Effects of Stimulus Type on 16-kHz Detection Thresholds

Emily Buss,[1] Stacey G. Kane,[1,2] Kathryn S. Young,[2] Chloe B. Gratzek,[2] Danielle M. Bishop,[3] Margaret K. Miller,[3] Heather L. Porter,[3] Lori J. Leibold,[3] G. Christopher Stecker,[3] and Brian B. Monson[4]

Objectives: Audiometric testing typically does not include frequencies above 8 kHz. However, recent research suggests that extended high-frequency (EHF) sensitivity could affect hearing in natural communication environments. Clinical assessment of hearing often employs pure tones and frequency-modulated (FM) tones interchangeably regardless of frequency. The present study was designed to evaluate how the stimulus chosen to measure EHF thresholds affects estimates of hearing sensitivity.

Design: The first experiment used standard audiometric procedures to measure 8- and 16-kHz thresholds for 5- to 28-year olds with normal hearing in the standard audiometric range (250 to 8000 Hz). Stimuli were steady tones, pulsed tones, and FM tones. The second experiment tested 18- to 28-year olds with normal hearing in the standard audiometric range using psychophysical procedures to evaluate how changes in sensitivity as a function of frequency affect detection of stimuli that differ with respect to bandwidth, including bands of noise. Thresholds were measured using steady tones, pulsed tones, FM tones, narrow bands of noise, and one-third-octave bands of noise at a range of center frequencies in one ear.

Results: In experiment 1, thresholds improved with increasing age at 8 kHz and worsened with increasing age at 16 kHz. Thresholds for individual participants were relatively similar for steady, pulsed, and FM tones at 8 kHz. At 16 kHz, mean thresholds were approximately 5 dB lower for FM tones than for steady or pulsed tones. This stimulus effect did not differ as a function of age. Experiment 2 replicated this greater stimulus effect at 16 kHz than at 8 kHz and showed that the slope of the audibility curve accounted for these effects.

Conclusions: Contrary to prior expectations, there was no evidence that the choice of stimulus type affected school-age children more than adults. For individual participants, audiometric thresholds at 16 kHz were as much as 20 dB lower for FM tones than for steady tones. Threshold differences across stimuli at 16 kHz were predicted by differences in audibility across frequency, which can vary markedly between listeners. These results highlight the importance of considering spectral width of the stimulus used to evaluate EHF thresholds.

Key words: Audibility, Extended high-frequency, Sloping loss.

(Ear & Hearing 2024;45;486–498)

INTRODUCTION

Hearing sensitivity from 250 to 8000 Hz is well recognized as a clinical benchmark for speech recognition and environmental awareness (ANSI 1997). In contrast, the clinical utility of measuring extended high frequencies (EHFs; >8 kHz) has historically been limited to special circumstances, such as ototoxicity monitoring for patients being treated with platinum-based chemotherapies (Knight et al. 2007) or aminoglycoside antibiotics (Garinis et al. 2017; Blankenship et al. 2021). In addition to early identification of iatrogenic hearing loss, EHF thresholds are sensitive to effects of noise exposure and aging, with age-related hearing loss beginning as early as the first decade of life (Hemmingsen et al. 2021). Audibility of EHFs plays a role in sound source localization (King & Oldfield 1997; Brungart & Simpson 2009), masked speech recognition (Levy et al. 2015; Monson et al. 2019; Yeend et al. 2019; Hunter et al. 2020; Blankenship et al. 2021; Flaherty et al. 2021; Braza et al. 2022), spatial awareness related to talker head orientation (Monson et al. 2019), and perceived sound quality (Moore & Tan 2003). Clinically, elevated EHF thresholds may provide an explanation for hearing deficits that are not captured in the standard audiogram (Shaw et al. 1996; Badri et al. 2011; Yeend et al. 2019; Drennan 2021; Petley et al. 2021; Lough & Plack 2022). Recent interest in the role of EHF hearing motivated the present study, which considered how the stimuli used to evaluate EHF thresholds might affect results in school-age children and adults.

Historically, the rationale for focusing on hearing from 250 to 8000 Hz in both clinical assessment and research is based on studies showing excellent intelligibility of speech when listeners have access to cues in this frequency region (Assmann & Summerfield 2004; Monson et al. 2014). More recently, however, several groups have shown that hearing loss at EHFs significantly reduces masked speech recognition (Monson et al. 2019; Trine & Monson 2020; Saxena et al. 2022). For example, Mishra et al. (2022a,b) reported that recognition of digits in a six-talker babble is poorer for children and adults with elevated thresholds at one or more EHFs between 10 and 16 kHz than for listeners with normal EHF sensitivity. Effects of EHF audibility may be most evident when the masker is low-pass filtered at 8 kHz (Motlagh-Zadeh et al. 2019; Polspoel et al. 2022) or when EHF content of the masker is reduced by rotating the masker talker's head away from the listener (Monson et al. 2019; Flaherty et al. 2021).

Although there is growing evidence that EHF sensitivity affects functional hearing abilities, protocols for measuring EHF thresholds are not well developed and vary widely across clinics and research laboratories. Test-retest reliability is relatively good for EHF thresholds in children (Beahan et al. 2012; John & Kreisman 2017) and adults (Frank 1990, 2001; Schmuziger et al. 2004). However, individual differences tend to be larger at EHFs than at standard audiometric frequencies for both age groups (Stelmachowicz et al. 1989; Schmuziger et al. 2004; Rodriguez Valiente et al. 2014; Hemmingsen et al. 2021). This variability could be due to calibration error, differences in transducer placement, individual differences in listening strategy, or

[1]Department of Otolaryngology/Head and Neck Surgery, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; [2]Department of Health Sciences, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; [3]Center for Hearing Research, Boys Town National Research Hospital, Omaha, Nebraska, USA; and [4]Department of Speech and Hearing Science, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA.

<zdoi; 10.1097/AUD.0000000000001446>

true variability in EHF sensitivity (Stelmachowicz et al. 1982, 1989; Stevens et al. 1987).

The stimuli commonly used to measure EHF thresholds include steady tones, pulsed tones, frequency-modulated (FM) tones (also referred to as warble tones), and narrow bands of noise (Prendergast et al. 2017; Petley et al. 2021; Polspoel et al. 2022). Clinically, it is a common practice to use complex test stimuli, such as bands of noise or FM tones, to maintain attention to the task when testing young children. This practice is based on clinical experience and on data showing lower thresholds for these stimuli in young children (Thompson & Thompson 1972; Orchik & Rintelmann 1978). Pediatric audiologists may also vary the stimulus type within a session to maintain interest with very young children, although data on this approach are mixed (Massie et al. 2006). Anecdotally, dynamic EHF stimuli may be easier to discriminate from tinnitus, and adults prefer listening for pulsed tones compared with steady pure tones at both the lower and upper ends of the standard audiometric range (Burk & Wiley 2004) and at EHFs (Lentz et al. 2017). The American Speech Language Hearing Association recommends the use of a pulsed- or FM-tone stimulus to differentiate between the test signal and tinnitus (ASHA 2005).

The spectral content of stimuli used to evaluate EHF thresholds could be particularly important due to the prevalence of sloping hearing loss in this range. Several studies have reported that EHF thresholds are similar for steady and pulsed tones, but that FM-tone thresholds tend to be lower (Hamill & Haas 1986; Lentz et al. 2017). Lentz et al. (2017) evaluated effects of stimulus type on thresholds in 18- to 88-year olds, including some participants with hearing loss and/or tinnitus. They found that mean thresholds at 12.5 and 16 kHz were approximately 5 dB better for FM tones than steady tones; smaller effects were observed at lower frequencies. Hamill and Haas (1986) studied a group of six 22- to 28-year olds with normal hearing up to 8 kHz, and they observed lower 14- and 16-kHz thresholds for FM tones than for steady or pulsed tones, with a mean difference of 5.4 dB. In both cases, the authors reported that these stimulus effects varied across listeners, with some obtaining similar thresholds with different stimuli and others providing thresholds that differed by 20 dB or more. Better thresholds with FM tones than with steady or pulsed tones is likely related to better audibility of stimulus components below the stimulus center frequency. For example, thresholds obtained using an FM tone could reflect hearing sensitivity at or near the low-frequency edge of the FM sweep rather than at the spectral center of the sweep. This effect is observed within the standard audiometric range for listeners with steeply sloping hearing loss (Walker et al. 1984), but it may be substantially more common at EHFs, where the normal sensitivity curve is steeply sloping.

EXPERIMENT 1

The first experiment was designed to evaluate the effects of stimulus type on EHF thresholds in school-age children and young adults with normal hearing in the standard audiometric range. Thresholds were measured at 8 and 16 kHz using steady tones, pulsed tones, and FM tones. The primary question of interest was how stimulus type affects EHF thresholds for school-age children as compared with adults. Common practice among clinical audiologists suggests that thresholds for young children are lower and more stable for dynamic stimuli (e.g., FM tones and pulsed tones) compared with steady tones, due to greater ease of attending to dynamic stimuli. Thresholds obtained at the edge of a steeply sloping hearing loss can be lower for FM tones than pulsed tones or steady tones, and this effect may be more common at EHFs than the standard audiometric range, due to the greater prevalence of hearing loss at EHFs. While EHF sensitivity is typically better in younger than older listeners, as many as 7% of school-age children and 18% of young adults with hearing thresholds ≤20 dB HL between 250 and 8000 Hz have at least one threshold >20 dB HL between 10 and 16 kHz (Mishra et al. 2022a,b). The a priori prediction was that children might have an easier time attending to dynamic stimuli (pulsed and FM tones) and therefore might have lower thresholds for dynamic stimuli than for steady tones at both 8 and 16 kHz. Further, audible cues below the stimulus center frequency might improve 16-kHz thresholds for FM tones as compared with pulsed or steady tones, particularly for adults who are more likely to have EHF threshold elevation; such a result would suggest that threshold estimation using FM tones does not provide an accurate estimate of sensitivity at 16 kHz.

