JASA ARTICLE

Sensory unpleasantness of very-high frequency sound and audible ultrasound

Ben Lineton,[1,a)] Rahma Abdul Rahman Al Balushi,[1,2] Sian Lloyd Jones,[3] Timothy G. Leighton,[1] and Mark D. Fletcher[1]
[1]*Faculty of Engineering and Physical Sciences, University of Southampton, Southampton, SO17 1BJ, United Kingdom*
[2]*Al Nahdha Hospital, Ministry of Health, Ruwi Street, Muscat, Sultanate of Oman*
[3]*Department of Audiology and Hearing Therapy, Royal South Hants Hospital, Brinton's Terrace, Southampton SO14 0YG, United Kingdom*

ABSTRACT:
Audible very-high frequency sound (VHFS) and ultrasound (US) have been rated more unpleasant than lower frequency sounds when presented to listeners at similar sensation levels (SLs). In this study, 17 participants rated the sensory unpleasantness of 14-, 16-, and 18-kHz tones and a 1-kHz reference tone. Tones were presented at equal subjective loudness levels for each individual, corresponding to levels of 10, 20, and 30 dB SL measured at 1 kHz. Participants were categorized as either "symptomatic" or "asymptomatic" based on self-reported previous symptoms that they attributed to exposure to VHFS/US. In both groups, subjective loudness increased more rapidly with sound pressure level for VHFS/US than for the 1-kHz reference tone, which is consistent with a reduced dynamic range at the higher frequencies. For loudness-matched tones, participants rated VHFS/US as more unpleasant than that for the 1-kHz reference. These results suggest that increased sensory unpleasantness and reduced dynamic range at high frequencies should be considered when designing or deploying equipment which emits VHFS/US that could be audible to exposed people.
https://doi.org/10.1121/10.0028380

(Received 31 December 2023; revised 14 June 2024; accepted 11 July 2024; published online 6 September 2024)

[Editor: William James Murphy] Pages: 1565–1574

I. INTRODUCTION

In previous studies, various adverse effects have been attributed to exposure to very-high frequency sounds (VHFS; 11.2–17.8 kHz) and ultrasound (US; >17.8 kHz), including fatigue, headache, fullness of the ears, and ear pain (reviewed in Lawton 2001; Leighton 2016, 2017). The frequency dividing VHFS from US is increasingly being taken to be 17.8 kHz (Paxton *et al.*, 2018; Fletcher *et al.*, 2018; van Wieringen and Glorieux, 2018; Scholkmann, 2019; Lubner *et al.*, 2020; Weichenberger *et al.*, 2022) because it is the lower limit of the third octave band (TOB) centered on 20 kHz, and for decades, the guidelines have been defined in terms of TOBs (Leighton, 2018).

The need to understand the relationship between exposure conditions and adverse effects increased with the report of airborne US in public places in the United Kingdom (UK) as a result of public address voice alarm (PAVA) systems and other technology and publication of citizen science techniques to detect the use of smartphones (Leighton, 2016). The occurrence of such exposure in public places has since been confirmed in North America, Germany, France, Italy, Australia, Japan, and Switzerland (Fletcher *et al.*, 2018; Paxton *et al.*, 2018; Mapp, 2018; Leighton *et al.*, 2019; Scholkmann, 2019; Alvares-Sanches *et al.*, 2019; Leighton *et al.*, 2020). Ueda *et al.* (2014) had previously reported pest deterrents which insonified specific public places, causing annoyance. The 21st century has subsequently witnessed new and emerging technologies, exposing members of the public [see examples in Figs. 1(a)–1(c)]. Exposures may be accidental, such as malfunctioning classroom equipment (Leighton, 2020), or an unintended by-product of their operation, such as supersonic hairdryers (Dolder *et al.*, 2018; Huang and Zheng, 2022) and PAVA systems (Mapp, 2018). In other cases, exposure is a key component to the operation of the technology, such as in acoustic spotlights (Dolder *et al.*, 2019) and haptic systems, or when exposure is deliberately designed to elicit an adverse human response as in commercial "pain generators" and "teen deterrents" (Leighton, 2017).

Although exposure to high sound pressure levels (SPLs) can cause temporary or permanent hearing threshold shifts, prolonged exposure to lower levels in public spaces may be of concern for different reasons, e.g., sustained annoyance or inability to perform tasks. Public exposures produce particular problems as in public places, we cannot know or control the ages, gender, aural diversity, or medical predispositions of those exposed, or the duration and amplitude of exposure, or offer hearing protection (IRPA, 1984). There is an imperative to establish which SPLs can cause adverse effects, such as annoyance and distraction from work, given that exposure of the public might occur not just

a)Email: bl1@soton.ac.uk

FIG. 1. (Color online) Example sources from a wide frequency range. (a) A bird deterrent, which can emit a wide range of frequencies (e.g., 16-40 kHz) depending on the setting, and (b) the output of ceiling-mounted PAVA speakers being measured is shown. These generally emit in the 18–25 kHz range. (c) An ultrasonic device is used to levitate microparticles, which operates around 20 kHz.

through their occupancy of public spaces but also as a result of the purchase of personal devices that emit US, such as entertainment, communication, and haptic technology [Fig. 1(b)].

Exposure to ultrasonic pest deterrents have been found to give rise to adverse effects such as annoyance or discomfort. Ueda *et al*. (2014) exposed individuals to a pest deterrent in a restaurant, which generated sound at 20 kHz at SPLs of 90–130 dB. Individuals who could hear the sounds gave high ratings for discomfort and irritation, amongst other attributes. Similarly, Glorieux and van Wieringen (2014) found that when participants could hear sounds from a pest deterrent (frequencies of 12.5–35 kHz and SPLs from 44 to 71 dB), they often described the sounds as "annoying" or "distressing." Unlike newer technologies (such as ultrasonic beacons and haptics), pest deterrents have been deployed around humans for decades and accrued anecdotal reports of adverse effects and observations of the wide variety in sensitivity to them: Herman and Powell (1981) reported significant person-to-person variations in reactions to a dog repeller that varied from "no perception or no symptoms at all, to expressions of severe discomfort 40 feet from the source, in another room."

Using the concept of "sensory unpleasantness" (see below), Kurakata *et al*. (2013) provided evidence that the usable dynamic range (the dB difference between the hearing threshold in an individual at a specific frequency and the lowest SPL at that frequency, which induces adverse effects in that individual) decreases with increasing frequency in the VHF/US range. This implies that considerable annoyance caused by VHFS/US may arise from exposure at low SLs. While attempting to study brain activity due to US exposure, Kühler *et al*. (2019) stated that "at ultrasound frequencies, the range of comfortable hearing is extremely narrow; if an ultrasound tone is heard, it is immediately perceived as unpleasant" (see also Leighton, 2017, who came to a similar conclusion). They also observed that "almost all of the test subjects described the hearing sensation as displeasing."

