Di Battista, A., Price, A., Malkin, R., Drinkwater, B. W., Kuberka, P. M., & Jarrold, C. (2022). The effects of high-intensity 40 kHz ultrasound on cognitive function. *Applied Acoustics*, *200*, [109051]. https://doi.org/10.1016/j.apacoust.2022.109051

Applied Acoustics 200 (2022) 109051

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

The effects of high-intensity 40 kHz ultrasound on cognitive function

Andrew Di Battista [a,*], Adam Price [a], Rob Malkin [a], Bruce W. Drinkwater [b], Paula Kuberka [c], Christopher Jarrold [c]

[a] Ultraleap Ltd., The West Wing, Glass Wharf, Bristol BS2 0EL, UK
[b] Department of Mechanical Engineering, University of Bristol, UK
[c] School of Psychological Science, University of Bristol, UK

ARTICLE INFO

Article history:
Received 25 May 2022
Received in revised form 23 September 2022
Accepted 25 September 2022
Available online 5 October 2022

Keywords:
Ultrasound
Noise exposure
Public health
Cognitive function

ABSTRACT

In this study we investigate the effects of short-term exposure to high-intensity airborne ultrasound on cognitive function. Test participants ($n = 40$) were asked to perform a go/no-go task (GNG) and continuous performance test (CPT) under baseline (no noise) conditions. The tests were also presented under exposure to high-intensity ultrasonic noise from a custom built ultrasonic array (40 kHz tone, 120 dB SPL re 20 μPa). GNG and CPT test results were analysed using a Bayesian ANOVA statistical model. The results provided clear positive evidence for no effect of ultrasound exposure on performance in each task, whether measured in terms of participants' ability to select the correct response or their reaction times when responding correctly. Participants were also not better than chance at stating when the ultrasound had been presented. These findings indicate that ultrasound exposure of this intensity and frequency has no detectable effect on cognitive task performance.

1. Introduction

1.1. Motivation

Off-the-shelf ultrasonic transducers are often associated with positioning sensor applications, most notably vehicle reversing/automated parking systems in addition to robotics and drone control [1,2]. By assembling an *array* of transducers, it is possible to generate acoustic fields powerful enough to levitate small particles [3,4] and produce haptic touch-feedback in mid air [5]. Some of these devices have even been promoted for use as school science projects[6]; yet they have the potential to expose their users to relatively high levels of ultrasound [7].

International guidelines regarding ultrasonic airborne emissions have been summarised on several occasions [8–14] [1] These have lead to the recommendations of maximum permissible levels (MPLs) for airborne ultrasonic emissions at frequencies in the range of 20–100 kHz. The current MPLs are influenced by the fact that no deleterious effects on humans have been observed at sound pressure levels (SPLs) below 110 dB. Conversely, levels above 145 dB SPL are associated with temporary thresh old shifts (TTSs); a loss in hearing sensitivity that, if recurrent, could lead to long-term permanent damage to auditory function. Between these two SPL limits, various *subjective effects* have been documented [10]. Subjective effects include headache, nausea, stress etc. They are 'subjective' because the severity varies greatly from one individual to another and depend on the circumstances under which they are exposed.

The international guidelines are motivated by *occupational* health and safety; they consider the well-being of predominantly factory workers with an assumed exposure of 8 h/day. This leaves a gap with regards to implications for *consumer devices*. Consumer products are targeted to (and operated by) potentially all demographics and are generally intended for casual use. Nonetheless, if a device is shown to generate personal exposure levels between 110–145 dB, it can be expected that some individual users may experience subjective effects.

1.2. Previous work

In previous work by [15] it was hypothesized that nuisance effects from ultrasonic noise may reduce one's ability to concentrate on a task. A noise source (20 kHz tone at 15 dB below hearing threshold) was presented to subjects while carrying out a sustained attention to response task (SART) in which they had to respond to a set of sequential stimuli while withholding a response to one particular stimulus. Investigations into the presence of the *nocebo* effect (when negative expectations of the participant exac-

---

* Corresponding author.
E-mail addresses: andrew.di.battista@ultraleap.com (A. Di Battista), c.jarrold@bistol.ac.uk (C. Jarrold).

[1] [8] is a good starting point for the reader looking for an insight into the history and current state of ultrasound exposure guidelines

https://doi.org/10.1016/j.apacoust.2022.109051
0003-682X/© 2022 The Author(s). Published by Elsevier Ltd.

erbate negative outcomes) were also carried out. Additionally, subjects were asked to complete a noise survey/ questionnaire, following guidelines defined by ICBEN [16]. Ultrasound was not found to provoke any symptoms, however, there was evidence of small *nocebo* effects. [15] warns against making generalized conclusions from the results due to these confounding factors.

Recent work by [17] exposed test subjects to a 40 kHz, $110-120 dBSPL$ tone for up to 30 min. Measurement of auditory function did not find evidence of threshold shifts or changes of behavioural or electrophysiological subclinical measures. A two-alternative forced-choice (2AFC) was implemented to detect if participants could perceive when the ultrasound was present. Though they scored no better than chance, the results were confounded by audible noise coming from the ultrasonic array (which required a masking noise to be used). Results were limited by a relatively small sample size ($n = 18$), but suggest any effects (if present) would be small.

1.3. Measuring distraction/cognitive function

The SART procedure employed by [15] is designed to test participants' sustained attention: the ability to maintain focus on a task for a prolonged period [18]. A SART is a specific example of the broader class of go/no-go tasks in which participants are presented with a randomly ordered sequence of stimuli and have to make a response ('go' trials) to one or more 'target' stimuli and withhold a response ('no-go' trials) to other less frequent 'non-target' stimuli [19]. For example, in a typical SART participants might be presented with a sequence that samples from the digits 1–9 and be told to respond with a single response (e.g., pressing the space bar) to all of the digits with the exception of the number '3'. The task therefore involves responding on the majority of trials (e.g., 8 in 9 trials or 89% of the time) thereby building a pre-potent tendency to respond that needs to be inhibited on the infrequent occasions when a non-target appears. SART or go/no-go tasks therefore require inhibition of this pre-potent response in addition to tapping participants' sustained attention [20]. Consequently, they provide a laboratory analogue of real-world behaviours that require an infrequent adaptation of a routine behaviour, for example stopping oneself from making a habitual manoeuvre when driving in response to a sudden change in the environment.

Go/no-go tasks (GNG) can be contrasted with continuous performance tests (CPTs); CPTs similarly involve responding to some stimuli but not others, however they include many more non-target than target trials to ensure that the participant only makes a response infrequently [21]. They therefore do not build up a pre-potent tendency to respond but rather test the participant's ability to stay on task and remain alert to the possibility of a target appearing. As a result, CPTs are arguably more direct measures of sustained attention than go/no-go tasks. They provide a good parallel to real-world situations where sustained vigilance is required during potentially monotonous and routine tasks, such as monitoring airport baggage scans for the rare cases where an inappropriate item is present. The current study took advantage of the parallels and differences between these two classes of tasks and presented both a GNG task and a CPT to participants, with each task being presented both with and without ultrasound exposure. This allowed us to make a comprehensive test of any effect of exposure on performance on two accepted measures of attentional control, one with a clear inhibitory component and one with a vigilance aspect [22], thereby substantially extending any previous work in this area.

1.4. Summary

In this experiment, we make use of a 120 dB SPL 40 kHz tone to examine the potential for inducing subjective effects in humans. This noise source was chosen because it is consistent with exposure levels of certain ultrasonic based consumers devices [23]. Further, it is not associated with any harmful effects to hearing (as mentioned above) but is sufficiently over the 110 dB SPL mark to begin exploring the possibility of inducing some form of subjective response. It is of particular interest to see if this response is sufficient to impair cognitive performance.

The following sections of this paper describe the experimental protocol (Section 2.1) and parameters used for the GNG/CPT tests (Section 2.2). Section 2.3 provides a description of the hardware design for a custom built ultrasonic noise source. Acoustic simulations and calibration measurements of the noise source are described in Section 2.5. Statistical analysis methods are introduced and presentation of the results are described in Section 2.6 and Section 3 respectively.