Methods

Participants • The population targeted for recruitment was 5- to 29-year olds with thresholds ≤20 dB HL bilaterally at octave frequencies between 250 and 8000 Hz (ANSI 2010), and parent or self-report of healthy middle ear status for the month preceding the test day. The decision to cap adult age at 29 years was based on the rationale that auditory performance at this age is fully mature, and age-related hearing loss is not yet prominent; this is also a cutoff commonly used when measuring normative thresholds by age (Rodriguez Valiente et al. 2014). Participants were recruited through the University of North Carolina Chapel-Hill (UNC; n = 20) and Boys Town National Research Hospital (BTNRH; n = 27). All participants provided informed consent and were compensated for their participation. The study sample included 47 participants in 4 groups: 12 children 5 to 7 years old (mean = 6.4 years; 6 females and 6 males), 13 children 8 to 11 years old (mean = 9.6 years; 7 females, 5 males, and 1 other), 11 adolescents 12 to 17 years old (mean = 14.5 years; 1 female and 10 males), and 11 adults 18 to 28 years old (mean = 23.5 years; 8 females and 3 males). Each participant was tested in either the left or right ear (n = 23 and n = 24, respectively), selected at random.

Stimuli • Stimuli were steady tones, pulsed tones, and FM tones presented at 8 and 16 kHz using professionally calibrated audiometers. Pulsed tones were 200 msec in duration excluding 60-msec onset and offset ramps, with pulses repeating every 400 msec (i.e., with 200-msec interstimulus intervals). Frequency modulation was 5% of the center frequency, with a rate of 5 Hz. Testing at UNC used the GSI AudioStar Pro audiometer (Eden Prairie, MN) and Sennheiser HDA 200 headphones (Wedemark, Germany); testing at BTNRH used a Madsen Astera^2 audiometer (Natus, Middleton, WI) with Sennheiser HDA 300 headphones for adults and a GSI 61 audiometer with RadioEar DD450 headphones (Middelfart, Denmark) for children and adolescents. All three audiometers and headphones had been professionally calibrated within the preceding year. Stimulus levels were also measured using a 6-cc flat-plate coupler and sound-level meter, confirming consistent calibration across test

sites. Procedures for calibration are described in more detail in the Methods section of experiment 2.

Procedure • Testing took place in a double-walled sound booth using the modified Hughson–Westlake procedure (ANSI 2004), with a final step size of 5 dB. For child participants, thresholds were obtained using either conventional or conditioned-play audiometry, as deemed appropriate by the tester. Thresholds were remeasured for adults and for a subset of children using 1-dB steps. The pattern of results obtained with 1-dB steps was not qualitatively different from that observed using 5-dB steps, but results appeared more variable, perhaps due to the longer test time associated with smaller step sizes. Thresholds obtained with 1-dB steps were therefore not considered in the results. The orders of frequencies and stimulus types were quasi-randomized for each participant. Thresholds for all three stimulus types were obtained for one frequency before proceeding to the next frequency. Two thresholds were obtained for each stimulus type at each frequency; the final threshold was the average of those two estimates. Audiometric testing was completed in one session, and breaks were given as necessary to minimize effects of fatigue. All procedures were reviewed and approved by the Institutional Review Boards at UNC and BTNRH.

Results

Figure 1 shows thresholds plotted as a function of participant age on a log scale for 8 and 16 kHz, with results shown separately by stimulus type. Line fits to the data in each panel are indicated with solid gray lines, and dotted lines indicate 0 dB HL. Test-retest reliability was excellent for all stimuli at both frequencies. Replicate thresholds were within 5 dB for all but three cases (99% of data),* and all replicate thresholds were within 10 dB. As expected based on the inclusion criteria, 8-kHz thresholds were at or below 20 dB HL for all participants and all stimulus conditions, with one exception; for 1 child (9.2 years), the mean steady-tone 8-kHz threshold was 22.5 dB HL. For all three stimulus types, there was a trend for lower 8-kHz thresholds in older participants. For example, the mean steady-tone 8-kHz threshold was 8.3 dB HL for 5- to 7-year olds and 1.7 dB HL for adults. Across age, mean 8-kHz thresholds were similar for the three stimuli, with values of 5.1 dB HL for steady tones, 4.9 dB HL for pulsed tones, and 4.1 dB HL for FM tones.

Similar to the data obtained at 8 kHz, most 16-kHz thresholds were at or below 20 dB HL. Exceptions were three steady-tone and two pulsed-tone thresholds from three participants (16.8, 26.3, and 28.3 years); those thresholds ranged from 25 to 37.5 dB HL. Individual differences were larger for thresholds at 16 kHz than at 8 kHz. Thresholds of −20 dB HL indicate detection at the lowest level presented by the audiometer. Whereas no participant obtained a mean threshold of −20 dB HL at 8 kHz, this occurred at 16 kHz for the steady tone (n = 4; 5.5, 15.3, 16.6, and 25.2 years), the pulsed tone (n = 5; 5.5, 10.0, 15.3, 16.6, and 25.2 years), and the FM tone (n = 6; 5.5, 6.5, 15.3, 16.6, 22.1, and 25.2 years). In contrast with the 8-kHz data, thresholds at 16 kHz tended to rise with increasing participant age. The mean steady-tone threshold at 16 kHz was −5.8 dB HL for 5- to 7-year olds and 6.6 dB HL for adults. Across age, mean 16-kHz thresholds as a function of stimulus type were −0.5 dB HL for steady tones, −1.8 dB HL for pulsed tones, and −6.2 dB HL for FM tones.

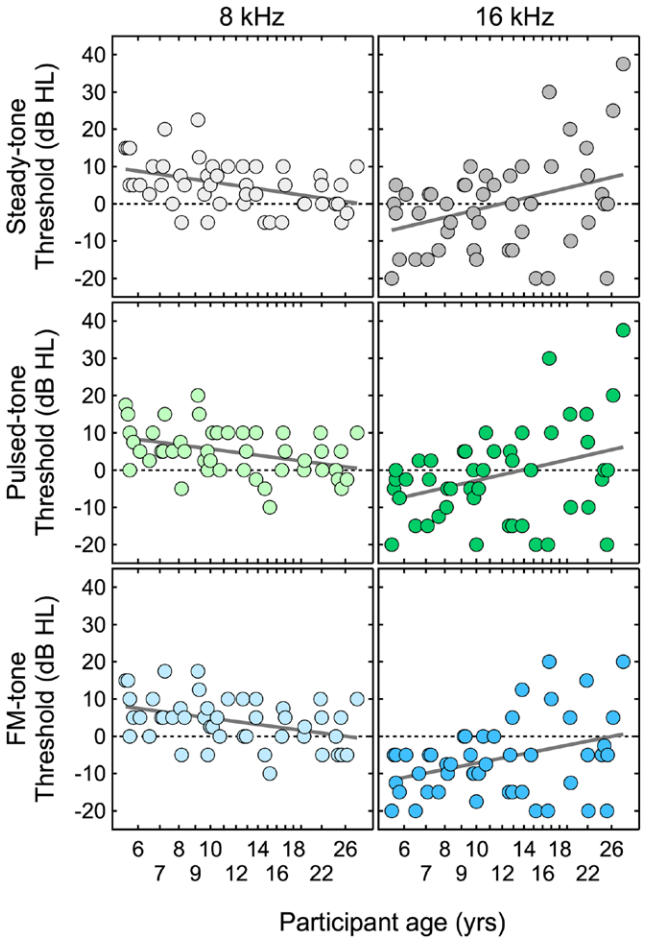

Fig. 1. Audiometric thresholds in dB HL as a function of age, plotted separately for each stimulus condition (by row) and frequency (by column). dB HL indicates decibels hearing level; FM, frequency-modulated.

These observations were confirmed using a linear mixed model. The dependent variable was threshold measured in dB HL. Independent variables were frequency (8 and 16 kHz), stimulus type (steady, pulsed, and FM tone), and participant age. Frequency and stimulus type were coded as categorical variables. To account for the decelerating effects of development in children as age increases, age was $\log_{10}$ transformed and centered on 10 years, the approximate center of the age range in log units. Reference conditions were 8 kHz and steady tones. On the basis of visual inspection of the data, variance was modeled separately for participants <12 and ≥12 years of age for each frequency. This analysis resulted in significant effects of frequency ($F_{1,225} = 34.09$, $p < 0.001$), age ($F_{1,45} = 10.17$, $p = 0.003$), and stimulus type ($F_{2,225} = 3.30$, $p = 0.039$). There were significant two-way interactions between age and frequency ($F_{1,225} = 23.65$, $p < 0.001$) and between stimulus type and frequency ($F_{2,225} = 3.53$, $p = 0.031$). There was no significant interaction between age and stimulus type ($F_{2,225} = 0.32$, $p = 0.726$)

*A 10-dB difference between replicate threshold estimates was observed at 16 kHz for a 7.7-year old (FM tone), 10.2-year old (pulsed tone), and 12.0-year old (pulsed tone).

or between age, stimulus type, and frequency ($F_{2,225}$ = 0.12, p = 0.885). Table 1 shows the parameters of this model fit. For 12- to 17-year olds and adults, the SD of thresholds for each stimulus was a factor of 2.3 to 3.7 higher at 16 kHz than at 8 kHz. For younger children, those factors were 0.8 to 1.5.

Two additional models were run to better understand the significant two-way interactions, one with 8-kHz data and one with 16-kHz data. The model of 8-kHz data confirmed significant improvements in thresholds with increasing age ($F_{1,45}$ = 9.90, p = 0.003), a significant effect of stimulus ($F_{2,90}$ = 3.76, p = 0.027), and absence of an age-by-stimulus interaction ($F_{2,90}$ = 0.36, p = 0.699). The model of 16-kHz data confirmed significant worsening in thresholds with increasing age ($F_{1,45}$ = 7.37, p = 0.009), a significant effect of stimulus ($F_{2,90}$ = 30.38, p < 0.001), and absence of an age-by-stimulus interaction ($F_{2,90}$ = 0.93, p = 0.398). Table 2 shows the parameters of these model fits. Individual differences in 16-kHz thresholds tended to grow with increasing age. For example, the SDs of steady-tone thresholds by age group were 8.9 dB (5 to 7 years), 7.4 dB (8 to 11 years), 15.2 dB (12 to 17 years), and 16.6 dB (18 to 29 years). These linear models captured data trends that are evident in Figure 1, but some of those effects are probably not entirely linear. For example, it is unlikely that 8-kHz thresholds improved with increasing age past early adolescence, and there appeared to be floor effects for the best performers at 16 kHz.

