Neuroscience

Auditory confounds can drive online effects of transcranial ultrasonic stimulation in humans

Reviewed Preprint

Revised by authors after peer review.

About eLife's process

Reviewed preprint version 2
February 13, 2024 (this version)

Reviewed preprint version 1
October 2, 2023

Sent for peer review
May 11, 2023

Posted to preprint server
April 17, 2023

Benjamin R. Kop, Yazan Shamli Oghli, Talyta C. Grippe, Tulika Nandi, Judith Lefkes, Sjoerd W. Meijer, Soha Farboud, Marwan Engels, Michelle Hamani, Melissa Null, Angela Radetz, Umair Hassan, Ghazaleh Darmani, Andrey Chetverikov, Hanneke E.M. den Ouden, Til Ole Bergmann, Robert Chen, Lennart Verhagen

Donders Institute for Brain, Cognition, and Behaviour; Radboud University Nijmegen, the Netherlands • Krembil Research Institute, University Health Network; University of Toronto, Canada • Neuroimaging Center; Johannes-Gutenberg University Medical Center Mainz, Germany • Department of Psychosocial Science, Faculty of Psychology, University of Bergen, Bergen, Norway • Leibniz Institute for Resilience Research Mainz, Germany

Abstract

Transcranial ultrasonic stimulation (TUS) is rapidly emerging as a promising non-invasive neuromodulation technique. TUS is already well-established in animal models, providing foundations to now optimize neuromodulatory efficacy for human applications. Across multiple studies, one promising protocol, pulsed at 1000 Hz, has consistently resulted in motor cortical inhibition in humans (Fomenko et al., 2020). At the same time, a parallel research line has highlighted the potentially confounding influence of peripheral auditory stimulation arising from TUS pulsing at audible frequencies. In this study, we disentangle direct neuromodulatory and indirect auditory contributions to motor inhibitory effects of TUS. To this end, we include tightly matched control conditions across four experiments, one preregistered, conducted independently at three institutions. We employed a combined transcranial ultrasonic and magnetic stimulation paradigm, where TMS-elicited motor-evoked potentials (MEPs) served as an index of corticospinal excitability. First, we replicated motor inhibitory effects of TUS but showed through both tight controls and manipulation of stimulation intensity, duration, and auditory masking conditions that this inhibition was driven by peripheral auditory stimulation, not direct neuromodulation. Further, we consider neuromodulation beyond driving overall excitation/inhibition and show preliminary evidence of how TUS might interact with ongoing neural dynamics instead. Primarily, this study highlights the substantial shortcomings in accounting for the auditory confound in prior TUS-TMS work where only a flip-over sham and no active control was used. The field must critically reevaluate previous findings given the demonstrated impact of peripheral confounds. Further, rigorous experimental design via (in)active control conditions is required to make substantiated claims in future TUS studies. Only when direct effects are disentangled from those driven by peripheral confounds can TUS fully realize its potential for research and clinical applications.

eLife assessment

This **important** multicenter study provides **convincing** evidence that the auditory noise emitted during online transcranial ultrasound stimulation (TUS) protocols can pose a considerable confound and is able to explain corticospinal excitability changes as measured with Motor Evoked Potentials (MEP). The findings lay the ground for future studies optimising protocols and control conditions to leverage TUS as a meaningful experimental and clinical tool. A clear strength of the study is the multitude of control conditions (i.e., control sites, acoustic masking, acoustic stimulation). These findings will be of interest to neuroscience researchers using brain stimulation approaches.

1. Introduction

Noninvasive neuromodulation is a powerful tool for causal inference that strengthens our understanding of the brain and holds great clinical potential (Bergmann & Hartwigsen, 2021; Bestmann & Walsh, 2017). Transcranial ultrasonic stimulation (TUS) is a particularly promising non-invasive brain stimulation technique, overcoming current limitations with high spatial resolution and depth range (Darmani et al., 2022). The efficacy of TUS is well-established in cell cultures and animal models (Menz et al., 2013; Mohammadjavadi et al., 2019; Murphy et al., 2022; Tyler et al., 2008, 2018; Yoo et al., 2022), and emerging evidence for the neuromodulatory utility of TUS in humans has been reported for both cortical and subcortical structures (cortical: Butler et al., 2022; Lee et al., 2016; Liu et al., 2021; Zeng et al., 2022; subcortical: Ai et al., 2016; Cain et al., 2021; Nakajima et al., 2022). Especially now, at this foundational stage of TUS in humans, it is essential to converge on protocols that maximize the specificity and efficacy of stimulation (Folloni et al., 2019; Verhagen et al., 2019).

Motor inhibitory effects of a commonly applied 1000 Hz pulsed TUS protocol are among the most robust and replicable human findings (Fomenko et al., 2020; Legon, Bansal, et al., 2018; Xia et al., 2021). Here, by concurrently applying transcranial magnetic stimulation (TMS), modulation of corticospinal excitability is indexed by motor-evoked potentials (MEPs). However, the mechanism by which TUS evokes motor inhibition has remained under debate (Xia et al., 2021).

Recent studies in both animal and human models demonstrate how electrophysiological and behavioral outcomes of TUS can be elicited by nonspecific auditory activation rather than direct neuromodulation (Airan & Butts Pauly, 2018; Braun et al., 2020; Guo et al., 2018; Sato et al., 2018). Indeed, there is longstanding knowledge of the auditory confound accompanying pulsed TUS (Gavrilov & Tsirulnikov, 2012). However, this confound has only recently garnered attention, prompted by a pair of rodent studies demonstrating indirect auditory activation induced by TUS (Guo et al., 2022; Sato et al., 2018). Similar effects have been observed in humans, where exclusively auditory effects were captured with EEG measures (Braun et al., 2020). These findings are particularly impactful given that nearly all TUS studies employ pulsed protocols, from which the pervasive auditory confound emerges (Johnstone et al., 2021).