2. Methods

2.1. Protocol overview

Ethical approval for this study was provided by the University of Bristol School of Psychological Science Human Research Ethics Committee (approval code 260421116004). Prior to testing a full risk assessment was carried out that included a careful analysis of any required Covid-19-related mitigations. All testing was conducted in line with this risk assessment and with government and institutional health and safety policies in operation at the time. Test subjects were required to complete a series of computer based cognitive tasks under both *experimental* (exposed to ultrasound) and *control* conditions. The experiment was implemented using a double-blinded design in which the control/exposed state was randomly allocated in custom software. The testing session was approximately 45 min long and involved two phases. First, as part of a screening process, participants' hearing was checked using a standardised audiometry assessment. Participants then proceeded to the central phase of the session involving the cognitive tasks. After each cognitive task, participants were asked if they could perceive when the array was actually switched on.

2.1.1. Participants

Data from a final sample of 40 participants are reported. Forty-three individuals took part in the study, but one participant was excluded for failing the initial hearing check (see Section 2.1.2 and a further two were excluded following data collection and prior to data analysis because they were older than the specified ages for inclusion (18 to 26 years of age). The remaining 40 participants (31 female, 9 male) had a mean age of 20 years (SD = 2; range 18–26).

Participants were recruited and tested individually in a small (3.5 m x 3 m) room at the University of Bristol. Full, informed consent was acquired from every participant prior to the study commencing. Participants took part either for course credits or were reimbursed €10 for their time.

2.1.2. Audiometry screening process

Participants were advised upon recruitment to avoid loud noises and activities (e.g. music concerts) 24 h prior to their scheduled testing session. Compliance with this instruction was checked at the start of the session and no participant reported having failed to meet this requirement. As part of the screening process, each participant completed a Pure Tone Audiometry (PTA) test using a clinically validated iPad based app (Shoebox Audiometry) with calibrated DD450 (Radioear) headphones to confirm they had no significant pre-existing hearing impairment [24,25]. This was established by testing across frequencies 250, 500,1 k,2 k,4 k,8 k

Hz. PTA scores were generated by averaging hearing loss (dB HL) values over the 200 Hz - 8 kHz range, for each ear. A threshold of $>$ 20 dB HL was used to define hearing impairment and formed the exclusion criteria. This occurred in one instance and that individual was advised to seek further testing and information from a medical practitioner. In addition to the standard PTA frequencies, the extended frequency of 16 kHz was also presented to participants during the hearing test [26]. It has been hypothesised that younger individuals could be at greater risk to ultrasonic noise due to their inherent greater sensitivity to higher frequencies [10]. Given the relatively young age group of test subjects, it was expected that a reasonable sensitivity to this higher frequency would be apparent; this was also confirmed by the test results (mean 18 dB HL @ 16 kHz, SD = 12, for both left and right ears).

2.2. Experimental tasks

The experimental component of the study involved two related cognitive tasks, GNG and a CPT. Each task consisted of two blocks of 270 trials, with the ultrasound exposure presented randomly in one of the two blocks (Fig. 1).

The order in which the two tasks were experienced by the participants was also randomly determined by the experimental software. Participants in Group A ($n = 20$) performed the two blocks of the GNG first, and the two blocks of the CPT second; participants in Group B ($n = 20$) completed the tasks in the reverse order. Each task involved the successive presentation of stimuli on a computer monitor that was placed approximately 50–60 cm away from the participant with stimuli appearing in the centre of their field of vision. The ultrasound device was situated just above the top of the monitor and in line with the participant's eye level (as depicted in Fig. 3). Each block of each task began with 18 practice trials in which the instructions for that particular block were explained. Additionally, participants were given a 30 second pause after every 90 trials (i.e. 2 breaks within each 270 trial block).

The GNG required the participant to press the spacebar on a computer keyboard whenever a stimulus appeared, with the exception of a rare target stimulus. In the first block of 270 trials the stimuli were the numbers 1 to 9, and the participant was required to withhold a response to the number '3'. Stimuli appeared equally often meaning that the target letter appeared on 30 of 270 trials (11.1% of trials). The second block of the GNG had a comparable structure but employed nine letters 'a' to 'i'. Participants were instructed to withhold their response whenever the letter 'c' appeared (again, 11.1% of trials).

The CPT was similar in structure but required the opposite response frequency (i.e., responding infrequently rather than frequently). The first block of the test involved nine shapes (e.g., diamond, triangle, cross, heart) all presented in grey. Participants were required to press the spacebar whenever a single target shape (the star) appeared. All stimuli occurred equally often meaning that in this task participants made a response on 11.1% of trials. The second block of the continuous performance test presented nine equally sized and shaped colour patches (e.g., purple, brown, pink, orange). Participants had to make a response only when the blue colour patch appeared (again, 11.1% of trials).

Stimuli in each task were shown for a duration of 250 ms. On any trial of each task a non-response was recorded, and the task moved on if participants failed to respond within 1 second of the stimulus onset (i.e., both a correct non-response to a stimulus that should not have been responded to and an incorrect non-response to a stimulus that should have been responded to was recorded after this time period). Accuracy of each response or non-response was recorded, as was the reaction time (RT) of any (correct or incorrect) response. In addition, at the end of each task participants were asked to state which of the two blocks of that task they thought the ultrasound exposure had occurred in.

2.3. Hardware

The design requirements for the ultrasonic noise source were as follows:

- deliver approximate 120 dB SPL at a distance of approximately 50–60 cm (the distance from the computer screen to the participant's head).
- produce an acoustic field to be as uniform as possible over as wide an area as possible (to account for variability in participant head movement, height, etc.).
- be software controllable in order to integrate into the GNG/CPT test protocol.
- no cues (audible or visible) from the array should alert the participant of its state (on/off).

The array was constructed from 9 Murata MA40S4S transducers. These were arranged into a subset of a sunflower (Fermat spiral) pattern (see A) and embedded into a Perspex frame (the array footprint was approximately a 7 cm × 7 cm region). A raspberry Pi 3B + coupled with a L298N amplifier circuit was used to drive the array with a DC power source (12 V) (B, Fig. 7). The open-source pigpio library [2] allows for control of the Pi's hardware generated pulse-width modulation (PWM) signal. Initial testing found that a sudden switch (on or off) of the array caused a subtle but discernible 'click' sound. Consequently, a ramp function was introduced to run the PWM duty cycle from 0–50% over a (default) 1 s upon start-up and shut-down of the ultrasonic noise source. This step effectively allowed for 'quiet' operation by reducing the unwanted transient response of the transducers.

2.3.1. Array function checks

An ammeter was connected in series with the array power source and the LED screen would display the current draw from the device when it was switched on. The LED display was hidden from view from participants (and the researcher) during testing so as not to provide a visual cue that the array was operating. However, pre-test checks could be carried out and the ammeter reading could be monitored to ensure the array still drew power consistently. Additionally, a final check of free-field acoustic measurements carried out prior to and at the end of all experiments showed no significant deviation in the output SPL.

2.4. Software

The cognitive tasks were implemented in Javascript and ran in a web browser on the Pi's desktop interface. PWM control of the Pi was accessed via a local webserver (implemented in C). All RTs and test scores were recorded and stored initially as JSON files and exported to an appropriate csv format for subsequent analysis.