One question of practical importance is the extent to which individual differences were preserved across stimulus types. Figure 2 shows thresholds for the dynamic stimuli (FM tones and pulsed tones) as a function of the steady-tone thresholds. Symbol size reflects the number of observations for each combination of values. Participant's age is not represented in this figure. At 8 kHz, individual participants' steady-tone thresholds were within ±5 dB of their pulsed-tone and FM-tone thresholds for the majority of cases (100% and 98%, respectively). At 16 kHz, pulsed-tone thresholds were within ±5 dB of the associated steady-tone thresholds in 91% of cases; this value fell to 60% of cases for the FM tones, with 17% of FM-tone thresholds being more than 10 dB below the associated steady-tone threshold. The maximum difference between FM- and steady-tone thresholds was 20 dB (26.3 years).

There is some evidence that the difference between FM- and steady-tone thresholds increases with increasing steady-tone threshold. The solid gray lines in Figure 2 indicate linear fits to pulsed- and FM-tone thresholds as a function of steady-tone thresholds. For the 16-kHz FM-tone data, the slope of this line is significantly less than 1 (β= 0.72, 95% confidence

TABLE 1. Linear mixed model evaluating thresholds as a function of participant age ($\log_{10}$ transformed and centered on 10 yrs), frequency (8 and 16 kHz), and stimulus type (steady, pulsed, or FM tone)

	Value	SE	Degrees of freedom	t	p
(Intercept)	5.92	0.90	225	6.56	<0.001
Age	**−12.83**	**4.02**	**45**	**−3.19**	**0.003**
PulsedTone	−0.26	0.42	225	−0.62	0.533
FMTone	**−1.03**	**0.42**	**225**	**−2.47**	**0.014**
Freq	**−7.53**	**1.29**	**225**	**−5.84**	**<0.001**
Age: PulsedTone	1.34	1.80	225	0.74	0.458
Age: FMTone	1.14	1.80	225	0.63	0.528
Age: Freq	**33.81**	**6.95**	**225**	**4.86**	**<0.001**
PulsedTone: Freq	−0.95	1.82	225	−0.52	0.603
FMTone: Freq	**−4.59**	**1.82**	**225**	**−2.52**	**0.013**
Age: PulsedTone: Freq	−2.39	9.82	225	−0.24	0.808
Age: FMTone: Freq	−4.86	9.82	225	−0.49	0.621

The reference condition was the 8 kHz steady tone.
FM indicates frequency-modulated.
Coefficients associated with p <0.05 are indicated with bold font.

TABLE 2. Linear mixed models evaluating 8- and 16-kHz thresholds as a function of participant age ($\log_{10}$ transformed and centered on 10 yrs) and stimulus type (steady, pulsed, or FM tone)

8-kHz Data	Value	SE	Degrees of freedom	t	p
(Intercept)	5.92	0.91	90	6.47	<0.001
Age	**−12.86**	**4.09**	**45**	**−3.15**	**0.003**
PulsedTone	−0.25	0.39	90	−0.64	0.522
FMTone	**−1.04**	**0.39**	**90**	**−2.63**	**0.010**
Age: PulsedTone	1.39	1.76	90	0.79	0.432
Age: FMTone	1.16	1.76	90	0.66	0.511
16-kHz Data	**Value**	**SE**	**Degrees of freedom**	t	p
(Intercept)	−1.91	1.72	90	−1.11	0.270
Age	**21.02**	**7.74**	**45**	**2.71**	**0.009**
PulsedTone	−1.21	0.73	90	−1.65	0.102
FMTone	**−5.42**	**0.73**	**90**	**−7.42**	**<0.001**
Age: PulsedTone	−1.02	3.43	90	−0.30	0.768
Age: FMTone	−4.46	3.43	90	−1.30	0.197

The reference condition was the steady tone.
FM indicates frequency-modulated.
Coefficients associated with p <0.05 are indicated with bold font.

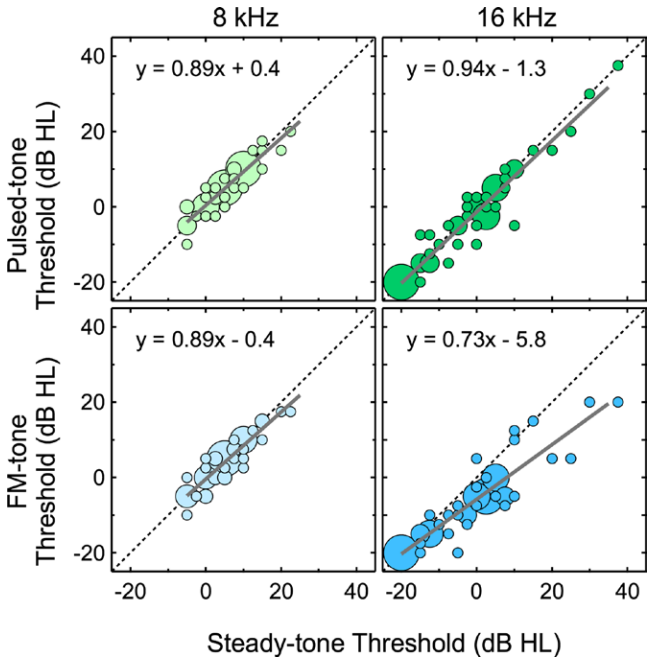

Fig. 2. Thresholds for dynamic stimuli (pulsed tones and FM tones) are plotted as a function of steady-tone thresholds. Thresholds for pulsed tones are shown in the top row, and those for FM tones are shown in the bottom row. Results for 8 kHz are shown on the left, and those for 16 kHz are shown on the right. Symbol size reflects the number of observations at each combination of values. The dotted line indicates identical thresholds for the steady and dynamic stimuli. Thick gray lines indicate data fits, with the associated equations in the upper left of each panel. FM indicates frequency-modulated.

interval = 0.62 to 0.83). One caveat when evaluating this result is the −20 dB HL limit of the audiometer; this limit could obscure variability in sensitivity for better-performing participants. On the basis of this consideration, a second line was fitted after removing the 13 cases with steady-tone thresholds below −10 dB HL (not shown). A slope of less than 1 was also observed in this subset of data ($\beta = 0.75$, 95% confidence interval = 0.58 to 0.93). This result is consistent with the idea that increasing hearing loss at 16 kHz is associated with a larger threshold reduction when using an FM tone as compared with a steady tone.

Discussion

The dataset from experiment 1 shows effects of age that differ for 8 and 16 kHz. Line fits to the data as a function of age at 8 kHz indicate that thresholds improved by 8.2 to 9.1 dB between 5 and 28 years of age. The opposite trend was observed at 16 kHz, where thresholds rose by 12.3 to 14.9 dB across the same age range. Differential effects of age across frequency have been documented in previous studies (Trehub et al. 1988; Rodriguez Valiente et al. 2014). Whereas maturation of attention tends to improve a listener's ability to make use of sensory input (Haapaniemi 1996), aging effects that are evident by 10 years of age (Buren et al. 1992; Hemmingsen et al. 2021; Mishra et al. 2022b) reduce the quality of sensory cues available at EHFs. Also consistent with previous literature (Frank 2001; Rodriguez Valiente et al. 2014), thresholds were more variable across participants at 16 kHz than 8 kHz. The range of mean steady-tone thresholds across all participants was −5 to 22.5 dB HL at 8 kHz and −20 to 37.5 dB HL at 16 kHz.

The primary goal of this study was to better understand how the choice of stimuli affects EHF detection thresholds obtained using standard clinical procedures in school-age children and adults. The expectation at the outset was that children might have lower thresholds for dynamic stimuli, including pulsed tones and FM tones, as compared with steady tones. Contrary to this expectation, mean 8 kHz thresholds were similar across all three stimulus types to within 1 dB. At 16 kHz, there were no differences observed between steady- and pulsed-tone thresholds; however, thresholds measured with FM tones were significantly lower than those measured with steady tones, regardless of participant age. For both children and adults, the mean difference between thresholds for FM and steady tones was approximately 5 dB, the smallest step used in standard clinical threshold estimation procedures. This finding is consistent with previous data from adults (Hamill & Haas 1986; Lentz et al. 2017). In contrast, mean thresholds for steady and pulsed tones were within 2.2 dB for all frequencies.

There are two possible explanations for the differences observed between steady and FM tones at 16 kHz. It is possible that FM tones were easier to attend to than steady tones at 16 kHz, either due to their dynamic temporal structure or the different sound quality compared with a participant's tinnitus. This possibility is undermined by the observation of similar thresholds for FM and steady tones at 8 kHz, and by the fact that thresholds were similar for steady tones and pulsed tones at 16 kHz, despite the dynamic nature of pulsed tones. A more likely explanation for the lower 16-kHz thresholds obtained with FM tones is based on the steep slope of the audibility curve at 16 kHz, even for participants with normal EHF hearing. It is possible that some participants relied on bursts of audible sound at or near the low end of the frequency sweep; this would be consistent with the spontaneous subjective report from several participants that the FM tone near threshold sounded like it was temporally pulsed. An FM tone with 5% modulation depth introduces energy approximately 0.84 semitones below and above the center frequency; at 16 kHz, this corresponds to 15.2 to 16.8 kHz. Differential stimulus effects as a function of frequency are illustrated in Figure 3, which shows the normal audibility curve (solid gray line) and the spectral width of an FM stimulus (in blue) at 8 and 16 kHz.