In a laboratory-controlled study using pure tones, Fletcher *et al*. (2018) categorized 42 participants according to whether or not they reported previously having experienced adverse symptoms that they attributed to exposure to VHFS/US: Those who did were categorized as "symptomatic"; those who did not were categorized as "asymptomatic." Audible tones were presented over headphones. Participants were exposed to a reference tone at 1 kHz or tonal VHFS/US between 13.5 and 20 kHz at a SL of 25 dB (i.e., 25 dB above their hearing threshold) to ensure that the sounds were clearly audible. The corresponding SPLs for high frequency tones measured in an ear simulator were between 82 and 92 dB. Both groups of participants reported greater discomfort for trials with the tonal VHFS/US compared to those with the 1-kHz reference tone. The symptomatic group also rated annoyance higher for the VHFS/US compared to that for the reference tone. The maximum frequency varied for each participant because it was set to be as high as possible for each individual participant while still being audible at the maximum intensity that the ethics protocols allowed use of. Consequently, all symptomatic participants were exposed to frequencies below 17.8 kHz (the maximum was 17.7 kHz). Only 20% of the asymptomatic participants were exposed to ultrasound frequencies (above 17.8 kHz).

The current study, similar to that of Fletcher *et al*. (2018), exposed subjects to audible US and conducted an analysis that subdivides the symptomatic (prior to exposure in this test, those who believe they have suffered adverse effects from airborne US) from the asymptomatic (those with no such belief). This is important because Ashihara *et al*. (2006) measured hearing thresholds up to 28 kHz, and

Ueda *et al.* (2016) measured hearing thresholds up to 26 kHz. This means that if the signal levels were sufficiently high, the 18–24 kHz direct beams from the PAVA and pest deterrent systems, subharmonics of a 40-kHz haptic system (Liebler *et al.*, 2019; Liebler *et al.*, 2020; Huang *et al.*, 2012), or an acoustic spotlight operating at 40 kHz (Yoneyama *et al.*, 1983; Kondo and Osanai, 2019) or 50 kHz (Wygant *et al.*, 2009) could fall within the potentially audible range of some of those who could be exposed.

The current study addresses the question of how subjective assessment of the annoyance and loudness of tonal VHFS/US is related to the frequency and SPL of the tone. The relationship between physical properties of a tone and the perception of annoyance is particularly difficult to quantify because the feelings of annoyance are affected by acoustic factors, such as SPL and frequency, and non-acoustic factors, such as the context in which the tone is presented and the attributes of the exposed individual, such as their personality and health condition (Pedersen, 2007). In laboratory studies in which participants are asked to rate the annoyance of a sound, it is difficult to reproduce the context in which the sound might be heard in real life (such as trying to sleep or concentrate on a task). For this reason, Pedersen (2007) describes ratings of sound delivered in the laboratory as ratings of "annoyance potential," whereas Fastl and Zwicker (2007) use the term "sensory pleasantness" in an attempt to isolate the sensory aspects of annoyance. Fastl and Zwicker (2007) report an increase in sensory unpleasantness with increases in the subjective loudness of the sound and the "sharpness" of the sound, which, in turn, increases with boosts in stimulus frequency (ISO 532-1:2017, 2017; DIN 45692:2009, 2009). These objective metrics have been incorporated into a predictive model of "perceived annoyance" that increases with metrics for subjective loudness and sharpness (Fastl, 2005). However, there is a paucity of data on perceptions of loudness, sharpness, or annoyance at frequencies above 12.5 kHz. Other properties of the sound related to frequency modulation (e.g., "roughness" and "fluctuation strength") have also been found to affect perceived annoyance (Fastl, 2005). However, to date, these properties have largely been ignored in studies of annoyance of VHFS/US, partly, because the slope of the threshold-frequency relationship leads to a strong coupling between frequency modulation and loudness modulation (or detectability) of the signal, a coupling which is highly variable between participants. The current study was restricted to pure tones of 1-s duration, which is long enough for the sound to establish the quality of a pure tone but short enough to prevent the institutional exposure limits from being exceeded.

As discussed above, researchers have speculated that the useable dynamic range may be greatly reduced in the ultrasonic range. This is supported partly by measurements of equal loudness contours, which plot, as a function of frequency, the SPL of tones that are judged to be equally loud. The lowest contour is close to the threshold of hearing, which is the highest measurable contour close to the uncomfortable loudness level, and the difference between the two gives one indication of the dynamic range. Normative data on equal loudness contours of pure tones presented in ISO 226:2023 (2023) show only a slight reduction in dynamic range at 12.5 kHz (the highest frequency reported) compared to that at 1 kHz. However, equal loudness contours up to 16 kHz have been measured in the freefield (Poulsen and Thøgersen, 1994; Takeshima *et al.*, 2001), which suggests that at 16 kHz, the dynamic range reduces still further above 12.5 kHz.

To extend the data on sensory unpleasantness to higher frequencies, Kurakata *et al.* (2013) presented participants with pure tones or narrowband noise between 1 and 18 kHz in a freefield at a number of SPLs between 0 and 80 dB and asked the participants to rate the sensory unpleasantness. They found that sensory unpleasantness at 18 kHz increased more rapidly with SPL than it did at 12.5 kHz and below, implying a reduction in dynamic range. Hence, loudness and unpleasantness appeared to show reduced dynamic ranges in the VHS/US frequency range.

The aim of the current study was to assess the ratings of sensory unpleasantness of audible US (tones at 18 kHz) and two VHF tones (at 14 and 16 kHz) while controlling for perceptions of loudness. Two different groups, symptomatic and asymptomatic, were assessed using the same criterion as employed Fletcher *et al.* (2018). The primary research question was whether and by how much unpleasantness ratings of VHFS/US were greater than those of a 1-kHz reference tone when stimulus levels were adjusted to achieve equal subjective loudness for all tones. A second aim was to assess the slope of subjective loudness rating vs SPL for VHFS/US tones.

II. METHODS

A. Participants

Participants were all acquaintances of the researchers. They were categorized as either symptomatic or asymptomatic based on their self-reported sensitivity to VHFS/US sound or US (Sec. II B 1). The aim was to recruit approximately equal numbers of symptomatic and asymptomatic participants. Initially 18 participants were recruited, but the results of 1 participant were excluded due to an inconsistency between their response to the screening questionnaire and the subsequent hearing assessment questionnaire. A total of 17 participants were tested (13 females and 4 males) aged 18–33 years old (mean 24.5 years old). The sample comprised nine asymptomatic participants (five females and four males, aged 18–33 years old; mean age 25.6 years old), and eight symptomatic participants (all female, aged 20–30 years old, mean age 23.3 years old). Participants were bilaterally normal on otoscopy and pure-tone audiometry between 0.125 and 8 kHz, free from tinnitus and hyperacusis, and had at least one ear with a measurable hearing threshold at either 16 or 18 kHz (Sec. II B 1). Testing was conducted on one ear only: the ear with the better hearing threshold at 16 or 18 kHz (Sec. II B 1). None of the participants were previously trained in the listening tasks performed in this study.

A pre-test sample-size calculation gave a required sample size of eight to detect an increased unpleasantness rating in the VHFS/US tone condition compared to the 1-kHz reference tone condition with a power of 80%. This assumed a mean increase of three-points and a standard deviation of three-points in the difference in ratings, estimated from the findings of Fletcher *et al.* (2018). and using a one-tailed paired *t*-test with a type-*I* error rate of 5%. The sample size was achieved in the symptomatic and asymptomatic groups for tones at 14 and 16 kHz, respectively, but not in all conditions at 18 kHz as some participants' could not be tested as a result of the institution's exposure limits (Sec. III).

The experiment complied with the declaration of Helsinki and was approved by University of Southampton Ethics Committee (No. 30365.A1).

