Noname manuscript No.
(will be inserted by the editor)

Optimizing auditory brainstem response acquisition using interleaved frequencies

Optimizing auditory brainstem response acquisition

Brad N. Buran · Sean Elkins · J. Beth Kempton · Edward V. Porsov · John V. Brigande · Stephen V. David

Received: date / Accepted: date

Abstract Auditory brainstem responses (ABRs) require averaging responses to hundreds or thousands of repetitions of a stimulus (e.g., tone pip) to obtain a measurable evoked response at the scalp. Fast repetition rates lead to changes in ABR amplitude and latency due to adaptation. To minimize the effect of adaptation, stimulus rates are sometimes as low as 10 to 13.3/s, requiring long acquisition times. The trade-off between reducing acquisition time and minimizing the effect of adaptation on ABR responses is an especially important consideration for studies of cochlear synaptopathy which use the amplitude of short latency responses (wave 1) to assess auditory nerve survival. It has been proposed that adaptation during ABR acquisition can be reduced by interleaving tones at different frequencies, rather than testing each frequency serially. With careful ordering of frequencies and levels in the stimulus train, adaptation in the auditory nerve can be minimized, thereby permitting an increase in the rate at which tone bursts are presented. However, widespread adoption of this paradigm has been hindered by lack of available software. Here, we develop and validate an interleaved stimulus paradigm to optimize the rate of ABR measurement while minimizing adaptation. We implement this method in an open-source data acquisition software tool, which permits either serial or interleaved ABR measurements. The software library, psiexperiment, is compatible with widely-used ABR hardware. Consistent with previous studies, careful design of an interleaved stimulus train can reduce ABR acquisition time by more than half, with minimal effect on ABR thresholds and wave 1 latency, while improving measures of wave 1 amplitude.

Keywords ABR · auditory brainstem response · ABR optimization · wave amplitude · tone burst

1 Introduction

Auditory brainstem responses (ABRs) are an essential tool for assessing peripheral auditory function in research animals and diagnosing auditory dysfunction

Oregon Hearing Research Center, Department of Otolaryngology, Oregon Health & Science University, Portland, Oregon, USA E-mail: bburan@alum.mit.edu

in humans. The ABR is measured using electrodes at the scalp and represents the far-field potential of the auditory nerve and several brainstem nuclei (Melcher et al., 1996). Average responses to hundreds or thousands of presentations of a tone pip or click are needed to obtain a measurable evoked response at the scalp. Because neural activity adapts during repeated sensory stimulation, sensitive ABR measurements require presenting stimuli at rates as low as 10 per second. Faster rates lead to changes in ABR amplitude and latency due to adaptation.

The trade-off between reducing ABR acquisition time and minimizing the effect of adaptation is an important consideration for some experiments. For example, studies of moderate noise exposure have found that ABR wave 1 amplitude is a sensitive measure of auditory nerve survival in animals (Furman et al., 2013; Lin et al., 2011; Kujawa and Liberman, 2009, 2006), and this measure has been used as an indirect assessment of hidden hearing loss in humans (Bramhall et al., 2017). Although it is desirable to maximize measures of wave 1 amplitude, many studies of auditory peripheral function in animals use presentation rates that drive adaptation of auditory nerve fibers in order to limit acquisition time (Spoor and Eggermont, 1971; Harris and Dallos, 1979; Burkard and Voigt, 1990). In humans, slow presentation rates of 10 to 13 tones/s are typically used to avoid adaptation, which limits the number of frequencies and levels that can be tested in a reasonable amount of time.

A number of studies have explored various time-saving strategies for ABR measurement. One approach is to randomize the stimulus timing and use deconvolution to reconstruct the ABR (e.g., Eysholdt and Schreiner, 1982; Polonenko and Maddox, 2019; Millan et al., 2006; Valderrama et al., 2012; Burkard et al., 1990). However, these studies showed evidence of adaptation, manifested as a decrease in ABR amplitude or increase in ABR latency. Recognizing this issue, Mitchell et al. (1999; 1996) designed a novel approach to minimize adaptation by interleaving multiple stimuli, which took advantage of the tonotopic tuning of auditory nerve fibers. Auditory nerve fibers have sharp frequency tuning, particularly at low stimulus levels, and do not respond robustly to frequencies above their characteristic frequency (Kiang, 1965), due to the asymmetric spread of basilar membrane excitation towards the base (i.e., low-frequency region) of the cochlea (Robles and Ruggero, 2001). Thus, careful ordering of the frequencies and levels in the stimulus train can minimize adaptation while increasing the presentation rate.

In the current study, we first confirm that an interleaved stimulus design results in less adaptation as compared to the conventional stimulus design when using the same presentation rate. We then demonstrate that interleaving five frequencies at a rate of 50 tones/s results in ABR amplitudes that are equivalent to those acquired using a conventional approach at a slower rate that does not drive adaptation (10 tones/s). Finally, we demonstrate that optimizing the ordering of frequencies and levels in the interleaved stimulus train yields additional increases in wave amplitude. We tested this approach over a stimulus frequency range of 2 to 32 kHz using mouse and gerbil as our primary model system. Corroborating data was generated in the ferret and rhesus macaque.

One reason why the interleaved approach has not been widely adopted is likely due to hardware and software limitations in most ABR measurement systems. To facilitate use of the interleaved stimulus design, we have written open-source data acquisition software for auditory experiments, psiexperiment, that implements both the conventional and interleaved stimulus designs described in this paper (Buran and David, 2018). This software runs on the same National Instruments hardware as the widely-used Eaton Peabody Laboratory Cochlear Function Test Suite (Hancock et al., 2015) and thus is readily available to many research groups.

2 Materials and Methods

2.1 Subjects

All procedures were performed in compliance with the Institutional Animal Care and Use Committee of Oregon Health & Science University and the Office of Laboratory Animal Welfare, Office of Extramural Research, National Institutes of Health.

The majority of auditory brainstem response (ABR) data were acquired from Mongolian gerbils (*Meriones unguiculatus*) and mice (*Mus musculus*). Data from ferret (*Mustela putorius*) and rhesus macaque (*Macaca mulatta*) are also shown for a more limited set of experimental conditions. The number of animals used is reported in the results on a per-experiment basis. Gerbils of either sex were used and spanned an age range of 8 to 16 weeks. Mice of either sex were used and spanned an age range of 4 to 20 weeks. For mouse, some data was from mice of the FVB strain and other data was from heterozygous Ush1C^{216GA} mice (Lentz et al., 2010). The ferret was a three year old spayed and descented male. Rhesus macaques were 5 month old females.

The sound system, consisting of two half-inch dome tweeters (Parts Express 275-010, now discontinued) and an electret microphone (Knowles FG-23329-P07) coupled to a probe tube, was positioned near the tragus of the ear canal. The sensitivity of the probe tube was measured between 0.1 and 100 kHz using a calibrated $\frac{1}{4}$ inch microphone (377B10 coupled with 426B03 preamplifier and 480M122 signal conditioner, PCB Piezotronics, Depew, NY). Acoustic stimuli were digitally generated (PXI data acquisition system with 24-bit analog to digital and digital to analog converter PXI-4461 card, National Instruments, Austin, TX) and amplified (SA1, Tucker-Davis Technologies, Alachua, FL).

Animals were anesthetized (gerbil: 100 mg/kg ketamine and 10 mg/kg xylazine; mouse: 65 mg/kg ketamine, 6 mg/kg xylazine and 1 mg/kg acepromazine; macaque: 10 mg/kg ketamine and 15 μg/kg dexmedetomidine with 15 μg/kg atipamezol used as a reversal agent; ferret: 5 mg/kg ketamine and 0.05 mg/kg dexmedetomidine). Three electrodes were inserted (gerbil and ferret: vertex and pinna with ground near the base of the tail; mouse: vertex and along the ipsilateral mandible with ground in the forepaw; macaque: midline halfway between the forebrow and the vertex of the skull with reference on the mandible ventral to the ear and ground in the shoulder). ABRs were evoked with tone pips. The voltage difference between pinna and vertex was amplified (gerbil and ferret: 100,000x; mouse and macaque: 10,000x), filtered (gerbil and ferret: 0.1 to 10 kHz, mouse and macaque: 0.3 to 3 kHz) and raw traces were digitized for subsequent analysis. For mouse, gerbil and ferret, an Astro-Med Grass P511 amplifier was used. For rhesus macaque, a Signal Recovery Model 5113 amplifier was used. Body temperature was maintained between 36 and 37°C using a homeothermic blanket (ferret, gerbil, mouse) or chemical heat packs (macaque).