If this explanation is correct, then the slope of the audibility curve determines the size of the stimulus effect. In the current dataset, the largest difference between 16-kHz thresholds obtained with steady and FM tones was 20 dB. This effect size exceeds the 10-dB criterion often adopted when evaluating whether changes in threshold following exposure to noise or ototoxic agents are clinically meaningful (Le Prell et al. 2022; OSHA 2022). A difference in sensitivity of this magnitude between 15.2 and 16 kHz implies that the audibility curve for these participants is quite steep, more than 100 dB per octave. While it is rare to observe audibility curves that steep below 8 kHz, there are reports in the literature of even steeper losses within the standard audiometric range. For example, Rosler and Anderson (1978) reported slopes as steep as 300 to 350 dB per octave between 2.5 and 5 kHz. The maximum slope may be even higher for EHFs, given the trend for steeper slopes at higher frequencies within the standard

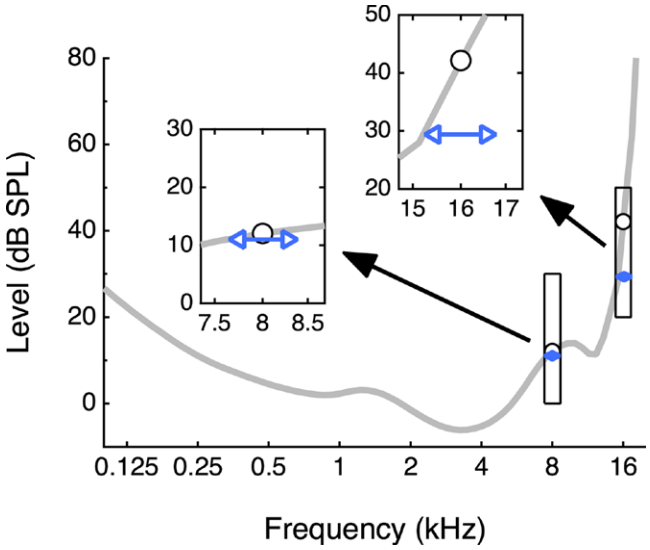

Fig. 3. Illustration showing how the normal sensitivity curve could result in lower thresholds for FM tones than steady tones at 16 kHz. Solid gray lines indicate thresholds in quiet for pure tones 20 to 20,000 Hz, plotted as a function of frequency. This curve is based on data from 12 recent studies, compiled by Suzuki and Takeshima (2004; Fig. 6). Boxes and inset panels illustrate stimulus features at 8 and 16 kHz. Circles indicate the frequency of a steady tone. The blue lines with arrows indicate the spectral extent of an FM tone with 5% modulation depth. FM indicates frequency-modulated.

audiometric range (Rosler & Anderson 1978). Experiment 2 was undertaken to evaluate whether slope of the audibility curve for individual listeners can predict their thresholds at 16 kHz for stimuli that differ in spectral width.

EXPERIMENT 2

The goal of this experiment was to evaluate the relationship between the slope of the audibility curve and effects of the spectral content of stimuli used to measure detection threshold. A modified swept-frequency Bekesy tracking procedure, described below, was used to estimate the audibility curve because it estimates detection thresholds at many points across a range of frequencies. This general approach for characterizing the audibility curve is not widely used in current clinical practice, or indeed in modern hearing research, but there is a large historical literature on this technique (Gelfand & Calandruccio 2023), and it is available in some modern clinical equipment (e.g., Madsen Astera). This experiment also considered a range of stimuli that might be used to evaluate thresholds, including the stimuli evaluated in experiment 1 and two types of band-pass noise that are available on clinical audiometers. Data obtained using a modified Bekesy tracking procedure were compared with thresholds obtained using standard clinical procedures and a two-alternative forced-choice (2AFC) adaptive procedure, with the goal of verifying results obtained with Bekesy tracking and facilitating comparisons to published clinical and laboratory data, respectively.

Methods

Participants • Inclusion criteria were the same as for the previous experiment with two exceptions. Children were excluded due to the greater time commitment and task demands associated with the psychophysical procedures, and data were collected at a single site (UNC). There were 16 participants between 18 and 28 years of age (mean = 22.9 years; 14 females and 2 males). Seven of these participants provided data for the full protocol, and nine of them provided 2AFC and Bekesy data only at 16 kHz due to time constraints and to the particular interest in EHFs. Whereas data collection at 16 kHz could be completed in approximately 1.5 hr, the full protocol included other frequencies and required five to six sessions of 1 hr each. Each participant was randomized to provide data in the left or right ear (n = 7 and n = 9, respectively).

Stimuli • Audiometric thresholds were measured at 4, 8, 12.5, and 16 kHz using pulsed tones, FM tones, and one-third-octave (1/3-oct) bands of noise, presented using the same professionally calibrated GSI AudioStar Pro audiometer and Sennheiser HDA 200 circumaural headphones used in experiment 1. Two-alternative forced-choice testing was carried out using five stimulus types, with parameters based on stimuli available in commercial audiometers: steady tones, pulsed tones, FM tones, narrowband (NB) noise, and 1/3-oct noise. All stimuli were 1 sec in duration, with 20-msec raised cosine ramps. The steady tone was a sinusoid. The pulsed tone was a sequence of three 200-msec tone bursts, each ramped on and off with 20-msec ramps and separated by 200-msec interburst gaps.† For the FM tone, frequency modulation was applied with a 5-Hz sinusoid that extended ±5% around the center frequency. The NB noise was a sample of Gaussian noise that was transformed into the frequency domain, restricted to a band ±5% around the center frequency, and transformed back into the time domain. The NB noise stimulus has the same spectral width as an FM tone, and it is similar to the "pediatric noise" or "fresh noise" that is available on some modern audiometers (Norrix & Anderson 2015). The 1/3-oct band of noise was generated using the same methods as NB noise, but with a bandwidth that spanned ±12% around the center frequency. A new random sample of noise was computed before each stimulus presentation. The modified Bekesy tracking procedure used pulsed-tone stimuli, as described for the forced-choice task. The full protocol included 2AFC and modified Bekesy testing at 4, 8, 12.5, and 16 kHz; some participants only completed testing at 16 kHz.

Stimuli for the 2AFC and modified Bekesy procedures were generated in MATLAB (Mathworks Inc, Natick, MA), played out of a U-phoria UMC202HD soundcard (Behringer, Willich, Nordrhein-Westfalen, Germany), with 24-bit depth and a rate of 88.2 kHz, and presented over one earphone of a RadioEar DD450 headset. The Sennheiser HDA 200 phones used for audiometry and the RadioEar DD450 phones used for the 2AFC and modified Bekesy procedures have functionally equivalent reference equivalent threshold sound pressure levels (RadioEar; Smull et al. 2019). Calibration of psychophysical stimuli was carried out using a sound level meter (System 824; Larson Davis, Depew, NY), a ½-in random incidence microphone (model 2559; Larson Davis), and a 6-cc flat-plate coupler (AEC201-A; Larson Davis). The frequency response of each earphone was characterized based on five recordings of white noise generated in Audacity® at 88.2 kHz, each >10 sec. The earphone was removed and replaced on the

†Pulsed tones in experiment 1 were gated with 60-msec ramps. For experiment 2, 20-msec ramps were used.

coupler before each recording. Recordings were made by routing the output of the sound level meter to an external soundcard (USB-SA; Andrea Communications, Farmingdale, NY), sampling at 88.2 kHz. The spectrum for each recording were characterized using the *pwelch* function in MATLAB, with 512-point windows spaced at 256-point intervals. The mean spectra for the left and right earphone were within ±1.5 dB at all frequencies between 3 and 18 kHz. The frequency response averaged across samples and across ears was used to set levels relative to the calibration value in decibels sound pressure level (dB SPL) based on a 16-kHz NB noise. Measurements were repeated using the professionally calibrated hardware (GSI AudioStar Pro and Sennheiser HDA 200 phones). For both sets of measurements, results at 16 kHz were within 0.5 dB of values expected based on the published data (RadioEar; Smull et al. 2019).

Stimuli were calibrated at their center frequencies, without additional spectral shaping. For example, the long-term average power spectrum of the 1/3-oct band of noise was approximately flat at the input to the headphones, but it was not perfectly flat at the output. This is typical for stimuli presented using commercially available audiometers. Deviations in calibration across frequency at the output of the headphones were relatively modest for the stimuli and hardware used in the present experiment. For the 1/3-oct noise, deviations in level in dB SPL across frequency relative to the center frequency were approximately −0.4 to 1.8 dB at 4 kHz, −3.4 to 1.6 dB at 8 kHz, −3.9 to 2.6 dB at 12.5 kHz, and −0.1 to 1.8 dB at 16 kHz.

Procedures • Procedures for audiometric testing were the same as described for experiment 1. Pulsed tones were used to establish normal hearing in the standard audiometric range. Audiometric thresholds at 12.5 and 16 kHz were then obtained for pulsed tones, followed by FM tones and 1/3-oct noise.

In the 2AFC threshold estimation procedure, each trial comprised two 1-sec intervals separated by 300 msec, with intervals marked visually on a computer screen. The signal was present in one of these intervals, selected at random, and the participant's task was to select the target interval. Visual feedback was provided following each response. The signal level was adapted using a 3-down, 1-up stepping rule that estimates the level associated with 79.4% correct (Levitt 1971). Adjustments in signal level were made in steps of 8 dB before the first reversal; step size was reduced by a factor of 2 following each of the first two reversals (i.e., from 8 to 4 dB, and from 4 to 2 dB). Tracks continued until eight reversals were obtained. The threshold estimate for each track was the mean signal level at the last six reversals. Two threshold estimates were obtained in each condition; a third estimate was collected if the first two differed by more than 3 dB, and the outlier estimate was dropped. The final threshold estimate was the mean of two estimates. Testing was blocked by frequency, with the order of frequencies randomly selected for each participant. Within a frequency, testing was blocked by stimulus type, with the order of stimuli randomized for each participant.

In the modified Bekesy tracking procedure, the participant heard a continuous train of pulsed-tone signals (200-msec tones separated by 200-msec interstimulus intervals). They were asked to press a button when they heard the signal and release the button when they no longer heard the signal. Presentation level adjusted adaptively every 1.2 sec, decreasing when the response button was depressed and otherwise increasing. Each track started with a block of trials at the initial signal frequency, which was 1/2 oct below (or above) the center frequency; the purpose of fixed-frequency trials was to ensure an accurate estimate of audibility at the outset of the Bekesy track. Subsequent signals swept up (or down) in frequency. The initial signal level was approximately 10 dB suprathreshold, based on each participant's 2AFC data. Initial changes in signal level were made in steps of 8 dB; step size reduced to 4 dB following the first level reversal. The signal frequency was fixed until the second signal level reversal. After this point in the track, signal level changed in steps of 2 dB, and signal frequency changed by 1/70th of an octave following each level reversal. A track ended once the signal frequency had traversed a full octave, which took 3 to 4 min. This modified Bekesy procedure can be thought of as a sequence of 1-down, 1-up adaptive tracks with a single reversal obtained at each of 70 frequencies. Obtaining a signal level reversal at each frequency was intended to accommodate very rapid changes in audibility across frequency, which could be underestimated using a single stimulus presentation at each frequency.