Indirect effects of stimulation are not unique to TUS, as transcranial magnetic and electric stimulation are also associated with auditory and somatosensory confounds. Indeed, the field of non-invasive brain stimulation as a whole depends on controlling for these confounding factors when present, to unveil the specificity of the neuromodulatory effects (Conde et al., 2019; Duecker et al., 2013; Polanía et al., 2018; Siebner et al., 2022). However, prior online TUS-TMS studies, including those exploring optimal neuromodulatory parameters to inform future

work, have considered some but not all necessary conditions to control for the salient auditory confound elicited by a 1000 Hz pulsed protocol (Fomenko et al., 2020; Legon, Bansal, et al., 2018; Xia et al., 2021).

In this multicenter study, we quantified the impact of the auditory confound to disentangle direct neuromodulatory and indirect auditory contributions to motor inhibitory effects of TUS. To this end, we substantially improved upon prior TUS-TMS studies implementing solely flip-over sham by including both (in)active controls and multiple sound-sham conditions. Further, we investigated dose-response effects through administration of multiple stimulus durations, stimulation intensities, and individualized simulations of intracranial intensity. Additionally, we considered the possibility that online TUS might not drive a global change in the excitation/inhibition balance but instead might interact with ongoing neural dynamics by introducing state-dependent noise. Finally, we interrogated sound-driven effects through modulation of auditory confound volume, duration, pitch, and auditory masking. We show that motor inhibitory effects of TUS are spatially nonspecific and driven by sound-cued preparatory motor inhibition. However, we do find preliminary evidence that TUS might introduce dose– and state-dependent neural noise to the dynamics of corticospinal excitability. The present study highlights the importance of carefully constructed control conditions to account for confounding factors while exploring and refining TUS as a promising technique for human neuromodulation.

2. Materials and methods

2.1. Participants

This multicenter study comprised of four experiments conducted independently across three institutions. Experiment I (N = 12, 4 female, M_{age} = 25.9, SD_{age} = 4.6; METC: NL76920.091.21) and Experiment II (N = 27, 13 female, M_{age} = 24.1, SD_{age} = 3.7; METC: NL76920.091.21) were conducted at the Donders Institute of the Radboud University (the Netherlands). Experiment III was conducted at the Krembil Research Institute (N = 16, 8 female, M_{age} = 31.4, SD_{age} = 7.9; Toronto University Health Network Research Ethics Board: 20-5740, Canada), and Experiment IV at the Neuroimaging Centre of the Johannes Gutenberg University Medical Centre Mainz (N = 12, 11 female, M_{age} = 23.0, SD_{age} = 2.7, Landesärztekammer Rheinland-Pfalz: 2021-15808_01, Germany). All participants were healthy, right-handed, without a history of psychiatric or neurological disorders, and provided informed consent. Ethical approval was obtained for each experiment.

2.2. Transcranial ultrasonic and magnetic stimulation

Ultrasonic stimulation was delivered with the NeuroFUS system (manufacturer: Sonic Concepts Inc., Bothell, WA, USA; supplier/support: Brainbox Ltd., Cardiff, UK). A radiofrequency amplifier powered a piezoelectric ultrasound transducer via a matching network. Transducers consisted of a two-element annular array. Further transducer specifications are reported in **Supplementary Table 1**. Ultrasonic stimulation parameters were based on those used in prior TUS-TMS studies (**Table 1**, **Fig. 1A**; Fomenko et al., 2020; Legon, Bansal, et al., 2018; Xia et al., 2021). While ramping the pulses can in principle mitigate the auditory confound (Johnstone et al., 2021; Mohammadjavadi et al., 2019), doing so for such short pulse durations (<= 0.3 ms) is not effective. Therefore, we used a rectangular pulse shape to match prior work.

Single-pulse TMS was delivered with a figure-of-eight coil held at 45° from midline to induce an approximate posterolateral to anteromedial current. The hand motor hotspot and required TMS intensity were determined using standard procedures as outlined in **Supplementary** Fig. 1. To apply TUS and TMS concurrently, the ultrasound transducer was affixed to the center of the TMS coil using a custom-made 3D-printed clamp (**Fig. 1B**; Experiments I, II, & IV; Experiment III: see Fomenko et al., 2020). TMS was triggered 10 ms prior to the offset of TUS (**Fig. 1C**). Muscular