B. Procedure

The experiment comprised four stages, screening, threshold measurement, loudness matching, and unpleasantness rating, and was conducted in that order.

1. Screening, ear selection, and the assignment of symptomatic status

Participants completed a questionnaire to ensure that they were free from tinnitus, hyperacusis, and other health-related exclusion criteria. Participants were categorized as either symptomatic or asymptomatic based on their self-reported sensitivity to VHFS/US sounds or US according to the participant's answer to the question "Have you ever experienced unpleasant symptoms that you believe were caused by exposure to very high frequency sound or ultrasound?" Hearing thresholds at 1, 14, 16, and 18 kHz were measured in both ears using the modified method of limits (BSA, 2017). The ear with the better hearing threshold at 18 kHz was selected as the test ear for further testing. If no threshold was obtainable at 18 kHz in either ear, then the threshold at 16 kHz was used to select the test ear. If no threshold at 16 kHz could be obtained in either ear, then the prospective participant was excluded from the study. A maximum stimulus level of 105 dB SPL was allowed for establishing hearing threshold levels to satisfy noise exposure limits set by the institution of an 8-h equivalent continuous *A*-weighted SPL of 76 dB with a maximum *A*-weighted SPL of 120 dB (ISVR, 1996).

In addition to the screening questionnaire, participants were also asked to complete a second questionnaire to assess other aspects of sensitivity to sound as part of a separate study with different goals (Ascone *et al.*, 2019). The questionnaire was unsuitable as a screening tool for the current study and, hence, the results are not used here.

2. Hearing threshold measurement

Once the test ear was established, more accurate pure-tone hearing threshold levels were measured at 1, 14, 16, and 18 kHz using a three-interval, three-alternative forced choice trial format with a stimulus duration of 1.5 s. For the four test frequencies, thresholds were obtained using an adaptive staircase with starting stimulus SPLs of 40, 80, 85, and 90 dB. The step size was 10 dB up to the first reversal, 5 dB up to the second reversal, and 2.5 dB for the remaining ten reversals. The first two reversals at this smallest step size were discarded, and the threshold was estimated from the mean of the final eight reversals. If the participant failed to correctly detect the signal three times in a row at the maximum allowable level of 105 dB SPL, the staircase was terminated, and the threshold was coded as >105 dB SPL. Of the 17 participants, 3 (all asymptomatic) had hearing thresholds >105 dB SPL at 18 kHz.

3. Subjective loudness matching of 14, 16, and 18-kHz tones to the 1-kHz reference tone

The 1-kHz tone was designated as the reference tone and three SLs of this tone were tested: 10, 20, and 30 dB SL. At each of these reference SLs, points on the equal loudness curve were obtained at the three VHFS/US test frequencies of 14, 16, and 18 kHz using a loudness-matching paradigm. This was achieved by fixing the level of the 1-kHz reference stimulus and asking the participant to adjust the level of the comparison tone until they judged that the subjective loudness of the comparison and reference stimuli to be equal. The participant could listen to the reference and comparison tones as many times as they wanted to achieve equality of subjective loudness. The level of the comparison tone could be adjusted up or down in either large (5 dB) or small (2.5 dB) steps, although the step sizes were not visible to the participant. For the three loudness levels, this resulted in nine loudness-matched stimulus levels in total for each participant. For each participant and each reference SL and comparison-tone frequency, the matching task was performed twice, once starting at a high comparison-tone level and once at a low comparison-tone level with a randomized offset in the starting level of the comparison tone each time. The two replicate loudness-matched stimulus levels were then averaged. The order of presentation of the 18 comparison trials (3 frequencies × 3 loudness levels × 2 repeats) was randomized for each participant.

The maximum stimulus level was 105 dB SPL (due to institutional safety limits on noise exposure); if the comparison-tone stimulus loudness reached 105 dB SPL without equal subjective loudness being achieved, the loudness-matched stimulus level was coded as missing. Out of the total of 153 cases (9 stimulus conditions × 17 participants), missing loudness-matched stimulus levels resulted for 28 cases (1 at 16 kHz and 27 at 18 kHz); details are given in Sec. III.

4. Subjective rating of sensory unpleasantness at three loudness levels

For each of the four test frequencies (1, 14, 16, and 18 kHz) and three loudness levels (SLs of the reference 1-

kHz tone of 10, 20, and 30 dB SL), participants were presented with a pure tone of 1-s duration and asked to rate the unpleasantness of the tone. The 1-s duration exceeds the temporal integration time of the auditory system (Stephens, 1973) while being short enough to avoid the institution's exposure limits from being exceeded (ISVR, 1996). For each presentation, the participant rated the unpleasantness on two different scales: an 11-point numerical rating scale from 0 to 10 (0 = not unpleasant at all and 10 = extremely unpleasant), and a 5-category verbal rating scale (categories, "not at all," "slightly," "moderately," "very," and "extremely"). The purpose of using the two scales was to reduce effects of bias in participants' interpretation of the scales and align with the recommendations of ISO/TS 15666:2021 (2021). In each condition, tones were presented 4 times, giving 48 presentations in total for each participant (4 frequencies × 3 loudness levels × 4 replicates). The order of presentation of the 48 test tones was randomized for each participant. Unpleasantness ratings were coded as "missing" for the same cases for which no loudness match was obtained (Sec. II B 3).

C. Equipment and stimuli

The experiment was conducted with the participant located in a sound-attenuated booth with background noise levels meeting the requirements of BSA (2017). Tones were presented via Sennheiser HDA-200 circumaural headphones (Sennheiser, Wedemark, Germany). The headphones were driven by an RME Babyface Pro soundcard (Haimhausen, Germany) connected to a laptop located in an observation room isolated from the test booth. Acoustic stimuli were generated with a sample rate of 96 kHz and 24-bit resolution using in-house software.

The stimuli were calibrated using a Bruel and Kjaer artificial ear (type 4152; Naerum, Denmark), conforming to IEC 60318-2 1998), with a flat-plate adaptor (DB0843). During calibration, the two earphones were separated by approximately 145 mm, giving a headband tension conforming to ISO 389-5:2006 (2006). No subharmonics were recorded, indicating that if there were any, they were below the noise floor of the calibration equipment. The participants registered their responses to the listening tasks via a mouse and computer monitor in the test booth, driven from the laptop in the observation room.

III. RESULTS

A. Hearing threshold levels

The median hearing threshold levels (Fig. 2) are broadly in agreement with the normative data tabulated in ISO 389-5 (2006) and ISO 389-8 (2004) for test frequencies of 1, 14, and 16 kHz and with data reported by Rodríguez Valiente *et al.* (2014) at 18 kHz. The increased intersubject variability observed in the hearing threshold levels for VHFS/US compared to that at the 1-kHz frequency is also analogous to previously reported normative data on variability (Han and Poulsen, 1998; Rodríguez Valiente *et al.*, 2014).

