Full title:

Effects of very high-frequency sound and ultrasound on humans part I: adverse symptoms after exposure to audible very-high frequency sound

Running title:

Effects of very high-frequency audible sound

Authors:

Mark D. Fletcher[1], Sian Lloyd Jones[2], Paul R. White[1], Craig N. Dolder[1], Timothy G. Leighton[1], Benjamin Lineton[1]

[1] Faculty of Engineering and the Environment, University of Southampton, University Road, Southampton, SO17 1BJ, United Kingdom

[2] Department of Audiology and Hearing Therapy, Royal South Hants Hospital, Brinton's Terrace, Southampton SO14 0YG, UK

Abstract

Various adverse symptoms resulting from exposure to very high-frequency sound (VHFS) and ultrasound (US) have previously been reported. This study aimed to establish whether these symptoms are experienced under controlled laboratory conditions and are specific to VHFS/US. To do this, participants were exposed to VHFS/US (at frequencies between 13.5 and 20 kHz and sound pressure levels between 82 and 92 dB) and to a 1 kHz reference stimulus, both at 25 dB above their hearing threshold. The VHFS/US and reference stimuli were presented 4 times, each time for 3 minutes, during which participants performed a sustained attention task, rated their symptom severity, and had their galvanic skin response (GSR) measured to assess their level of anxiety. Prior to exposure, participants were assigned either to a symptomatic or an asymptomatic group, based on their prior history of symptoms that they attributed to VHFS/US. In both groups, overall discomfort ratings were higher in the VHFS/US condition than the reference condition. In the symptomatic group only, difficulty concentrating and annoyance were also rated higher in the VHFS/US than the reference condition. No difference between the two stimulus conditions was seen in performance on the attention task or on average GSRs for either group.

I INTRODUCTION

For more than 70 years there have been reports of adverse symptoms resulting from exposure to very high frequency sound (VHFS; 11.2- 17.8 kHz) and ultrasound (US; > 17.8 kHz, for reasons discussed by Leighton, 2017). These reports are reviewed by Lawton (2001) and Leighton (2016a, 2017). Reported symptoms causing discomfort have included headache, tinnitus, fatigue, and fullness or pain in the ears. Recently, there have been media reports of discomfort and headaches resulting from exposure to VHFS/US in public places, from devices such as the 'Mosquito' youth repellent (Ebelthite, 2016; British Broadcasting Corporation, 2017). Our research group has also been contacted by several members of the public who report finding VHFS/US "extremely annoying", "unpleasant" and "hard to ignore" (Fletcher, 2016). These latter symptoms, unlike some of the physiological symptoms, are often reported to arise as soon as, or very soon after, VHFS/US is first perceived.

Some previous studies have attempted to measure effects of exposure to VHFS/US. Ueda *et al.* (2014) assessed symptoms experienced by a number of members of the public visiting a restaurant where they were exposed to US from a rodent repeller, which produced sound at a frequency of 20 kHz and at levels of between 90 and 130 dB sound pressure level. This study only included people who reported hearing the sound source. The respondents gave high ratings on a number of items, including pain in the ear, discomfort, and irritation. Another study, reported by Glorieux (2014) and van Wieringen (2014) exposed participants to a pest deterrent under more controlled conditions, in a semi-anechoic chamber. The deterrent had settings that produced both audible and inaudible VHFS/US ranging in frequency from 12.5 to 35 kHz and ranging in level from 44 to 71 dB sound pressure level.

After exposures of 20 minutes, participants rated their symptoms. Average ratings of severity were generally close to the lowest rating possible, suggesting very minor or no adverse effects. However, the authors stated that they avoided using some of the device settings because they were found to be “so loud and annoying” that they would be inappropriate for use with participants.

No previous study has determined whether the range of reported adverse symptoms can be provoked under controlled laboratory conditions by exposing participants to audible VHFS/US at the levels that may be encountered in public spaces and that approach the limit of what is deemed safe. The aim of the present study was to assess the degree to which exposure to audible VHFS/US may lead to adverse physiological and psychological effects compared to exposure to a reference 1 kHz tone for the same sensation level and exposure time. A participant’s symptoms were assessed using three different types of outcome measure: their subjective ratings of psychological and physiological symptoms, their performance on an attention task, and their galvanic skin responses (GSR). The primary outcome measure was their subjective rating of overall discomfort. Two different participant groups were studied: symptomatic participants who, prior to recruitment, reported experiencing some adverse effects which they attributed to VHFS/US, and asymptomatic participants who did not report any previous adverse effects of VHFS/US. The adverse symptoms reported by symptomatic participants in the group selection phase included nausea, pain or pressure in the ears or head, a feeling of light-headedness or dizziness, anxiety, annoyance, tiredness, and inability to concentrate. No previous study known to the authors has separated those who have complained of public nuisance from VHFS/US and those who have not, and so may have missed symptoms only experienced by the most vulnerable individuals. The aim of the study was the same for both groups: to

assess the difference in outcome measures between exposure to VHFS/US and exposure to a 1 kHz reference tone at the same sensation level.

The attention task was included as one outcome measure since "inability to concentrate" was one of the symptoms previously attributed to VHFS/US exposure. GSR was included as an objective measure that is affected by the participant's psychological state. GSR measures skin conductivity, which varies with changes in the state of sweat glands. This is known to be correlated with psychological state of arousal and therefore GSR has frequently been used as a measure of stress and anxiety (Homma *et al.*, 1998; Homma *et al.*, 2001; Kobayashi *et al.*, 2003 Kucera *et al.*, 2004). GSR has previously been used to assess emotional responses to sound including whether or not they are unpleasant or annoying (Khalfa *et al.,* 2002; Dimberg, 1990; Björk, 1986).

II METHODS

A. Procedure

1. Experimental session structure

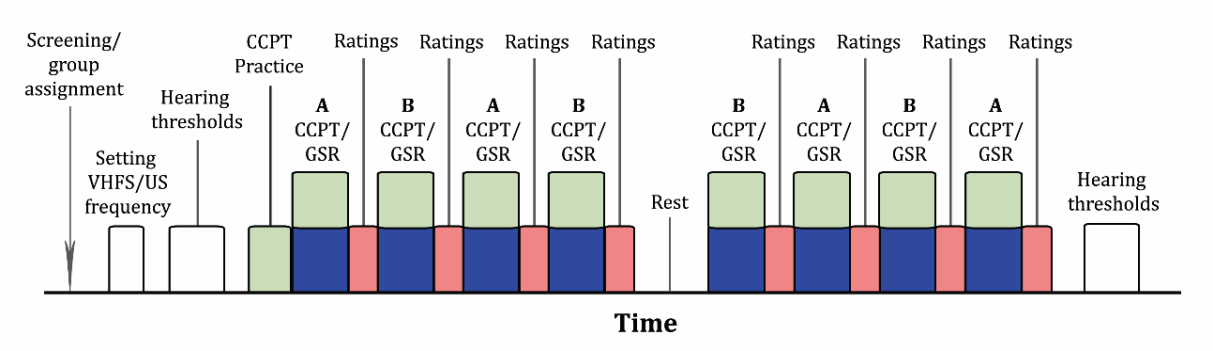

Figure 1 (color online): schematic (not to scale), showing the timeline of a session. The condition label (A/B) is shown in bold (see section II.A.8).