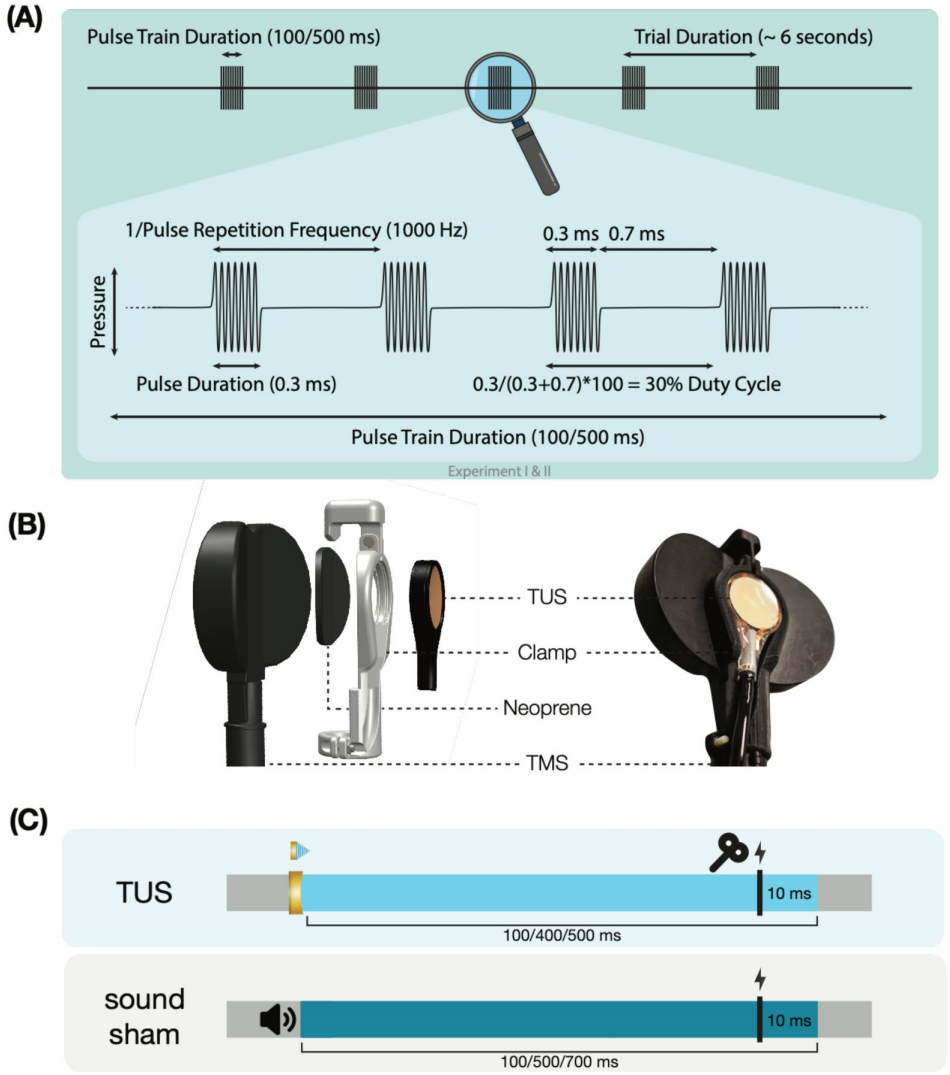

Fig. 1

| Experimental procedures.

(A) Ultrasonic stimulation protocol for Experiments I & II. In Experiment III a duty cycle of 10% was used. In Experiment IV a stimulus duration of 400 ms was used. **(B)** TUS-TMS clamp (DOI: 10.5281/zenodo.6517599). **(C)** Experimental timing. Detailed experimental timing for each experiment is reported in **Supplementary** Fig. 3.

Exp.	*F* (kHz)	depth (mm)	PD (ms)	PRF (kHz)	DC	PTD (ms)	I_{sppa} (W/cm^2)	P (Mpa)	MI_{tc}
I	500	35	0.3	1000	30%	100/500	32.5/65	1.02/1.44	0.99/1.40
II	250	28	0.3	1000	30%	500	6.35/19.06	0.45/0.78	0.65/1.12
III	500	30	0.1	1000	10%	500	9.26	0.54	0.53
IV	250	50[†]	0.3	1000	30%	400	4.34/8.69/10.52	0.37/0.53/0.58	0.53/0.76/0.83

[†]Note: Experiments I-III targeted the hand motor area. Experiment IV targeted the corticospinal white matter.

Table 1

| Ultrasonic stimulation parameters.

f = fundamental frequency, depth = TPO focus setting for distance of free-water full-width half-maximum from transducer exit plane, PD = pulse duration, PRF = pulse repetition frequency, DC = duty cycle, PTD = pulse train duration, I_{sppa} = spatial-peak pulse-average intensity in free-water, P = pressure, MItc = transcranial derated mechanical index. The ramp shape for all experiments was rectangular. For estimated intracranial indices for Experiments I & II see **Supplementary** Figure 2.

activity was recorded in the first dorsal interosseous (FDI; Experiments I-III) or in the abductor pollicis brevis (APB; Experiment IV) via electromyography with surface adhesive electrodes using a belly-tendon montage (**Supplementary Table 1**).

In Experiments I, II, and IV, we used online neuronavigation with individual anatomical scans to support target selection and consistent TMS and TUS placement (Localite Biomedical Visualization Systems GmbH, Sankt Augustin, Germany; MRI specifications: **Supplementary Table 2**). Further, we recorded the position of TUS in Experiments I and II for post-hoc acoustic and thermal simulations.

2.3. Experiment I

On-target TUS was delivered to the left-hemispheric hand motor area to determine the effect of ultrasonic stimulation on corticospinal excitability. We introduced controls that improve upon the sole use of flip-over sham conditions used in prior work. First, we applied active control TUS to the right-hemispheric face motor area, allowing for the assessment of spatially specific effects while also better mimicking on-target peripheral confounds. In addition, we also included a sound-only sham condition that closely resembled the auditory confound (**Fig. 2**). Specifically, we generated a 1000 Hz square wave tone with 0.3 ms long pulses using MATLAB. We then added white noise at a signal-to-noise ratio of 14:1. This stimulus was administered to the participant via bone-conducting headphones (AfterShockz Trekz, TX, USA). Finally, we incorporated a baseline condition consisting solely of TMS.

Ultrasonic stimulation was delivered at two pulse train durations (100/500 ms) and at two intensities (32.5/65 W/cm^2 I_{sppa}) to probe a potential dose-response effect. Additionally, with consideration of potentially audible differences between on-target and active control stimulation sites, we applied these conditions both with and without masking stimuli identical those used during sound-only sham. Auditory stimuli used for sound-sham and/or masking for each experiment are accessible here: *https://doi.org/10.5281/zenodo.8374148*. See **Supplementary** Fig. 3 for an overview of conditions and experimental timing for each experiment.