2.5. Ultrasound array testing and calibration

2.5.1. Simulation

To simulate the acoustic pressure, we used a simple Huygen's model implemented in Matlab (we have included the acoustic model in the Supplementary Materials). The model is a linear, frequency domain representation of an acoustic field generated by discreet point like acoustic sources (the elements of our array). As each of the acoustic elements has an angular dependant amplitude we use a weighting function (also known as a directivity

---

[2] (https://abyz.me.uk/rpi/pigpio/)

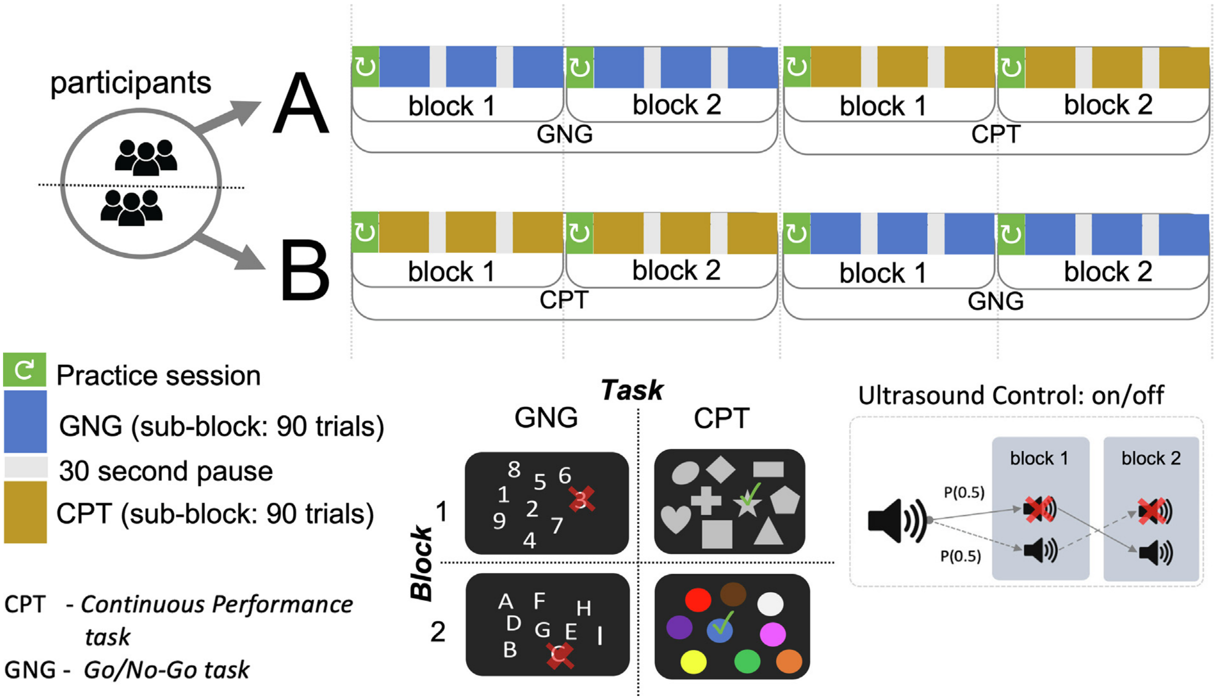

Fig. 1. Experimental Tasks Participants were randomly assigned to two groups: Group A (n = 20) and Group B (n = 20). Group A began with the GNG tasks before proceeding to the CPT; Group B did the opposite. In each task, the ultrasonic noise source was randomly assigned to either the first or second block. Each 270 trial block was split into three sub-blocks of 90 trials, with a 30 s pause/break for participants from the task. When a group had completed each of their GNG/CPT tasks, they were asked if the could determine in which block the ultrasound was present.

function). The experimental directivity of the transducers used are found to be very accurately represented by the weighting function. The complex pressure, $p_j(\alpha)$, at a given point in space, α, produced by the j^{th} emitter located at β_j is given by Eq. 1.

$$p_j(\alpha) = \frac{A_j}{d(\alpha,\beta_j)} \frac{2J_1(ka\sin\theta)}{(ka\sin\theta)} e^{-i(\phi_j + kd(\alpha,\beta_j))} \tag{1}$$

Where A_j is the amplitude of the j^{th} element at the transducer surface, $d(\alpha,\beta_j)$ is the Euclidean distance between points α and β_j,J_1 is the Bessel function of the first kind (accounting for directivity), k is the wave-number $(2\pi/\lambda)$, λ is the wavelength, a is the transducer radius, θ is the polar angle between points α and β_j, i is the root of negative 1 and ϕ_j is the phase delay applied to the j^{th} element. The value of A_j was determined by matching the experimentally measured acoustic pressure from a single transducer driven at the same voltage as used in the physical array. The axial symmetry of the transducers directivity along its acoustic axis makes for an efficient simulation. It should be noted that as this model is linear it is incapable of capturing very high amplitude acoustic phenomena (e.g. wave-steepening, harmonic generation, heating). However from experimental experience these non-linear effects only occur for well focussed acoustic fields where the acoustic pressure exceeds 145 dB SPL.

The complex total pressure, P_T, at point α generated by N transducers is given by Eq. 2.

$$P_T = \sum_{j=1}^{N} p_j(\alpha) \tag{2}$$

When displaying the instantaneous acoustic pressure we either take the real part of the complex sum or when displaying the time-averaged acoustic pressure we take the absolute value of the sum.

2.5.2. Measurement

The array was characterised in an anechoic chamber, that was ISO 3745 compliant up to 40 kHz. Scans of the acoustic field were taken at 40, 50 and 60 cm away from the centre of the array, whilst all ultrasonic transducers were driven in phase with a 50% duty. The scans covered an area of 0.09 m^2 dimensions 0.3 m x 0.3 m, with the plane of the scan parallel to the array's surface.

A Brüel & Kjaer (B&K) Microphone Unit Type 4138-L-006 was attached to a robot arm with the data acquired with a NI PXIe-4310 analog input module at 390Ksamples/s. A comparison of simulated and measurement results are depicted in Fig. 2.

Uncertainties in measurement data are represented in Fig. 2 as error bars. These were calculated from systematic and random errors, the largest source of systematic error being microphone calibration and microphone-source separation error. Measurements of 0.01 s were taken 10,000 times at each location quantifying repeatability.

2.5.3. HATS in situ measurement

In-situ measurements were carried out using a B&K HATS (Human & Torso Simulator) with a Type 4191 microphone (Fig. 3). A data acquisition system consisting of a digital oscilloscope (DrDAQ,Pico Technology) collected real-time SPL data filtered through a digital 1/3-octave band-pass filter centred at 40 kHz. A digital equalization filter was also implemented to convert the free-field response of the microphone to the pressure-field response.

It was not practicable to carry out all possibe permutations of HATS orientations. The goal was to examine and verify the ~ 120 dB SPL output of the array. HATS free-field calibration measurements for 40 kHz ultrasound (see C (Table 4) indicate that exposure *at the ear* can vary greatly due to the shadowing effects of the human pinna (outer ear). Table 1 lists HATS measurements from two orientations (x2 for each ear). They are consistent with the calibration carried out on the HATS (namely an approximate 7 $-$10 dB

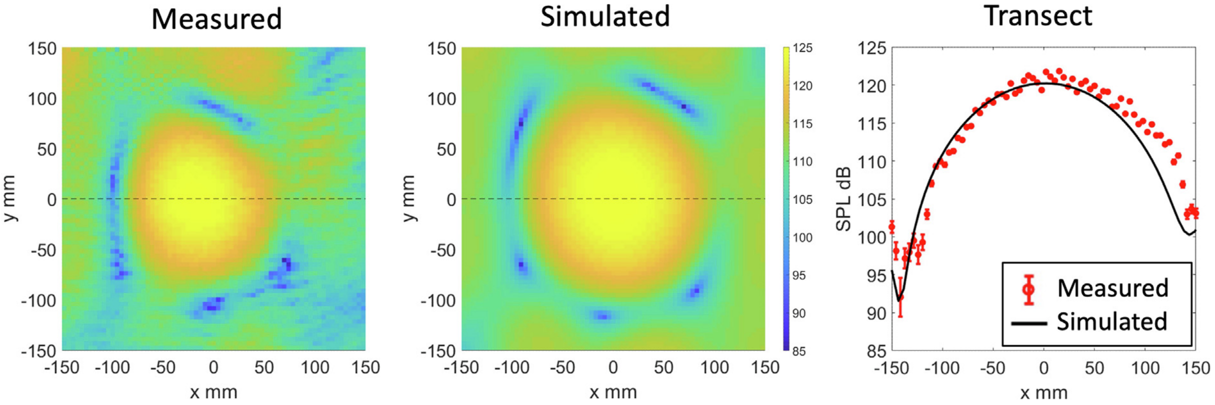

Fig. 2. Simulation and Measurement 2D scan at 60 cm from array (resolution 4 mm). The array was driven at 12 V. A transection (dotted line in images) is plotted separately (right-most image).