All participants completed two Bekesy tracks centered on 16 kHz; one started 1/2 oct below 16 kHz and swept up in frequency, and the other started 1/2 oct above 16 kHz and swept down in frequency. A subset of participants also provided a pair of Bekesy tracks with center frequencies at 4, 8, and 12.5 kHz. The interoctave frequency (12.5 kHz) was included to ensure accurate characterization of hearing sensitivity at and below 16 kHz. In several cases, the first track for a participant was replaced based on subjective evidence of practice effects (e.g., >10 dB deviation across tracks at one or more frequencies). A progress bar indicated the number of frequencies remaining in the Bekesy track, but participants did not receive any other feedback.

Results

Audiometric thresholds for the FM tone and 1/3-oct noise are missing for two participants due to scheduling restrictions. Data from the remaining 14 participants are plotted in Figure 4. Thresholds for the FM tone and the 1/3-oct noise are plotted as a function of the pulsed-tone threshold, with results for each frequency shown in separate panels. Mean differences between pulsed-tone thresholds and FM-tone thresholds were 1.8 dB (4 kHz), 2.1 dB (8 kHz), 2.9 dB (12.5 kHz), and 6.8 dB (16 kHz). For 1/3-oct noise, those values were 2.9 dB (4 kHz), 3.2 dB (8 kHz), 6.4 dB (12.5 kHz), and 18.9 dB (16 kHz). Larger discrepancies for the 1/3-oct noise than the FM tone were predicted based on the wider bandwidth of the noise.

Figure 5 shows data for the seven participants who provided Bekesy data across the full frequency range. Symbols show the mean 2AFC thresholds. Symbol shape and color reflect the stimulus type, as defined in the legend. Solid gray lines show the mean of Bekesy track reversals as a function of frequency, combined across tracks‡ and smoothed with a five-point Hann window. Smoothing removes the variability associated with obtaining a single reversal at each frequency. Panels with data

‡Reversal values at each frequency were averaged for each participant and stimulus. There were two reversal values at most frequencies. However, there were four reversal values per frequency within the octave centered on 12.5 kHz due to overlap between octaves centered on 8, 12.5, and 16 kHz.

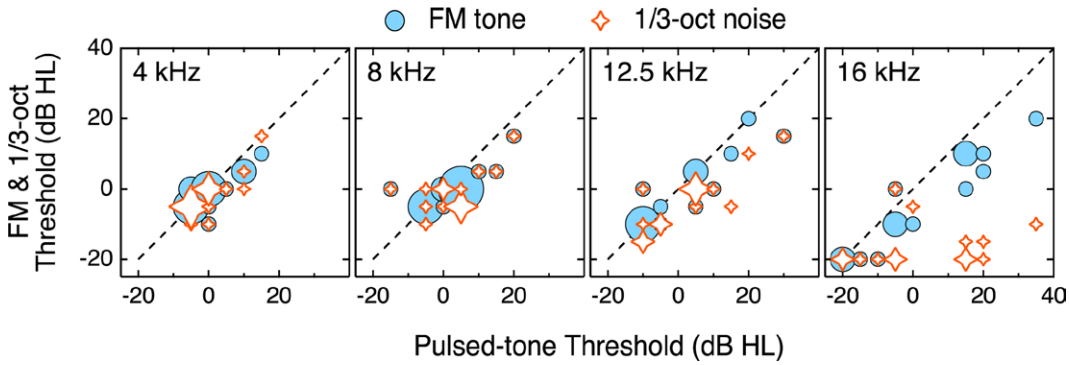

Fig. 4. Audiometric thresholds for FM-tone and 1/3-oct noise plotted as a function of the pulsed-tone thresholds, all in dB HL. Results are shown separately by frequency, as indicated in the upper left of each panel. Symbol shape and color reflect stimulus type, as defined in the legend. Symbol size reflects the number of observations for each combination of values. Diagonal lines indicate perfect correspondence of thresholds across stimuli. 1/3-oct indicates one-third-octave; dB HL, decibels hearing level; FM, frequency-modulated.

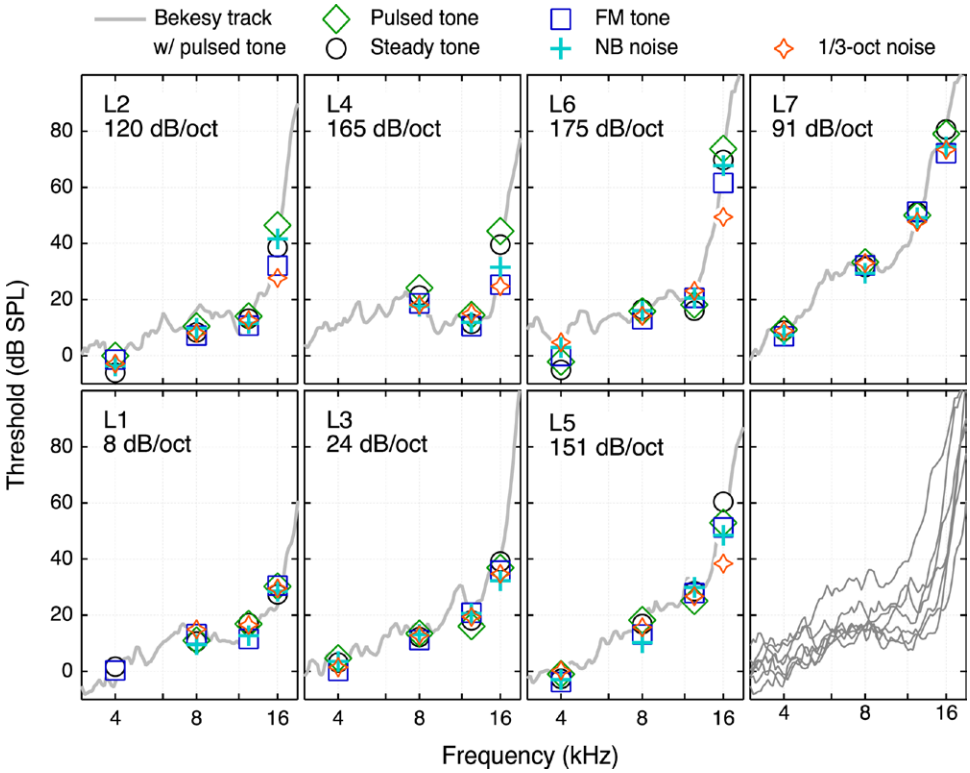

Fig. 5. Data for the seven participants providing Bekesy data at multiple frequencies. Gray lines show smoothed Bekesy track data, and symbols show 2AFC thresholds as a function of frequency, with stimulus condition defined in the key at the top of the figure. The slope of Bekesy data from 14.3 to 16 kHz is indicated at the top of each panel. Bekesy data for all seven participants are plotted together in the lower right panel. 2AFC indicates two-alternative forced-choice; FM, frequency-modulated; NB, narrowband.

for individual participants are ordered by the Bekesy threshold at 16 kHz (lower thresholds for panels on the left). A line was fitted to Bekesy track data between 14.3 and 16 kHz, which is the lower half of the 1/3-oct band of noise; the resulting slope estimates are shown at the top of each panel in Figure 5. The slope of the sensitivity curve in this frequency region was expected to roughly capture the benefit of increased stimulus bandwidth relative to the pulsed tone. The lower right panel of Figure 5 shows Bekesy data for all seven participants, highlighting the large individual differences in thresholds and slope at EHFs compared with lower frequencies.

Visual inspection of Figure 5 indicates that 2AFC steady- and pulsed-tone thresholds correspond closely to Bekesy data, and thresholds obtained using 2AFC adaptive methods are more similar across stimuli below 16 kHz than at 16 kHz. On average, 2AFC pulsed-tone thresholds agree with Bekesy estimates to within 1.5 dB at 4 kHz, 2.3 dB at 8 kHz, 2.1 dB at 12.5 kHz, and 4.9 dB at 16 kHz. On average, the range of thresholds obtained for each participant with the five stimuli agree to within 4.7 dB at 4 kHz, 4.7 dB at 8 kHz, 4.8 dB at 12.5 kHz, and 14.7 dB at 16 kHz. The difference between the minimum and maximum 2AFC threshold at 16 kHz varies markedly across individual

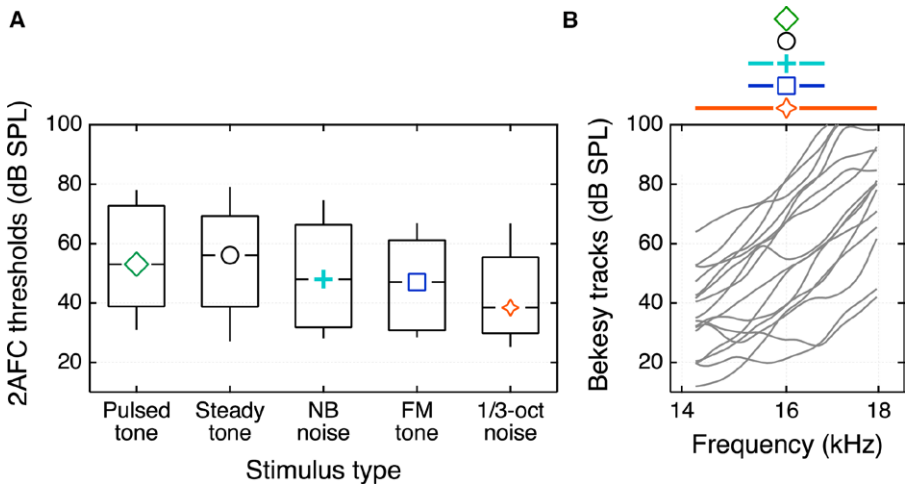

Fig. 6. Sensitivity at and around 16 kHz. A, The distributions of 16-kHz 2AFC thresholds for each of the five stimuli, indicated on the *x* axis. Horizontal lines and symbols indicate the median, boxes span the 25th to 75th percentiles, and vertical lines span the 10th to 90th percentiles. B, Smoothed Bekesy data in the one-third-oct band around 16 kHz for individual participants. The key above (B) illustrates the spectral extent of each stimulus; colors and symbols follow the plotting conventions of (A). 2AFC indicates two-alternative forced-choice; FM, frequency-modulated; NB, narrowband.

participants, with values of 3.2 to 24.3 dB. The slope of the Bekesy track between 14.3 and 16 kHz tends to be steeper for listeners with greater spread in the 2AFC thresholds, as assessed via Spearman correlation (r_s = 0.96, p = 0.003, n = 7). Notice that the slope of the audibility curve can differ for participants with the same 2AFC steady-tone threshold; for example, L3 and L4 both have 16-kHz 2AFC steady-tone thresholds near 40 dB SPL, but L3 has a relatively shallow slope (24 dB/oct) and a 6.8-dB difference in 2AFC thresholds across stimuli, whereas L4 has a steep slope (165 dB/oct) and a 19.7-dB difference in 2AFC thresholds.