Conditions were administered in a single-blind inter-subject counterbalanced blocked design while participants were seated at rest. Ultrasound gel was used to couple both transducers to the participant's scalp (Aquasonic 100, Parker Laboratories, NJ, USA). In total, participants completed 14 blocks of 20 trials each. Each trial lasted 6 ± 1 seconds. Two baseline measurements were completed, the first occurring as one of the first four blocks, and the second as one of the last four, to capture any general shift in excitability throughout the experiment. TMS was administered on every trial for a total of 280 single pulses.

2.4. Experiment II

To confirm and expand upon our findings from Experiment I we conducted a second, preregistered, experiment using the same main conditions and procedures, with a few adaptations (*https://doi.org/10.17605/OSF.IO/HS8PT*). The 2×2×2 design comprised of stimulation site (on-target/active control), stimulation intensity (6.35/19.06 W/cm^2), and auditory masking (no mask/masked). We applied ultrasonic stimulation exclusively at an effective 500 ms pulse train duration. In this experiment, the same 1000 Hz square wave auditory stimulus was used for sound-only sham and auditory masking conditions. This stimulus was administered to the participant over in-ear headphones (ER-3C Insert Earphones, Etymotic Research, Illinois, USA). To better capture any baseline shift in excitability during the experiment, we presented conditions in a single-blind pseudorandomized order in which each consecutive set of 10 trials included each of 10 conditions once. Participants completed 25 trials per condition, resulting in 250 trials total.

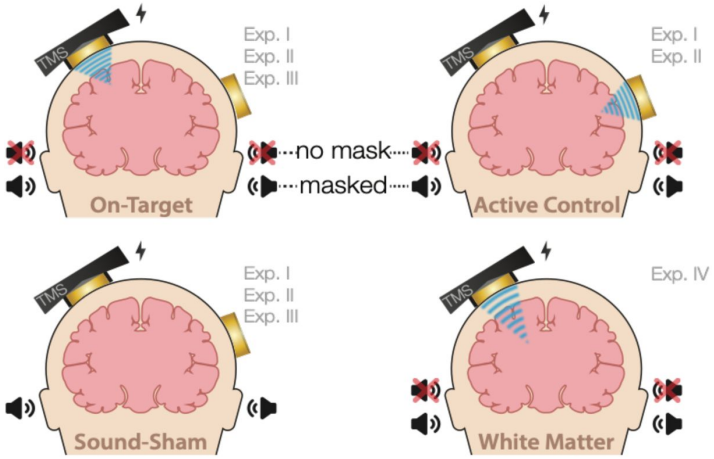

Fig. 2

| Experimental conditions.

On-target TUS of the left-hemispheric hand motor area (Exp. I-III), active control TUS of the right-hemispheric face motor area (Exp. I-II), sound-only sham (Exp. I-III), and inactive control TUS of the white matter ventromedial to the hand motor area (Exp. IV). Conditions involving TUS were presented both with and without auditory masking stimuli.

To further probe a potential dose-response effect of stimulation intensity, we ran acoustic and thermal simulations (**Supplementary** Fig. 4). Here, we assessed the relationship between estimated intracranial intensities and perturbation of corticospinal excitability. While simulations were also run for Experiment I, its sample size was insufficient to test for intracranial dose-response effects.

Following the main experiment, we tested the efficacy of our masking stimuli with a forced-choice task wherein participants reported if they had received TUS for each condition, excluding baseline. Additionally, we investigated whether audible differences between stimulation sites were present during auditory masking (**Supplementary** Fig. 5).

2.5. Experiment III

We further characterized possible effects of auditory confounds on motor cortical excitability by administering varied auditory stimuli, both alongside on-target TUS and without TUS (i.e., sound-only sham). Auditory stimuli were either 500 or 700 ms in duration, the latter beginning 100 ms prior to TUS (**Supplementary** Fig. 3.**3**). Both durations were presented at two pitches. Using a signal generator (Agilent 33220A, Keysight Technologies), a 12 kHz sine wave tone was administered over speakers positioned to the left of the participant as in Fomenko and colleagues (2020). Additionally, a 1 kHz square wave tone with 0.5 ms long pulses was administered as in Experiments I, II, IV, and prior research (Braun et al., 2020) over noise-cancelling earbuds.

First, we investigated changes in corticospinal excitability from baseline following these auditory stimuli. Participants received 15 trials of baseline (i.e., TMS only) and 15 trials of each of the four sound-only sham stimuli. Conditions were presented in a blocked single-blind randomized order with participants seated at rest. An inter-trial interval of 5 seconds was used.

Next, we assessed whether applying on-target TUS during these auditory stimuli affected motor excitability. Here, TMS intensity was set to evoke a ~1 mV MEP separately for each of the four sound-only sham conditions (**Supplementary** Fig. 1.**2**). To account for different applied TMS intensities between baseline and these conditions, we calculated *Relative MEP amplitude* by multiplying each trial by the ratio of applied TMS intensity to baseline TMS intensity. Participants received 15 trials of each auditory stimulus, once with on-target TUS and once as a sound-only sham. Ultrasound gel (Wavelength MP Blue, Sabel Med, Oldsmar, FL) and a 1.5 mm thick gel-pad (Aquaflex, Parker Laboratories, NJ, USA) were used to couple the transducer to the participants' scalp. Conditions were presented in pairs of sound-sham and TUS for each auditory stimulus, counterbalanced between subjects. The order of the different auditory stimuli was randomized across participants.

2.6. Experiment IV

We further investigated the role of TUS audibility on motor excitability by administering stimulation to an inactive control site – the white matter ventromedial to the hand motor area. In doing so, TUS is applied over a homologous region of the scalp and skull without likely direct neuromodulation, thus allowing us to closely replicate the auditory confound while simultaneously isolating its effects.