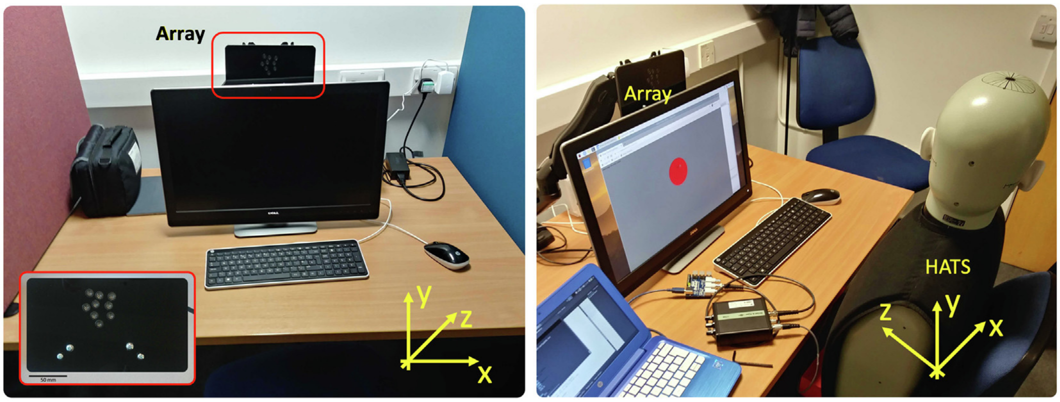

Fig. 3. Experimental setup & **HATS measurements** The experiment involved a computer based cognitive task(s). (left) A custom built ultrasound array was placed at the top of the screen, similar to a (classic) webcam. The extra hardware/ equipment (not shown) was hidden from the participants' view behind the screen. (right) HATS setup for in situ measurements of ultrasound exposure *at the ear*. Measurements we carried out in both left and right ears at 0° (i.e. head-on) and 90° incidence (depicted). The HATS, was positioned so that its face aligned to the height of the array (through an adjustable chair). Real/ human participants were also asked to position themselves in a similar fashion; the reflective Perspex of the array frame helpfully serving as a mirror to facilitate this task.

drop from 90 to 0 degree incidence). Moreover, the near parity between left and right ear measurements were expected from (and required of) the array design.

2.6. Statistical analysis – Bayesian ANOVA

Consider that the outcome measures of this experiment i.e. test scores and reaction times, may be affected by several factors. The most obvious and important consideration in this work is the effect of ultrasound exposure (denoted as C - for the *condition* under which participants were tested). However, cognitive test performance can also be affected by the test order and/or duration which can lead to effects associated with participant fatigue or boredom. Thus, we also examine the effects of group allocation (G), whether ultrasound exposure occurred in the first or second block of the task (E_b) and whether there was any difference between the block types within a task i.e. letters vs numbers and shape vs colors (B_t). A statistical model of the results may consist of some or all of these factors (including interaction between them). The aim of a traditional ANOVA study is to determine which model(s) (and associated factors) 'best' fit the data.

Table 1
Acoustic HATS measurement in situ. Units are in dB SPL averaged (Mean ± SD) over a 40 s acquisition.

Left(0°)	Left(90°)	Right(0°)	Right(90°)
115.2 ± 0.6	124.1 ± 0.5	116.7 ± 0.2	123.2 ± 0.5

All of the experimental results were analysed using Bayesian Analysis of Variance (BANOVA) ([27]) with the open-source software JASP 0.16.3 ([28]). This produced as an output several Bayes Factors. The Bayes Factor BF_{10} is defined as the ratio of probability of the observed data, D, under two models, e.g. M_0 and M_1 (Eq. 3). Bayes Factors above 3 can be taken as providing substantial evidence for preferring M_1 over M_0 [29]. Thus, if M_0 denotes the null-hypothesis, a $BF_{10} > 3$ would be cause to reject it.

$$BF_{10} = \frac{P(D|M_1)}{P(D|M_0)} \tag{3}$$

Exhaustive pairwise comparisons of models, can be used to determine the 'best' fitting one. However, this best fitting model will generally contain several factors; how do we determine which one(s) are the most important? By combining (i.e. summing the probabilities of) all models that depend on a particular *factor* and compare that to all those that do not, [3] it is possible to separate out the contribution from a particular effect. This results in another Bayes Factor, (BF_{incl}) which can be used as a grounds to include or exclude

---

[3] With respects to interaction effects, only models that follow the *principle of marginality* are considered

Table 2
Analysis of Effects – Section 3.3 and 3.4.

Effect	GNG d'		CPT d'		GNG RT		CPT RT	
	BF_{incl}	BF_{excl}	BF_{incl}	BF_{excl}	BF_{incl}	BF_{excl}	BF_{incl}	BF_{excl}
B_t	–	2.25	–	3.56	–	3.34	78.92	–
G	–	1.66	–	2.95	–	2.08	1.39	–
E_b	–	2.15	–	2.93	–	1.80	–	1.93
$B_t * G$	–	1.49	–	2.94	–	3.00	–	1.05
$B_t * E_b$	–	3.05	–	1.58	–	2.94	–	3.46
$G * E_b$	–	1.83	–	2.36	–	1.27	–	1.29
$B_t * G * E_b$	1.69	–	–	2.12	–	2.27	1.12	–

B_t - block type, G - group, E_b - exposure block. Note: either BF_{inc} or BF_{excl} is listed depending on which was > 1

Table 3
Analysis of Effects – Section 3.6 and 3.7.

Effect	GNG d'		CPT d'		GNG RT		CPT RT	
	BF_{incl}	BF_{excl}	BF_{incl}	BF_{excl}	BF_{incl}	BF_{excl}	BF_{incl}	BF_{excl}
C	–	4.08	–	4.02	–	3.62	–	4.62
G	–	1.07	–	1.71	–	2.09	2.77	–
E_b	–	2.44	–	1.99	–	1.77	–	1.54
$C * G$	1.19	–	–	1.43	–	2.85	2.61	–
$C * E_b$	–	3.17	–	2.04	–	2.58	4.79	–
$G * E_b$	–	1.86	–	1.50	–	1.30	–	1.25
$C * G * E_b$	2.26	–	–	2.56	–	2.19	–	1.04

C – condition(ultrasound exposure), G – group, E_b – exposure block. Note: either BF_{inc} or BF_{excl} is listed depending on which was > 1

($BF_{excl} = (BF_{incl})^{-1}$) that factor from consideration. For example, a $BF_{incl} = 4$ for the factor C implies that models that included C are 4 times better odds of explaining the data than those models that omit C. Conversely, a $BF_{incl} = 0.2$ would not be strong evidence to consider C and one could use $BF_{excl} = 1/0.2 = 5$ to say that models that excluded C had 5 times better odds.

The combination of BF_{10} and BF_{incl} allow inference on the best models and most relevant factors. Bayes Factors for the inclusion or exclusion of any effects and interactions were calculated using the 'across matched models' option in JASP [30].

2.7. Test scores parametrization – d-prime

Test accuracy scores were parametrized using signal-detection parameter d-prime ($d\prime$) which is calculated as $z(hits) - z(falsealarms)$, where $z()$ is the inverse cumulative normal distribution function [31]. In cases where either hit rate or false alarm rate was 1 or 0 respectively these values were subject to a loglinear correction [32].[4]

3. Results

3.1. Subjective detection of the ultrasound exposure

Of the 80 opportunities that participants had to say which block of each task the ultrasound exposure occurred in, participants stated that they were unable to determine this or were unwilling to guess on 34 occasions. Of the remaining 46 choices, 18 were correct (39% vs. chance level accuracy of 50%). Neither the Score nor Wald test indicate this as a significant result, (single-tail p-value of 0.07 and 0.065 respectively). These p-values both increase to 0.13 if the 34 omissions are included and weighted at 50% chance.

3.2. Summary of test scores and reaction times

Participants were highly accurate on each block of the two tasks. Accuracy levels (averaged across cases when the ultrasound exposure occurred or not) were 96.2% and 95.9% for the number and letter blocks respectively of the GNG task. They were 99.6% and 99.7% for the shape and colour blocks respectively of the continuous performance test. Given these high levels of accuracy, subsequent analyses focussed on a more sensitive signal-detection based measure of accuracy ($d\prime$).