There were six main stages to the experiment (the structure of a single session is shown in Figure 1), which are discussed in detail in the following sections:

1. Participants were screened and assigned to either the symptomatic or asymptomatic group based on their self-reported sensitivity to VHFS/US (section II.A.2).
2. The frequency at which the VHFS/US stimulus would be presented was determined (section II.A.3).
3. Hearing thresholds were measured to determine the levels at which the VHFS/US and reference stimuli would be presented (section II.A.4). Hearing thresholds for VHFS/US are known to vary with removal and replacement of the headphones (e.g. Stelmachowicz et al., 1989) and therefore once the headphones were positioned over the participant's ears, they were not removed for the duration of the experiment.
4. To familiarize participants with the Conjunctive Continuous Performance Task (CCPT), which was used to measure sustained attention, participants had a practice session where their performance was monitored to ensure they had understood the task (section II.A.5).
5. Participants sat through eight exposure periods, alternating between the VHFS/US and reference conditions (section II.A.8). Participants performed the CCPT task (section II.A.5) and had their GSR measured (section II.A.6) while the sounds were presented. After each exposure period, the CCPT and GSR were stopped and participants completed the subjective ratings of symptoms (section II.A.7). Following the forth exposure period, once the symptom ratings had been submitted, there was a 1-minute rest period before the experiment continued.

6. Hearing thresholds for the VHFS/US and reference stimuli were measured again to ensure that the level reaching the participant's ear had not changed significantly.

2. Screening and group assignment

The inclusion criteria for the symptomatic group were that they reported symptoms (including annoyance, distraction, anxiety, ear pain, headache, or fatigue) that they believed were caused by audible VHFS/US. Participants were excluded if they had hearing threshold levels that exceeded 20 dB hearing level (HL) at any of the standard audiometric frequencies of 0.25, 0.5, 1, 2, 4 and 8 kHz in either ear. These thresholds were measured following the recommended procedure of the British Society of Audiology (British Society of Audiology, 2017). In addition, participants were required to have a hearing threshold at 13 kHz that did not exceed 63 dB sound pressure level (SPL; all SPLs stated re 20 µPa), as described in section II.A.3.

Participants who reported no adverse symptoms that they attributed to VHFS/US were recruited to an asymptomatic group. Participants were excluded if they did not have hearing threshold levels less than or equal to 20 dB HL at standard audiometric frequencies in either ear or had a hearing threshold at 14 kHz that exceeded 63 dB SPL. The purpose of this exclusion criterion was to ensure that the stimulus frequency was similar to that typically found in devices found in public spaces that that have led to the reports of the symptoms detailed in section I (Leighton, 2016a; Ueda *et al.*, 2014; van Wieringen, 2014). The criterion was set initially at 14 kHz but was relaxed to 13 kHz for symptomatic participants to facilitate recruitment. Full details of the questionnaires used for screening and for group membership can be found in Appendix 1. A discussion section was included at the end of the questionnaire to give participants the opportunity to report additional symptoms that

they experienced after exposure to VHFS/US, which might be useful for future studies. No such symptoms were identified.

For all participants, additional exclusion criteria were troublesome hyperacusis (the inability to tolerate sounds that are not uncomfortably loud for most people) and tinnitus (for safety reasons), color blindness (as color blindness would have left them unable to do the attention task, see section II.A.4), epilepsy (as the attention task included flashing images which could trigger an epileptic fit), recent alcohol or recreational drug use and recent strenuous physical or mental activity (as these could affect performance on the attention task). All participants were also asked to limit their caffeine intake to no more than one cup of coffee on the day of testing to avoid artificially enhanced performance on the sustained attention task (see Foxe *et al.*, 2012). They were also excluded if their subjective rating of any symptom prior to testing exceeded a pre-set threshold of 2, on an 11-point rating scale (section II.A.7).

3. Setting the VHFS/US frequency

For each participant, it was desired to find the frequency at which their hearing threshold level was 63 dB SPL. This was so that the VHFS/US tone could be set to a sensation level of 25 dB without the maximum allowable level of 92 dB SPL being exceeded, which was set to meet safety requirements for allowable daily noise exposure. The ethical framework of this study would not permit louder or longer exposures than the maximum ones used here and the symptoms studied only apply for this level of exposure. Devices operating at levels higher than those used here fall outside the scope of this study and additional data at those exposure levels would be required to determine the safety and presence of adverse symptoms.

The frequency was determined using an automated three-interval, three-alternative forced choice paradigm. In these threshold measurements, and in all subsequent tests, the stimulus was presented diotically in order to more closely reproduce real-world exposure conditions than would be achieved by monaural stimulation. Each trial comprised three listening intervals. One interval, chosen randomly with equal *a priori* probability, contained the signal, and the other two contained silence. The participant selected the response using the mouse and visual feedback was given indicating whether the response was correct or incorrect. The listening intervals were 550 ms in duration, separated by 300 ms of silence. The stimulus steady-state duration was 500 ms, with 25 ms quarter-sine and quarter-cosine ramps at the beginning and end of the stimulus (making the total duration between 0 volt points 550 ms). The stimulus level was fixed at 63 dB SPL and the frequency was adapted using a two-up one-down procedure in steps of 1 kHz up to the first reversal, 0.5 kHz up to the second reversal, and 0.25 kHz for the remaining eight reversals. The starting frequency was 12 kHz. The resulting frequency was estimated as the average of the last 8 reversals. The VHFS/US tonal stimulus used in the main experiment was set individually to the frequency determined using this task. For the asymptomatic group, participants were only included if this frequency was greater than 14 kHz, and for the symptomatic group if this threshold was greater than 13 kHz (section II.A.2). The mean VHFS/US tone frequency for the symptomatic group was 15.5 kHz [± 0.4 kHz (standard error of the mean), ranging from 13.4 to 17.7 kHz], and for the asymptomatic group was 16.8 kHz (± 0.26 kHz, ranging from 14.1 to 20 kHz).

4. Estimation of pure-tone hearing thresholds

Having established the desired VHFS/US stimulus frequency (section II.A.3), the hearing threshold at that frequency was re-measured using a three-alternative forced-choice procedure, with stimulus level adapted. Although this was expected to be close to 63 dB SPL, re-measuring ensured that sensation levels were set in an identical way for both VHFS/US and the 1 kHz reference condition. For these measurements, 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 eight reversals. The starting level of the signal was 30 dB SPL for the 1 kHz threshold measurement and 83 dB SPL for the tonal VHFS/US measurement. The staircase terminated early if there were three incorrect responses at the maximum stimulus level of 105 dB SPL and the detection threshold was assumed to be above this level. The signal exposure levels for the VHFS/US and 1 kHz stimuli were set at 25 dB above the detection thresholds determined using the level adaptation task. The mean VHFS/US tone level for the symptomatic group was 88.2 dB SPL (± 0.97 dB, ranging from 81.8 to 91.8 dB SPL), and for the asymptomatic group was 87.9 dB SPL (± 0.38 dB, ranging from 83.9 to 91.8 dB SPL). The reference tone level for the symptomatic group was 24 dB SPL (± 1.11 dB, ranging from 19.1 to 28.1 dB SPL), and for the asymptomatic group was 24.8 dB SPL (± 0.68 dB, ranging from 18.4 to 33.8 dB SPL). The maximum allowable unweighted stimulus sound pressure level for the high-frequency tone, when set at 25 dB above the participant's detection threshold, was 92 dB SPL, in order to meet safety requirements for allowable daily noise exposure in this frequency range. The level setting procedure maximized the frequency and level of the VHFS/US stimulus, without exceeding this limit.