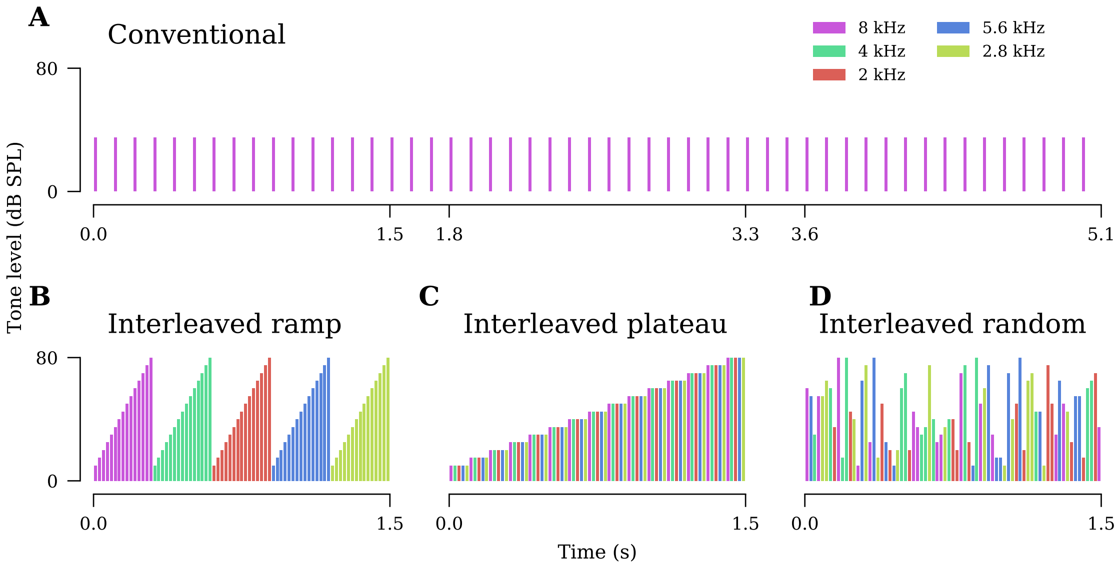

Fig. 1: Schematic of conventional and interleaved ABR stimulus designs tested. **A)** The conventional approach presents tone pips at a single frequency and level until the desired number of averages are acquired. **B-D)** In contrast, the interleaved approach presents a stimulus train containing a single tone pip of each tested frequency and level. This train is then presented repeatedly until the desired number of averages are acquired. For details regarding the ordering of frequencies and levels in each protocol, see text. Note that the horizontal time axis is plotted at the same scale for all four sequences shown, highlighting differences in tone presentation rate, 10 tones/s for conventional and 50 tones/s for interleaved. These rates were used for the data shown in figure 4.

2.2 Stimulus Design

ABRs were generated using 5 ms tone pips with an 0.5 ms cosine-squared envelope. Levels were incremented in 5 dB steps from 10 to 80 dB SPL. The order of tone pip presentation depended on the stimulus design (Fig. 1):

- **conventional**: Tone pips were repeated at a fixed frequency and level until the desired number of artifact-free trials was acquired (Fig. 1A).
- **interleaved**: A train of tone pips containing a single presentation of each level and frequency was constructed. This train was then presented repeatedly until the desired number of artifact-free trials was acquired (Fig. 1B-D).

For the interleaved stimulus designs, the presentation rate was defined as the rate at which individual tones appear in the train (i.e., a 5 frequency, 15 level train with a presentation rate of 50 tones/s would result in a train that was 1.5 seconds long). We tested three different rules for constructing the interleaved stimulus train:

- **ramp**: Levels were swept from low to high before advancing to the next frequency (Fig. 1B).
- **plateau**: All frequencies were presented at a fixed level before advancing to the next higher level (Fig. 1C).
- **random**: The set of levels and frequencies was shuffled randomly on each presentation of the train (Fig. 1D).

In the conventional stimulus design, the polarity of tone pips was alternated on each presentation of the tone pip to remove frequency-following responses. In the interleaved stimulus designs, the polarity of all tone pips was alternated between each presentation of a train.

All interleaved experiments tested a sequence of five frequencies. For the interleaved ramp and interleaved plateau designs, frequencies were arranged in decreasing order while maintaining a minimum spacing of one octave between adjacent frequencies. The exact order was 8, 4, 2, 5.6 and 2.8 kHz for gerbil and 32, 16, 8, 22.6 and 11.3 kHz for mouse. For ferret, we acquired 2 to 45.2 kHz in half-octave steps using the interleaved random protocol. In macaque, 0.5, 2, 4, 8, 16, 22.6 and 32 kHz were tested. Since auditory nerve fibers are preferentially tuned to frequencies within half an octave of the characteristic frequency of the fiber (Kiang, 1965), tones falling outside of this range should not drive much adaptation of the fiber (Harris and Dallos, 1979).

When comparing two or more stimulus designs (e.g., conventional at 10/s vs. interleaved ramp at 50/s), all permutations were tested within a single animal during a single session (i.e., no re-positioning of the electrodes and/or acoustic system). To avoid biases introduced by variations in anesthesia depth, the ordering of the stimulus designs and presentation rates were randomized for each ear. Anesthetic boosters dosing was only necessary when acquiring data from gerbil and care was taken to perform the injection without altering the position of the animal's head. For mouse, all data was acquired without using an anesthetic booster.

2.3 Artifact rejection

For interleaved studies, only the segment of the train containing the artifact was rejected, rather than rejecting the entire stimulus train. This means that we acquired a variable number of artifact-free averages for each frequency and level tested, but every frequency and level had at least 512 averages. When generating waveforms for analysis, only the first 512 averages were included to ensure that the number of averages were identical across all experiments.

2.4 Analysis

ABR waveforms were extracted (-1 to 10 ms re tone pip onset) and averaged. To match the filter settings for mice, gerbil waveforms were digitally filtered (0.3 to 3 kHz) prior to averaging. Thresholds were identified via visual inspection of stacked waveforms by two trained observers, each blind to the stimulus design. Results from the two observers were compared and discrepancies of greater than

10 dB reconciled. Wave amplitude and latency were identified using a computer-assisted peak-picking program (Buran, 2015). Wave amplitude was defined as the difference between the peak and the following trough. Input-output functions were generated by plotting the wave amplitude as a function of stimulus level.

Mixed linear models Differences in ABR threshold, wave amplitude and wave latency were assessed using a general mixed linear model. For wave amplitude and latency, the measured value at 80 dB SPL was used. Intercept (β_i) allowed for a constant offset. Frequency (β_f), stimulus design (β_c) and repetition rate (β_r) were fixed effects. All two-way (β_{fc}, β_{rf} and β_{rc}) and three-way (β_{rfc}) interactions between the fixed effects were included. Frequency, f, and stimulus design, c, were treated as categorical variables and rate, r, as a continuous variable. Dummy (i.e., treatment) coding was used for all categorical variables. Since both mouse and gerbil were tested at 8 kHz, data from each species were coded separately at this frequency to avoid introducing an additional effect for species. Ear was treated as a random effect and coded as U_e where e represents index of the ear.

$$\begin{aligned} y_i =& U_e + \beta_i + \beta_f + \beta_c + \beta_{fc} + \\ & (\beta_r + \beta_{rf} + \beta_{rc} + \beta_{rfc}) \times r_i \end{aligned} \tag{1}$$

For comparing sequential at 10 tones/s versus interleaved ramp at 50 tones/s, the rate term was dropped, simplifying the model to:

$$y_i = U_e + \beta_i + \beta_f + \beta_c + \beta_{fc} \tag{2}$$

Bayesian regression All models were fit using Bayesian regression to maximize the Normal likelihood of free parameters using pyMC3 (Salvatier et al., 2016). In contrast to conventional model fitting, Bayesian analysis allows for simple calculations of credible intervals on derived parameters (e.g., parameters that are mathematical functions of fitted coefficients), offers simple construction of realistic hierarchical models (Gelman et al., 2013) and avoids a number of problems with conventional p-values derived from null hypothesis significance testing (Szucs and Ioannidis, 2017). Another advantage of Bayesian analysis is that it determines the probability that the model coefficients take on a particular value or range of values (McMillan and Cannon, 2019).

Each model was fit four times for 2000 samples following a 1000 sample burn-in period using the No U-Turn Sampler (Hoffman and Gelman, 2014). Posterior samples were combined across all fits (i.e., chains) for inference. Gelman-Rubin statistics were computed to ensure that the four fits, each of which started with a random estimate for each parameter, converged to the same final estimate ($\hat{R} < 1.1$).

On p-values Unlike conventional (i.e., frequentist) approaches, Bayesian analysis does not offer p-values. Instead, Bayesian analysis quantifies the probability that the true value for a parameter falls between two points. These distributions can be used to calculate the probability that there is a true difference between groups, which is typically the information people incorrectly attempt to glean from p-values (Nuzzo, 2014). In our analyses, we report the mean and 95% credible interval (CI) for the difference between groups (e.g., interleaved ramp vs conventional). The

CI should be interpreted as the interval in which we are 95% certain contains the true value. Therefore, if the 95% CI does not bracket 0, we can assume the value is significantly different from 0. To further aid in interpretation of our results, we calculate the probability of a true difference in the parameter between groups by integrating over the portion of the posterior distribution where the value was greater than a reference value.

