The perceptual cues that govern the discrimination of tinnitus from objective sound cannot be eliminated, but tinnitus tests can be made more resistant to the problematic effects of discrimination training. Our approach was to create a discrete behavioral classification, namely, the safe sound of tinnitus. Under these conditions, accurate tinnitus discrimination improves the reliability of the behaviors that indicate a positive test. In addition, long-term behavioral stability is maintained through the reinforcement of drinking in the presence of tinnitus.

Tinnitus classification training begins immediately after animals are exposed to the induction procedure. If the induction method is successful, the sound of tinnitus becomes an auditory cue for safe drinking. This new behavioral classification is expected to generalize to objective sounds with similar perceptual features. As a result, tinnitus-positive animals show stimulus-specific drinking when classification behavior is measured with unreinforced probe sounds. One advantage to using tinnitus as the cue for safe drinking is that classification behavior is based exclusively on each subject's tinnitus percept. The investigator does not need to know the characteristics of the tinnitus percept and subjects do not need to share the same tinnitus experience.

Our experimental animals maintained high drinking rates throughout tinnitus classification training. If animals have a natural bias to group tinnitus with objective sound, these results suggest that the disorder develops gradually smoothing the transition from warning signal to safe signal. Alternatively, the unique perceptual properties of tinnitus may facilitate the rapid extinction of previously learned suppression behavior. In either instance, animals learn to discriminate tinnitus from broadband noise because the two stimulus conditions are separable and differentially reinforced. Unlike conventional screening procedures, our classification paradigm benefits from training because accurate discrimination improves the selectivity of the probe sounds that elicit the behavioral indications of tinnitus.

Generalization Testing

Stimulus generalization is the tendency to group stimuli that share the salient features of a conditioned stimulus. Response tendencies are measured by observing how animals apply previously learned classifications to probe sounds. Tinnitus-negative animals should classify all probes as warning sounds because they never experienced tinnitus as a safe sound. Tinnitus-positive animals should classify select probes as safe sounds, depending on the acoustic properties

of the stimulus and the perceptual cues that govern the separate classifications for tinnitus and sound.

The goal of generalization testing is to observe the behavioral indications of tinnitus without changing previously learned classification behaviors. This objective is met by relaxing the contingencies of reward and punishment. Drinking in the presence of tinnitus is still reinforced. Drinking in the presence of broadband noise is still suppressed. Pure tones and narrow bands of noise are reintroduced as unreinforced generalization probes.

The salient perceptual features of tinnitus are revealed by the selective classification of generalization probes. In this study, sound-exposed rats showed a strong tendency to associate 16-kHz tones with the sound of tinnitus. Salicylate-treated rats showed stronger generalization when probed with noise bands, which implies a roaring or buzzing percept (Cazals, 2000). Any modification of the methods outlined here have the potential to generate additional variations in the pitch and bandwidth characteristics of tinnitus, making it necessary to explore a broad assortment of generalization probes when introducing a new animal model or induction protocol.

Generalization probes are never reinforced in our paradigm. Experimental animals are highly trained psychophysical observers by the time they are screened for tinnitus. If responses to probes are suppressed with shocks, tinnitus-positive animals will learn to classify all probes as warning sounds (false-negative screening errors). If responses to probes are rewarded with water, tinnitus-negative animals will learn to classify all probes as safe sounds (false-positive screening errors).

Conditioned responses extinguish without reinforcement (Brennan and Jastreboff, 1991; Jastreboff and Sasaki, 1994), potentially disrupting the behavioral patterns that indicate a positive tinnitus test. In our paradigm, the extinction of unreinforced drinking behavior would be incorrectly interpreted as a negative test. Extinction would elevate false-positive screening errors if the opposite training contingencies were in effect. Extinction was avoided during 2–3 months of generalization testing by limiting an animal's exposure to probes. Each generalization session was preceded by 1–2 weeks of tinnitus classification training. Reinforced training sounds outnumbered unreinforced probes during generalization tests.

Tinnitus Simulations

Tinnitus simulations are not a requirement for screening. The objective of the simulation is to confirm the interpretation of a tinnitus test by producing the same generalization profile with a known sound (Brennan and Jastreboff, 1991).