At the end of the experiment, hearing thresholds for the VHFS/US and 1 kHz reference tone were re-measured to ensure that the level reaching the participant's ear had not significantly changed over the course of the experiment. The unsigned average difference

between thresholds before and after the experiment for the symptomatic group was 2.1 dB (± 0.6 dB) for the VHFS/US tone and 2 dB (± 0.6 dB) for the 1 kHz tone and for the asymptomatic group was 3.9 dB (± 0.8 dB) for the VHFS/US tone and 2.1 dB (± 0.5 dB) for 1 kHz tone.

5. Assessment of sustained attention: Conjunctive Continuous Performance Task (CCPT)

One of the symptoms that was assessed using both subjective ratings and a separate performance measure was distraction during exposure to either VHFS/US or the 1 kHz reference tone. To assess distraction during sound exposure, the participant's ability to sustain attention was measured using a continuous performance task. In this task, a visual stimulus was presented, comprising two objects with different shapes and colors. In every trial, there was a large outer square shape on a black background, within which there was a smaller colored shape that was either a triangle, square, or pentagon. The inner and outer shapes could be red, blue, or green (with the inner and outer shapes never having the same color). The task was to click when the outer square shape was blue and the inner shape was a green pentagon. This is an adaptation of the Conjunctive Continuous Performance Task (CCPT) method used by Shalev *et al.* (2011), where only one shape and color combination was used. The task was altered here in order to make the task more difficult, as ceiling effects were observed in some participants during the piloting phase of this study.

The target stimuli appeared on 15 % of trials, which were selected at random. Stimuli were displayed for 100 ms, after which no stimulus (a black screen) was displayed for a duration of between 1000 and 1500 ms, which were randomly varied. Participant responses were recorded from the onset of the stimulus until the end of the trial (i.e. any time before the next image was displayed). CCPTs lasting 3 minutes were completed four times for each

condition (section II.A.8). A five second countdown was given before the start of each CCPT so that the trial start was not unexpected.

At the start of the session, participants completed a familiarization session with the CCPT for 40 trials in the absence of sound presentation. Instructions were given before the start of the familiarization session. As part of these instructions, the participant was shown an image and text description of the target shape. Participants were instructed to give equal priority to speed and accuracy in their responses. To ensure they had understood the task, this familiarization was repeated until they achieved a performance level of greater than 25 % correct identification of target trials and less than 10 % incorrect identification of non-target trials. No participant required more than two training sessions to achieve the required level of performance.

6. Galvanic skin responses

The two electrodes used to measure GSR were attached to the proximal phalanges of the index and middle fingers of either the left or right hand (the opposite hand to that which the participant would normally use to operate a mouse). Recordings were made throughout each sound exposure. Participants were asked to keep the hand with the GSR sensors attached as still as possible throughout the experiment.

7. Subjective ratings of symptoms

At nine times during the experiment (including once as part of the screening session), the participant was asked to give a subjective rating of the severity of the following 10 items: overall discomfort, nausea, pain, pressure, or fullness in one or both ears, headache/ pain or pressure somewhere other than the ears, dizziness or light-headedness, tinnitus, anxiety,

annoyance or irritation, fatigue, inability to concentrate, and other symptoms. The format of the questions posed to participants is given in Appendix 2. Participants were asked to give a rating for each item on an 11-point scale from 0 to 10, with “0” and “10” given the descriptors of “not at all” and “severe”, respectively. The items were selected because they have been previously associated with VHFS/US (Maccà *et al.*, 2014; Ueda *et al.*, 2014; van Wieringen, 2014), or had been reported by symptomatic people who had previously made contact with our research group. The “overall discomfort” item was included as the primary outcome measure to allow comparisons across participants who may experience different symptoms, or may select different descriptions of similar symptoms. Overall discomfort was always rated first and “other symptoms” was always rated last, but all other symptoms were rated in a random order each time that ratings were given. At the beginning of the experiment, participants were given the opportunity to have the meaning of any of the symptoms listed clarified.

8. Exposure conditions

Participants experienced two acoustic exposure conditions: a reference condition, where the frequency of a pure tone stimulus was set at 1 kHz, and a VHFS/US condition where the pure tone frequency was set using the protocol described in section II.A.3. For both conditions, the hearing thresholds were measured (section II.A.4) and these thresholds were used to set the stimuli at a level of 25 dB above the participant’s detection threshold in the exposure conditions.

A total of eight three-minute exposure periods were used in an alternating A-B-A-B pattern for the first half of the session and a B-A-B-A pattern for the second half, where “A” used one of the exposure frequencies and “B” the other. After the first half of the session, a break

of 60 seconds was given [during which the headphones were kept on and the participant remained seated to avoid headphone repositioning errors (section II.A.1)]. Whether "A" was the reference stimuli or the VHFS/US was counterbalanced across participants. After each exposure period, the acoustic stimulus was stopped, and participant was asked to rate their symptoms (section II.A.7), after which the next condition began. This gave a typical duration of approximately 1 minute between exposure periods.

During each exposure period, the sound was presented diotically as a pulsed pure tone, with a pulse duration that varied randomly between 500, 1000, 1500 and 2000 ms, and an inter-pulse silent interval that varied randomly between 250, 500, 750 and 1000 ms (stated sound levels refer to the time that the tone was on). Interval and pulse durations were selected independently. The pulse was gated on and off with 25 ms quarter-sine and quarter-cosine ramps, respectively. One reason for using pulsed pure tones was to minimize the rapid loudness adaptation that has been reported particularly for high-frequency tones (Bacon and Viemeister, 1994). A second was to make the sound less predictable, reducing habituation, and hence potentially increasing annoyance and any adverse effect on concentration. Participants could immediately stop the sound at any time by pressing a stop button displayed on the screen. No participants stopped the experiment.

The experiment was approved by the Ethics Committee of the Faculty and Engineering and the Environment at the University of Southampton (submission number 26450).

B. Participants

Two groups of participants were recruited. The group labelled "symptomatic" comprised participants who had either responded to a call for people who experienced some adverse symptoms that they attributed to VHFS/US, or who had contacted the researchers due to

their symptoms. The call was put out in the form of posters displayed around the University of Southampton, and through social media (Fletcher, 2016) and a website (Leighton, 2016b), which was given publicity via the national media (e.g. Gallagher, 2016).