Averaging across frequencies To estimate the mean and credible interval for values that were mathematical functions of the fitted parameters, we computed the calculation using the posterior samples from the fitted parameters (Gelman et al., 2013). As an example, we will illustrate how the average threshold across all frequencies was calculated from the fit to equation 1. Since dummy (i.e., treatment) coding was used for the categorical levels, β_i (the intercept) represents the value for the first frequency and β_c represents the difference in value between each frequency and the first frequency in the set. Therefore, if we have four frequencies (2.8, 5.6, 11.3 and 22.6 kHz), then the predicted threshold for 2.8 kHz is β_i, 5.6 kHz is $\beta_i + \beta_{f[5.6]}$, 11.3 kHz is $\beta_i + \beta_{f[11.3]}$ and 22.6 kHz is $\beta_i + \beta_{f[22.6]}$. It follows that the mean threshold, $\bar{y}$, across all frequencies is:

$$\bar{y} = (4\beta_i + \beta_{f[5.6]} + \beta_{f[11.3]} + \beta_{f[22.6]})/4 \quad (3)$$

3 Results

We first test whether ABR data acquired using an interleaved stimulus design undergoes less adaptation than data acquired using a conventional stimulus design at the same presentation rate. Next, we assess whether interleaving five frequencies at a rate of 50 tones/s can produce results equivalent to a conventional approach at a rate of 10 tones/s that does not drive adaptation. Finally, we test whether we can reduce adaptation even further by modifying the order of the stimuli within the interleaved stimulus design.

3.1 Conventional vs. interleaved ramp at matched rates

We first assessed whether the interleaved ramp stimulus design offers an advantage over the conventional design for stimuli presented at the same rate. We compared two rates, 40/s, which is common in the animal literature, and 80/s, to assess whether doubling the presentation rate yielded additional benefits. Measurements were compared for both mouse and gerbil. A set of five frequencies in each species was assessed when using the interleaved ramp design, but time constraints from anesthesia duration limited the number of frequencies measured in the conventional protocol (2.8 and 5.6 kHz for gerbil, 11.3 and 22.6 kHz for mouse). Thus, data is only shown for the subset of frequencies common to both stimulus designs.

Both the interleaved ramp and conventional stimulus designs yielded clean ABR waveforms, with waves 1 through 5 easily identifiable (Fig. 2A). At presentation rates of 40/s and 80/s, the interleaved ramp design had ABR thresholds that were at least as low as thresholds acquired using a 40/s conventional design (Fig. 2B, Table 1). Regardless of presentation rate or species, ABR thresholds

Table 1: Change in ABR threshold relative to 40/s conventional stimulus design. Units are dB. Negative values indicate that the threshold was less than for 40/s conventional. Measurements are averaged across all stimulus frequencies for the given species. Significance of differences is indicated by the posterior probability that the difference in threshold does not exceed ±2.5 or ±5 dB relative to 40/s conventional.

			95% CI			
	Species	Mean	lower	upper	Pr($< \pm 2.5dB$)	Pr($< \pm 5dB$)
40/s ramp	gerbil	-0.8	-2.3	0.8	0.99	1.0
	mouse	-1.8	-3.7	0.1	0.78	1.0
80/s ramp	gerbil	-0.1	-1.8	1.6	1.00	1.0
	mouse	-1.3	-3.4	0.7	0.86	1.0
80/s conv.	gerbil	0.5	-0.9	1.9	1.00	1.0
	mouse	1.9	0.1	3.8	0.74	1.0

Table 2: Change in wave 1 amplitude relative to 40/s conventional stimulus design. Units are percent change. Negative values indicate the amplitude measurement was less than for 40/s conventional. Measurements are averaged across all stimulus frequencies for the given species. Significance of differences is indicated by the posterior probability that the increase in wave 1 amplitude exceeds 10% and 20% relative to 40/s conventional.

			95% CI			
	Species	Mean	lower	upper	Pr($>$ 10%)	Pr($>$ 20%)
40/s ramp	gerbil	23	12	33	0.99	0.66
	mouse	26	15	38	1.00	0.85
80/s ramp	gerbil	12	2	23	0.67	0.08
	mouse	16	6	26	0.87	0.21
80/s conv.	gerbil	-23	-32	-14	0.00	0.00
	mouse	-17	-26	-8	0.00	0.00

in the interleaved stimulus designs were within ±5 dB of the conventional stimulus design (Table 1). Wave 1 amplitude, defined as the difference between the first peak and following trough (Fig. 2A), was larger in both the 40 and 80/s interleaved ramp stimulus design as compared to 40/s conventional for all stimulus levels tested (Fig. 2C). In particular, wave 1 amplitude for 40/s interleaved was 23-26% larger than 40/s conventional and 80/s interleaved was 12-16% larger than 40/s conventional (Fig. 2D, Table 2). In contrast, wave 1 amplitude in 80/s conventional was 17-23% smaller than 40/s conventional. Waves 2 through 5 had amplitudes in the 40 and 80/s interleaved design that were at least as large as 40/s conventional (Fig. 3A,C,E,G). Although there were some significant differences in wave latency, they were generally small (Figs. 2E,F and 3B,D,F,H, Table 3).

Taken together, these results demonstrate that an interleaved ramp design that doubles stimulation rate to 80 tones/s yields ABR results that are equivalent to, and often of greater amplitude than, data acquired using a conventional design at 40 tones/s.

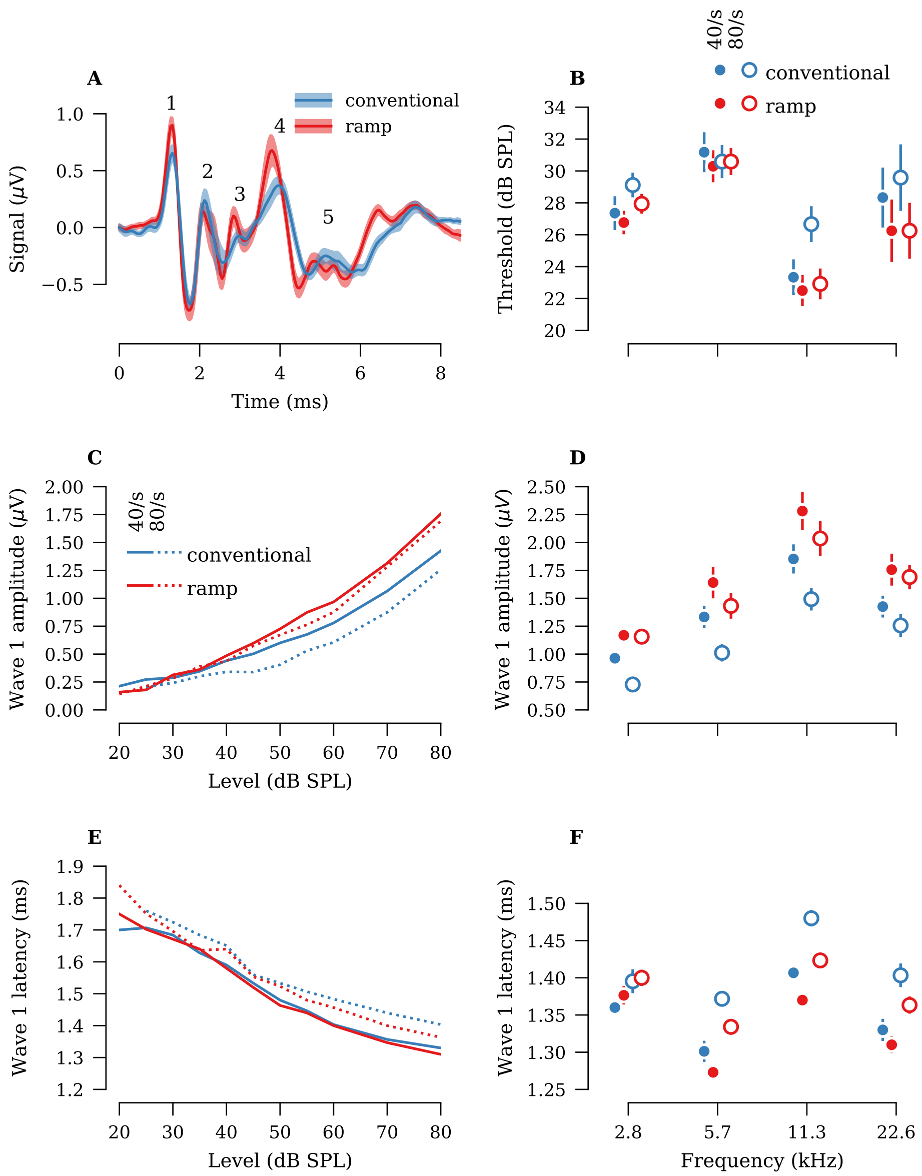

Fig. 2: The interleaved ramp stimulus design yields greater wave 1 amplitudes than the conventional stimulus design regardless of presentation rate (n=17 ears from gerbil, 8 ears from mouse). **A)** Comparison of ABR waveforms (ensemble average across all ears) acquired using the interleaved and conventional stimulus designs. Data shown are in response to 40/s 22.6 kHz, 80 dB SPL tone pips in mouse. Shaded area indicates ±SEM. Numbers indicate wave. **B)** Average ABR thresholds. **C)** Average wave 1 amplitude vs. stimulus level for 40/s 22.6 kHz tones in mouse. **D)** Average wave 1 amplitude at 80 dB SPL. **E)** Average wave 1 latency vs. stimulus level for 40/s 22.6 kHz tones in mouse. **F)** Average wave 1 latency at 80 dB SPL. Error bars in all panels indicate ±SEM.