Beyond verifying a tinnitus test, our simulations have revealed an important and unexpected effect of context on classification behavior. Pure tones elicited pitch-based classifications for tone and noise band simulations. Noise bands elicited classifications that were never constrained by frequency, presumably because a more generalized bandwidth cue was being used. These task-specific classification schemes suggest that the selection of probes ultimately determines the sensitivity of a generalization test to the perceptual features of tinnitus.

If the most salient cue for tinnitus classification is context dependent, animals may optimize performance by forming more efficient categories during generalization testing (Beecher et al., 1979; May et al., 1988). For example, if pure tones convey a strong pitch cue, animals minimize unreinforced drinking behavior by responding only to sounds with the pitch of tinnitus. If noise bands convey a weak pitch, animals maintain access to water by responding less selectively.

Behavioral screening can be made more sensitive by selecting probes that complement the anticipated tinnitus percept. The best probes are likely to change with induction procedure. Generalization tests with pure tones adequately described the pitch of tinnitus in sound-exposed rats, presumably because the exposure-induced tinnitus with a dominant, perhaps tonal pitch. The same tests proved less definitive in salicylate-treated rats, implying a tinnitus that was poorly matched to pure tones or noise bands. The roaring, buzzing, or humming of tinnitus is expected to elicit equally robust classification behavior when generalization tests are conducted using the appropriate probes. Simulations provide an objective method for identifying the most sensitive probes for the many possible sounds of tinnitus.

Future Considerations

Screening procedures promoted today may fall out of favor tomorrow. Controversial procedures persist when there is no clear alternative (Luo et al., 2012; Norman et al., 2012; Pace and Zhang, 2013; Park et al., 2013; Hu et al., 2014). Given the contentious state of tinnitus screening, investigators must learn to recognize the limitations of their behavioral assays (Lobarinas et al., 2013) and how those limitations impact experimental outcomes. The misclassification of tinnitus-positive animals compromises test sensitivity. The misclassification of tinnitus-negative animals limits test specificity. Classifications reverse in animals that have not reached behavioral stability. While recent advances in our understanding of the neuropathology of tinnitus have been remarkable (Allman et al., 2013; Moller, 2016), these

descriptions cannot be more rigorous than the screening procedures upon which they are based.

Screening procedures fail for a reason. Our findings illustrate how critical behaviors may be eroded by training, the contingencies of reward and punishment, or the selection of auditory stimuli. Although our behavioral observations were made in the context of a classification procedure, they serve as useful guidelines for assessing the reliability of any tinnitus test. Until there is a universally accepted method for tinnitus screening, investigators should regard the results of tinnitus testing with healthy skepticism.

Problematic screening procedures can be corrected. If tinnitus does not fill in the gap, is the GPIAS method flawed beyond redemption or simply in need of revision (Berger et al., 2013; Lobarinas et al., 2013; Galazyuk and Hebert, 2015)? Human studies have pointed out that a major limitation with the procedure is the potential for false-negative screening errors. This misclassification of tinnitus-positive animals is particularly problematic when the disorder is experimentally induced with sound exposure because tinnitus-negative animals provide a critical control for separating the specific effects of tinnitus from the general effects of hearing loss. If GPIAS screening errors arise from tinnitus discrimination, as our results suggest, the procedure could be made more reliable simply by the eliminating the strict temporal pairing of the gap prepulse (CS) and the startle-eliciting stimulus (UCS). Applying stricter criteria to a positive test will only exacerbate the problem.

The more robust screening procedures of future tinnitus studies are likely to come from the refinement of existing methods, not the introduction of radically new methodologies. Until a coherent system of best practices is established, qualified laboratories should consider testing animals with multiple screening procedures. A greater emphasis on cross-validation would immediately improve the credibility of current behavioral assessments, and it would expose the limitations of competing procedures. Our classification method offers a rigorous standard for the critical appraisal of alternative approaches that may prove to be more tractable and equally reliable.

ACKNOWLEDGMENTS

This project was initiated in response to the ARO symposium on Assessing Tinnitus in Animals: Progress and Pitfalls (36th Midwinter Meeting, MC Liberman, Chair). The authors thank BL Allman, JI Berger, TJ Brozoski, A Galazyuk, MC Liberman, DB Moody, AR Palmer, MN Wallace, and ED Young for their intellectual contributions during manuscript preparation. Support for this research was provided by the Tinnitus Research Consortium, Action on Hearing Loss, and NIDCD grant P30 DC005211.