Box-plots of experimental results in terms of $d\prime$ and RTs are depicted in Fig. 4. Data can be compared in terms of exposure and control condition both within and between factors such as group A & B and test block.

3.3. Block differences in accuracy

Fig. 4(a), (b) present average $d\prime$ values for each of the blocks of the two tasks by group (Group A vs. B) and *'exposure block'*, which is the block in which the ultrasound exposure took place (i.e., 1 = first block, which would be the number block of the GNG task or the shape block of the continuous performance test).

These data were analysed by a pair of Bayesian ANOVAs using JASP (Table 2). One compared the number and letter blocks of the GNG task, the other compared the shape and colour blocks of the continuous performance test. Each had the within-participant factor of block type (number or letter/ shape or colour), group (A or B), and exposure block (1 or 2). The best fitting model in each case was the null model. There was anecdotal evidence to support the exclusion of the main effect of block type in the analysis of the GNG task, $BF_{excl} = 2.25$, and substantial evidence against the need to include it in the continuous performance test analysis, $BF_{excl} = 3.56$

3.4. Block differences in reaction times

Fig. 4(c), (d) present reaction time data (for correct responses only) for each block of the two tasks, again broken down by group

---

[4] Specifically, values of 0 or 1 were adjusted by adding either 0.11 or 0.89 to the number of hits or false alarms (0.11 for conditions where the type of trial occurred 11% of the time and 0.89 for conditions where the type of trial occurred 89% of the time) and then dividing by n + 1 where n is the number of trials of that type.

Table 4
40 kHz calibration corrective gain for HATS model, compared to free-field measurement at different HATS orientations.

Azimuth (deg.)	Gain (dB)
0	−10
90	−3
75	+8

and exposure block. These data were subjected to two Bayesian ANOVAs, one for each task, with the factors of block type, group, and exposure block (Table 2).

The null model was the preferred model for the analysis of the GNG task reaction time data, with substantial evidence against the need to include the main effect of block type, $BF_{excl} = 3.34$. In the analysis of the CPT the best fitting model included the main effect of block type and group, and was preferred over the null model by a Bayes factor (BF_{10}) of 155.85. However, analysis of effects showed only meaningful (very strong) evidence of the inclusion of the main effect of block type, $BF_{incl} = 78.92$, with the evidence for the inclusion of the main effect of group being only anecdotal, $BF_{incl} = 1.39$. The main effect of block type was due to faster responses in the shape block than in the colour block.

3.5. Interim summary of block differences

The above analyses provide positive evidence to indicate that the two blocks of the GNG task were well matched for difficulty. In contrast, the shape block of the CPT was somewhat easier than the colour block, as evidenced by the faster responses on the former. However, although these preceding analyses did account for which block the ultrasound exposure occurred in, with interactions with this factor never appearing in a preferred model, they do not provide a direct test of the effect of ultrasound exposure. This was tested in the following analyses that contrasted the 'control' vs. 'ultrasound exposed' conditions of each task regardless of the specific block in which that condition occurred.

3.6. Exposure differences in accuracy

Fig. 4(a), (b) summarise d' values for condition (control vs. ultrasound exposed) by group and exposure block. These data were

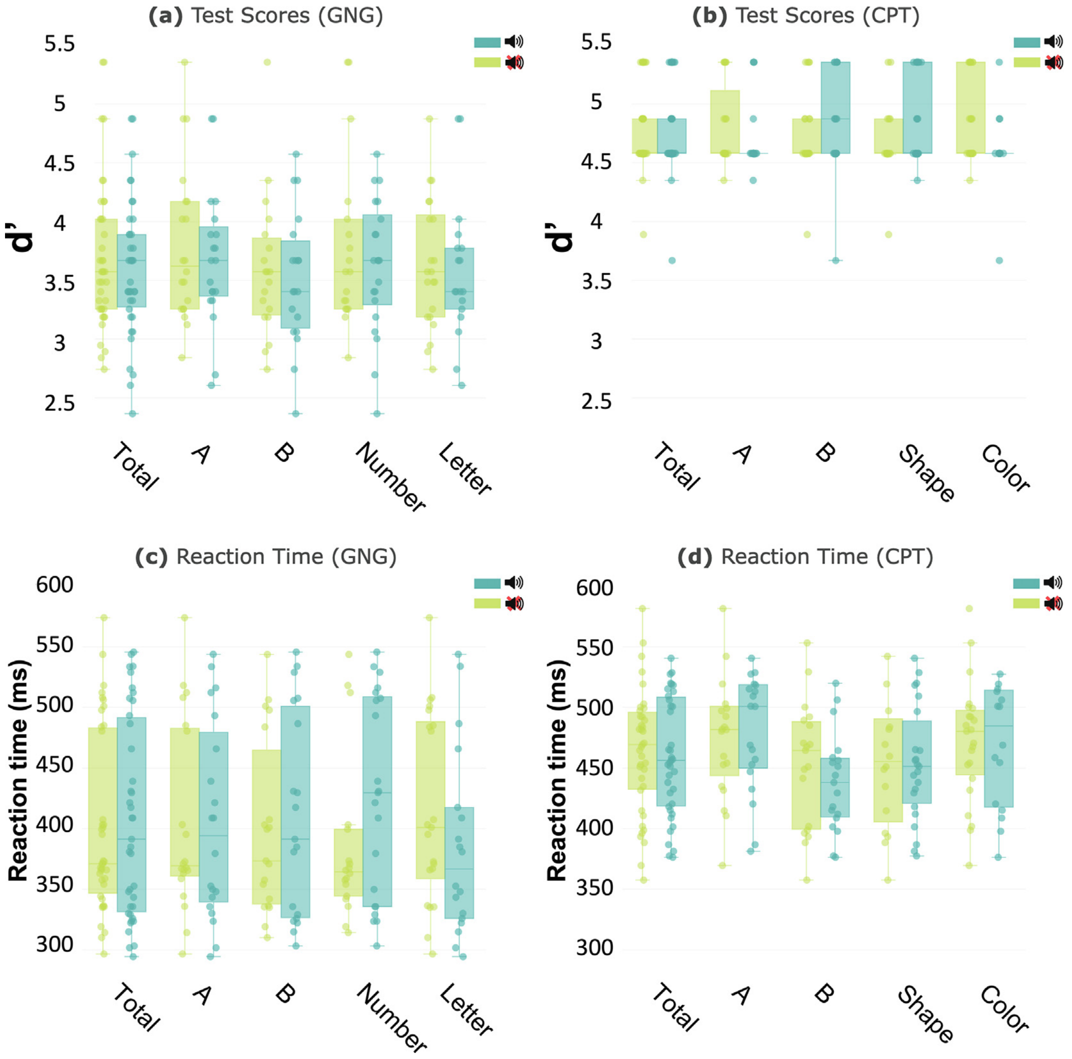

Fig. 4. Results - d' test scores and *reaction times* by task, block type, group and exposure block. For each task type, both d' scores and reaction times are arranged under *control* and *exposure* conditions for multiple groupings: *Total* refers to the cumulative data points across all groups and test block types. *A* and *B* are the cumulative data for each group respectively, across all block types. *Number*, *Letter*, *Shape* and *Color* are the cumulative data for each block type, across all groups. Data pertaining to each task type is arranged vertically with GNG data in plots (a) and (c), and CPT data in (b) and (d).

again subjected to two Bayesian ANOVAs, one for each task, which in this case had the within-participants repeated measures factor of condition and the between-participants factors of group and exposure block (Table 3). The null model emerged as the preferred model in each analysis, with substantial positive evidence against the need to include the main effect of condition in both the GNG task, $BF_{excl} = 4.08$, and the continuous performance test, $BF_{excl} = 4.02$.

3.7. Exposure differences in reaction time

Fig. 4(c), (d) provide details of correct RTs for the control and exposed conditions of each task, split by group and exposure block. Two Bayesian ANOVAs were conducted on these data, one for each task, including each of these factors (Table 3).