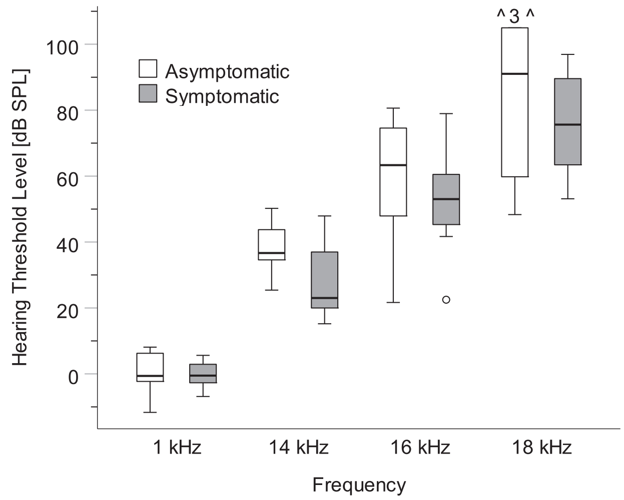

FIG. 2. Boxplots of hearing threshold levels expressed in dB SPL at four test frequencies for asymptomatic subjects ($n = 9$) and symptomatic participants ($n = 8$). At 18 kHz, three asymptomatic participants had unobtainable thresholds at SPLs ≤ 105 dB SPL, which were coded as 105 dB SPL for the purposes of the boxplot. The box indicates the interquartile range, the horizontal bar indicates the median, and the whiskers indicates the maximum and minimum values excluding outliers. Where present, outliers marked with an open circle indicate values that are 1.5 times the interquartile range larger than the third quartile or 1.5 times the interquartile range smaller than the first quartile. Extreme outliers, if present, are marked with an asterisk and indicate values that are three times the interquartile range larger than the third quartile or three times the interquartile range smaller than the first quartile.

B. SLs for equal subjective loudness

The SL that gave equal loudness was calculated by subtracting the hearing threshold levels from the loudness levels in dB SPL obtained in Sec. II B 3. The pattern of missing data resulting from the limit placed on the SPL is indicated in Fig. 3. The results (Fig. 3) suggest that lower SLs are required at the VHF/US frequencies to achieve the same subjective loudness as the reference 1-kHz tone, as predicted by the hypothesis that the dynamic range (based on subjective loudness) reduces in the VHF/US range. To test this for the combined sample of asymptomatic and symptomatic participants, an average over the three SLs was calculated at each of the four test frequencies. For the 1-, 14-, and 16-kHz conditions, a Friedman test showed a significant effect of stimulus frequency on the average loudness-matched SL [$n = 16$, $Q = 9.88$, degrees of freedom (df) = 2, $p < 0.01$; the 18-kHz condition was excluded due to a substantial number of participants being excluded as described in Sec. II B 3]. Three Wilcoxon signed-rank tests comparing the 14, 16, and 18-kHz conditions to the 1-kHz reference condition were conducted, which showed that the average loudness-matched SL in three VHFS/US conditions was statistically significantly lower than that at the 1-kHz reference condition after a Bonferroni correction for three tests (respectively, $n = 17$, 16, 9; Wilcoxon's $Z = -2.9$, -3.1, -2.4; $p = 0.004$, 0.002, 0.015; critical $p = 0.016$, with the caveat

https://doi.org/10.1121/10.0028380

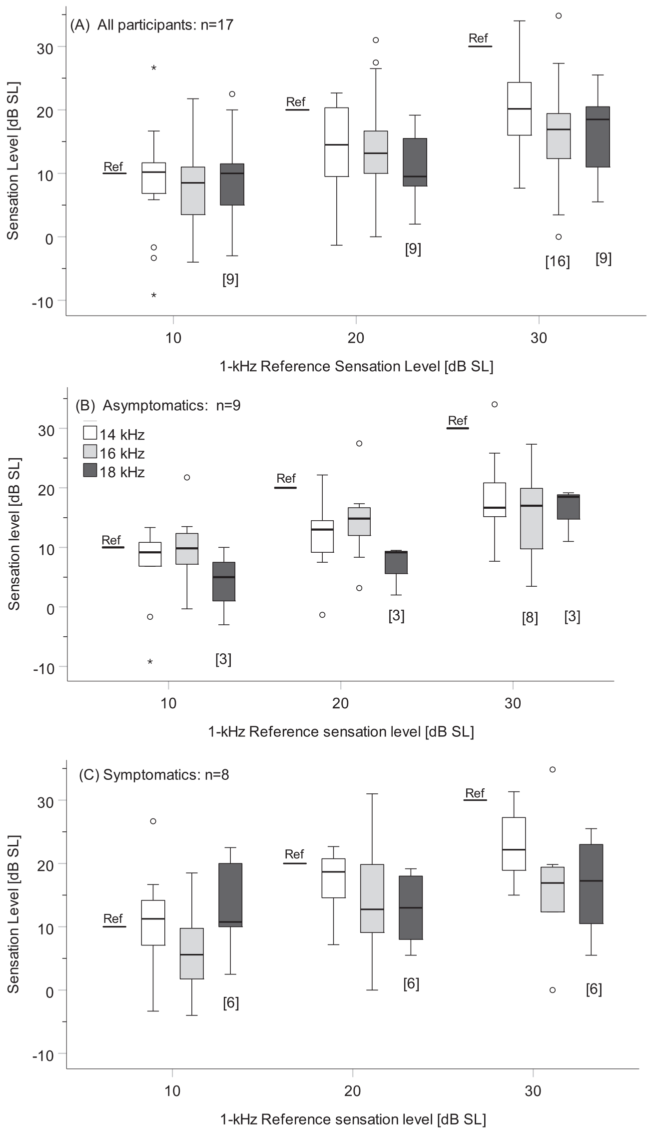

FIG. 3. Boxplots of SLs at 14, 16, and 18 kHz achieving equal subjective loudness to a 1-kHz reference tone at SLs of 10, 20, and 30 dB SL, showing (A) all participants ($n = 17$), (B) asymptomatic participants ($n = 9$), and (C) symptomatic participants ($n = 8$). Where there is missing data, the number of participants represented in the box plot is shown in square brackets beneath the box plot. The bar marked "Ref" indicates the SL of the 1-kHz reference tone. The key to the boxplot display is given in the caption for Fig. 2.

that the Bonferroni assumption of independence of the three Z-values is invalidated by the repeated-measures design). The results indicate that subjective loudness grew more rapidly with SL at the three VHF/US frequencies than at the 1-kHz reference frequency as discussed further in Sec. IV.

C. Ratings of sensory unpleasantness

The verbal unpleasantness ratings on the five-point Likert scale were converted first to a five-integer scale, 0–4, and then scaled by a factor of 2.5 to obtain a scale ranging from zero to ten (corresponding to "not at all" to "extremely"). For each participant and each of the 12 stimulus conditions, the unpleasantness ratings were averaged over the 4 replicates separately for the verbal and numerical ratings. As expected, a scatterplot of the two rating scales showed that the participants' ratings on the numerical scale were highly correlated with those on the verbal scale for both symptomatic-status groups (Fig. 4), suggesting a high degree of participant engagement in the task. To reduce measurement error, the average of the values on the two rating scales was calculated; this average is termed the "unpleasantness rating" in the following analyses.

Boxplots of the unpleasantness ratings within the sample (Fig. 5) suggest increasing median unpleasantness with stimulus frequency and subjective loudness. Intersubject variability also appears to increase with stimulus frequency. To illustrate these statistics more clearly, the median and interquartile interval are plotted in Fig. 6. As the data were not normally distributed, nonparametric tests of the statistical significance of the effects on the central tendencies of the unpleasantness ratings were conducted.