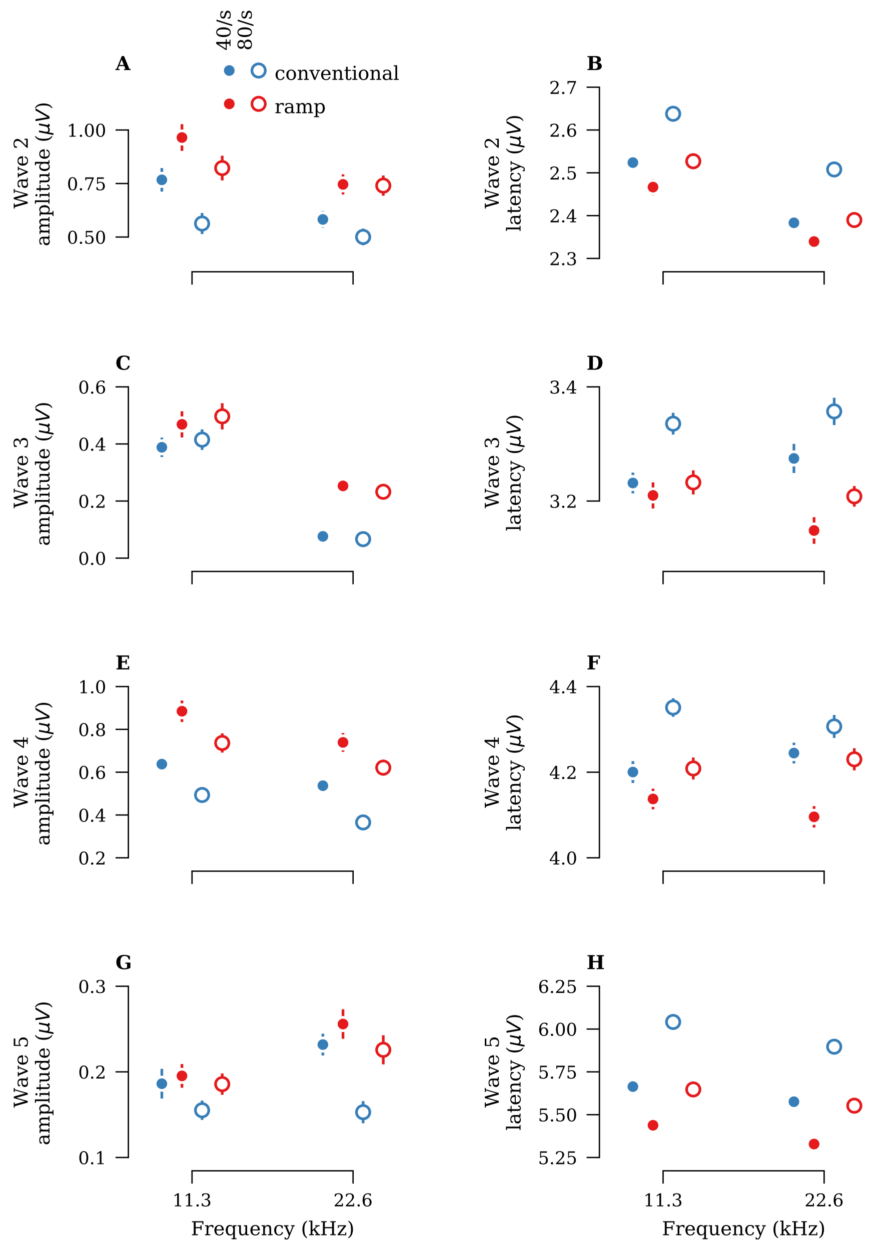

Fig. 3: Comparison of ABR metrics acquired using conventional and interleaved ramp stimulus designs for waves 2 through 5 at all frequencies and presentation rates tested (see Fig. 2 for wave 1 data; $n = 8$ ears in mouse). **A,C,E,G)** Average wave amplitude at 80 dB SPL. **B,D,F,H** Average wave latency at 80 dB SPL. Error bars indicate ±SEM.

Table 3: Change in wave 1 latency relative 40/s conventional stimulus design. Units are ms. Negative values indicate the latency was shorter than for 40/s conventional. Measurements are averaged across all stimulus frequencies for the given species. Significance of differences is indicated by the posterior probability that the difference in wave 1 latency is less than ± 0.25 and ± 0.5 ms.

			95% CI			
	Species	Mean	lower	upper	Pr($< \pm 0.25ms$)	Pr($< \pm 0.5ms$)
40/s ramp	gerbil	0.26	0.15	0.37	0.41	1.00
	mouse	0.38	0.24	0.51	0.03	0.96
80/s ramp	gerbil	0.14	0.03	0.26	0.97	1.00
	mouse	0.23	0.10	0.37	0.59	1.00
80/s conv.	gerbil	-0.27	-0.37	-0.16	0.36	1.00
	mouse	-0.25	-0.38	-0.12	0.48	1.00

Table 4: ABR threshold for 50/s ramp relative to 10/s conventional. Units are dB. Negative values indicate threshold for 50/s ramp was less than 10/s conventional. Measurements are averaged across all stimulus frequencies for the given species. Significance of differences is indicated by the posterior probability that the difference in threshold does not exceed ±2.5 or ±5 dB relative to 10/s conventional.

		95% CI			
	Species	lower	upper	Pr($< \pm 2.5dB$)	Pr($< \pm 5dB$)
gerbil	0.4	-1.8	2.5	0.97	1.0
mouse	0.0	-2.2	2.2	0.97	1.0

3.2 Conventional at 10/s vs. interleaved ramp at 50/s

Slower stimulus rates of 10 to 13/s are commonly used for ABR measurement in human subjects to minimize adaptation. Routine measurements at these slow rates is typically not feasible in anesthetized animals since prolonged anesthesia can have adverse effects on the subjects' metabolism. Here, we assess whether the interleaved ramp design allows rapid acquisition of ABR data equivalent to that acquired using a conventional stimulus design at 10/s. Since we used five frequencies in the interleaved ramp protocol, a rate of 50/s results in an effective rate of 10/s for each frequency.

We measured ABR at one frequency using the slower-rate conventional stimulus design (4 kHz in gerbil, 16 kHz in mouse) and compared results to those for the same tone frequency using the 50/s interleaved design (Fig. 4A-C). ABR thresholds between the two stimulus designs were within ±2.5 dB (Fig. 4D, Table 4). For wave amplitude and latency, measurements were similar between designs (Fig. 4B,C,E,F, Table 5 and 6). Although there was a 48% chance wave 1 amplitude in interleaved was at least 5% smaller than conventional in gerbil, there was only a 10% chance the difference was at least 10%.

Taken together, these results demonstrate that using an interleaved ramp design with a presentation rate of 50 tones/s minimizes adaptation of auditory nerve responses, and produces results nearly identical to ABR measurements using a 10 tone/s conventional design.