The preferred model of the GNG task reaction time data was the null model. The evidence against the need to include the main effect of condition was substantial, $BF_{excl} = 3.62$. The best fitting model of the continuous performance test data included all three main effects and their two-way interactions (but not the three-way interaction) and was preferred over the null model by a Bayes factor (BF_{10}) of 5.45. However, the only model term for which there was substantial rather than anecdotal evidence to support its inclusion was the interaction between condition and exposure block, $BF_{incl} = 4.79$. There was substantial positive evidence against the need to include the main effect of condition, $BF_{excl} = 4.62$. The interaction between condition and exposure block is plotted in Fig. 5.

Post-hoc analysis of the interaction reflected in Fig. 5 first examined the effect of exposure block at the two levels of condition separately. Each of these Bayesian ANOVAs included the factors of exposure block and group. The analysis of the control condition data indicated that the null model was the best model of the data, with substantial evidence against the need to include the main effect of exposure block, $BF_{excl} = 3.21$. The best model from the analysis of the exposed condition data included just the main effect of group and was preferred over the null model by a Bayes Factor of 8.23. There was substantial evidence for the need to include the main effect of group, $BF_{incl} = 9.12$, but no evidence to support the inclusion of the main effect of exposure block, $BF_{incl} = 0.72$.

A further pair of post hoc Bayesian ANOVAs decomposed the interaction shown in Fig. 5 by examining the effect of condition separately for each exposure block; these analyses therefore included the factors of condition and group. The null model provided the best fit to the data for participants who received the exposure in the first block, with no evidence to support the need to include the main effect of condition, $BF_{incl} = 0.73$. The corresponding analysis of the data from participants who received the exposure in the second block produced a best fitting model that included both main effects of condition and group (preferred over the null model by a Bayes Factor of 5.73). However, the evidence for the need to include each specific effect was anecdotal, $BF_{incl} = 2.20$ for condition, $BF_{incl} = 2.53$ for group.

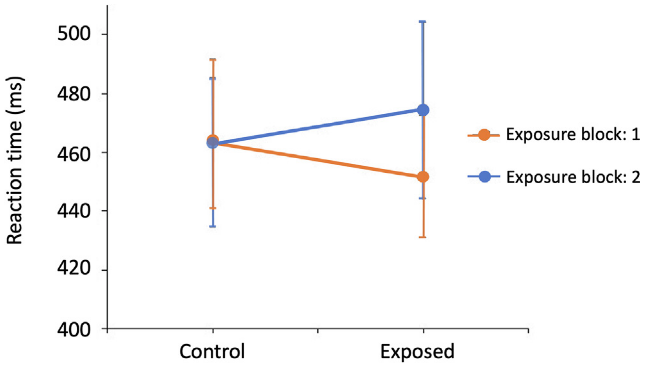

Fig. 5. Interaction The interaction between condition and exposure block (1^{st} block vs. 2^{nd} block) on correct reaction times in the continuous performance test (95% confidence intervals shown).

4. Discussion

The aim of the current study was to examine whether 40 kHz ultrasound exposure at ~ 120 dB SPL had any meaningful effect on cognitive task performance. In line with previous work, this question was assessed using a GNG task that required participants to inhibit a pre-potent response action on a minority of trials. However, in an extension of previous studies this was complemented by a CPT that required sustained attention in order to make a motor response to an infrequent target stimulus. Although these two tasks therefore have parallels to one another, and share a common task structure, they tap separable and complementary cognitive components of everyday functioning.

Participants exhibited high levels of accuracy on both tasks. While this might raise concerns about ceiling levels of performance that could mask any experimental effect, this concern was mitigated by the use of a d' measure of performance that combines hit rates and false alarms. Although Fig. 4(a), (b) show that d' values were also high, they are nevertheless associated with a reasonable degree of variation. Indeed, although a direct comparison between the GNG and continuous performance test was not a focus of our analyses, a comparison of Fig. 4a) and Fig. 4(b) clearly indicates that the GNG task was less easy for participants than the continuous performance test. Another consequence of generally high levels of task accuracy is that the vast majority of RTs recorded during each task were associated with correct responses. This makes reaction time data readily interpretable, and RTs are not subject to floor or ceiling effects in the way that accuracy data can be. In the current task participants did have a fixed time window of 1 s in which to respond, but, as Fig. 4(c), (d) show, average RTs fell well below that upper limit. In addition, any failure to respond within 1 s on a trial when a response was required would have been coded as an error and, as already noted, error rates were very low.

A further strength of the current study is that, broadly speaking, the two component blocks of each task were of equal difficulty. The comparison of the number and letter blocks of the GNG tasks produced positive, though anecdotal, evidence for the null hypothesis of comparable block d' scores (see Fig. 4)) and substantial positive evidence for the comparability of correct RTs (see Fig. 4(c)). In the case of the CPT the strength of evidence for the null hypothesis of equivalent d' scores across the shape and colour blocks was substantial. However, the shape block was clearly associated with faster correct responses than the colour block. Presumably this reflects the fact that the various different shapes were more discriminable from one other than was the case for the colours used in the colour block, perhaps particularly so for the one target stimulus (the star) that the participants had to respond to in the shape block. Although this difference in difficulty is worth noting, and confirms that RTs are able to capture reliable cross-condition effects, it raises no concerns for our interpretation of the key experimental effects because the presence or absence of ultrasound exposure was randomly varied, by participant, across the two blocks of each task.

Across all our analyses there was clear evidence that ultrasound exposure had no meaningful effect on task performance. The analyses of the data shown in Fig. 4 provide indirect evidence to support this statement. The factor of exposure block employed in these analyses reflects whether the exposure took place on the first or the second block of a given task. As the order of the two component blocks of each task was fixed (the number block preceding the letter block and the shape block preceding the colour block in the GNG task and continuous performance test respectively), any effect of ultrasound exposure would manifest itself as a exposure block by block type interaction. This interaction term never appeared in a preferred model in these analyses.

More direct support for the claim of no effect of ultrasound exposure comes from the corresponding analyses of the data which directly contrasted the control and exposed blocks of each task, regardless of which particular block that was, though with the factor of exposure block again included (see the 'Total' column in Fig. 4 for a representation of the main effect of exposure block). The data from the GNG task provided clear and unequivocal evidence that ultrasound exposure had no effect on either dependent measure derived from the task. On the continuous performance test there was similarly clear evidence that exposure had no overall effect on either d' scores or correct RTs. However, there was evidence for an interaction between condition (control condition vs. exposed condition) and exposure block for correct RTs on this task. As Fig. 5 shows, there is a suggestion in the data that participants who received the ultrasound exposure in the second block of the continuous performance test were slower on that second block than they were in the first, unexposed block. While this might appear to imply a detrimental effect of ultrasound on cognitive task performance, there are a number of reasons to reject that interpretation.

First, our other analyses have shown that the colour block, which was the second block of the CPT, was generally associated with slower RTs than the shape block which could explain this pattern. Second, while the overall interaction between condition and exposure block was associated with a Bayes Factor of greater than 3, further post hoc analysis showed that the evidence for a condition effect was only anecdotal ($BF_{incl} = 2.20$). Similarly, there was no good evidence that the RTs under ultrasound exposure of participants who received that exposure in the second block of the CPT were slower than RTs under exposure among those who received exposure in the first block ($BF_{incl} = 0.72$). Finally, and as already noted, there was substantial evidence for there being no overall effect of ultrasound exposure in the main analysis of CPT RTs ($BF_{excl} = 4.62$).

The proper functioning of the ultrasonic array was checked prior, during and after experimentation. Exposure levels were consistent with a free-field output of $\sim$ 120 dB SPL at distance of 50–60 cm. These measurements ultimately characterise the array and are reproducible (provided an anechoic chamber is available). In this study, effects of ultrasound exposure were correlated to this free-field measurement for consistency. In reality, HATS in situ measurements verified that head position has considerable influence over exposure *at the ear*. This raises some considerations for any attempt to standardize airborne ultrasound devices; SPL output alone does not describe the total exposure risk, how users interact with a device (and move within its acoustic field) plays an important role as well.