A total of 10 participants were recruited for the symptomatic group, comprising 3 males and 7 females, with a mean age of 30 years (range 19 to 48 years) and 32 other participants (18 males and 14 females, with a mean age of 23 years, ranging from 18 to 34 years), who reported no adverse symptoms that they attributed to VHFS/US, were recruited to an "asymptomatic" group. The participants were screened following the criteria and procedures detailed in section II.A.2.

For both groups, the primary outcome measure was the difference in overall discomfort rating between the reference and VHFS/US conditions. An increase in rating of 3 points would be considered important, as it represents an increase from no noticeable effect to a clearly noticeable adverse effect. A pilot study indicated that the measurement test-retest error (standard deviation of measurement in identical exposure conditions) was typically <1 point in asymptomatic participants who scored near the bottom of the rating scale throughout the test. However, this is likely to be an underestimate for cases where symptoms arise. For the purposes of this study, the target sample size was set using the much more conservative estimate of the standard deviation of 6 points.

The pre-test sample-size calculation was based on the directional alternative hypothesis. This was, specifically, that the effect of VHFS/US would be a three point increase on the response scale for the primary outcome measure compared to the reference condition. The standard deviation of the difference measure was assumed to be 6 points. For a type 1 error rate of 5%, and a power of 80%, this gave a required sample size of 27 for a paired t-test.

Although the distribution of the outcome measure was known to be positively skewed, violating the assumption of normality required for the t-test, the test was deemed appropriate given the conservative assumptions regarding standard deviation.

The sample size was exceeded in the asymptomatic group (32 participants), though not in the symptomatic group (10 participants). The number of symptomatics was reduced by the fact that several of those who contacted the group were not willing to undertake the journey to the test site at Southampton.

C. Equipment

Participants were seated in a sound-attenuated booth with a background noise level conforming to the recommendations of the British Society of Audiology (2017). Acoustic stimuli were generated by a laptop located in a separate observation room, and played out via an RME Babyface Pro soundcard, a Creek OBH-21 headphone amplifier, and Sennheisser HDA 200 circumaural headphones. A sample rate of 96 kHz and a bit depth of 24 bits was used. The stimuli were calibrated using a Bruel and Kjaer artificial ear (type 4152) with a flat-plate adaptor (DB0843), in a configuration such that the two earphones were separated by approximately 145 mm, as specified in ISO 389-5:2006. The headband tension was measured and found to comply with the requirement of ISO 389-5:2006. The level of any subharmonics was below the noise floor of the artificial ear.

The participants sat facing a computer screen (at a distance of 1 m from the eyes) operating a mouse. The screen and mouse were connected to the laptop located in the observation room. The screen and mouse were used for measurements of hearing threshold, for administering a visual task used for assessing performance on a sustained attention task, and for collecting subjective ratings.

GSR was recorded using an Edu-lab Galvanic Skin Response Logger system, which was connected to a second laptop located in the observation room. GSR recordings gave a time-history of skin conductance, with a sample rate of 20 Hz.

III RESULTS

A. Primary outcome measure

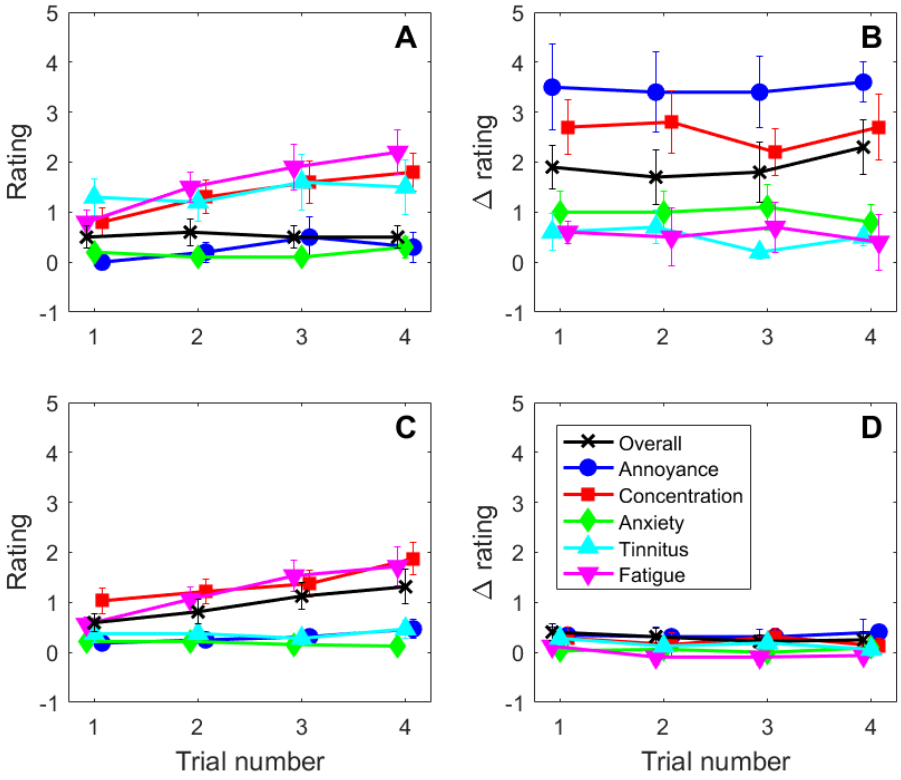

Figure 2 (color online): Variation in ratings over the four trials for a subset of symptoms, averaged over participants for the symptomatic (10 participants) and asymptomatic (32 participants) groups. The mean subjective ratings for the symptomatic (panel A) and asymptomatic (panel C) groups for the 1 kHz reference condition, and the change in rating between the reference and VHFS/US condition for the symptomatic (panel B) and asymptomatic (panel D) groups are shown. In panels B and D, a positive value indicates that the rating was higher in VHFS/US condition than in the reference condition. The markers are horizontally offset or clarity. Error bars show the standard error of the mean.

The primary outcome measure was the difference in the overall discomfort rating between the VHFS/US condition and the reference 1 kHz condition (Figure 2), analyzed separately for the two groups of participants.

For both groups, the high number of low overall discomfort ratings led to a strongly positively skewed distribution that approximately followed a gamma distribution. One approach to analyzing the results was to use generalized estimating equations (e.g. Hardin and Hilbe, 2003), which allow a repeated-measures analysis of a gamma-distributed random variable. Using generalized estimating equations to model the overall discomfort rating with stimulus condition as a factor and time as covariate showed that VHFS/US led to greater discomfort than the 1 kHz reference stimulus, with a high statistical significance both for the asymptomatic group (*Wald* $\chi^2(1) = 13.7$; $p < 0.0005$; mean effect: 0.3 rating points) and the symptomatic group (*Wald* $\chi^2(1) = 7.5$, $p < 0.01$; mean effect: 1.9 rating points). There was a statistically significant increase in overall discomfort over time in the asymptomatic group, but not in the symptomatic group. There was no statistically significant interaction between time period and exposure condition.