Here, we probed whether the varying volume of the auditory confound at different stimulation intensities might itself impact motor cortical excitability. To this end, we applied stimulation at 4.34, 8.69, and 10.52 W/cm^2 I_{sppa}, or in effect, at three auditory confound volumes. We additionally applied stimulation both with and without a continuous auditory masking stimulus that sounded similar to the auditory confound. The stimulus consisted of a 1 kHz square wave with 0.3 ms long pulses. This stimulus was presented through wired bone-conducting headphones (LBYSK Wired

Bone Conduction Headphones). The volume and signal-to-noise ratio of the masking stimulus were increased until the participant could no longer hear TUS, or until the volume became uncomfortable.

We administered conditions in a single-blind inter-subject randomized blocked design. Two blocks were measured per condition, each including 30 TUS-TMS trials and an additional 30 TMS-only trials to capture drifts in baseline excitability. These trials were applied in random order within each block with an inter-trial interval of 5 ± 1 seconds. Ultrasound gel (Aquasonic 100, Parker Laboratories, NJ, USA) and a ~2-3 mm thick gel-pad were used to couple the transducer to the participant's scalp (Aquaflex, Parker Laboratories, NJ, USA). During blocks with auditory masking, the mask was played continuously throughout the block. Following each block, participants were asked whether they could hear TUS (yes/no/uncertain).

2.7. Analysis

Raw data were exported to MATLAB, where MEP peak-to-peak amplitude was calculated using a custom script and confirmed by trial-level visual inspection. Trials where noise prevented an MEP to be sufficiently quantified were removed. Given the right-skewed nature of the raw MEP values, we performed a square root transformation to support parametric statistics. For visualization purposes, baseline corrected MEP amplitudes were also calculated.

Linear mixed-effects models (LMMs) were fitted using the lme4 package in R (Bates et al., 2015; R core team, 2021). Intercepts and condition differences (slopes) were allowed to vary across participants, including all possible random intercepts, slopes, and correlations in a maximal random effects structure (Barr et al., 2013). Statistical significance was set at two-tailed $\alpha = 0.05$ and was computed with t-tests using the Satterthwaite approximation of degrees of freedom. For direct comparisons to a reference level (e.g., baseline), we report the intercept (b), standard error (SE), test-statistic (t), and significance (p). For main effects and interactions, we report the F statistic, significance, and partial eta squared. LMMs included square root transformed MEP peak-to-peak amplitude as the dependent variable, with the relevant experimental conditions and their interactions as predictors. Given the large number of baseline trials in Experiment IV (50% of total), the LMM testing effects of stimulation intensity and auditory masking instead included baseline corrected MEP amplitude as the dependent variable.

3. Results

3.1. Motor cortical inhibition is not specific to on-target TUS

We first corroborate previous reports of MEP suppression following 500 ms of TUS applied over the hand motor area (Experiments I-III; Fomenko et al., 2020; Legon, Bansal, et al., 2018; Xia et al., 2021). A LMM revealed significantly lower MEP amplitudes following on-target TUS as compared to baseline for Experiment I ($b = –0.14$, $SE = 0.06$, $t(11) = –2.23$, $p = 0.047$), Experiment II ($b = –0.18$, $SE = 0.04$, $t(26) = –4.82$, $p = 6{\cdot}10^{-5}$), and Experiment III ($b = –0.22$, $SE = 0.07$, $t(15) = –3.08$, $p = 0.008$).

However, corticospinal inhibition from baseline was also observed following control conditions. LMMs revealed significant attenuation of MEP amplitude following active control stimulation of the right-hemispheric face motor area (Experiment I: $b = –0.12$, $SE = 0.05$, $t(11) = –2.29$, $p = 0.043$; Experiment II: $b = –0.22$, $SE = 0.04$, $t(26) = –5.60$, $p = 7{\cdot}10^{-6}$), as well as after inactive control stimulation of the white matter ventromedial to the left-hemispheric hand motor area (Experiment IV: $b = –0.14$; $SE = 0.04$; $t(11) = –3.09$; $p = 0.010$). The same effect was observed following sound-only sham (Experiment I: $b = –0.14$; $SE = 0.05$; $t(11) = –3.18$; $p = .009$; Experiment II:

b = –0.22; SE = 0.04; $t(26)$ = –5.38; $p = 1 \cdot 10^{-5}$; Experiment III: 500ms-1kHz; b = –0.24; SE = 0.08; $t(15)$ = –2.86; p = 0.012). These results suggest a spatially non-specific effect of TUS that is related to the auditory confound (**Fig. 3**).

3.2. No dose-response effects of TUS on corticospinal inhibition

We further tested for direct ultrasonic neuromodulation by investigating a potential dose-response effect of TUS intensity (I_{sppa}) on motor cortical excitability. First, we applied TUS at multiple free-water stimulation intensities (**Fig. 4C**). In Experiment I, a linear mixed model with the factor 'intensity' (32.5/65 W/cm^2) did not reveal a significant effect of different on-target TUS intensities on motor excitability ($F(1,11)$ = 0.47, p = 0.509, η_p^2 = 0.04). In Experiment II, a linear mixed model with the factors 'stimulation site' (on-target/active control), 'masking' (no mask/masked), and 'intensity' (6.35/19.06 W/cm^2) similarly did not reveal an effect of stimulation intensity ($F(1,50)$ = 1.29, p = 0.261, η_p^2 = 0.03). Importantly, there was no effect of stimulation site ($F(1,168)$ = 1.75, p = 0.188, η_p^2 = 0.01), nor any significant interactions (all p-values > 0.1; all $\eta_p^2 < 0.06$). These results provide neither evidence for spatially specific neuromodulation when directly comparing stimulation sites, nor evidence for a dose-response relationship within the range of applied intensities.