The larger set of data obtained at 16 kHz was used to better understand how the audibility curve, characterized by the Bekesy track, affects 2AFC thresholds obtained with each of the five stimuli for individual participants. Those results are shown in Figure 6. Across the 16 participants, mean 2AFC thresholds were 55.4 dB SPL for pulsed tones (SD = 17.2 dB), 53.6 dB SPL for steady tones (SD = 18.4 dB), 49.7 dB SPL for NB noise (SD = 17.3 dB), 46.3 dB SPL for FM tones (SD = 15.6 dB), and 42.9 dB SPL 1/3-oct noise (SD = 15.3 dB). Line fits to smoothed Bekesy data between 14.3 and 16 kHz resulted in slope estimates of −35 to 235 dB/oct (median of 131 dB/oct).

Figure 7 shows 16-kHz thresholds from the 2AFC task plotted as a function of predictions based on Bekesy track data. Small black circles show predictions based on the Bekesy track value at the stimulus center frequency of 16 kHz. Open symbols show predictions based on Bekesy track values at all frequencies contained in the stimulus. Audibility across stimulus frequency was defined by fitting a spline function to the smoothed Bekesy track reversals at each frequency in the 1/3-oct band around 16 kHz. The frequency components comprising the long-term power spectrum of each stimulus were then scaled to account for differential audibility, summed in power, and represented in dB SPL using the calibration value at 16 kHz.§ This approach assumes that listeners integrate cues across frequency when they are available.

Predictions based on the stimulus spectrum accounted for 90 to 94% of variance for all five stimuli. The difference between behavioral thresholds and those predicted based on the stimulus spectrum was not significantly different from zero for the pulsed tone, steady tone, NB noise, or FM tone (p > 0.05 without correction for multiple tests).¶ For the 1/3-oct noise, behavioral thresholds were on average 2.6 dB higher than predicted based on the stimulus spectrum (t_{15} = 2.73, p = 0.016). For FM tones and the two noise stimuli, audibility across the stimulus spectrum was a significantly better predictor than audibility at the 16-kHz center frequency (p < 0.001 two-tailed). Overall, accurate predictions of behavioral responses based on the stimulus spectra provide strong support for the idea that audibility over the spectral range of the stimulus impacts threshold.

Previous data indicate that the magnitude of threshold change between octave frequencies is associated with the magnitude of the stimulus effect, with steeper audiograms predicting larger differences between steady- and FM-tone thresholds (Walker et al. 1984; Lentz et al. 2017). This also appears to be the case in the present dataset. For example, the difference between audiometric thresholds at 8 and 16 kHz correlated with the difference between thresholds for the steady tone and the 1/3-oct band of noise (r = −0.58, p = 0.019) and the difference between the steady tone and the FM-tone thresholds (r = −0.60, p = 0.013). While this result demonstrates an association between the audiometric slope (8 to 16 kHz) and the magnitude of the stimulus effect, there are striking exceptions to this trend. The participant with the largest difference between 8- and 16-kHz audiometric thresholds (30 dB) had a

§The procedure for estimating 2AFC thresholds based on Bekesy data represented levels in terms of voltage applied to the headphones as opposed to dB SPL. Integrating across frequency for data represented in volts rather than dB SPL accommodates the fact that the stimuli were generated without consideration of the frequency response of the headphones.

¶For the steady tone, the long-term spectrum contains a single component, but its level is 2.4-dB higher than the pulsed tone, which was used for the Bekesy track. Incorporating this lower level into the threshold prediction has a nonsignificant effect on the fit, decreasing percent of variance accounted from 95 to 94%; in neither case was the difference between behavioral thresholds and the predictions significantly different from zero (M = −0.9 dB, t_{15} = −0.86, p = 0.403; M = 1.6 dB, t_{15} = 1.55, p = 0.143).

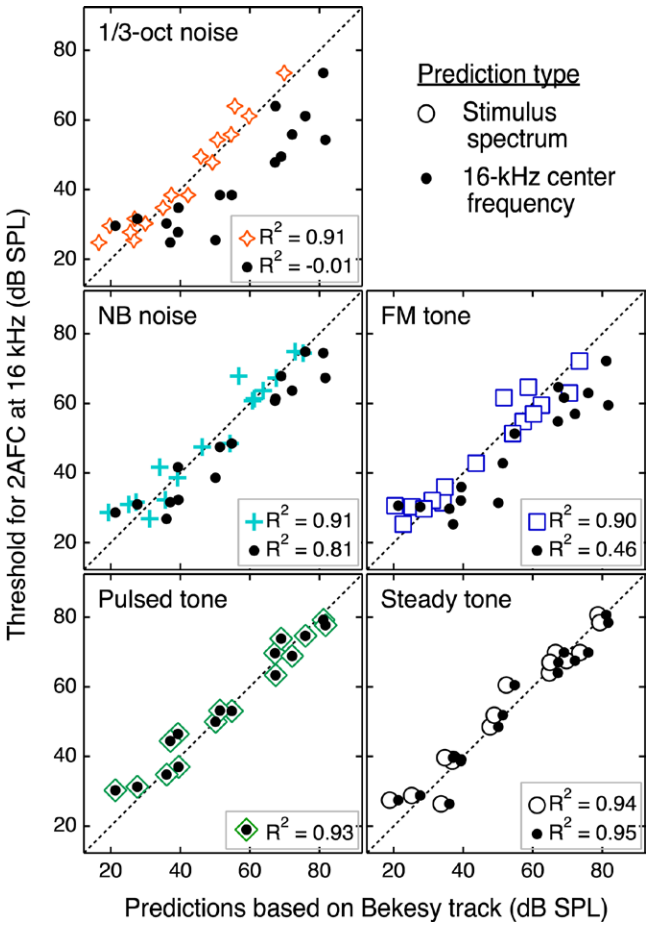

Fig. 7. Thresholds at 16 kHz for the 2AFC task, plotted as a function of predictions based on Bekesy track data for each participant. Results are shown in separate panels for each of the five stimulus conditions. Filled black circles show predictions based on audibility at 16 kHz, and open symbols show predictions based on audibility across the long-term power spectrum of each stimulus. Text in the lower right of each panel indicates the percent of variance accounted for by each prediction type. Dotted lines indicate perfect correspondence of thresholds and predictions. 2AFC indicates two-alternative forced-choice; FM, frequency-modulated; NB, narrowband.

very modest difference between 2AFC steady tone and 1/3-oct noise thresholds (3.9 dB), but threshold was predicted to within 1 dB of the observed threshold using the power spectrum model for this datapoint. These results suggest that the threshold difference between octave frequencies is only partly successful at capturing effects related to stimulus bandwidth and audibility.

One somewhat surprising result from this experiment is the very good sensitivity demonstrated at 16 kHz by some participants. Whereas the mean 2AFC threshold was 53.6 dB SPL for the steady tone, close to the published reference equivalent threshold sound pressure level of 56 dB (RadioEar), the best performer had a threshold of 26.4 dB SPL (approximately −30 dB HL). This raises the possibility that the output limit of the clinical hardware at −20 dB HL could preclude accurate threshold estimation for some participants. Omitting data for the three best performers, audiometric pulsed-tone thresholds are on average 4 dB higher than the associated 2AFC thresholds when both measures are represented in the same units (either SPL or HL). A 4 dB discrepancy in audiometric and psychophysical thresholds might be expected based on differences in measurement procedures. The forced-choice structure of the 2AFC task eliminates the tendency for participants to adopt a conservative response bias (Marshall & Jesteadt 1986), which is particularly pronounced at EHFs (Stelmachowicz et al. 1989), and the adaptive track with feedback provided opportunities for participants to optimize their listening strategies. If the trend for 4 dB lower 2AFC thresholds than audiometric thresholds holds across participants, then the three performers with the lowest 2AFC thresholds (26.4, 27.4, and 28.8 dB SPL) would have audiometric thresholds of −25.6, −24.6, and −23.2 dB HL, if not for the output limits of the audiometer. One practical consequence of this result is that the lower output limit of −20 dB HL imposed by clinical hardware could miss individual differences or changes in EHF sensitivity over time within participants.

Discussion

Experiment 2 evaluated audiometric thresholds at 4, 8, 12.5, and 16 kHz in a group of young adults with normal hearing in the standard audiometric range. Mean audiometric thresholds at 4 and 8 kHz were within 3 dB for pulsed tones, FM tones, and 1/3-oct bands of noise, but stimulus effects were evident at higher frequencies, with mean differences relative to pulsed tones at 16 kHz of 6.8 dB (FM tone) and 18.9 dB (1/3-oct noise). At 16 kHz, stimulus effects met or exceeded the 10-dB criterion for a clinically meaningful change in thresholds for 21% (FM tone) and 57% (1/3-oct noise) of participants. These results are consistent with those reported in experiment 1.

For the seven participants providing data at all four frequencies, 2AFC thresholds at 4, 8, and 12.5 kHz were similar for steady tones, pulsed tones, FM tones, NB noise, and 1/3-oct noise, and those 2AFC thresholds were similar to Bekesy track data at the associated frequencies. Whereas 2AFC thresholds differed by <5 dB within participants at 4, 8, and 12.5 kHz, stimulus effects of up to 24 dB were observed at 16 kHz. As observed for audiometric thresholds, 2AFC thresholds tended to be better for spectrally broader stimuli, with the lowest mean thresholds obtained with 1/3-oct noise. The range of 16-kHz thresholds across stimuli was correlated with the local slope of Bekesy track data between 14.3 and 16 kHz, the lower half of the 1/3-oct noise bandwidth. These results are consistent with the hypothesis that the stimulus effects observed at 16 kHz are related to the slope of the audibility curve and off-frequency listening in this frequency region.