The statistical results were verified using a more conservative non-parametric Friedman's analysis, which also showed a statistically significant difference in overall discomfort rating between the VHFS/US and reference conditions (symptomatic group: $Q(1) = 8.0$, $p = 0.008$; asymptomatic group: $Q(1) = 6.5$, $p = 0.02$), with no change in the difference between conditions over time (symptomatic group: $Q(3) = 2.8$, $p = 0.44$; asymptomatic group: $Q(3) = 2.5$, $p = 0.48$). Figure 2 shows the variation of mean discomfort over the four time points (trials) for each group. Though statistically significant, the overall discomfort ratings remain

relatively low even in the VHFS/US condition for both groups of participants; recall that the scale runs from 0 (“not at all”) to 10 (“severe”).

B Secondary outcome measures.

Statistical analysis of secondary outcome measures were made without any additional correction for multiple hypothesis tests, using the more conservative non-parametric Friedman’s analysis. Though less statistically reliable than the primary measure, they nonetheless provide insight into the underlying symptoms that contribute to the increased discomfort in the VHFS/US condition.

1. *Subjective Ratings*

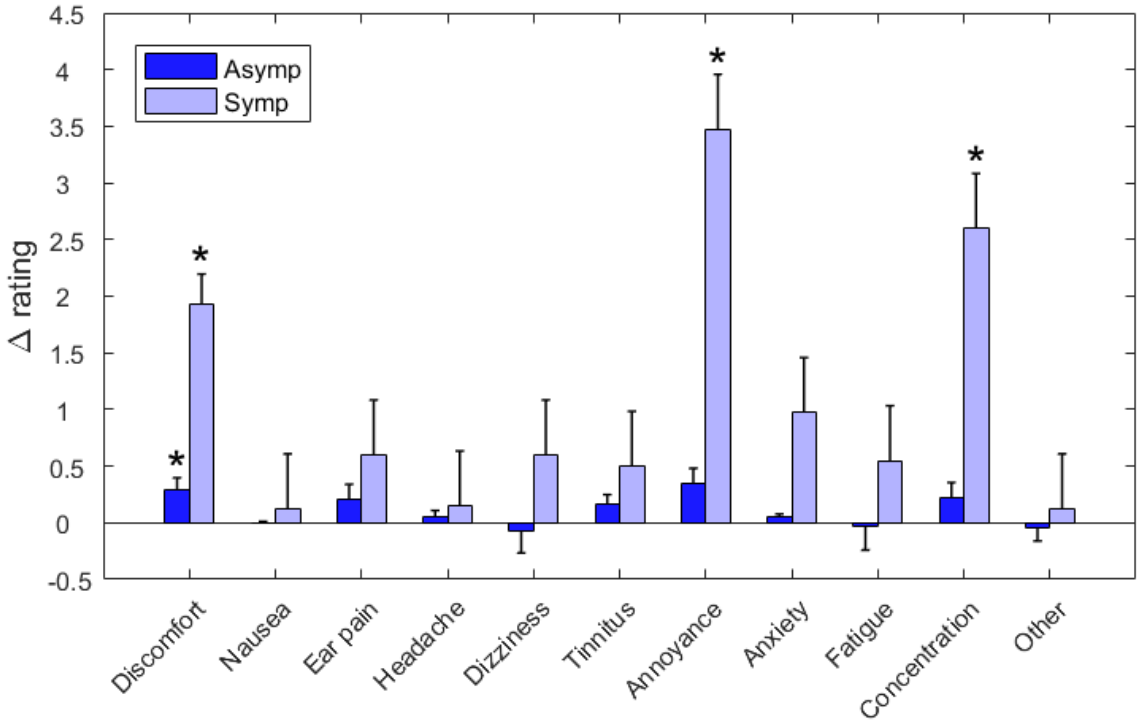

Figure 3 (color online): Mean difference in rating between the VHFS/US and 1 kHz reference conditions for all ratings. Ratings are averaged over all four time periods, and over all participants in the group (32 asymptomatic; 10 symptomatic). For all items, a positive effect means that the VHFS/US condition produced a higher symptom rating than the 1 kHz reference condition. Error bars show the standard error of the mean across participants. Statistically significant differences ($p < 0.05$) are marked with an asterisk.

The differences between the VHFS/US and the reference conditions were analyzed for each of the individual items rated by participants using a non-parametric Friedman's analysis. Highly significant effects of condition were found for annoyance ($Q(1)$ = 10.0, p = 0.002; 2.6 mean difference in rating points) and for inability to concentrate ($Q(1)$ = 10.0, p = 0.002; 3.5 mean difference in rating points) for the symptomatic group, as shown in Figures 2 and 3. As can be seen in Figure 2, symptoms such as fatigue and concentration tended to be rated as more severe at later time points, as may be expected, but the difference between the VHFS/US and 1 kHz reference conditions did not show a statistically significant change with time.

2. Performance on sustained attention task

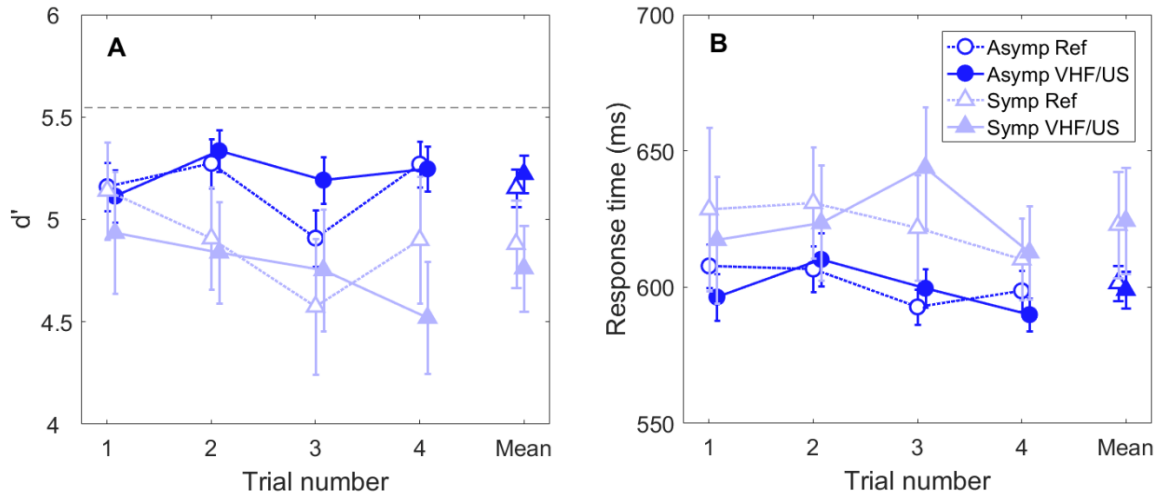

Figure 4 (color online): Variation over the four trials of performance metrics for the conjunctive continuous performance task (CCPT), averaged over participants for the two groups (32 in asymptomatic group; 10 in symptomatic group). Panel A shows the d' measure of discriminability and panel B shows the mean reaction time over participants. The dashed line in panel A shows the maximum possible d' value. Markers are horizontally offset for clarity. Error bars show the standard error of the mean across participants.