For the 1-, 14-, and 16-kHz conditions, a Friedman test showed a significant effect of stimulus frequency on the average unpleasantness ($n = 16$, $Q = 16.1$, df $= 2$, $p < 0.001$; the 18-kHz condition was excluded as in Sec. III B). Three subsequent Wilcoxon signed-rank tests comparing the 14-, 16-, and 18-kHz conditions to the 1-kHz reference condition showed significantly higher average unpleasantness ratings at the three VHFS/US conditions compared to the 1-kHz reference condition after a three-test Bonferroni correction; respectively, $n = 17$, $n = 16$, and $n = 9$; Wilcoxon's $Z = -3.6$, -3.1, and -2.4; $p < 0.001$, $p = 0.002$, and $p = 0.015$; critical $p = 0.016$.

The analysis of average unpleasantness ratings was repeated for the two subgroups. Comparing the 1-, 14-, and 16-kHz conditions, the Friedman test showed a statistically significant effect of stimulus frequency for the asymptomatic group ($n = 8$, $Q = 9.25$, df $= 2$, $p < 0.01$) and

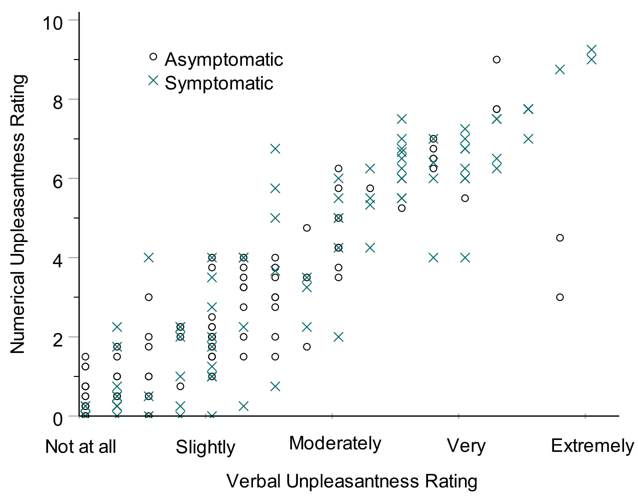

FIG. 4. (Color online) Scatterplot of unpleasantness rating on the numerical scale against unpleasantness ratings on the verbal scale. Each data point represents the rating for 1 subject in 1 of the 12 conditions (4 frequencies × 3 loudness levels) after averaging over the 4 replicate presentations of the tone. Filled circles indicate asymptomatic participants, and crosses indicate symptomatic participants.

JASA
https://doi.org/10.1121/10.0028380

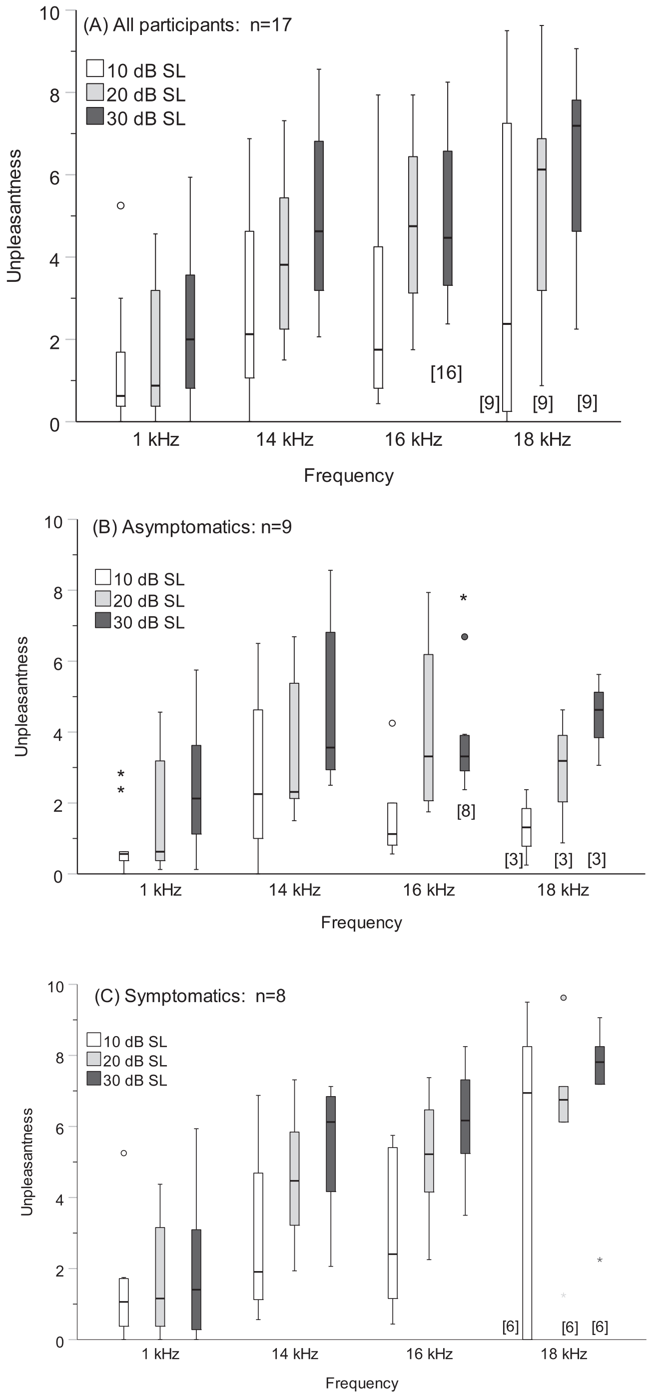

FIG. 5. Boxplots of unpleasantness ratings of the loudness-matched tones at four frequencies and three loudness levels corresponding to SLs of 10, 20, and 30 dB SL of the reference 1-kHz tone. The ratings are the average of the numerical and transformed verbal ratings after also averaging over four replicate presentations, showing (A) all participants ($n=17$), (B) asymptomatic participants ($n=9$), and (C) symptomatic participants ($n=8$). Where there is missing data, the number of participants represented in the box plot is shown in square brackets beneath the box plot.