Data at 16 kHz from a larger group of 16 participants were used to model these effects. Thresholds in the 2AFC task were predicted based on Bekesy data in two ways: based on audibility at the stimulus center frequency and based on the combination of the long-term power spectrum of the stimulus and audibility at the frequencies contained in the signal. Predictions based on the spectral distribution of stimulus energy accounted for 90 to 94% of variance. These predictions were significantly better than predictions based on audibility at the center frequency for the 1/3-oct noise, NB noise, and FM tone.

One unexpected finding of experiment 2 was evidence of sensitivity better than −20 dB HL at 16 kHz in the psychophysical data for three of the 16 participants. Attempting to characterize individual differences or within-listener changes in sensitivity could be problematic when using clinical hardware if some listeners have thresholds <−20 dB HL. This problem would be

exacerbated when using FM tones or noise-band stimuli, due to cues provided at frequencies below the center frequency. For this reason, it may be advantageous to use experimental procedures and hardware capable of measuring performance at lower levels when characterizing EHF sensitivity for children and young adults with normal hearing in the standard audiometric range.

GENERAL DISCUSSION AND CONCLUSIONS

Hearing in the EHF range shares many of the features of hearing in the standard audiometric frequencies, but there are also differences. For example, there is more loudness adaptation at EHFs than in the standard audiometric range (Miskiewicz et al. 1993; Hellman et al. 1997; Wynne et al. 2015), and a more conservative response bias at EHFs (Stelmachowicz et al. 1989). Conversely, auditory filters at EHFs are consistent with extrapolated parameters based on lower-frequency data (Zhou 1995). Results from the present experiments support the idea that stimulus effects related to the slope of the audibility curve, which have been observed in the standard audiometric frequencies (Walker et al. 1984; Lentz et al. 2017), also occur at 16 kHz. However, in contrast with those previous datasets, the majority of participants in the present study had 16-kHz steady-tone thresholds within the normal range (≤20 dB HL; 44/47 in experiment 1 and 15/16 in experiment 2). This suggests that the loss of sensitivity with increasing frequency in the EHF range for normal-hearing listeners has the same effect as a steeply sloping loss in the standard audiometric range. As a result, stimulus effects are more likely and more pronounced at EHFs than within the standard audiometric range.

In experiment 1, the majority of thresholds provided at 16 kHz for children were below 0 dB HL, and in experiment 2 there was evidence of steady-tone thresholds <−20 dB HL for several adults. These results suggest that interpretation of audiometric data at EHFs may differ from the standard audiometric range. For young adults with no evidence of hearing loss, we expected thresholds to cluster around 0 dB HL, and that is approximately the case at 16 kHz for adults. Median thresholds defining 0 dB HL may not be appropriate for children, however, as significant effects of age on EHF sensitivity begin at or below 10 years of age (Trehub et al. 1988; Buren et al. 1992; Hemmingsen et al. 2021; Mishra et al. 2022b). This has prompted some researchers to suggest development of a separate set of HL values for younger listeners (Hemmingsen et al. 2021). There are also larger individual differences at EHFs compared with lower frequencies (Rodriguez Valiente et al. 2014). If characterizing the full range of individual differences is a priority, then it may be necessary to use stimulus levels <−20 dB HL. Extending the range of signal levels below −20 dB HL in clinical equipment could be particularly helpful when using FM tones or bands of noise, as floor effects would likely be observed for a larger proportion of listeners using these stimuli.

Results of these two experiments have implications for the measurement of EHF thresholds in both clinical settings and in research. Whereas steady tones and pulsed tones provide frequency-specific estimates of sensitivity, FM tones and bands of noise are typically a better choice for testing in the free-field (Walker et al. 1984), to avoid standing waves, and temporally dynamic stimuli may be preferred when testing young children or adults with tinnitus. The present dataset illustrates how choice of stimuli could limit the ability to characterize good performance at EHFs with clinical hardware. Whichever stimulus is selected, care should be taken to consider stimulus characteristics when evaluating results; for example, comparing thresholds obtained with different stimuli over time or across listeners could give a false impression of changes in hearing or differences in sensitivity across individuals.

Large individual differences in the local slope and shape of the EHF audibility curve could also thwart attempts to characterize hearing by measuring thresholds at octave or half-octave frequencies, as is commonly done in the standard audiometric range. Approaches taken for monitoring ototoxicity include measuring EHF thresholds at 1/6th-oct intervals (Fausti et al. 1999), or fixing stimulus level and adaptively changing its frequency (Rieke et al. 2017). Frequency-swept Bekesy tracking could also be informative if detailed information about the shape of the audibility curve is of value. Frequency-sweep stimuli have been proposed for avoiding misleading results associated with spectral fine structure (Lee & Long 2012), although this consideration may be less of a concern at EHFs, where spectral fine structure is reduced relative to lower frequencies (Alenzi & Lineton 2021). The best approach for characterizing EHF sensitivity may require balancing the competing priorities discussed here for each application, including the possible value of characterizing sensitivity below −20 dB HL.

ACKNOWLEDGMENTS

This project was funded by the June and James Ficklen Professorship and the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health (R01 DC019745 and R01 DC011038). Some of these data were presented at the 2022 and 2023 meetings of the American Auditory Society. Support for this project was provided by Selena Marcum, Jenna Van Bosch, and Olivia Zimmerman. Statistical support was provided by Hashem Amireh (Odum Institute, UNC).

Experiments were designed in collaboration with E.B., S.G.K., K.S.Y., D.M.B., and H.L.P. E.B. and L.J.L. oversaw data collection. E.B. and S.G.K. analyzed the data and wrote the first draft of the article. All authors contributed to interpretation of data and article revision.

Protocol approved by UNC's Behavioral research IRB. All data reported are available at https://osf.io/n9zfy/.

The authors have no conflicts of interest to disclose.

Address for correspondence: Emily Buss, Department of Otolaryngology/Head and Neck Surgery, The University of North Carolina at Chapel Hill, 170 Manning Drive, CB7070, Chapel Hill, NC 27599, USA. E-mail: ebuss@med.unc.edu

Received March 30, 2023; accepted October 4, 2023; published online ahead of print January 5, 2024.

REFERENCES

Alenzi, H., & Lineton, B. (2021). Transient otoacoustic emissions and audiogram fine structure in the extended high-frequency region. *Int J Audiol*, *60*, 985–994.

ANSI. (1997). *ANSI S3.5-1997 (R2017), American National Standard Methods for Calculation of Speech Intelligibility Index*. American National Standards Institute.

ANSI. (2004). *ANSI S3.21-2004, American National Standard Specification Methods for Manual Pure-tone Audiometry*. American National Standards Institute.

ANSI. (2010). *ANSI S3.6-2010, American National Standard Specification for Audiometers*. American National Standards Institute. [ASHA Practice Policy]. 2023. https://www.asha.org/policy/gl2005-00014/.

ASHA. (2005). Guidelines for manual pure-tone threshold audiometry. [ASHA Practice Policy]. 2023. https://www.asha.org/policy/gl2005-00014/.

Assmann, P., & Summerfield, A. Q. (2004). The perception of speech under adverse conditions. In S. Greenberg, W. A. Ainsworth, R. R. Popper, R. R. Fay (Eds.), *Speech Processing in the Auditory System* (pp. 231–308). Springer.

Badri, R., Siegel, J. H., Wright, B. A. (2011). Auditory filter shapes and high-frequency hearing in adults who have impaired speech in noise performance despite clinically normal audiograms. *J Acoust Soc Am*, *129*, 852–863.

Beahan, N., Kei, J., Driscoll, C., Charles, B., Khan, A. (2012). High-frequency pure-tone audiometry in children: A test-retest reliability study relative to ototoxic criteria. *Ear Hear*, *33*, 104–111.

Blankenship, C. M., Hunter, L. L., Feeney, M. P., Cox, M., Bittinger, L., Garinis, A. C., Lin, Li, McPhail, G., Clancy, J. P. (2021). Functional impacts of aminoglycoside treatment on speech perception and extended high-frequency hearing loss in a pediatric cystic fibrosis cohort. *Am J Audiol*, *30*, 834–853.

Braza, M. D., Corbin, N. E., Buss, E., Monson, B. B. (2022). Effect of masker head orientation, listener age, and extended high-frequency sensitivity on speech recognition in spatially separated speech. *Ear Hear*, *43*, 90–100.

Brungart, D. S., & Simpson, B. D. (2009). Effects of bandwidth on auditory localization with a noise masker. *J Acoust Soc Am*, *126*, 3199–3208.

Buren, M., Solem, B. S., Laukli, E. (1992). Threshold of hearing (0.125-20 kHz) in children and youngsters. *Br J Audiol*, *26*, 23–31.

Burk, M. H., & Wiley, T. L. (2004). Continuous versus pulsed tones in audiometry. *Am J Audiol*, *13*, 54–61.

Drennan, W. R. (2021). Identifying subclinical hearing loss: Extended audiometry and word recognition in noise. *Audiol Neurootol*, *27*, 217–226.

Fausti, S. A., Henry, J. A., Helt, W. J., Phillips, D. S., Frey, R. H., Noffsinger, D., Larson, V. D., Fowler, C. G. (1999). An individualized, sensitive frequency range for early detection of ototoxicity. *Ear Hear*, *20*, 497–505.

Flaherty, M., Libert, K., Monson, B. B. (2021). Extended high-frequency hearing and head orientation cues benefit children during speech-in-speech recognition. *Hear Res*, *406*, 108230.

Frank, T. (1990). High-frequency hearing thresholds in young adults using a commercially available audiometer. *Ear Hear*, *11*, 450–454.

Frank, T. (2001). High-frequency (8 to 16 kHz) reference thresholds and intrasubject threshold variability relative to ototoxicity criteria using a Sennheiser HDA 200 earphone. *Ear Hear*, *22*, 161–168.

Garinis, A. C., Cross, C. P., Srikanth, P., Carroll, K., Feeney, M. P., Keefe, D. H., Hunter, L. L., Putterman, D. B., Cohen, D. M., Gold, J. A., Steyger, P. S. (2017). The cumulative effects of intravenous antibiotic treatments on hearing in patients with cystic fibrosis. *J Cyst Fibros*, *16*, 401–409.