However, it is likely that the effectiveness of TUS depends primarily on realized intracranial intensity, which we estimated with individualized 3D simulations (**Fig. 4A**). Yet, testing the relationship between estimated intracranial intensity and MEP amplitude change following on-target TUS similarly did not yield evidence for a dose-response effect (**Fig. 4B**, **Supplementary** Fig. 6).

Prior work has primarily focused on probing facilitatory or inhibitory effects on corticospinal excitability. Here, we also explored an alternative: how TUS might introduce noise to ongoing neural dynamics, rather than a directional modulation of excitability. Indeed, human TUS studies have often failed to show a global change in behavioral performance, instead finding TUS effects primarily around the perception threshold where noise might drive stochastic resonance (Butler et al., 2022; Legon et al., 2018). Whether the precise principles of stochastic resonance generalize from the perceptual domain to the current study is an open question, but it is known that neural noise can be introduced by brain stimulation (Van Der Groen & Wenderoth, 2016). It is likely that this noise is state-dependent and might not exceed the dynamic range of the intra-subject variability (Silvanto et al., 2007). Therefore, in an exploratory analysis, we exploited the natural structure in corticospinal excitability that exhibits as a strong temporal autocorrelation in MEP amplitude. Specifically, we tested how strongly the MEP on test trial t is predicted by the previous baseline trial t-1. As such, we quantified state-dependent autocorrelation between baseline MEP amplitude and MEP amplitude following on-target TUS, active control TUS, and sound-sham conditions (**Supplementary** Fig. 7). In brief, we found a significant interaction between previous baseline (t-1), stimulation site (on-target/active control), and intensity (6.35/19.06 W/cm^2; $F(1,30)$ = 12.10, p = 0.002, η_p^2 = 0.28) during masked trials. This interaction exhibited as increased autocorrelation for on-target TUS compared to active control TUS at 6.35 W/cm^2 (i.e., lower TUS-induced noise; $F(1,1287)$ = 13.43, $p = 3 \cdot 10^{-4}$, η_p^2 = 0.01), and reduced autocorrelation at 19.06 W/cm^2 (i.e., higher noise; $F(1,1282)$ = 5.76, p = 0.017, $\eta_p^2 = 4 \cdot 10^{-3}$; **Fig. 4D**). This effect was not only dependent upon intensity and stimulation site, but also dependent on the presence of auditory masking. As such, the effect was also observed in a four-way interaction of the previous baseline, site, intensity, and masking (**Supplementary** Fig. 7). These preliminary results might suggest that ultrasound stimulation can interact with ongoing neural dynamics by introducing temporally specific noise, rather than biasing the overall excitation/inhibition balance beyond its natural variation, but further work specifically designed to detect such effects is required.

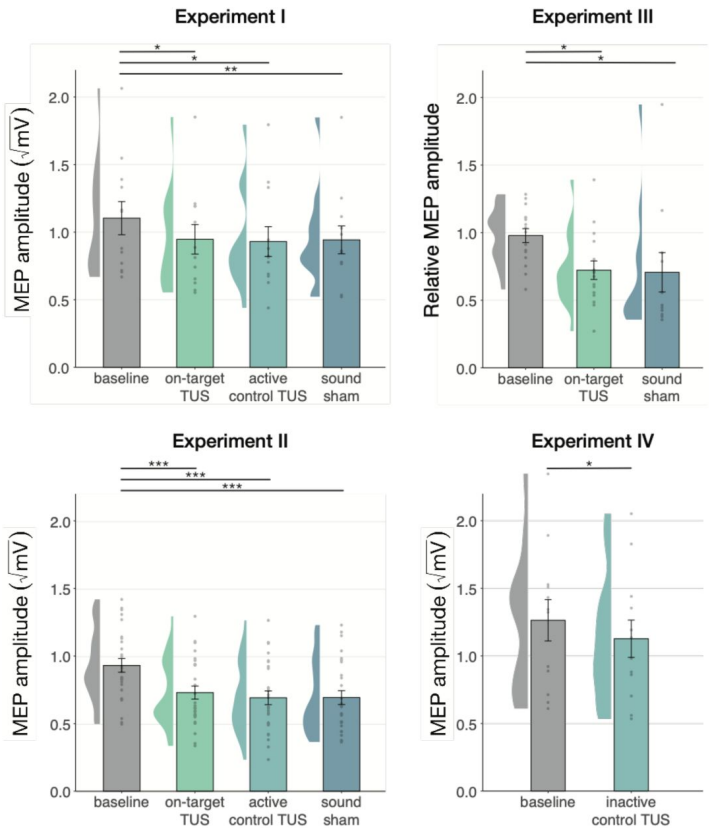

Fig. 3

| Non-specific motor inhibitory effects of TUS.

A significant suppression of MEP amplitude relative to baseline (gray) was observed for on-target TUS (green), but also for stimulation of a control region (cyan), and presentation of a sound alone (sound-sham; blue) indicating a spatially non-specific and sound-driven effect on motor cortical excitability. There were no significant differences between on-target and control conditions. Bar plots depict condition means, error bars represent standard errors, clouds indicate the distribution over participants, and points indicate individual participants. Square-root corrected MEP amplitudes are depicted for Experiments I, II, and IV, and *Relative MEP amplitude* is depicted for Experiment III (see ***Methods***). *p < 0.05, **p < 0.01, ***p < 0.001.