Three quantities characterized the performance on each 3-minute trial: the hit rate (i.e., the percentage of correct detections of the target), the false-alarm rate, and the reaction time.

The hit rate and false alarm rate were combined into the single measure of discriminability, *d'*, with corrections applied for ceiling and floor effects (Macmillan and Creelman, 2005). In the asymptomatic group, in 70 % of trials performance was at 100 % correct and on 37 % of trials the false alarm rate was 0 %. In the symptomatic group, in 58 % of trials performance was at 100% correct and on 33 % of trials the false alarm rate of 0 %. This indicates that there were substantial ceiling and floor effects.

No statistically significant differences were found between the VHFS/US and reference condition for either reaction time or for *d'*, in either the asymptomatic or symptomatic group, despite a significant increase in the subjective rating of difficulty concentrating for the symptomatic group. Means across participants are shown in Figure 4. Differences in the response times between the VHFS/US and 1 kHz conditions were not shown to correlate with the differences in subjective ratings for any of the symptoms for either group.

3. Galvanic skin responses

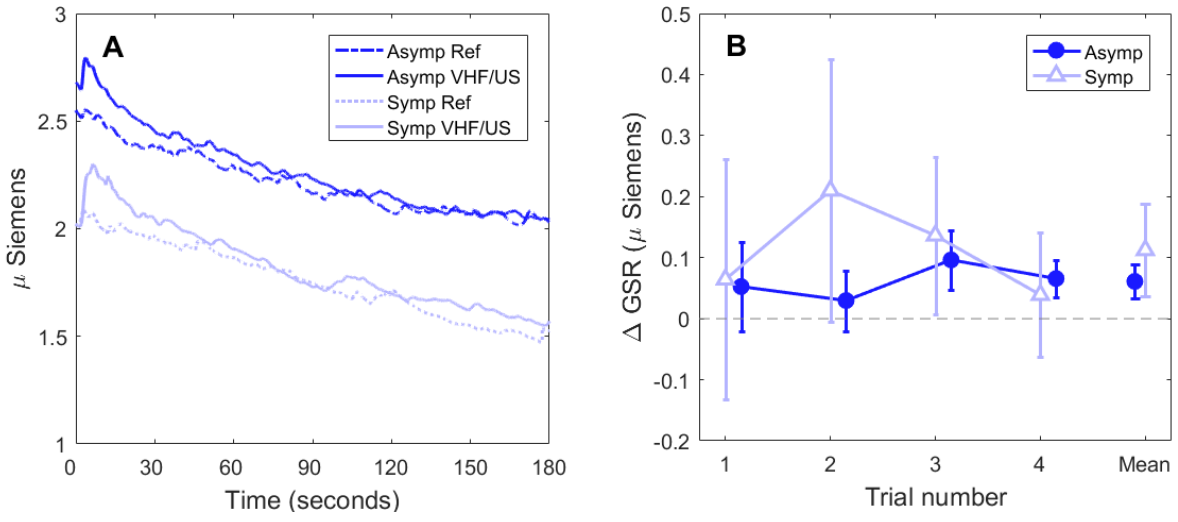

Figure 5 (color online): Galvanic skin response time series for the first exposure period (panel A) and mean difference in galvanic skin response between the VHFS/US and reference conditions for each trial (panel B) for the asymptomatic (32 participants) and symptomatic (10 participants) groups. A positive effect in panel B means that the galvanic skin response

was largest in the VHFS/US condition. Markers are horizontally offset for clarity. Error bars show the standard error of the mean across participants.

The GSR recordings give a time-history of skin conductance. The difference in skin conductance between the VHFS/US condition and the 1 kHz reference condition was calculated, averaged over all time points in each 3 minute trial, and then averaged over all participants in the two groups (Figure 5B). No statistically significant effect of stimulus condition or time period was found using a Friedman test.

Though the overall effect of stimulus condition on GSR was not statistically significant, the averaged time histories consistently showed a response within the VHFS/US condition in the first 20 seconds of the exposure, with the peak occurring ~2 seconds after the onset of the sound. This increased response was not present in the 1 kHz reference condition (Figure 5A).

A test of Spearman's rank order correlation showed that in the symptomatic group, the difference in the mean GSR between the VHFS/US and reference stimulus conditions was positively correlated with difference in subjective ratings of difficulty in concentration (r_2 = 0.72, p = 0.02) and anxiety (r_2 = 0.64, p = 0.04). No other statistically significant correlations were found. Note that the correlation statistics have not been corrected for multiple comparisons.

IV DISCUSSION

The aim of the present study was to investigate whether exposure to VHFS/US provokes adverse symptoms to a greater extent than exposure to a reference sound of 1 kHz. Symptoms were assessed using the three measures discussed below.

1. Subjective ratings

In both the symptomatic and asymptomatic groups, subjective ratings of overall discomfort were higher in the VHFS/US condition than the reference condition. In both groups the difference between conditions was modest, although the difference was larger in the symptomatic than the asymptomatic group. In the symptomatic group only, ratings of difficulty in concentrating and annoyance were also higher for the VHFS/US condition than the 1 kHz reference condition.

Direct comparisons of the subjective ratings with previous findings are difficult given the differences in rating scales and stimuli used. However, as a proportion of the full rating scale range, the effects of VHFS/US measured in the present study were considerably larger than those found by van Wieringen (2014) for similar symptoms, although their sound source had a much lower sound pressure level (up to 71 dB at 12-15 kHz for one exposure condition and up to 64.6 dB at 24-29 kHz for another). Conversely, the effects measured in the present study were much smaller than ratings for similar symptoms reported in Ueda *et al.* (2014) as a proportion of the full rating scale range, but their sound source produced much higher levels (up to 130 dB SPL) and was at a higher frequency than was used on average in this experiment (the peak energy for their device was at 20 kHz). As well as being lower in sound pressure level than some previous studies, the stimulus duration in the current study was much shorter than in many older studies (e.g. Acton and Carson, 1967), where reports were obtained from workers who experienced high sound levels for several hours each day. The levels and durations used in the present study were subject to ethical restrictions due to the different ethical considerations applied to experimental provocation studies compared to observational studies, and to the lack of clear evidence regarding the safety of exposure to VHFS/US. The effects measured may well have been larger had higher sound levels been used. Another ethical consideration, which may have reduced the effect size in this and

other lab-based studies, was the knowledge on the part of the participant that they could stop the stimulus at any time. This may have particularly reduced symptoms such as anxiety.