Gelfand, S. A., & Calandruccio, L. (2023). *Essentials of Audiology*. Thieme.

Haapaniemi, J. J. (1996). The hearing threshold levels of children at school age. *Ear Hear*, *17*, 469–477.

Hamill, T. A., & Haas, W. H. (1986). The relationship of pulsed, continuous, and warble extended-high frequency thresholds. *J Commun Disord*, *19*, 227–235.

Hellman, R., Miskiewicz, A., Scharf, B. (1997). Loudness adaptation and excitation patterns: Effects of frequency and level. *J Acoust Soc Am*, *101*, 2176–2185.

Hemmingsen, D., Stenklev, N. C., Klingenberg, C. (2021). Extended high frequency audiometry thresholds in healthy school children. *Int J Pediatr Otorhinolaryngol*, *144*, 110686.

Hunter, L. L., Monson, B. B., Moore, D. R., Dhar, S., Wright, B. A., Munro, K. J., Zadeh, L. M., Blankenship, C. M., Stiepan, S. M., Siegel, J. H. (2020). Extended high frequency hearing and speech perception implications in adults and children. *Hear Res*, *397*, 107922.

John, A. B., & Kreisman, B. M. (2017). Equivalence and test-retest reproducibility of conventional and extended-high-frequency audiometric thresholds obtained using pure-tone and narrow-band-noise stimuli. *Int J Audiol*, *56*, 635–642.

King, R. B., & Oldfield, S. R. (1997). The impact of signal bandwidth on auditory localization: Implications for the design of three-dimensional audio displays. *Hum Factors*, *39*, 287–295.

Knight, K. R., Kraemer, D. F., Winter, C., Neuwelt, E. A. (2007). Early changes in auditory function as a result of platinum chemotherapy: Use of extended high-frequency audiometry and evoked distortion product otoacoustic emissions. *J Clin Oncol*, *25*, 1190–1195.

Le Prell, C. G., Brewer, C. C., Campbell, K. C. M. (2022). The audiogram: Detection of pure-tone stimuli in ototoxicity monitoring and assessments of investigational medicines for the inner ear. *J Acoust Soc Am*, *152*, 470.

Lee, J., & Long, G. (2012). Stimulus characteristics which lessen the impact of threshold fine structure on estimates of hearing status. *Hear Res*, *283*, 24–32.

Lentz, J. J., Walker, M. A., Short, C. E., Skinner, K. G. (2017). Audiometric testing with pulsed, steady, and warble tones in listeners with tinnitus and hearing loss. *Am J Audiol*, *26*, 328–337.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J Acoust Soc Am*, *49*, 467–477.

Levy, S. C., Freed, D. J., Nilsson, M., Moore, B. C. J., Puria, S. (2015). Extended high-frequency bandwidth improves speech reception in the presence of spatially separated masking speech. *Ear Hear*, *36*, e214–e224.

Lough, M., & Plack, C. J. (2022). Extended high-frequency audiometry in research and clinical practice. *J Acoust Soc Am*, *151*, 1944.

Marshall, L., & Jesteadt, W. (1986). Comparison of pure-tone audibility thresholds obtained with audiological and two-interval forced-choice procedures. *J Speech Hear Res*, *29*, 82–91.

Massie, R., Dillon, H., Ching, T., Birtles, G. (2006). Does a diversity of nontonal stimuli at each test frequency assist in the hearing assessment of young children? *Aust N Z J Audiol*, *28*, 1–10.

Mishra, S. K., Saxena, U., Rodrigo, H. (2022a). Extended high-frequency hearing impairment despite a normal audiogram: Relation to early aging, speech-in-noise perception, cochlear function, and routine earphone use. *Ear Hear*, *43*, 822–835.

Mishra, S. K., Saxena, U., Rodrigo, H. (2022b). Hearing impairment in the extended high frequencies in children despite clinically normal hearing. *Ear Hear*, *43*, 1653–1660.

Miskiewicz, A., Scharf, B., Hellman, R., Meiselman, C. (1993). Loudness adaptation at high frequencies. *J Acoust Soc Am*, *94*, 1281–1286.

Monson, B. B., Hunter, E. J., Lotto, A. J., Story, B. H. (2014). The perceptual significance of high-frequency energy in the human voice. *Front Psychol*, *5*, 587.

Monson, B. B., Rock, J., Schulz, A., Hoffman, E., Buss, E. (2019). Ecological cocktail party listening reveals the utility of extended high-frequency hearing. *Hear Res*, *381*, 107773.

Moore, B. C. J., & Tan, C. T. (2003). Perceived naturalness of spectrally distorted speech and music. *J Acoust Soc Am*, *114*, 408–419.

Motlagh-Zadeh, L., Silbert, N. H., Sternasty, K., Swanepoel, D. W., Hunter, L. L., Moore, D. R. (2019). Extended high-frequency hearing enhances speech perception in noise. *Proc Natl Acad Sci USA*, *116*, 23753–23759.

Norrix, L. W., & Anderson, A. (2015). Audiometric thresholds: Stimulus considerations in sound field and under earphones. *Am J Audiol*, *24*, 487–493.

Orchik, D. J., & Rintelmann, W. F. (1978). Comparison of pure-tone, warble-tone and narrow-band noise thresholds of young normal-hearing children. *J Am Audiol Soc*, *3*, 214–220.

OSHA. (2022). *Occupational Safety and Health Standards: Occupational Noise Exposure (Standard No. 1910.95)*. Occupational Safety and Health Administration, U.S. Department of Labor.

Petley, L., Hunter, L. L., Motlagh Zadeh, L., Stewart, H. J., Sloat, N. T., Perdew, A., Lin, Li, Moore, D. R. (2021). Listening difficulties in children with normal audiograms: Relation to hearing and cognition. *Ear Hear*, *42*, 1640–1655.

Polspoel, S., Kramer, S. E., van Dijk, B., Smits, C. (2022). The importance of extended high-frequency speech information in the recognition of digits, words, and sentences in quiet and noise. *Ear Hear*, *43*, 913–920.

Prendergast, G., Millman, R. E., Guest, H., Munro, K. J., Kluk, K., Dewey, R. S., Hall, D. A., Heinz, M. G., Plack, C. J. (2017). Effects of noise exposure on young adults with normal audiograms II: Behavioral measures. *Hear Res*, *356*, 74–86.

RadioEar. DD450 Technical Specifications. 2023, January 22. https://www.radioear.us/-/media/radio-ear/main/datasheets/datasheet_dd450.pdf.

Rieke, C. C., Clavier, O. H., Allen, L. V., Anderson, A. P., Brooks, C. A., Fellows, A. M., Brungart, D. S., Buckey, J. C. (2017). Fixed-level frequency threshold testing for ototoxicity monitoring. *Ear Hear*, *38*, e369–e375.

Rodriguez Valiente, A., Trinidad, A., Garcia Berrocal, J. R., Górriz, C., Ramírez Camacho, R. (2014). Extended high-frequency (9-20 kHz) audiometry reference thresholds in 645 healthy subjects. *Int J Audiol*, *53*, 531–545.

Rosler, G., & Anderson, H. (1978). Maximum steepness of slopes in hearing-threshold curves. *Audiology*, *17*, 299–316.

Saxena, U., Mishra, S. K., Rodrigo, H., Choudhury, M. (2022). Functional consequences of extended high frequency hearing impairment: Evidence from the speech, spatial, and qualities of hearing scale. *J Acoust Soc Am*, *152*, 2946.

Schmuziger, N., Probst, R., Smurzynski, J. (2004). Test-retest reliability of pure-tone thresholds from 0.5 to 16 kHz using Sennheiser HDA 200 and Etymotic Research ER-2 earphones. *Ear Hear*, *25*, 127–132.

Shaw, G. M., Jardine, C. A., Fridjhon, P. (1996). A pilot investigation of high-frequency audiometry in obscure auditory dysfunction (OAD) patients. *Br J Audiol*, *30*, 233–237.

Smull, C. C., Madsen, B., Margolis, R. H. (2019). Evaluation of two circumaural earphones for audiometry. *Ear Hear*, *40*, 177–183.

Stelmachowicz, P. G., Beauchaine, K. A., Kalberer, A., Kelly, W. J., Jesteadt, W. (1989). High-frequency audiometry: Test reliability and procedural considerations. *J Acoust Soc Am*, *85*, 879–887.

Stelmachowicz, P. G., Gorga, M. P., Cullen, J. K. (1982). A calibration procedure for the assessment of thresholds above 8000 Hz. *J Speech Hear Res*, *25*, 618–623.

Stevens, K. N., Berkovitz, R., Kidd, G., Jr., Green, D. M. (1987). Calibration of ear canals for audiometry at high frequencies. *J Acoust Soc Am*, *81*, 470–484.

Suzuki, Y., & Takeshima, H. (2004). Equal-loudness-level contours for pure tones. *J Acoust Soc Am*, *116*, 918–933.

Thompson, M., & Thompson, G. (1972). Response of infants and young children as a function of auditory stimuli and test methods. *J Speech Hear Res*, *15*, 699–707.

Trehub, S. E., Schneider, B. A., Morrongiello, B. A., Thorpe, L. A. (1988). Auditory sensitivity in school-age children. *J Exp Child Psychol*, *46*, 273–285.

Trine, A., & Monson, B. B. (2020). Extended high frequencies provide both spectral and temporal information to improve speech-in-speech recognition. *Trends Hear*, *24*, 1–8.

Walker, G., Dillon, H., Byrne, D. (1984). Sound field audiometry: Recommended stimuli and procedures. *Ear Hear*, *5*, 13–21.

Wynne, D. P., George, S. E., Zeng, F. G. (2015). Amplitude modulation reduces loudness adaptation to high-frequency tones. *J Acoust Soc Am*, *138*, 279–283.

Yeend, I., Beach, E. F., Sharma, M. (2019). Working memory and extended high-frequency hearing in adults: Diagnostic predictors of speech-in-noise perception. *Ear Hear*, *40*, 458–467.

Zhou, B. (1995). Auditory filter shapes at high frequencies. *J Acoust Soc Am*, *98*, 1935–1942.