Given the findings reported above, the current study provides clear evidence that the ultrasound exposure (40 kHz tone, 120 dB SPL re 20 μPa) had no effect on task performance. Importantly this has been demonstrated using a series of Bayesian analyses. A Bayesian approach has the advantage of allowing one to examine the evidence in favour of both the experimental and the null hypotheses. Our results therefore go substantially further than those provided by any preceding study that simply demonstrated an absence of evidence for an effect of ultrasound exposure on performance. Instead, in this work we have provided meaningful positive evidence for there being no effect of ultrasound exposure on cognitive function. Another important finding of this work concerns participants' ability to detect when the ultrasound exposure was taking place. Although some previous work has suggested that there might be detectable effects of exposure on individuals subjective experience, this was not the case in this study. Participants were generally reluctant to even guess which block exposure had occurred in, reflecting their lack of confidence in this decision. When they did make a choice they did not exceed 50% accuracy, indicating that they were unable to reliably detect the presence of ultrasound.

It should also be noted that human auditory sensitivity to high-frequency noise diminishes rapidly with age. Thus, participants in this study (mean age: 20 years), represent some of the youngest (and potentially most sensitive) *adults* available to ethically experiment with. So far, the proliferation of ultrasonic devices such as DIY acoustic levitators has not resulted in a public wave of complaints from vulnerable groups (representing children or otherwise). The apparent benign nature of these devices may, in part, be inferred by the results from this study. If further investigations are to be carried out on (or to identify) vulnerable groups, the evidence from this trial will serve to inform a safe and ethical protocol.

Supplementary material

Source code for the cognitive task/array driver is available open-source at:https://github.com/andydiba/sart_cpt.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Andrew Di Battista reports financial support was provided by Ultraleap Limited. Adam Price reports financial support was provided by Ultraleap Limited. Rob Malkin reports financial support was provided by Ultraleap Limited. Christopher Jarrold reports financial support was provided by Ultraleap Limited. The research was sponsored by Ultraleap Ltd and carried out in collaboration with the University of Bristol. AD, AP and RB are employees of Ultraleap. CJ, BD and PK are researchers at University of Bristol.

Acknowledgement

We thank Rafael Morales González for assistance in the design/-construction of the ultrasonic array. Funding: this work was sponsored by Ultraleap Ltd in collaboration with the Univeristy of Bristol.

Appendix A. Array design

When using arrays of transducers a possibility exists to create unwanted grating effects (energy directed in more than 1 direction). Grating lobes are caused where the distance between neighbouring elements (known as the pitch, D_{el}) in a rectilinear array is greater than the critical pitch, $D_{crit} \approx \lambda/2$. The transducers in our study have a physical outer diameter of $\approx$10 mm and a wavelength of 8.6 mm, thus cannot meet the critical pitch needed to avoid grating lobes. In order to avoid unwanted grating lobes which would complicate the acoustic environment needlessly we use a non-regular array layout, specifically a Fibonacci spiral derived layout (Fig. 6).

Fig. 6

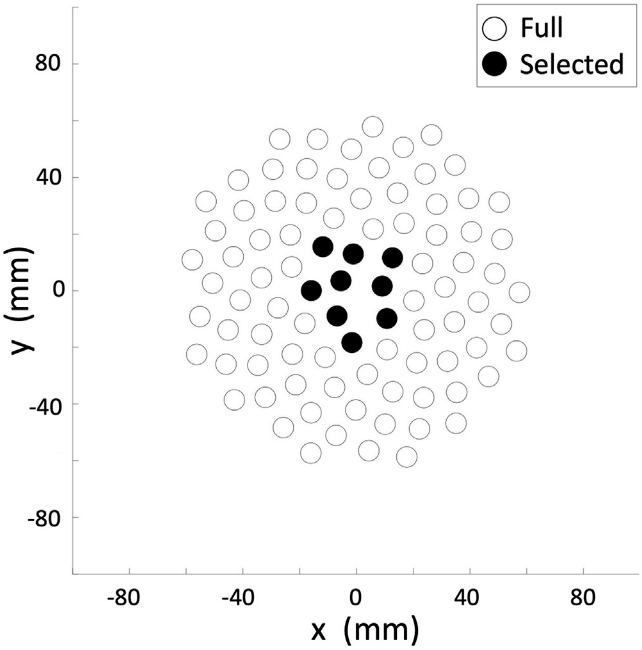

Fig. 6. Sunflower Array The 9 element array was selected from the subset of a ‘sunflower’ (Fermat spiral) design. Each transducer (n = 1,2…,9,…N) has polar coordinates $r = c\sqrt{n}, \theta = n \times 137.508^\circ$ where 137.508° is the *golden angle*. *c* is a constant scale factor that was heuristically adjusted to account for the transducer 10 mm diameters.

Appendix B. Array hardware

Fig. 7

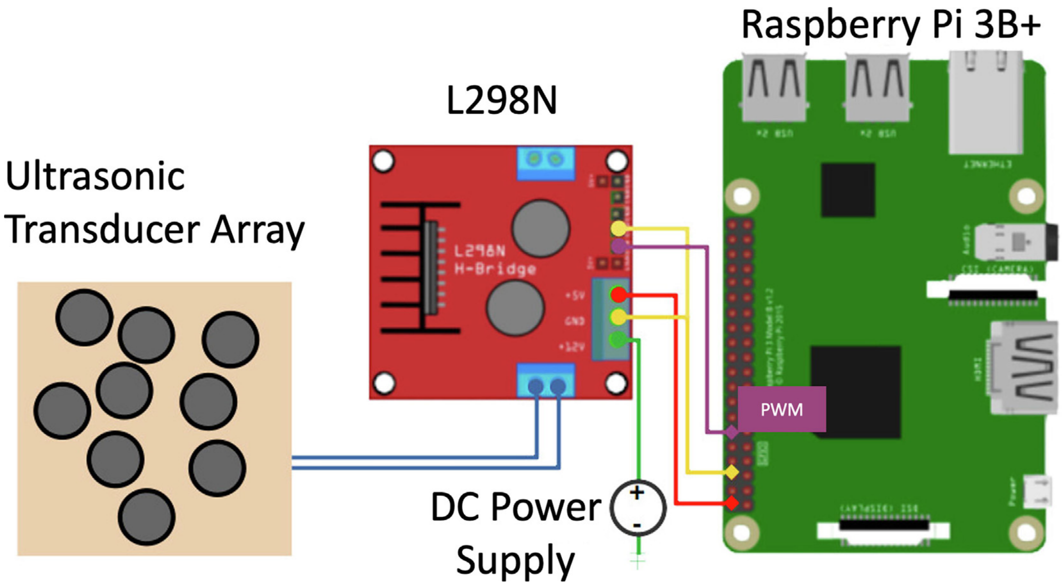

Fig. 7. Ultrasonic Array – Hardware overview A raspberry Pi 3B + provides the programming logic and PWM signal source (40 kHz) to both drive the ultrasonic array and run the cognitive test software. The L298N DC motor drive circuit serves as an PWM amplifier. The array itself consists of 9 Murata MA40S4S transducers arranges in a pseudo sunflower (Fermat spiral) pattern. The Pi can be connected to a computer monitor and keyboard or accessed remotely via a LAN. The Pi can also produce its own access point (AP) for remote access in the absence of any other network infrastructure.

Appendix C. HATS free-field calibration measurements

Fig. 8

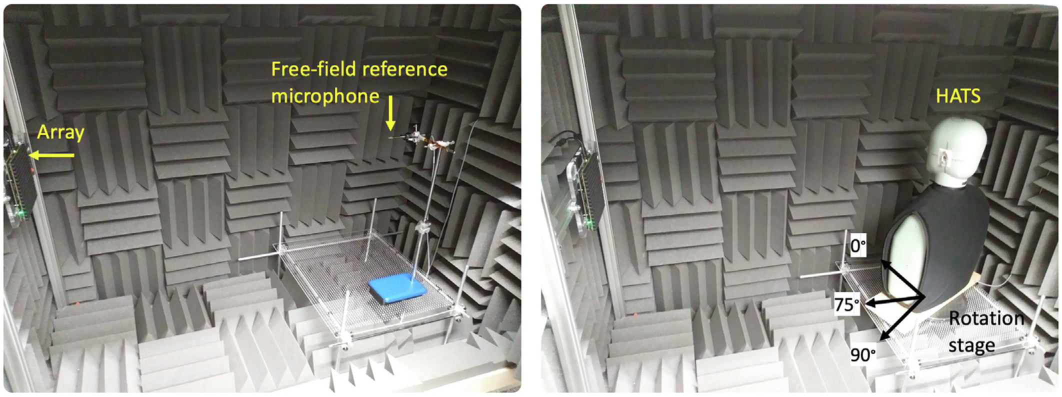

Fig. 8. HATS Free-Field Calibration HATS free-field-equivalent calibration. (left) Free-Field reference measurement with TYPE 4939 microphone. (right) HATS setup on ration stage to measure SPL at various azimuthal angles (currently showing 75°). Note: Azimuths are defined relative to the HATS i.e. 0° corresponds to head-on incidence.