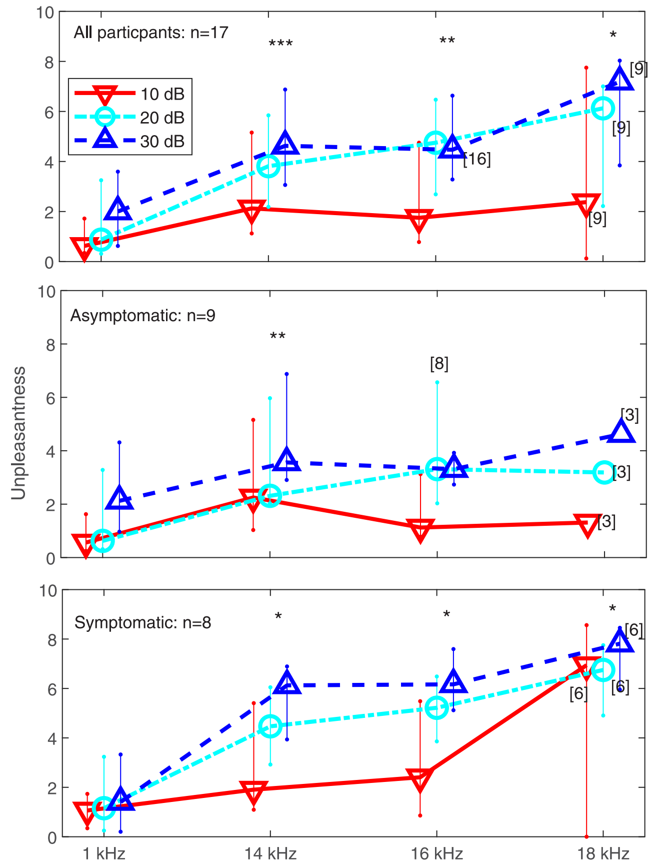

FIG. 6. (Color online) Median values of unpleasantness ratings as a function of stimulus frequency and equal loudness levels. The ratings are the average of the numerical and transformed verbal ratings after also averaging over four replicate presentations. Error bars indicate the upper and lower quartiles around the median except at 18 kHz in the middle panel, where there was insufficient data (i.e., $n=3$) to calculate quartiles meaningfully. The frequency values are plotted with slight jitter to increase clarity. The loudness levels are quantified by the SL at the 1-kHz reference frequency, which are 10, 20, and 30 dB SL as indicated by downward triangles, circles, and upward triangles, respectively. The upper, middle, and lower panels, respectively, present the data for all participants ($n=17$), asymptomatic subjects alone ($n=9$), and symptomatic subjects alone ($n=8$). Where there is missing data, the number of participants included in the statistics is indicated in square brackets next to the corresponding median. Asterisks at 14, 16, and 18 kHz indicate the significance of the Wilcoxon signed-rank test of unpleasantness at that frequency compared to the unpleasantness at the 1-kHz reference frequencies without Bonferroni correction: ***, < 0.001; **, <0.01; *, <0.05.

symptomatic group ($n=8$, $Q=7.75$, $\mathrm{df}=2$, $p<0.05$). For the asymptomatic group, two Wilcoxon signed-rank tests comparing the 14- and 16-kHz to the 1-kHz reference condition only showed statistically significant higher average unpleasantness ratings for the 14-kHz condition (respectively, $n=9$ and $n=8$; Wilcoxon's $Z=-2.7$ and -1.7; $p<0.01$ and $p>0.05$; critical $p=0.025$). A test at 18-kHz was inappropriate given that $n=3$. For the symptomatic group, the three Wilcoxon signed-rank tests showed a statistically significant higher average unpleasantness ratings at the 5% level of significance for 14-, 16-, and 18-kHz conditions compared to the 1-kHz condition but significance did not survive Bonferroni correction (respectively, $n=8$, $n=8$, and $n=6$; Wilcoxon's $Z=-2.4$, -2.4, and -2.0; $p=0.017$, $p=0.017$, and $p=0.05$; critical $p=0.016$). These results are summarized in Fig. 6.

A further objective of the study was to assess how problematic the measured degree of annoyance might be. A metric that has been widely used to assess the likelihood of complaints is the percentage of people who are highly annoyed by a sound source (Schultz, 1978). In this study, this was assessed by counting the number of participants whose rating of the unpleasantness (after averaging across the three loudness levels) were in or exceeded the "very unpleasant" category (unpleasantness rating > 6). The number of participants meeting this criterion in the 1-, 14-, 16-, and 18-kHz conditions was 0/17, 6/17, 5/16, and 4/9, respectively, corresponding to 0%, 35%, 31%, and 44% after accounting for missing data at 16 and 18 kHz. To test the statistical significance of the differences in these numbers between the VHFS/US conditions and 1-kHz reference condition, three McNemar tests were conducted to compare the numbers in the 1-kHz condition with those in the 14-, 16-, and 18-kHz conditions giving $p = 0.03$, $p = 0.06$, and $p = 0.13$, respectively. Thus, only the result in the 14-kHz condition was significant at the 0.05 level, although this did not survive Bonferroni correction for three tests. The comparison at 18 kHz was underpowered due to the small number of participants for whom the tone was audible.

To test whether the average unpleasantness ratings were higher for the symptomatic group than the asymptomatic group, one-tailed Mann-Whitney U tests were conducted at all four test frequencies. Only the unpleasantness ratings at 16 kHz were significantly higher in the symptomatic group than the asymptomatic group, but this result does not survive Bonferroni correction for four tests. For the 1-, 14-, 16-, and 18-kHz conditions, the results of the four tests were, respectively, $U = 33$, 27, 11, and 3; $p = 0.41$, $p = 0.21$, $p = 0.014$, and $p = 0.06$. Because the interpretation of the Mann-Whitney U test in terms of median differences is only valid when the two distributions are identical, the significance of the result at 16 kHz was further assessed using a bootstrap test for median differences between groups which showed a statistically significant difference in medians (Chernick, 2008). It should be noted that at 18 kHz, where there are only three asymptomatic participants, statistical power was severely limited.

IV. SUMMARY AND DISCUSSION

The results of the loudness-matching task indicate that for a given SL of a tone (SPL relative to hearing threshold), the listeners judged VHFS/US sounds to be louder than the 1-kHz reference tone at the same SL. This is consistent with a compression in dynamic range in the VHF/US frequency range, which manifests as subjective loudness increasing more rapidly with SPL than is observed at lower frequencies. Normative equal loudness contours in ISO 226:2023 (2023) are only tabulated up to 12.5 kHz; at 12.5 kHz, the equal loudness contours show slight compression but not to the extent found in the current study at 14, 16, and 18 kHz. However, equal loudness contours up to 16 kHz have been measured in two of the studies contributing to ISO 226:2023 (2023), where considerably greater compression was reported (Poulsen and Thøgersen, 1994; Takeshima *et al.*, 2001). A similar degree of compression based on loudness perception was observed in the current study whereby a change in SPL of 20 dB at 1 kHz resulted in the same median change in subjective loudness as that which resulted from a change in SPL of only around 10 dB at 16 kHz.

The unpleasantness of the tones was assessed using subjective ratings after adjusting the stimulus levels of the tones to achieve three specified loudness levels. The loudness-matched VHFS/US sounds were rated as significantly more unpleasant than the 1-kHz reference tone. This trend agrees with the findings of Kurakata *et al.* (2013), who assessed sensory unpleasantness at frequencies up to 18 kHz using tones presented in the freefield. The increase in unpleasantness with frequency for loudness-matched tones is qualitatively predicted by sound-quality models that relate perceived unpleasantness to the loudness and sharpness of a sound, which increases for the frequency of a tone, although these models were not developed using such high frequency stimuli (Fastl, 2005; Fastl and Zwicker, 2007). Similar trends of increasing adverse effects with stimulus frequency were also reported by Fletcher *et al.* (2018), although, in this case, the SLs were held constant rather than the subjective loudness, and ratings were for "discomfort" and "annoyance" rather than "unpleasantness."

In the current study, despite their relatively low SLs (median SLs were $<$20 dB), a considerable proportion of participants rated the unpleasantness of the VHFS/US tones as either very unpleasant or "extremely unpleasant," whereas no participants did so for the 1-kHz tone at the similar or higher SLs.