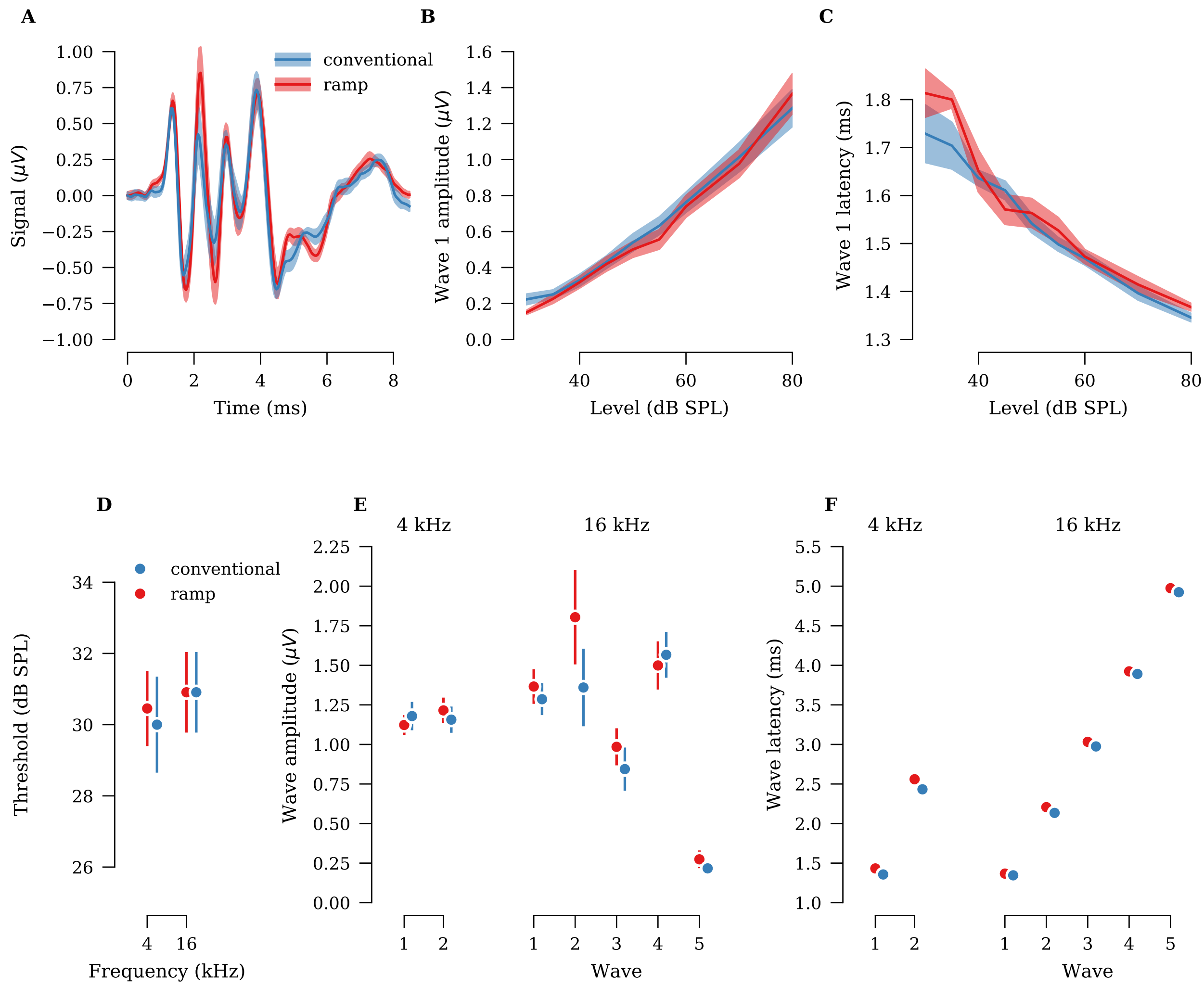

Fig. 4: ABR metrics acquired using a five-frequency 50/s interleaved ramp protocol are comparable to a 10/s conventional protocol for both low (4 kHz) and high (16 kHz) frequencies (n = 11 ears from gerbil, n = 11 ears from mouse). **A)** ABR waveforms (ensemble average across all ears) in response to 80 dB SPL tone pips at 16 kHz in mouse. Shaded area indicates ±SEM. **B)** ABR wave 1 amplitude as a function of stimulus level. **C)** ABR wave latency as a function of stimulus level. **A, B, C)** Shaded area indicates ±SEM. **D)** Average ABR thresholds. **E)** Average wave amplitude at 80 dB SPL. **F)** Average wave latency at 80 dB SPL. **D-F)** Error bars indicate ±SEM. Error bars in F are too small to be visible.

Table 5: ABR wave 1 amplitude for 50/s ramp relative to 10/s conventional. Units are percent. Negative values indicate the amplitude for 50/s ramp was smaller than 10/s conventional. Measurements are averaged across all stimulus frequencies for the given species. Significance of differences is indicated by the posterior probability that wave 1 amplitude in the 50/s ramp is at least 5 and 10% smaller than 10/s conventional.

		95% CI			
	Species	lower	upper	Pr($< -5\%$)	Pr($< -10\%$)
gerbil	-5	-13	4	0.48	0.1
mouse	6	-2	14	0.00	0.0

Table 6: ABR wave 1 latency for 50/s ramp relative to 10/s conventional. Units are ms. Negative values indicate the latency for 50/s ramp was shorter than 10/s conventional. Measurements are averaged across all stimulus frequencies for the given species. Significance of differences is indicated by the posterior probability that the difference in wave 1 latency is less than ± 0.05 and ± 0.1 ms.

		95% CI			
	Species	lower	upper	Pr($< \pm 0.05ms$)	Pr($< \pm 0.1ms$)
gerbil	0.07	0.04	0.10	0.07	0.97
mouse	0.02	-0.01	0.05	0.96	1.00

3.3 Refining the ordering of interleaved stimuli

In the experiments described above, the interleaved ramp design grouped stimuli by frequency within a single stimulus train (Fig. 1B). This grouping may still drive some adaptation since it rapidly sweeps through the sequence of levels for a single frequency before advancing to the next frequency. To test whether we can obtain further improvements, we assessed two alternative approaches to stimulus ordering: interleaved plateau and interleaved random (Fig. 1C,D). The interleaved plateau design groups stimuli by level, thereby sweeping through all frequencies at a particular level before moving to the next highest level. In this design, tones at each frequency are presented at 20% of the overall rate. For example, a 16 kHz tone in a five-frequency, 80 tones/s interleaved plateau train would appear at a rate of only 16 tones/s. However, at the highest stimulus levels there may be some overlap in excitation patterns along the cochlear partition (Kiang, 1965; Robles and Ruggero, 2001). The interleaved random orders tone frequency and level randomly. It imposes no constraints on the grouping of stimuli and may offer a compromise between the interleaved ramp (grouped by frequency) and interleaved plateau (grouped by level) stimulus designs.

To emphasize potential differences between the three stimulus designs, we measured ABR using a presentation rate of 80 tones/s (Fig. 5A). There was no difference in response threshold between the three stimulus designs, with the exception that threshold for random in mouse was 2.8-3.0 dB greater than ramp (Fig. 5B, Table 7). For ABR amplitude, wave 1 was approximately 11-15% greater in random compared to ramp (Fig. 5C-D, Table 8). There was no difference in wave 1 amplitude between plateau and ramp. Results for later waves were more variable

Table 7: ABR threshold for 80/s plateau and random stimulus designs relative to 80/s ramp. Units are dB. Measurements are averaged across all stimulus frequencies for the given species. Significance of differences is indicated by the posterior probability that the difference in threshold does not exceed ±2.5 or ±5 dB relative to ramp.

			95% CI			
	Species	Mean	lower	upper	$\Pr(< \pm 2.5dB)$	$\Pr(< \pm 5dB)$
plateau	gerbil	-0.7	-2.3	0.7	0.99	1.0
	mouse	0.9	-0.3	2.0	1.00	1.0
random	gerbil	3.0	1.4	4.5	0.28	1.0
	mouse	2.8	1.7	4.1	0.28	1.0

Table 8: ABR wave 1 amplitude for 80/s plateau and random stimulus designs relative to 80/s ramp. Units are percent. Measurements are averaged across all stimulus frequencies for the given species. Significance of differences is indicated by the posterior probability that wave 1 amplitude is 10 and 20% greater than ramp.

			95% CI			
	Species	Mean	lower	upper	$\Pr(> 10\%)$	$\Pr(> 20\%)$
plateau	gerbil	1	-6	7	0.00	0.00
	mouse	-2	-6	2	0.00	0.00
random	gerbil	15	8	23	0.92	0.11
	mouse	11	6	15	0.63	0.00

Table 9: ABR wave 1 latency for 80/s plateau and random stimulus designs relative to 80/s ramp. Units are ms. Measurements are averaged across all stimulus frequencies for the given species. Significance of differences is indicated by the posterior probability that the difference in wave 1 latency is less than ± 0.05 and ± 0.1 ms.

			95% CI			
	Species	Mean	lower	upper	$\Pr(< \pm 0.05ms)$	$\Pr(< \pm 0.1ms)$
plateau	gerbil	0.05	0.04	0.06	0.62	1.0
	mouse	-0.05	-0.06	-0.04	0.82	1.0
random	gerbil	-0.04	-0.05	-0.02	0.98	1.0
	mouse	-0.05	-0.06	-0.04	0.59	1.0

(Fig. 6A,C,E,G). For latency, all waves generally occurred slightly earlier (0.03 to 0.14 ms) in both random and plateau relative to ramp (Fig. 5E-F, Table 9).

Overall, we observe relatively small differences between the different interleaved configurations. Wave 1 amplitude is slightly larger for random compared to plateau and ramp, and threshold is slightly lower for ramp and plateau compared to random. Thus, among the three options, there is a trade-off between optimizing for threshold and wave amplitude.