A comparison of free-field and HATS measurements were compared to explore the transfer function from air-to-ear. The shadowing effects from the head and pinna (i.e. outer-ear) are readily apparent.

References

[1] Wu T-F, Tsai P-S, Hu N-T, Chen J-Y. Research and implementation of auto parking system based on ultrasonic sensors. In: International Conference on Advanced Materials for Science and Engineering (ICAMSE) IEEE. p. 643–5.

[2] Li T-HS, Chang C-C, Ye Y-J, Tasi G-R. Autonomous parking control design for car-like mobile robot by using ultrasonic and infrared sensors. In: Bredenfeld A, Jacoff A, Noda I, Takahashi Y, editors. RoboCup 2005: Robot Soccer World Cup IX. Berlin Heidelberg, Berlin, Heidelberg: Springer; 2006. p. 472–9.

[3] Marzo A, Barnes A, Drinkwater BW. Tinylev: A multi-emitter single-axis acoustic levitator. Rev Sci Instrum 2017;88:085105.

[4] E. Chołodowicz, P. Orłowski, Low-cost air levitation laboratory stand using matlab/simulink and arduino, Pomiary Automatyka Robotyka 21 (2017) 33–39. 10.14313/PAR_226/33.

[5] T. Carter, S.A. Seah, B. Long, B. Drinkwater, S. Subramanian, Ultrahaptics: multi-point mid-air haptic feedback for touch surfaces, in: Proceedings of the 26th annual ACM symposium on User interface software and technology, 2013, pp. 505–514.

[6] UpnaLab, Acoustic levitator, 2022. URL: https://www.instructables.com/Acoustic-Levitator/, accessed: 2022-05-24.

[7] Liebler M, Kling C, Best B, Gerlach A, Koch C. Quantitative characterization of high-intensity focused airborne ultrasonic fields. In: Proceedings of the 23rd International Congress on Acoustics.

[8] B. Lawton, Exposure limits for airborne sound of very high frequency and ultrasonic frequency, 2013. URL: https://eprints.soton.ac.uk/351902/.

[9] Acton W. The effects of industrial airborne ultrasound on humans. Ultrasonics 1974;12:124–8. https://doi.org/10.1016/0041-624X(74)90069-9.

[10] Leighton T. Are some people suffering as a result of increasing mass exposure of the public to ultrasound in air?. Proceedings of the Royal Society 2016;472. https://doi.org/10.1098/rspa.2015.0624.

[11] M. Lenhardt, Airborne ultrasonic standards for hearing protection, 9th International Congress on Noise as a Public Health Problem (ICBEN), 2008. URL: http://icben.ethz.ch/2008/PDFs/Lenhardt.pdf.

[12] Smagowska B, Pawlaczyk-Łuszczyńska M. Effects of ultrasounic noise on the human body - A bibliographic review. Int J Occup Saf Ergon 2013;19:195–202. https://doi.org/10.1080/10803548.2013.11076978.

[13] Wiernicki C, Karoly W. Ultrasound: Biological effects and industrial hygiene concerns. Am Ind Hygeine Assoc J 1985;46:488–96. https://doi.org/10.1080/15298668591395229.

[14] Parrack H. Effect of air-borne ultrasound on humans. Int Audiol 1966:294–308.

[15] Fletcher M, Jones SL, White PR, Dolder CN, Leighton TG, Lineton B. Effects of very high-frequency sound and ultrasound on humans. part ii: A double- blind randomized provocation study of inaudible 20-khz ultrasound. J Acoust Soc Am 2018;144:2521–31. https://doi.org/10.1121/1.5063818.

[16] Flindell I, Hooper P. Noise annoyance: What does it mean?. In: Proceedings of the 23rd International Congress on Acoustics. p. 346–52. URL: http://pub.dega-akustik.de/ICA2019/data/articles/001547.pdf..

[17] Carcagno S, Battista AD, Plack C. Effects of high-intensity airborne ultrasound exposure on behavioural and electrophysiological measures of auditory function. Acta Acustica United With Acustica 105 2019:1183–97. https://doi.org/10.3813/AAA.919395.

[18] T. Manly, A.M. Owen, L. McAvinue, A. Datta, G.H. Lewis, S.K. Scott, C. Rorden, J. Pickard, I.H. Robertson, Enhancing the sensitivity of a sustained attention task to frontal damage: Convergent clinical and functional imaging evidence, Neurocase 9 (2003) 340–349. URL: doi: 10.1076/neur.9.4.340.15553.

[19] Donders F. On the speed of mental processes. Acta Psychol 1969;30:412–31. https://doi.org/10.1016/0001-6918(69)90065-1.

[20] Verbruggen F, Logan GD. Automatic and controlled response inhibition: associative learning in the go/no-go and stop-signal paradigms. J Exp Psychol 2008;137:649–72. https://doi.org/10.1037/a0013170.

[21] Mackworth N. The breakdown of vigilance during prolonged visual search. Q J Exper Psychol 1948;1:6–21. https://doi.org/10.1080/17470214808416738.

[22] Huang-Pollock C, Karalunas S, Tam H, Moore A. Evaluating vigilance deficits in adhd: A meta-analysis of cpt performance. J Abnorm Psychol 2012;121:360–71. https://doi.org/10.1037/a0027205.

[23] Di Battista A. The effect of 40 khz ultrasonic noise exposure on human hearing. In: Proceedings of the 23rd International Congress on Acoustics.

[24] A. Frank, S. Goldlist, M. Fraser, E. Amy, M. Bromwich, Validation of shoebox quicktest hearing loss screening tool in individuals with cognitive impairment, Frontiers in Digital Health 3 (2021). URL: https://www.frontiersin.org/article/10.3389/fdgth.2021.724997. DOI: 10.3389/fdgth.2021.724997.

[25] Saliba J, Al-Reefi M, Carriere J, Verma N, Provencal C, Rappaport J. Accuracy of mobile-based audiometry in the evaluation of hearing loss in quiet and noisy environments. Otolaryngol - Head Neck Surgery 2017;156:706–11. https://doi.org/10.1177/0194599816683663.

[26] Plack C, Guest H, Carcagno S. Understanding extended high-frequency thresholds. In: Proceedings of the 23rd International Congress on Acoustics.

[27] van den Bergh D, van Doorn J, Marsman M, Draws T, van Kesteren E, Derks K, Wagenmakers E. A tutorial on conducting and interpreting a bayesian anova in jasp. L'Année Psychologique 2020;120:73–96. https://doi.org/10.3917/anpsy1.201.0073.

[28] JASP Team, Jasp, 2022. URL: https://jasp-stats.org/.

[29] Jefferys H. Theory of probability. 1 ed. Oxford University Press; 1939.

[30] Keysers C, Gazzola V, Wagenmakers EJ. Using bayes factor hypothesis testing in neuroscience to establish evidence of absence. Nat Neurosci 2020;23:788–99. https://doi.org/10.1038/s41593-020-0660-4.

[31] Haatveit BC, Sundet K, Hugdahl K, Ueland T, Melle I, Andreassen OA. The validity of d prime as a working memory index: Results from the 'bergen n-back task'. J Clin Exp Neuropsychol 2010;32:871–80. https://doi.org/10.1080/13803391003596421.

[32] Hautus MJ. Corrections for extreme proportions and their biasing effects on estimated values of d-prime. Behavior Research Methods Instruments and Computers 1995;27:46–51. https://doi.org/10.3758/BF03203619.