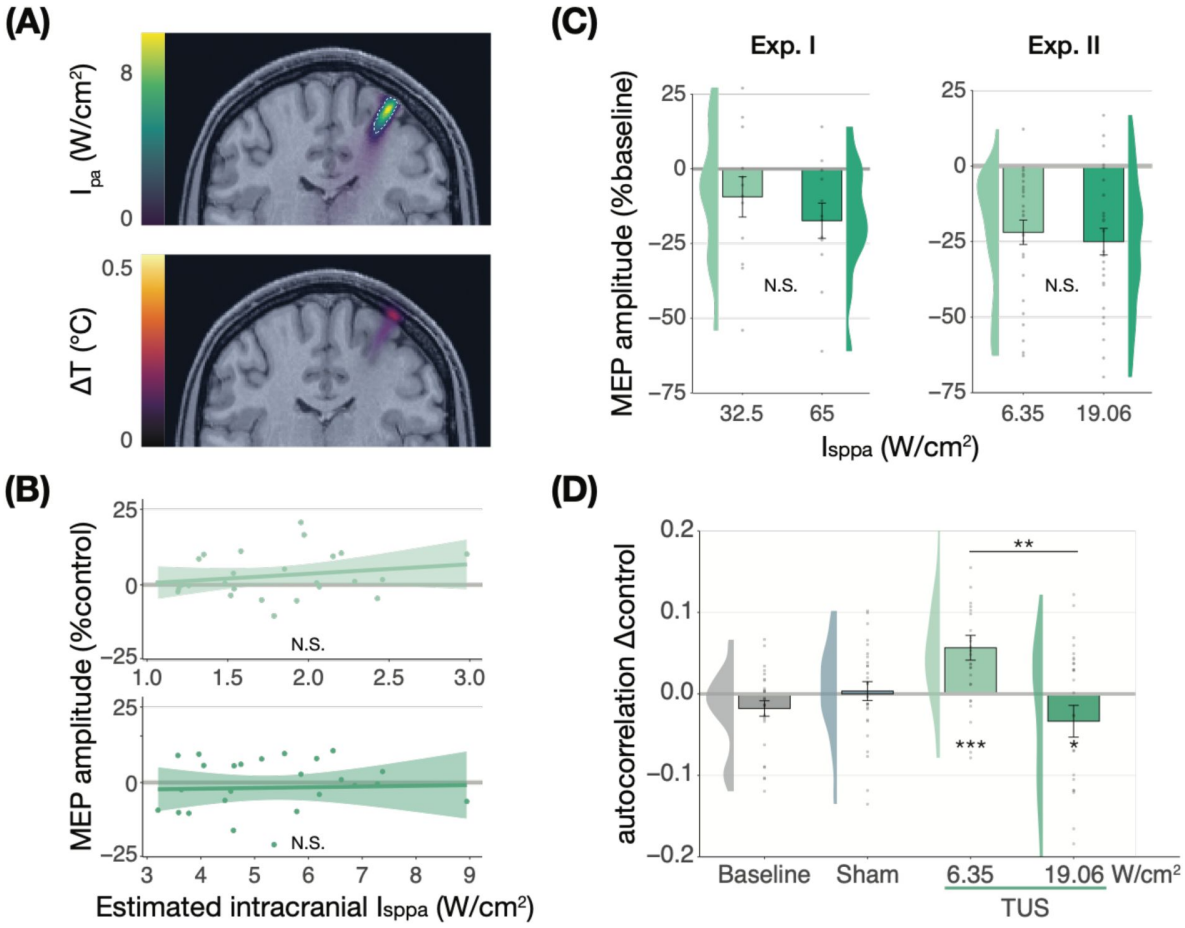

Fig. 4

| No significant dose-response effects of TUS.

(A) Acoustic (top) and thermal (bottom) simulations for a single subject in Experiment II. The acoustic simulation depicts estimated pulse-average intensity (I_{pa}) above a 0.15 W/cm^2 lower bound, with the dotted line indicating the full-width half-maximum of the pressure. The thermal simulation depicts maximum estimated temperature rise. **(B)** On-target TUS MEP amplitude as a percentage of active control MEP amplitude against simulated intracranial intensities at the two applied free-water intensities: 6.35 W/cm^2 (top) and 19.06 W/cm^2 (bottom). The shaded area represents the 95% CI, points represent individual participants. No significant intracranial dose-response relationship was observed. **(C)** There is no significant effect of free-water stimulation intensity on MEP amplitude. Values are expressed as a percentage of baseline MEP amplitude (square root corrected). Remaining conventions are as in **Fig. 3**. **(D)** Temporal autocorrelation, operationalized as the slope of the linear regression between trial *t* and its preceding baseline trial *t*-1, differed significantly as a function of stimulation site and intensity for masked trials. Individual points represent the differential autocorrelation compared to the active control site. Autocorrelation was not modulated during baseline or sound-only sham, but was significantly higher for on-target TUS at 6.35 W/cm^2, and significantly lower for on-target TUS at 19.06 W/cm^2 compared to active control TUS. *p < 0.05, **p < 0.01, ***p < 0.001.

3.3. Audible differences between stimulation sites do not underlie nonspecific inhibition

Stimulation over two separate sites could evoke distinct perceptual experiences arising from bone-conducted sound (Braun et al., 2020). To account for possible audibility differences between stimulation of on-target and active control sites in Experiments I and II, we also tested these conditions in the presence of a time-locked masking stimulus (**Fig. 5**). Following Experiment II, we additionally assessed the blinding efficacy of our masking stimuli, finding that the masking stimulus effectively reduced participant's ability to determine whether TUS was administered to approximately chance level (**Supplementary** Fig. 5).