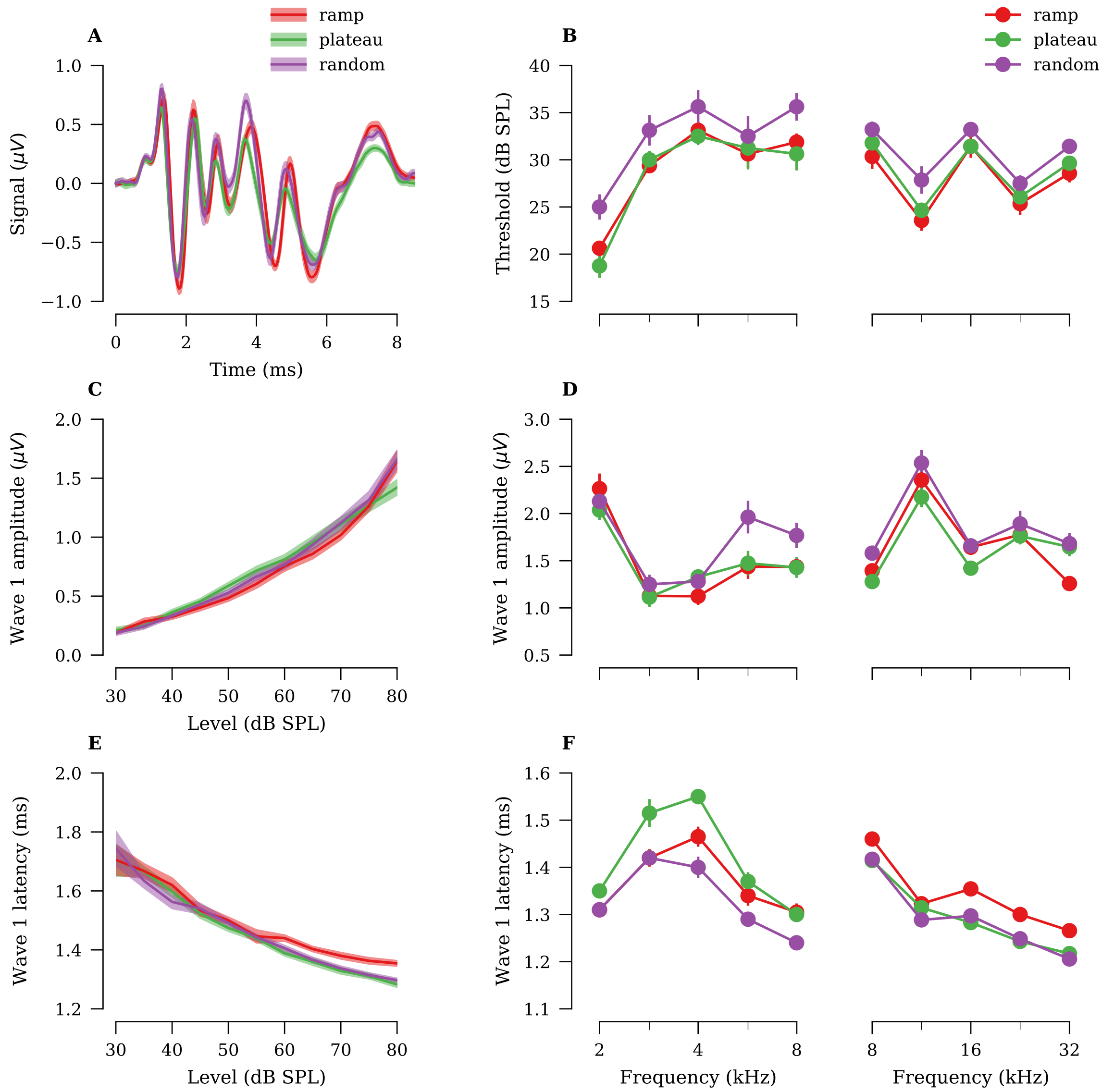

Fig. 5: Comparison of interleaved stimulus designs ($n = 8$ ears from gerbil, $n = 10$ ears from mouse). **A)** ABR waveforms (ensemble average across all ears) in response to 80/s 16 kHz, 80 dB SPL tone pips in mouse. Shaded area indicates ±SEM **B)** Average ABR thresholds. **C)** Average ABR wave 1 amplitude vs. stimulus level for 80/s 16 kHz tone pips. **D)** Average wave 1 amplitude at 80 dB SPL. **E)** Average ABR wave 1 latency vs. stimulus level for 80/s 16 kHz tone pips. **F)** Wave 1 latency at 80 dB SPL. Errorbars in all panels indicate ±SEM

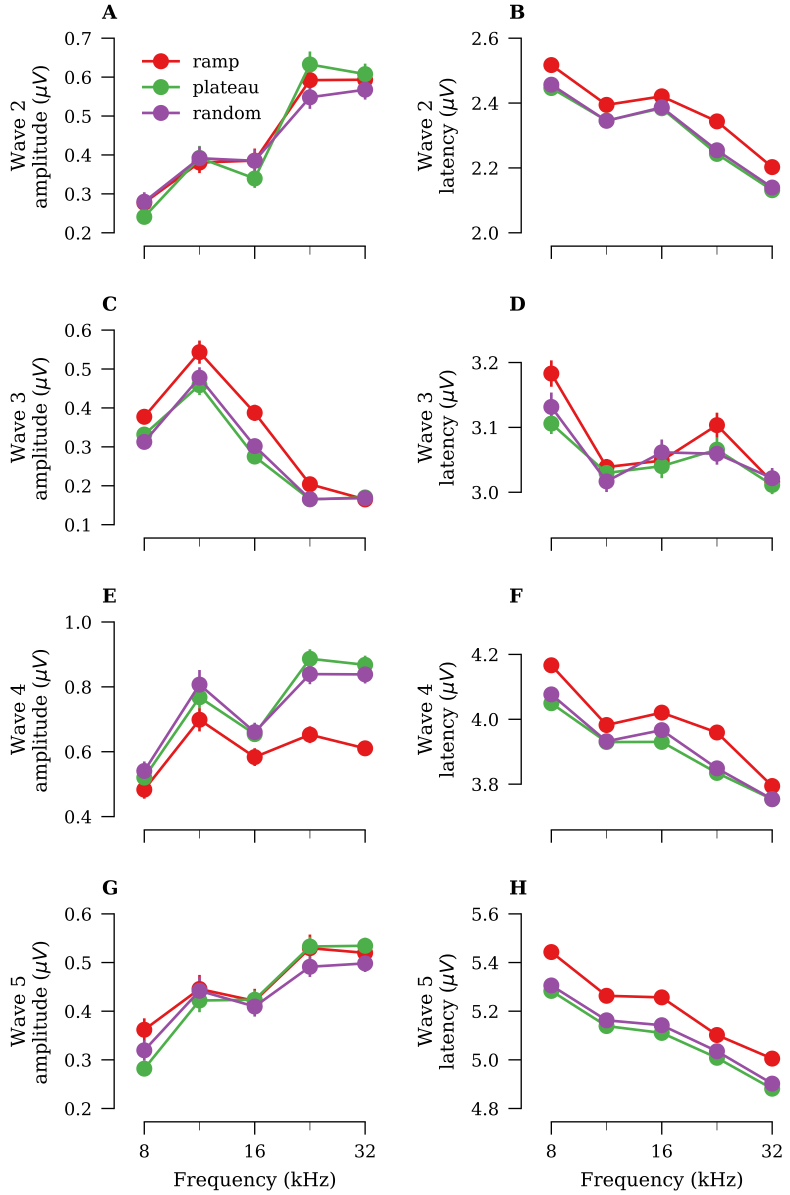

Fig. 6: Comparison of ABR metrics acquired using different interleaved designs for waves 2-5 at all frequencies and presentation rates tested ($n = 8$ ears in mouse). **A,C,E,G)** Wave amplitude at 80 dB SPL. **B,D,F,H)** Wave latency at 80 dB SPL. Error bars indicate ±SEM

3.4 ABR measurements in larger species

Experiments in larger species are often time-constrained. Thus, efficient ABR protocols may be especially beneficial for work with these animals. To test feasibility of the interleaved design in other species, we collected ABR data from ferret ($n = 2$ ears, Fig. 7A) and rhesus macaque ($n = 4$ ears, Fig. 7B). Average waveforms were clearly defined and permitted straightforward threshold measurement in both species. For macaque, data collected using the a 60/s interleaved ramp design were compared to data collected using a 30/s conventional design (Fig. 7B). Due to time constraints associated with experiments in nonhuman primates, we only had time to assess a limited number of stimulus levels using the conventional design. In the ranges tested with both designs in macaque, results were comparable (Fig. 7C-E). Wave 1 amplitude and latency showed some variability across tone frequency (Fig. 7D-E). This variability likely results from the low number of ears and, because of limited data, the comparison of responses at 20 dB sensation level (SL), rather than more standard 80 dB SPL. Aside from these relatively small differences, the more rapid interleaved protocol produced consistent results with the slower conventional design.

4 Discussion

We demonstrate that interleaving stimulus frequencies provides substantially more efficient ABR measurements than conventional designs. Interleaved stimuli reduce adaptation effects, which can affect response amplitude measurements. The benefits are greatest when ABRs are required for multiple frequencies.

Advantages of interleaved over conventional stimulus configurations In the comparison of interleaved to conventional ABR stimulus designs, interleaving tones of different frequencies reduces the effective rate at which individual frequencies are presented, thereby minimizing the effect of adaptation. This reduction is proportional to the number of frequencies included in the stimulus train. Specifically, we found:

- Interleaving five frequencies at 40 tones/s results in wave 1 amplitudes 20% larger than a conventional approach at 40 tones/s.
- Interleaving five frequencies at 80 tones/s produces wave 1 amplitudes comparable to a conventional approach at 40 tones/s while cutting acquisition time by 50%.
- Interleaving five frequencies at 50 tones/s produces wave 1 amplitudes equivalent to a conventional approach at 10 tones/s, while cutting acquisition time by 80% (i.e., when interleaving, five times as many frequencies can be acquired in the same amount of time as the conventional design).
- The interleaved stimulus design allows us to assess a larger range of frequencies and levels in time-sensitive experiments in rhesus macaque and ferret.