The results of the study have implications for guidelines on permissible SPLs (e.g., those in IRPA, 1984, which suggests exposure limits for the TOB, which is centered on 20 kHz and spans the range 17.8–22.4 kHz), that the public could be exposed to by devices emitting VHF/US. Because the equivalent freefield correction factors for the Sennheiser HDA 200 headphones is not known at 18 kHz, the SPL measurements in the current study made using the ear simulator cannot be directly converted to the equivalent freefield levels that are relevant to the guidelines in IRPA (1984). However, at 18 kHz, normative values for the median adult hearing threshold levels in the freefield are approximately 70 dB SPL (ISO 389-7, 2019), which is equal to the IRPA guidance for the maximum permissible levels for public exposure. Given the finding in the current study that SL of only 10–20 dB can cause very high ratings of unpleasantness and there is a wide intersubject variation in hearing thresholds at 18 kHz, it seems likely that a significant proportion of adults may find tones presented in the freefield at 18 kHz at 70 dB SPL to be very unpleasant. The consequence for a future noise reduction or safety strategy could be to base the absolute upper limit of exposure at US frequencies on a point on the hearing threshold distribution to avoid a hearing sensation altogether in the majority of the population. The results of the study also have implications for understanding the perception of VHF/US. Further studies are required to extend the normal equal loudness contours in ISO 226:2023 (2023) to higher frequencies and develop models of perceived annoyance that are applicable to VHF/US, including investigating any as-yet unknown effects of stimulus duration and order of presentation. Ideally, annoyance should also be studied in more realistic stimulus conditions and listening situations.

ACKNOWLEDGMENTS

This research was funded by the European Metrology Programme of Innovation and Research (EMPIR; Grant No. 15HLT03 EARII). The EMPIR Programme was cofinanced by the participating states and the European Union's Horizon 2020 Research and Innovation Programme.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

All experiments complied with the declaration of Helsinki and Acoustical Society of America (ASA) ethical principles. Ethics application 30365.A1 was approved by the University of Southampton ethics committee and informed consent was obtained from all participants.

DATA AVAILABILITY

The data that support the findings of this study are openly available from the University of Southampton repository, http://doi/10.5258/SOTON/D3167.

Alvares-Sanches, T., Osborne, P. E., White, P., and Bahaj, A. (**2019**). "Spatial variation in sound frequency components across an urban area derived from mobile surveys," Future Cities Environ. **5**, 1–17.

Ascone, L., Uppenkamp, S., Behler, O., Lineton, B., Burke, E., Koch, C., Kühn, S., and Geršak, G. (**2019**). "Deductive development and validation of a questionnaire to assess sensitivity to very low and very high frequency sounds: SISUS-Q (Sensitivity to Infra-Sound and Ultra-Sound Questionnaire)," Noise Health **21**, 173–182.

Ashihara, K., Kurakata, K., Mizunami, T., and Matsushita, K. (**2006**). "Hearing threshold for pure tones above 20 kHz," Acoust. Sci. Tech. **27**, 12–19.

BSA (**2017**). *Recommended Procedure: Pure-Tone Air-Conduction and Bone-Conduction Threshold Audiometry with and without Masking* (British Society of Audiology, Reading, UK).

Chernick, M. R. (**2008**). *Bootstrap Methods: A Guide for Practitioners and Researchers*, 2nd ed. (Wiley-Interscience, Hoboken, NJ).

DIN 45692:2009 (**2009**). "Measurement Technique for the Simulation of the Auditory Sensation of Sharpness" (German Institute for Standardization, Berlin, Germany).

Dolder, C. N., Fletcher, M. D., Jones, S. L., Lineton, B., Dennison, S., Symmonds, M., White, P. R., and Leighton, T. G. (**2018**). "Measurements of ultrasonic deterrents and an acoustically branded hairdryer: Ambiguities in guideline compliance," J. Acoust. Soc. Am. **144**, 2565–2574.

Dolder, C. N., Fletcher, M., Jones, S. L., Lineton, B., and Leighton, T. (**2019**). "Exposure levels for parametric arrays in light of guideline ambiguities," in *Proceedings of the 23rd International Congress on Acoustics: Integrating 4th EAA Euroregio 2019, 9-13 September 2019*, Aachen, Germany, edited by M. Ochmann, M. Vorlander, and J. Fels (Deutsche Gesellschaft für Akustik, Germany).

Fastl, H. (**2005**). "Psycho-acoustics and sound quality," in *Communication Acoustics*, edited by J. Blauert (Springer, Berlin), Chap. 6.

Fastl, H., and Zwicker, E. (**2007**). *Psychoacoustics: Facts and Models*, 3rd ed. (Springer, Berlin), Chap 9.

Fletcher, M. D., Jones, S. L., White, P. R., Dolder, C. N., Leighton, T. G., and Lineton, B. (**2018**). "Effects of very high-frequency sound and ultrasound on humans. Part I: Adverse symptoms after exposure to audible very-high frequency sound," J. Acoust. Soc. Am. **144**, 2511–2520.

Glorieux, C., and van Wieringen, A. (**2014**). "Undesirable effects as a result of short-term exposure to an ultrasonic repellent device. Part I—Acoustic measurements" (Federal Public Service for Health, Food Chain Safety and Environment, Brussels, Belgium), Assignment No. DGS/PB_PP/IVC/13026.

Han, L. A., and Poulsen, T. (**1998**). "Equivalent threshold sound pressure levels for Sennheiser HDA 200 earphone and Etymotic Research ER-2 insert earphone in the frequency range 125 Hz to 16 kHz," Scand. Audiol. **27**, 105–112.

Herman, B. A., and Powell, D. (**1981**). *Airborne Ultrasound: Measurement and Possible Adverse Effects* (Food and Drug Administration, U.S. Department of Health and Human Resources, Rockville, MD), FDA: 81-8163.

Huang, S., Dong, W., and Olson, E. S. (**2012**). "Subharmonic distortion in ear canal pressure and intracochlear pressure and motion," J. Assoc. Res. Otolaryngol. **13**, 461–471.

Huang, Y., and Zheng, Q. (**2022**). "Sound quality modelling of hairdryer noise," Appl. Acoust. **197**, 108904.

IEC 60318-2 (**1998**). *Electroacoustics—Simulators of Human Head and Ear—Part 2: An Interim Acoustic Coupler for the Calibration of Audiometric Earphones in the Extended High-Frequency Range* (International Electrotechnical Commission, Geneva, Switzerland).

IRPA (**1984**). "Interim guidelines on limits of human exposure to airborne ultrasound," International Non-Ionizing Radiation Committee of the International Radiation Protection Association, Health Phy. **46**, 969–974.

ISO 389-8:2004 (**2004**). "Acoustics—Reference zero for the calibration of audiometric equipment—Part 8: Reference equivalent threshold sound pressure level for pure tones and circumaural earphones" (International Organization for Standardization, Geneva, Switzerland).

ISO 389-5:2006 (**2006**). "Acoustics—Reference zero for the calibration of audiometric equipment—Part 5: Reference equivalent threshold sound pressure level for pure tones in the frequency range 8 kHz to 16 kHz" (International Organization for Standardization, Geneva, Switzerland).

ISO 532-1:2017 (**2017**). "Acoustics—Methods for calculating loudness—Part 1: Zwicker method" (International Organization for Standardization, Geneva, Switzerland).

ISO 389-7:2019 (**2019**). "Acoustics—Reference zero for the calibration of audiometric equipment-Part 7: Reference threshold of hearing under free-field and diffuse-field listening conditions" (International Organization for Standardization, Geneva, Switzerland).