2. Sustained attention task

The absence of a measurable difference between the VHFS/US and reference conditions on the attention task, despite an increased rating of difficulty in concentrating for the VHFS/US condition in the symptomatic group, may have been due to the relatively short exposure time. Most previous work on the effects of noise on attention, as assessed through the CCPT or a similar task, has studied effects of much longer noise exposure (see Szalma and Hancock, 2011). Given that the participants in the symptomatic group reported struggling to concentrate more in the VHFS/US condition than the reference condition, it is possible that the stimulus duration was not long enough to produce measurable effects on the sustained attention task used. Despite attempts to make the attention task more difficult (section II.A.4), ceiling effects were still observed (a significant proportion of trials yielding 100 % correct or 0 % false alarms) which may have reduced the sensitivity of these performance metrics to more subtle effects of noise exposure on sustained attention. Because of this, some studies that use the CCPT have focused on response times as their primary outcome measure (e.g. Shalev *et al.*, 2011). However, as for the *d'* outcome measurement, no significant effect of condition was observed on the average response times and no evidence of a correlation between response times and subjective ratings was observed, even for the rating of ability to concentrate on the task.

3. Galvanic skin responses

No mean change in the GSR between conditions was found. However, in the symptomatic group, the difference in GSR response in the VHFS/US and reference conditions was strongly

positively correlated with the difference in subjective ratings for both difficulty in concentration and anxiety, suggesting that this variance did reflect differences in psychological state between the conditions. Furthermore, there was evidence of a peak in the GSR shortly after the onset of the VHFS/US stimulus. This was most marked in the first of the four trials (see Figure 5A), suggesting that this onset effect may habituate. The short delay between the onset of the stimulus (0 seconds) and the peak of the GSR effect (around 2 seconds) matches the time course previously measured for other effects on GSR in healthy volunteers, such as the use of a click sound or a concurrent electrical and acoustic pulse to create a startle response (see Kucera *et al.*, 2004 for a review). Further study is required to establish whether this onset response might represent a useful metric of increased sensitivity to VHFS/US.

V CONCLUSIONS

In both participant groups, exposure to VHFS/US led to greater subjective ratings of overall discomfort than exposure to a 1 kHz reference stimulus at the same level above the participant's detection threshold. In the symptomatic group, the overall discomfort was mainly associated with subjective ratings of difficulty in concentrating, and with annoyance. No significant physiological symptoms such as nausea, pain or tinnitus were found in either group for either stimulus condition.

No statistically significant effect on the performance for the sustained attention task was found in the VHFS/US condition compared to the 1 kHz condition. Time-averaged GSRs also showed no statistically significant difference between the two stimulus conditions when averaged over all four time points in the 3 minute exposure conditions. However, the GSR time-histories showed a clear increase in the VHFS/US condition above the reference, which

peaks at around 2 seconds after the stimulus onset for both participant groups. Additionally, for the symptomatic group, the mean difference in GSR for the VHFS/US and reference conditions was positively correlated with the difference in subjective ratings of difficulty concentrating and anxiety for the corresponding conditions.

It should be emphasized that the sound pressure levels and exposure durations were limited in this study by the application of the precautionary principle, and the researcher's duty of care to participants. These findings cannot be used to predict outcomes from exposures at higher sound pressure levels or longer durations.

Acknowledgements

This research was primarily funded by the Colt Foundation (ref: CF/03/15). Additional financial support from the European Metrology Programme of Innovation and Research (EMPIR, grant no. 15HLT03 EARII) for Timothy Leighton's and Ben Lineton's time is gratefully acknowledged. The EMPIR programme is co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme. The datasets generated and analysed during the current study are openly available from the University of Southampton repository at http://dx.doi.org/10.5258/SOTON/ [Note to editors: The policy of the Southampton University is that they grant a link for insertion once the paper is fully accepted to avoid their repository referring to papers that were not published].

APPENDIX 1: PARTICIPANT SCREENING AND GROUP ALLOCATION

Please fill in the following questionnaire to determine your eligibility for this experiment. If yes to any of the questions 2 to 15, please give additional details below. All data will be kept confidential.

1. What is your age in years? _____ Years

2. Do you have a hearing impairment that you are aware of? Yes / No

3. Do you have, or have you recently had any pain, tenderness, infections, discharge, surgery or bleeding in either of your ears? Yes / No

4. Do you have a history of frequent exposure to loud noise? Yes/No

5. Do you take any ototoxic medications (for e.g. aminoglycoside antibiotics, such as gentamicin)? Yes/No

6. Do you experience tinnitus (ringing, buzzing, whistling or any other sounds in either of your ears)? Yes / No

7. Do you suffer from hyperacusis (reduced tolerance and increased sensitivity to everyday sounds)? Yes / No

8. Have you been exposed to loud sounds in the past 24 hours? Yes / No

9. Do you expect to be exposed to loud sounds in the next 24 hours (e.g. visiting a night club, or concert, or taking part in an experiment involving high levels of sound presentation?) Yes / No

10. Are you colour blind? Yes / No

11. Do you suffer from epilepsy? Yes / No

12. Have you ingested a significant amount of caffeine in the last two hours e.g. drank more than one cup of tea or coffee, drunk an energy drink, or taken pro plus)? Yes / No

13. Have you taken recreational drugs in the last week? Yes / No

14. Have you drunk more than 6 units of alcohol (more than 2 pints of beer or 2 standard glasses of wine) in the last 24 hours? Yes / No

15. Have you undertaken strenuous physical or mental activity today? Yes / No

If you have answered "yes" to any of questions 2-15, please give further details below.

Details for Question number (s): _____:

Please turn over to complete the questionnaire overleaf

16. Have you ever had experienced unpleasant symptoms that you believe were caused by exposure to very high frequency sound or ultrasound? Yes / No

If you have answered "yes", please give further details below. If possible, include answers to the following:

a) What is the nature of these symptoms?
b) How long ago, approximately, did you first experience them?
c) In general, do/did the symptoms arise as soon the exposure began, or only after a period of time?
d) In general, how long did/do the symptoms endure after exposure has ceased?
e) What type of device/devices do you suspect have caused the symptoms, if known? (e.g. pest scarers).

Details (if "Yes"):

17. Do you have any expectation of the symptoms that you might experience during testing at high frequencies, or expectation of your hearing abilities at very-high frequencies (e.g. do you believe you can hear sounds that most people cannot)? Yes / No

Details (if "Yes"):

APPENDIX 2: SUBJECTIVE RATING QUESTIONNAIRE

Please rate your overall discomfort level

0	1	2	3	4	5	6	7	8	9	10
None										*Severe*

Over the last 4 minutes I experienced[1]…

Nausea

0	1	2	3	4	5	6	7	8	9	10
Not at all										*Severe*

Pain, pressure, or fullness in one or both ears

0	1	2	3	4	5	6	7	8	9	10
Not at all										*Severe*

Headache/ pain or pressure somewhere other than my ears

0	1	2	3	4	5	6	7	8	9	10
Not at all										*Severe*

Dizziness or light-headedness

0	1	2	3	4	5	6	7	8	9	10
Not at all										*Severe*

Tinnitus (ringing, buzzing, or other sounds in my ears)

0	1	2	3	4	5	6	7	8	9	10
Not at all										*Severe*

[1] This phrase was changed to "over the last hour I experienced…" for the screening phase.