While our initial focus was on comparing the interleaved ramp design to the conventional approach, we found that the interleaved random design produces wave 1 amplitudes 11-15% larger than the interleaved ramp. Although we did not directly compare interleaved plateau with conventional in a single experiment,

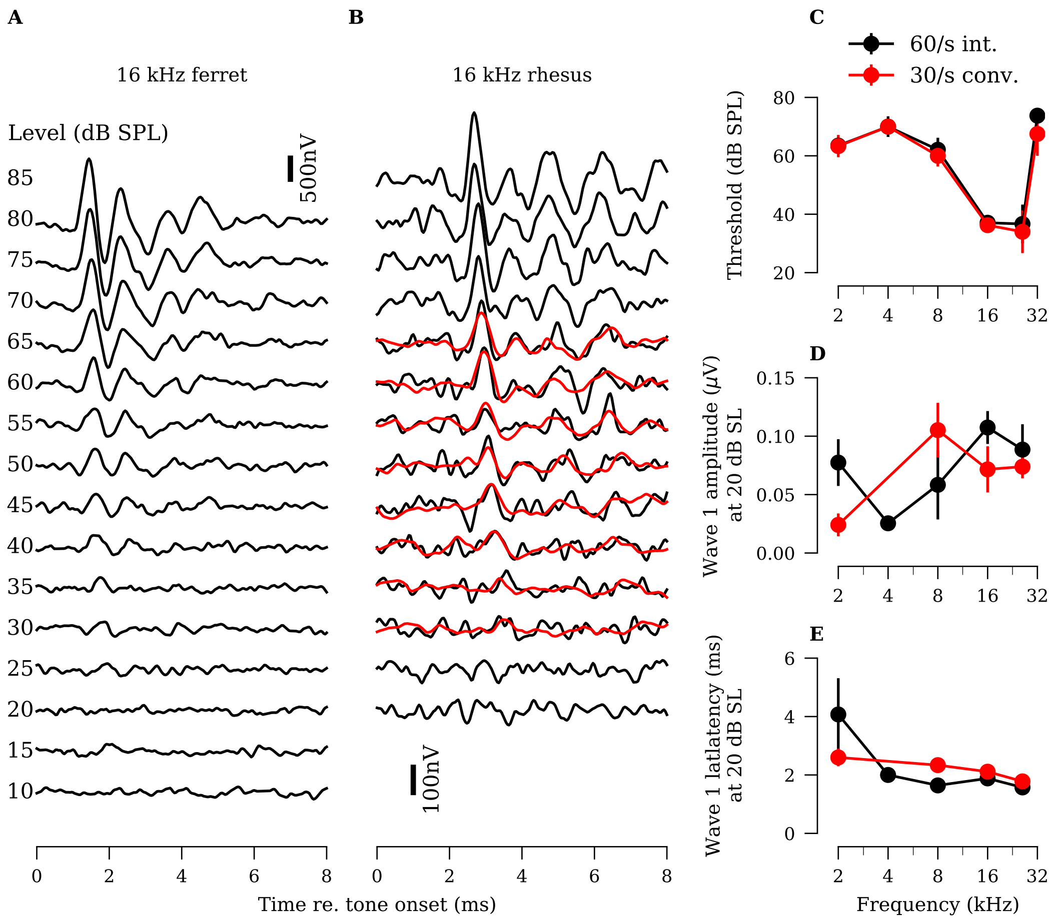

Fig. 7: ABR data acquired using the interleaved stimulus protocol in other species. **A)** ABR waveforms from a single ferret ear using a 10 frequency (2 to 45.2 kHz), 15 level (10 to 80 dB SPL) 80/s interleaved random protocol. Data from 16 kHz is shown. **B)** ABR waveforms from a single rhesus macaque run on a 7 frequency (0.5 to 32 kHz), 14 level (20 to 85 dB SPL) 60/s interleaved ramp stimulus design. Overlaid are waveforms acquired using a single-frequency 30/s conventional approach. Data from 16 kHz is shown. **C-E)** Average ABR measurements from rhesus macaque ($n = 4$ ears), comparing 60/s interleaved ramp with 30/s conventional. Errorbars indicate ±SEM.

comparing Fig. 2D with Fig. 5D suggests that we can expect to see larger wave 1 amplitudes using a 80/s interleaved random stimulus design as compared to 40/s conventional. However, the the increased response amplitude in the interleaved random design comes with the trade-off of a small threshold elevation of 2.8 to 3.0 dB as compared to the interleaved ramp design.

Although observed differences in wave latency were sometimes significant, they were small and clinically insignificant (e.g., less than 0.5 ms). Changes in wave latency can potentially result from small shifts in the place of excitation along the cochlea due to the other frequencies presented in the stimulus train (Polonenko and Maddox, 2019).

Parallels with other studies A study comparing a 9 tones/s conventional stimulus design with an 83 tones/s four-frequency interleaved ramp stimulus design (1 octave frequency spacing, 10 dB level spacing) found no significant difference in ABR threshold, wave amplitude or latency (Mitchell et al., 1996). A follow-up study expanded on this approach by increasing the number of frequencies in the interleaved ramp to 7, decreasing the level spacing to 5 dB and increasing the presentation rate to 100 tones/s. Although there was no difference in ABR threshold, this study showed a slightly reduced wave 1 amplitude in the interleaved ramp compared to conventional (Mitchell et al., 1999). Since we used a slower presentation rate of 50 tones/s for the interleaved ramp stimulus design when comparing with 10/s conventional, these studies are consistent with our data (Fig. 4). Although their data suggests that one can go up to 83 tones/s without adaptation, they introduced a 108 ms intertrain interval in the first study and 88 ms intertrain interval in the second study. This means that the average presentation rate was reduced from 83 to 59 tones/s and 100 to 88 tones/s, respectively. In our study, there was no intertrain interval. Taken together, this indicates that 50 to 59 tones/s likely represents the upper limit at which one can acquire data using an interleaved stimulus design while minimizing adaptation.

Traditionally, the maximum presentation rate for any ABR design is set by the analysis window used to measure evoked responses. For example, if the analysis window is 10 ms, then the maximum presentation rate is 100 tones/s (i.e., 10 ms separation between tone onsets). Faster tone presentations can result in multiple tones falling in the analysis window. More recently, a study tested an approach in which tone pips of five frequencies are presented randomly to each ear (Polonenko and Maddox, 2019). The interval between tone pips was generated using a Poisson process with an average presentation rate of 200 tones/s across all frequencies in a single ear. Since this study was performed in humans, only wave V was analyzed. At these high presentation rates, they showed a reduction of up to 50% in wave V amplitude and an increase in wave V latency compared to an approach where a single frequency was presented at an average rate of 40 tones/s using a Poisson process to randomize the interval between tone pips. The changes in response amplitude and latency were attributed to the enhanced place specificity of the ABRs due to other frequencies acting as a notched-noise masker. While there may be faster approaches to acquiring ABRs, they come with the trade-off of changes in wave amplitude and latency. These dense stimulation approaches also increase the computational complexity analysis, requiring deconvolution to measure unbiased ABR waveforms from responses that may be evoked by tones closely spaced in time.

Later waves in the auditory brainstem response are typically used to assess central auditory processing. Amplitude across the later waves was significantly enhanced by the interleaved stimulus design as compared to conventional when matched for rate (Fig. 3). Taken together, these data suggest that experiments assessing later waves will benefit from the interleaved stimulus design.

Interleaved configurations and mechanisms of adaptation The phenomenon of forward masking limits the rate at which the ABR and related measures, such as the compound action potential (CAP), can be measured without affecting the amplitude of the response (Spoor and Eggermont, 1971; Harris and Dallos, 1979; Burkard and Voigt, 1990). In our study, the rapidly presented 5 ms tone pips form an effective masker preceding the response to each pip. The different interleaved configurations determine the details of what tones frequencies and levels form the masker for each response. Based on the forward masking recovery equation developed by Harris and Dallos (1979), repetition rates of 10, 40, 50 and 80 tones/s suppress auditory nerve fiber activity by 0, 6, 9 and 15%, respectively. However, the actual suppression will be greater due to the cumulative effect of multiple tone pips acting as the masker. The reduced adaptation in the interleaved stimulus designs, as compared to conventional at the same presentation rate, is likely due to the increased recovery time between presentations of tones with the same frequency and level. For example, at 40 tones/s using 5 frequencies and 12 levels, an 80 dB SPL, 4 kHz tone is presented only once every 1.5 s.

When considering the effective spacing of same-frequency masking tones, there is a caveat to consider with the interleaved ramp design. In this design, tones are grouped by frequency (Fig. 1B) such that a sequence of tones at the same frequency are presented in rapid succession from low to high intensity. For this configuration, it might be surprising that we saw an average increase in wave 1 amplitude of 20% relative to the slower conventional stimulus design, considering an 80 dB SPL tone is immediately preceded by a 70 (mouse) or 75 (gerbil) dB SPL tone. However, lower level maskers have less suppression than high level maskers (Spoor and Eggermont, 1971; Harris and Dallos, 1979), thereby resulting in a relatively small cumulative masking effect.