REFERENCES

- ALLMAN BL, BAIZER JS, SALVI RJ, LOBARINAS E (2013) Special issue in hearing research: neuroscience of tinnitus. *Hear Res* 295:1–2
- BAUER CA, BROZOSKI TJ, ROJAS R, BOLEY J, WYDER M (1999) Behavioral model of chronic tinnitus in rats. *Otolaryngol Head Neck Surg* 121:457–462
- BEECHER MD, PETERSEN MR, ZOLOTH SR, MOODY DB, STEBBINS WC (1979) Perception of conspecific vocalizations by Japanese macaques. Evidence for selective attention and neural lateralization. *Brain Behav Evol* 16:443–460
- BERGER JI, COOMBER B, SHACKLETON TM, PALMER AR, WALLACE MN (2013) A novel behavioural approach to detecting tinnitus in the Guinea pig. *J Neurosci Methods* 213:188–195
- BOYEN K, BASKENT D, VAN DIJK P (2015) The gap detection test: can it be used to diagnose tinnitus? *Ear Hear* 36:e138–e145
- BRENNAN JF, JASTREBOFF PJ (1991) Generalization of conditioned suppression during salicylate-induced phantom auditory perception in rats. *Acta Neurobiol Exp* 51:15–27
- BROZOSKI TJ, BAUER CA (2016) Animal models of tinnitus. *Hear Res* 338:88–97
- CAMPOLO J, LOBARINAS E, SALVI R (2013) Does tinnitus “fill in” the silent gaps? *Noise Health* 15:398–405
- CAZALS Y (2000) Auditory sensori-neural alterations induced by salicylate. *Prog Neurobiol* 62:583–631
- CHAO Z, QIUJU W, WEI S (2014) Animal behavioral models of tinnitus. *J Otol* 9:58–63
- EGGERMONT JJ, ROBERTS LE (2015) Tinnitus: animal models and findings in humans. *Cell Tissue Res* 361:311–336
- FOURNIER P, HEBERT S (2013) Gap detection deficits in humans with tinnitus as assessed with the acoustic startle paradigm: does tinnitus fill in the gap? *Hear Res* 295:16–23
- GALAZYUK A, HEBERT S (2015) Gap-prepulse inhibition of the acoustic startle reflex (GPIAS) for tinnitus assessment: current status and future directions. *Front Neurol* 6:88
- HEFFNER HE, HEFFNER RS (1995) Conditioned avoidance. In: Klump GM, Dooling RJ, Fay RR, Stebbins WC (eds) *Methods in comparative psychoacoustics*. Birkhauser Verlag, Basel, pp. 79–94
- HEFFNER HE, HARRINGTON IA (2002) Tinnitus in hamsters following exposure to intense sound. *Hear Res* 170:83–95
- HEFFNER HE, KOAY G, HILL EM, HEFFNER RS (2013) Conditioned suppression/avoidance as a procedure for testing hearing in birds: the domestic pigeon (*Columba livia*). *Behav Res Methods* 45:383–392
- HICKOX AE, LIBERMAN MC (2014) Is noise-induced cochlear neuropathy key to the generation of hyperacusis or tinnitus? *J Neurophysiol* 111:552–564
- HU SS, MEI L, CHEN JY, HUANG ZW, WU H (2014) Expression of immediate-early genes in the inferior colliculus and auditory cortex in salicylate-induced tinnitus in rat. *Eur J Histochem: EJH* 58:2294
- JASTREBOFF PJ, SASAKI CT (1994) An animal model for tinnitus: a decade of development. *Otol Neurotol* 15:19–27
- JASTREBOFF PJ, SASAKI CT, BRENNAN JF (1988) An animal model for tinnitus. *Laryngoscope* 98:280–286
- KALTENBACH JA, ZACHAREK MA, ZHANG J, FREDERICK S (2004) Activity in the dorsal cochlear nucleus of hamsters previously tested for tinnitus following intense tone exposure. *Neurosci Lett* 355:121–125
- LAUER AM, MAY BJ, HAO ZJ, WATSON J (2009) Analysis of environmental sound levels in modern rodent housing rooms. *Lab Anim (NY)* 38:154–160
- LOBARINAS E, HAYES SH, ALLMAN BL (2013) The gap-startle paradigm for tinnitus screening in animal models: limitations and optimization. *Hear Res* 295:150–160