Please turn over for more…

Anxiety

0	1	2	3	4	5	6	7	8	9	10
Not at all										*Severe*

Annoyance or irritation

0	1	2	3	4	5	6	7	8	9	10
Not at all										*Severe*

Fatigue

0	1	2	3	4	5	6	7	8	9	10
Not at all										*Severe*

Other symptoms

0	1	2	3	4	5	6	7	8	9	10
Not at all										*Severe*

REFERENCES

Acton, W. I., and Carson, M. B. (**1967**). "Auditory and subjective effects of airborne noise from industrial ultrasonic sources," Br J Ind Med **24**, 297-304.

Bacon, S. P., and Viemeister, N. F. (**1994**). "Intensity discrimination and increment detection at 16 kHz," Journal of the Acoustical Society of America **95**, 2616-2621.

Björk, E. A. (**1986**). "Laboratory annoyance and skin conductance responses to some natural sounds," Journal of Sound and Vibration **109**, 339-345.

International Standards Organization (**2003**). "Acoustics — Normal equal-loudness-level contours," BS-ISO-226.

British Broadcasting Corporation (**2017**). "Anger over Hamilton station 'mosquito' device," [online]. British Broadcasting Corporation. Available from: http://www.bbc.co.uk/news/uk-scotland-glasgow-west-40715995 [Accessed 02 August 2017].

British Society of Audiology (**2017**). " Recommended Procedure: Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking," British Society of Audiology, Reading UK.

Dimberg, U. (**1990**). "Facial electromyographic reactions and autonomic activity to auditory stimuli," Biological Psychology **31**, 137-147.

Ebelthite, H. (**2016**). "Could your health be ruined by noises you can't hear? Some gadgets emit silent ultra-high whines that may hurt you," [online]. The Daily Mail. Available from: http://www.dailymail.co.uk/femail/article-3527060/Could-health-ruined-noises-t-hear-gadgets-emit-silent-ultra-high-whines-hurt-you.html [Accessed 02 August 2017].

Fletcher, M. D. (**2016**). "Effects of ultrasound and very high-frequency sound," [online]. Facebook (public group). Available from: https://www.facebook.com/groups/249882768679771/ [Accessed 29 July 2017].

733 Formby, C., Gold, S.L., Keaser, M.L., Block, K.L. and Hawley M.L. (**2007**) “Secondary benefits from
734 tinnitus retraining therapy: Clinically significant increases in loudness discomfort level and
735 expansion of the auditory dynamic range,” Semin Hear **28**, 227–260.
736 Foxe, J. J., Morie, K. P., Laud, P. J., Rowson, M. J., de Bruin, E. A., and Kelly, S. P. (**2012**). "Assessing
737 the effects of caffeine and theanine on the maintenance of vigilance during a sustained
738 attention task," Neuropharmacology **62**, 2320-2327.
739 Gallagher, P. (**2016**). "Public being increasingly exposed to airborne ultrasound, research says,"
740 [Online]. The Independent. Available from:
741 http://www.independent.co.uk/news/science/public-being-increasingly-exposed-to-
742 airborne-ultrasound-research-says-a6822076.html [Accessed 29 July 2017].
743 Glorieux, C. (**2014**). "Undesirable effects as a result of short-term exposure to an ultrasonic repellent
744 device. Part I—acoustic measurements.," Assignment no. DGS/PB_PP/IVC/13026 from the
745 Federal Public Service in Belgium for Health, Food Chain Safety and Environment (28 July
746 2014).
747 Hardin, J. W. and Hilbe, J. M. (**2003**). “Generalized Estimating Equations,” Chapman and Hall, New
748 York.
749 Homma, S., Nakajima, Y., Toma, S., Itob, T. and Shibatab, T. (**1998**). “Intracerebral source localization
750 of mental process-related potentials elicited prior to mental sweating response in humans,”
751 Neurosci. Lett. **247**, 25–28.
752 Homma, S., Matsunami, K., Han, X.Y. and Deguchi, K. (**2001**). “Hippocampus in relation to mental
753 sweating response evoked by memory recall and mental calculation: a human
754 electroencephalography study with dipole tracing,” Neurosci. Lett. **305**, 1–4.
755 Khalfa, S. Peretz, I., Blondin J., and Manon, R. (**2002**). “Event-related skin conductance responses to
756 musical emotions in humans,” Neuroscience Letters **328**, 145–149.

Kobayashi, M., Tomioka, N., Ushiyama, Y. and Ohhashi, T. (**2003**) “Arithmetic calculation, deep inspiration or handgrip exercise-mediated pre-operational active palmar sweating responses in humans,” Auton Neurosci **104**, 58–65.

Kucera, P., Goldenberg, Z., and Kurca, E. (**2004**). "Sympathetic skin response: review of the method and its clinical use," Bratisl Lek Listy **105**, 108-116.

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.

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

Leighton, T. (**2016b**). "Health Effects of Ultrasound in Air (HEFUA)," [Online]. University of Southampton. Available from: https://sites.google.com/site/hefua2/ [Accessed 29 July 2017].

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. Royal Soc. A 473(2199), 20160828.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," The Journal of the Acoustical Society of America **49**, Suppl 2:467.

Maccà, I., Scapellato, M. L., Carrieri, M., Maso, S., Trevisan, A., and Bartolucci, G. B. (**2014**). "High-frequency hearing thresholds: effects of age, occupational ultrasound and noise exposure," International Archives of Occupational and Environmental Health **88**, 197-211.

Macmillan, N. A., and Creelman, D. C. (**2005**). *Detection Theory: A User's Guide* (Lawrence Erlbaum Associates).

Meeus, O.M., Spaepen, M., Ridder, D.D. and Heyning P.H. (**2010**). “Correlation between hyperacusis measurements in daily ENT practice,” Int J Audiol **49**, 7–13.

Shalev, L., Ben-Simon, A., Mevorach, C., Cohen, Y., and Tsal, Y. (**2011**). "Conjunctive Continuous Performance Task (CCPT)--a pure measure of sustained attention," Neuropsychologia **49**, 2584-2591.

783 Sheldrake, J., Diehl, P. U., and Schaette, R. (**2015**). "Audiometric characteristics of hyperacusis
784 patients," Front Neurol **6**, 105.
785 Stelmachowicz, P. G., Beauchaine, K. A., Kalberer, A., Kelly, W. J., and Jesteadt, W. (**1989**). "High-
786 frequency audiometry: test reliability and procedural considerations," The Journal of the
787 Acoustical Society of America **85**, 879-887.
788 Szalma, J. L., and Hancock, P. A. (**2011**). "Noise effects on human performance: a meta-analytic
789 synthesis," Psychol Bull **137**, 682-707.
790 Ueda, M., Ota, A., and Takahashi, H. (**2014**). "Investigation on high-frequency noise in public space. -
791 We tried noise abatement measures for displeasure people," In Proc. of the 7th Forum
792 Acusticum, Krakow, Poland, 7–12 September 2014.
793 van Wieringen, A. (**2014**). "Undesirable effects as a result of short-term exposure to an ultrasonic
794 repellent device. Part II—exposure of volunteers," Assignment no. DGS/PB_PP/IVC/13026
795 from the Federal Public Service in Belgium for Health, Food Chain Safety and Environment
796 (28 July 2014).