In contrast, the interleaved plateau design groups tones by level (Fig. 1C), resulting in a relatively large gap between tones of the same frequency, proportional to the number of frequencies tested. For example, in a five-frequency 80 tones/s interleaved plateau design, each individual frequency is presented at a rate of 16 tones/s for an interstimulus interval of 62.5 ms. Despite this modified ordering, there was no difference in wave 1 amplitude between the ramp and plateau designs. This may be due to the larger spread of excitation along the cochlear partition at higher stimulus levels. Since the tones are grouped by stimulus level, the spread of excitation at 80 dB SPL may result in a moderate level of masking comparable to that seen in the ramp design.

The interleaved random design had wave 1 amplitudes that were larger than for the interleaved ramp; however, there was a small increase in ABR threshold of approximately 3 dB (Table 7). This is likely due to the possibility of high-intensity tones immediately preceding a low-intensity tone. If the frequency of the high-intensity tone falls within the receptive field of the cochlear region activated by the low-intensity tone, forward masking may suppress the neural response to the low-intensity tone thereby slightly increasing estimated threshold. The lack of a commensurate reduction in wave 1 amplitude despite the increase in threshold is likely due to forward masking having a greater effect on low-intensities as compared to high (Spoor and Eggermont, 1971).

All the interleaved stimulus designs offer an additional advantage over the conventional stimulus design. Since an animal's anesthetic plane can fluctuate during the course of the experiment, this fluctuation may adversely affect a subset of frequencies and levels in the conventional stimulus design. In contrast, the interleaved stimulus designs average out these fluctuations since all frequencies and levels are interrogated throughout the full course of the experiment.

4.1 Epilogue: technical considerations

Acoustic stimuli are typically generated digitally and converted to an analog waveform using a digital to analog converters (DAC) in the data acquisition system (e.g., internal sound card or National Instruments card). Thus, the resolution of the DAC is important for handling stimulus waveforms that contain both low and high-intensity tones. The 12-bit DAC used by Mitchell et al. (1996) has a theoretical dynamic range of 74 dB (Kester, 2005). Although a standard ABR experiment may require only 70 dB of dynamic range for a given frequency, the calibration required by some closed-field speakers used in animal studies can vary by up to 20 dB across the frequency range of interest due to resonances shaped by the acoustic cavity (Hancock et al., 2015), requiring DACs that support dynamic ranges of at least 90 dB. Further, the effective dynamic range of the DAC is lower than the theoretical dynamic range. For example, modern 24-bit DACs have a theoretical dynamic range of 146 dB, but ambient noise limits it to approximately 120 dB in practice (Fujimori et al., 2000). Fortunately, 120 dB of dynamic range is sufficient for implementing the interleaved stimulus design across a large range of levels and frequencies. Therefore, equipment with 24-bit DACs are highly recommended when implementing this stimulus paradigm.

The software used in this study, psiexperiment, is available under the BSD three-clause license on Github and is free for anyone to download and modify. It currently is designed to work with National Instruments hardware and has been tested on the PXI system configuration recommended by Eaton-Peabody Laboratories (Hancock et al., 2015). Instructions for installing the software and configuring it to run the various stimulus protocols are posted on the website. Although other hardware platforms are not supported as of publication time, the modular nature of psiexperiment will simplify the process of getting it to run on other hardware platforms (e.g., high-quality 24-bit sound cards, TDT System 3, etc.).

Acknowledgements The authors would like to acknowledge Jesyin Lai's assistance in acquiring data from ferret and Garnett McMillian for tutorials on implementing Bayesian models. This work was supported by the following grants:

– NIDCD R21 DC016969 to BNB
– NIDCD R01 DC014160 to JVB
– NIDCD R01 DC014950 to SVD
– Hearing Health Foundation Emerging Research Grant to BNB
– Pilot Program support from the Oregon National Primate Research Center, NIH award P51 OD011092 to Barr-Gillespie.

Conflict of interest

Brad Buran receives financial compensation for programming, data analysis, experiment design and tutoring services for government agencies, academic institu-

tions and private companies in addition to his part-time work at Oregon Health & Science University.

References

Bramhall NF, Konrad-Martin D, McMillan GP, Griest SE (2017) Auditory Brainstem Response Altered in Humans With Noise Exposure Despite Normal Outer Hair Cell Function. Ear and hearing 38(1):e1–e12

Buran BN (2015) Auditory-wave-analysis: V1.1 — Zenodo. https://zenodo.org/record/17365

Buran BN, David SV (2018) ψ - psiexperiment. Zenodo

Burkard R, Voigt HF (1990) Stimulus dependencies of the gerbil brain-stem auditory-evoked response (BAER). III: Additivity of click level and rate with noise level. The Journal of the Acoustical Society of America 88(5):2222–2234

Burkard R, Shi Y, Hecox KE (1990) A comparison of maximum length and Legendre sequences for the derivation of brain-stem auditory-evoked responses at rapid rates of stimulation. The Journal of the Acoustical Society of America 87(4):1656–1664

Eysholdt U, Schreiner C (1982) Maximum length sequences – a fast method for measuring brain-stem-evoked responses. Audiology: Official Organ of the International Society of Audiology 21(3):242–250

Fujimori I, Nogi A, Sugimoto T (2000) A multibit delta-sigma audio DAC with 120-dB dynamic range. IEEE Journal of Solid-State Circuits 35(8):1066–1073

Furman AC, Kujawa SG, Liberman MC (2013) Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. Journal of Neurophysiology 110(3):577–586

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013) Bayesian Data Analysis, 3rd edn. Chapman and Hall/CRC, Boca Raton

Hancock K, Stefanov-Wagner IJ, Ravicz ME, Liberman MC (2015) The Eaton-Peabody Laboratories Cochlear Function Test Suite. In: Association for Research in Otolaryngology 38th Annual Midwinter Meeting, Baltimore, MD

Harris DM, Dallos P (1979) Forward masking of auditory nerve fiber responses. Journal of Neurophysiology 42(4):1083–1107

Hoffman MD, Gelman A (2014) The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research 15:1593–1623

Kester W (ed) (2005) The Data Conversion Handbook. Newnes, Englewood Cliffs, NJ

Kiang NYS (1965) Discharge Patterns of Single Fibers in the Cat's Auditory Nerve. MIT Press, Cambridge, MA

Kujawa SG, Liberman MC (2006) Acceleration of Age-Related Hearing Loss by Early Noise Exposure: Evidence of a Misspent Youth. J Neurosci 26(7):2115–2123

Kujawa SG, Liberman MC (2009) Adding Insult to Injury: Cochlear Nerve Degeneration after "Temporary" Noise-Induced Hearing Loss. The Journal of neuroscience : the official journal of the Society for Neuroscience 29(45):14077–14085

Lentz JJ, Gordon WC, Farris HE, MacDonald GH, Cunningham DE, Robbins CA, Tempel BL, Bazan NG, Rubel EW, Oesterle EC, Keats BJ (2010) Deafness and

Retinal Degeneration in A Novel USH1C Knock-In Mouse Model. Developmental neurobiology 70(4):253–267

Lin HW, Furman AC, Kujawa SG, Liberman MC (2011) Primary Neural Degeneration in the Guinea Pig Cochlea After Reversible Noise-Induced Threshold Shift. JARO: Journal of the Association for Research in Otolaryngology 12(5):605–616

McMillan GP, Cannon JB (2019) Bayesian Applications in Auditory Research. Journal of speech, language, and hearing research: JSLHR 62(3):577–586

Melcher JR, Knudson IM, Fullerton BC, Guinan Jr JJ, Norris BE, Kiang NY (1996) Generators of the brainstem auditory evoked potential in cat. I. An experimental approach to their identification. Hearing Research 93(1–2):1–27

Millan J, Ozdamar O, Bohórquez J (2006) Acquisition and analysis of high rate deconvolved auditory evoked potentials during sleep. Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 1:4987–4990

Mitchell C, Kempton JB, Creedon T, Trune D (1996) Rapid acquisition of auditory brainstem responses with multiple frequency and intensity tone-bursts. Hearing Research 99(1):38–46

Mitchell CR, Kempton JB, Creedon TA, Trune DR (1999) The Use of a 56-Stimulus Train for the Rapid Acquisition of Auditory Brainstem Responses. Audiology and Neurotology 4(2):80–87

Nuzzo R (2014) Scientific method: Statistical errors. Nature News 506(7487):150

Polonenko MJ, Maddox RK (2019) The parallel auditory brainstem response. bioRxiv p 648659

Robles L, Ruggero MA (2001) Mechanics of the Mammalian Cochlea. Physiological reviews 81(3):1305–1352

Salvatier J, Wiecki TV, Fonnesbeck C (2016) Probabilistic programming in Python using PyMC3. PeerJ Computer Science 2:e55

Spoor A, Eggermont JJ (1971) Action potentials in the cochlea. Masking, adaptation and recruitment. Audiology: Official Organ of the International Society of Audiology 10(5):340–352

Szucs D, Ioannidis JPA (2017) When Null Hypothesis Significance Testing Is Unsuitable for Research: A Reassessment. Frontiers in Human Neuroscience 11

Valderrama JT, Alvarez I, de la Torre A, Carlos Segura J, Sainz M, Luis Vargas J (2012) Recording of auditory brainstem response at high stimulation rates using randomized stimulation and averaging. The Journal of the Acoustical Society of America 132(6):3856–3865