ISO/TS 15666:2021 (**2021**). "Acoustics—Assessment of Noise Annoyance by Means of Social and Socio-Acoustic Surveys" (International Organization for Standardization, Geneva, Switzerland).

ISO 226:2023 (**2023**). "Acoustics–Normal equal-loudness-level contours" (International Organization for Standardization, Geneva, Switzerland).

ISVR (**1996**). "Guide to experimentation involving human subjects," ISVR Technical Memorandum No. 808, ISVR Human Experimentation Safety and Ethics Committee (Institute of Sound and Vibration Research, University of Southampton, Southampton, UK).

Kondo, K., and Osanai, S. (**2019**). "Audio spotlight using subdivided AM sideband wave delivery from separate ultrasonic array speakers," in *Proceedings of the 23rd International Congress on Acoustics*, 9–13 September, Aachen, Germany (Deutsche Gesellschaft für Akustik, Berlin, Germany), pp. 1037–1044.

Kühler, R., Weichenberger, M., Bauer, M., Hensel, J., Brühl, R., Ihlenfeld, A., Ittermann, B., Sander, T., Kühn, S., and Koch, C. (**2019**). "Does airborne ultrasound lead to activation of the auditory cortex?," Biomed. Tech. **64**, 481–493.

Kurakata, K., Mizunami, T., and Matsushita, K. (**2013**). "Sensory unpleasantness of high-frequency sounds," Acoust. Sci. Technol. **34**, 26–33.

Lawton, B. W. (**2001**). "Damage to human hearing by airborne sound of very high frequency or ultrasonic frequency," Health and Safety Executive Contract Research Report No. 343/2001 (Health and Safety Executive, Bootle, UK).

Leighton, T. G. (**2016**). "Are some people suffering as a result of increasing mass exposure of the public to ultrasound in air?," Proc. R. Soc. A **472**(2185), 20150624.

Leighton, T. G. (**2017**). "Comment on 'Are some people suffering as a result of increasing mass exposure of the public to ultrasound in air?,' " Proc. R. Soc. A **473**(2199), 20160828.

Leighton, T. G. (**2018**). "Ultrasound in air—Guidelines, applications, public exposures, and claims of attacks in Cuba and China," J. Acoust. Soc. Am. **144**, 2473–2489.

Leighton, T. G. (**2020**). "Ultrasound in air—Experimental studies of the underlying physics are difficult when the only sensors reporting contemporaneous data are human beings," Phys. Today **73**(12), 38–43.

Leighton, T. G., Currie, H. A. L., Holgate, A., Dolder, C. N., Jones, S. L., White, P. R., and Kemp, P. S. (**2019**). "Analogies in contextualizing human response to airborne ultrasound and fish response to acoustic noise and deterrents," Proc. Mtgs. Acoust. **37**, 010014.

Leighton, T. G., Lineton, B., Dolder, C. N., and Fletcher, M. D. (**2020**). "Public exposure to airborne ultrasound and very high frequency sound," Acoust. Today **16**, 17–26.

Liebler, M., Kling, C., Best, B., Gerlach, A., and Koch, C. (**2019**). "Quantitative characterization of high-intensity focused airborne ultrasonic fields," in *Proceedings of the 23rd International Congress on Acoustics*, 9–13 September 2019, Aachen, Germany (Deutsche Gesellschaft für Akustik, Berlin, Germany), pp. 6373–6380.

Liebler, M., Kling, C., Gerlach, A., and Koch, C. (**2020**). "Experimental characterization of high-intensity focused airborne ultrasound fields," J. Acoust. Soc. Am. **148**, 1713–1722.

Lubner, R. J., Kondamuri, N. S., Knoll, R. M., Ward, B. K., Littlefield, P. D., Rodgers, D., Abdullah, K. G., Remenschneider, A. K., and Kozin, E. D. (**2020**). "Review of audiovestibular symptoms following exposure to acoustic and electromagnetic energy outside conventional human hearing," Front Neurol. **11**, 234.

Mapp, P. (**2018**). "Potential audibility of ultrasonic signal monitoring of public address and life safety sound systems," J. Acoust. Soc. Am. **144**, 2539–2547.

Paxton, B., Harvie-Clark, J., and Albert, M. (**2018**). "Measurements of ultrasound from public address and voice alarm systems in public places," J. Acoust. Soc. Am. **144**, 2548–2553.

Pedersen, T. (**2007**). "The 'Genlyd' Noise Annoyance Model; Dose-response relationships modelled by logistic functions" (Danish Electronics, Light & Acoustics, DELTA, Copenhagen, Denmark), Journal No. AV 1102/07; Project No. FR13402-01.

Poulsen, T., and Thøgersen, L. (**1994**). "Hearing threshold and equal loudness level contours in a free sound field for pure tones from 1 kHz to 16 kHz," in *Proceedings of the 20th Nordic Acoustic Society Meeting*, 6–8 June, Arhus, Denmark (Nordic Acoustic Association, Denmark), pp. 195– 198.

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

Scholkmann, F. (**2019**). "Exposure to high-frequency sound and ultrasound in public places: Examples from Zurich, Switzerland," Acoustics **1**, 816–824.

Schultz, T. (**1978**). "Synthesis of social surveys on noise annoyance," J. Acoust. Soc. Am. **64**, 377–405.

Stephens, S. D. G. (**1973**). "Auditory temporal integration as a function of intensity," J. Sound and Vib. **30**, 109–126.

Takeshima, H., Suzuki, Y., Fujii, H., Kumagai, M., Ashihara, K., Fujimori, T., and Sone, T. (**2001**). "Equal-loudness contours measured by the randomized maximum likelihood sequential procedure," Acust. Acta Acust. **87**, 389–399.

Ueda, M., Ashihara, K., and Takahashi, H. (**2016**). "How high-frequency do children hear?," in *Internoise 2016, 45th International Congress and Exposition of Noise Control Engineering*, August 21–24, Hamburg, Germany (Deutsche Gesellschaft für Akustik, Berlin, Germany).

Ueda, M., Ota, A., and Takahashi, H. (**2014**). "Investigation on high frequency noise in public space—We tried noise abatement measures for displeasure people," in *Proceedings of the 7th Forum Acusticum*, 7–12 September, Krakow, Poland (Promet - Traffic and Transportation, Poland).

van Wieringen, A., and Glorieux, C. (**2018**). "Assessment of short-term exposure to an ultrasonic rodent repellent device," J. Acoust. Soc. Am. **144**, 2501–2510.

Weichenberger, M., Bug, M. U., Brühl, R., Ittermann, B., Koch, C., and Kühn, S. (**2022**). "Air-conducted ultrasound below the hearing threshold elicits functional changes in the cognitive control network," PLoS One **17**(12), e0277727.

Wygant, I. O., Kupnik, M., Windsor, J. C., Wright, W. M., Wochner, M. S., Yaralioglu, G. G., Hamilton, M. F., and Khuri-Yakub, B. T. (**2009**). "50 kHz capacitive micromachined ultrasonic transducers for generation of highly directional sound with parametric arrays," IEEE Trans. Ultrason., Ferroelect., Freq. Contr. **56**, 193–203.

Yoneyama, M., Fujimoto, J., Kawamo, Y., and Sasabe, S. (**1983**). "The audio spotlight: An application of nonlinear interaction of sound waves to a new type of loudspeaker design," J. Acoust. Soc. Am. **73**, 1532–1536.