- LOBARINAS E, SUN W, CUSHING R, SALVI R (2004) A novel behavioral paradigm for assessing tinnitus using schedule-induced polydipsia avoidance conditioning (SIP-AC). *Hear Res* 190:109–114
- LONGENECKER RJ, CHONKO KT, MARICICH SM, GALAZYUK AV (2014) Age effects on tinnitus and hearing loss in CBA/CaJ mice following sound exposure. *Springerplus* 3:542
- LUO H, ZHANG X, NATION J, PACE E, LEPZYK L, ZHANG J (2012) Tinnitus suppression by electrical stimulation of the rat dorsal cochlear nucleus. *Neurosci Lett* 522:16–20
- MAY B, MOODY DB, STEBBINS WC (1988) Significant features of Japanese macaque coo sounds: a psychophysical study. *Anim Behav* 36:1432–1444
- MAY B, MOODY DB, STEBBINS WC (1989) Categorical perception of conspecific communication sounds by Japanese macaques, *Macaca fuscata*. *J Acoust Soc Am* 85:837–847
- MAY BJ, LAUER AM, ROOS MJ (2011) Impairments of the medial olivocochlear system increase the risk of noise-induced auditory neuropathy in laboratory mice. *Otol Neurotol* 32:1568–1578
- MOLLER AR (2016) Sensorineural tinnitus: its pathology and probable therapies. *Int J Otolaryngol* 2016:2830157
- MOODY DB (1995) Classification and categorization procedures. In: Klump GM, Dooling RJ, Fay RR, Stebbins WC (eds) *Methods in comparative psychoacoustics*. Birkhauser Verlag, Basel, pp. 293–306
- NGAN EM, MAY BJ (2001) Relationship between the auditory brainstem response and auditory nerve thresholds in cats with hearing loss. *Hear Res* 156:44–52
- NORMAN M, TOMSCHA K, WEHR M (2012) Isoflurane blocks temporary tinnitus. *Hear Res* 290:64–71
- PACE E, ZHANG J (2013) Noise-induced tinnitus using individualized gap detection analysis and its relationship with hyperacusis, anxiety, and spatial cognition. *PLoS One* 8:e75011
- PARK YM, NA WS, PARK IY, SUH MW, RHEE CK, CHUNG PS, JUNG JY (2013) Trans-canal laser irradiation reduces tinnitus perception of salicylate treated rat. *Neurosci Lett* 544:131–135
- PUEL JL, GUITTON MJ (2007) Salicylate-induced tinnitus: molecular mechanisms and modulation by anxiety. *Prog Brain Res* 166:141–146
- RADZIWIW KE, STOLZBERG DJ, URBAN ME, BOWLER RA, SALVI RJ (2015) Salicylate-induced hearing loss and gap detection deficits in rats. *Front Neurol* 6:31
- RÜTTIGER L, CIUFFANI J, ZENNER H-P, KNIPPER M (2003) A behavioral paradigm to judge acute sodium salicylate-induced sound experience in rats: a new approach for an animal model on tinnitus. *Hear Res* 180:39–50
- STOLZBERG D, SALVI RJ, ALLMAN BL (2012) Salicylate toxicity model of tinnitus. *Front Syst Neurosci* 6:28
- TURNER JG, BROZOSKI TJ, BAUER CA, PARRISH JL, MYERS K, HUGHES LF, CASPARY DM (2006) Gap detection deficits in rats with tinnitus: a potential novel screening tool. *Behav Neurosci* 120:188–195
- WILSON PN, PEARCE JM (1989) A role for stimulus generalization in conditional discrimination learning. *Q J Exp Psychol B* 41:243–273
- YANG G, LOBARINAS E, ZHANG L, TURNER J, STOLZBERG D, SALVI R, SUN W (2007) Salicylate induced tinnitus: behavioral measures and neural activity in auditory cortex of awake rats. *Hear Res* 226:244–253
- ZHANG J, LUO H, PACE E, LI L, LIU B (2016) Psychophysical and neural correlates of noise-induced tinnitus in animals: intra- and inter-auditory and non-auditory brain structure studies. *Hear Res* 334:7–19
- ZOU D, HUANG J, WU X, LI L (2007) Metabotropic glutamate subtype 5 receptors modulate fear-conditioning induced enhancement of prepulse inhibition in rats. *Neuropharmacology* 52:476–486