In Experiment I, a linear mixed model with factors 'masking' (no mask/masked) and 'stimulation site' (on-target/active control) did not reveal a significant effect of masking ($F(1,11) = 0.01$, $p = 0.920$, $\eta_p^2 = 1{\cdot}10^{-5}$), stimulation site ($F(1,11) = 0.15$, $p = 0.703$, $\eta_p^2 = 0.01$), nor their interaction ($F(1,11) = 1{\cdot}10^{-3}$, $p = 0.971$, $\eta_p^2 = 1{\cdot}10^{-4}$). Similarly, in Experiment II, the linear mixed model described under the previous section revealed no significant main effect of masking ($F(1,30) = 1.68$, $p = 0.205$, $\eta_p^2 = 0.05$), nor any interactions (all p-values > 0.1; all $\eta_p^2 < 0.06$). These results indicate that an underlying specific neuromodulatory effect of TUS was not being obscured by audible differences between stimulation sites.

3.4. Sound-driven effects on corticospinal excitability

3.4.1. Duration and pitch

Prior research has shown that longer durations of TUS significantly inhibited motor cortical excitability (i.e., ≥400 ms; Fomenko et al., 2020), while shorter durations did not. In Experiment I, we applied on-target, active control, and sound-sham conditions at shorter and longer durations to probe this effect. When directly comparing these conditions at different stimulus durations (100/500 ms), no evidence for an underlying neuromodulatory effect of TUS was observed, in line with our aforementioned findings. Instead, a linear mixed model with factors 'condition' (on-target/active control/sound-sham) and 'stimulus duration' (100/500 ms) revealed only a significant main effect of (auditory) stimulus duration, where longer stimulus durations resulted in stronger MEP attenuation ($F(1,11) = 10.07$, $p = 0.009$, $\eta_p^2 = 0.48$). There was no significant effect of condition ($F(2,11) = 1.30$, $p = 0.311$, $\eta_p^2 = 0.19$), nor an interaction between stimulus duration and condition ($F(2,11) = 0.65$, $p = 0.543$, $\eta_p^2 = 0.11$). These results further show that the auditory confound and its timing characteristics, rather than ultrasonic neuromodulation, underlies the observed inhibition of motor cortical excitability (**Fig. 6A**).

We further tested auditory effects in Experiment III, where we administered sound-sham stimuli at four combinations of duration and pitch. A LMM with factors 'duration' (500/700 ms) and 'pitch' (1/12 kHz) revealed significantly lower MEPs following 500 ms auditory stimuli (**Fig. 6B**; duration: $F(1,15) = 7.12$, $p = 0.017$, $\eta_p^2 = 0.32$; pitch: $F(1,15) = 0.02$, $p = 0.878$, $\eta_p^2 = 2{\cdot}10^{-3}$; interaction: $F(1,15) = 2.23$, $p = 0.156$, $\eta_p^2 = 0.13$), supporting the role of auditory stimulus timing in perturbation of MEP amplitude.

Subsequently, ultrasonic stimulation was also administered alongside these four auditory stimuli. Here, a LMM with factors 'auditory stimulus duration' (500/700 ms), 'pitch' (1/12 kHz), and 'ultrasonic stimulation' (yes/no) revealed no significant effect of auditory stimulus duration in contrast to the first test ($F(1,15) = 0.44$, $p = 0.517$, $\eta_p^2 = 0.03$). However, a 1 kHz pitch resulted in significantly lower MEP amplitudes than a 12 kHz pitch (**Fig 6C**; $F(1,15) = 4.94$, $p = 0.042$, $\eta_p^2 = 0.25$). Importantly, we find no evidence for ultrasonic neuromodulation, where both on-target TUS and sound-sham reduced MEP amplitude from baseline (**Fig. 3**), and where applying on-target TUS did not significantly affect MEP amplitude as compared to sound-sham ($F(1,15) = 0.42$, $p = 0.526$, $\eta_p^2 = 0.03$; **Fig. 6C**). We observed a nonsignificant trend for the interaction between

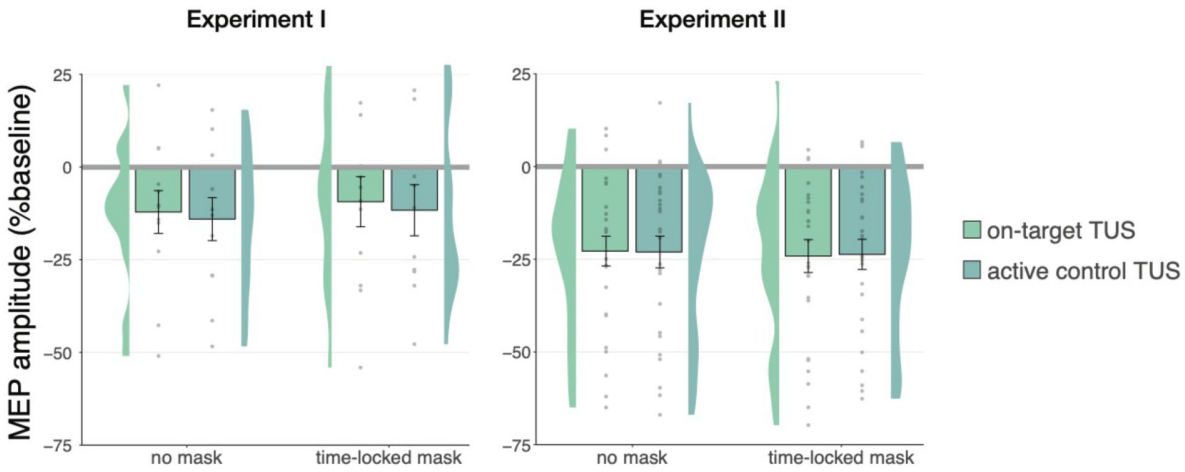

Fig. 5

| No effects of time-locked masking.

There were no significant effects of time-locked masking, indicating that audible differences between stimulation sites did not obscure or explain the absence of direct neuromodulation. Conventions are as in Figs. **3** and **4C**.

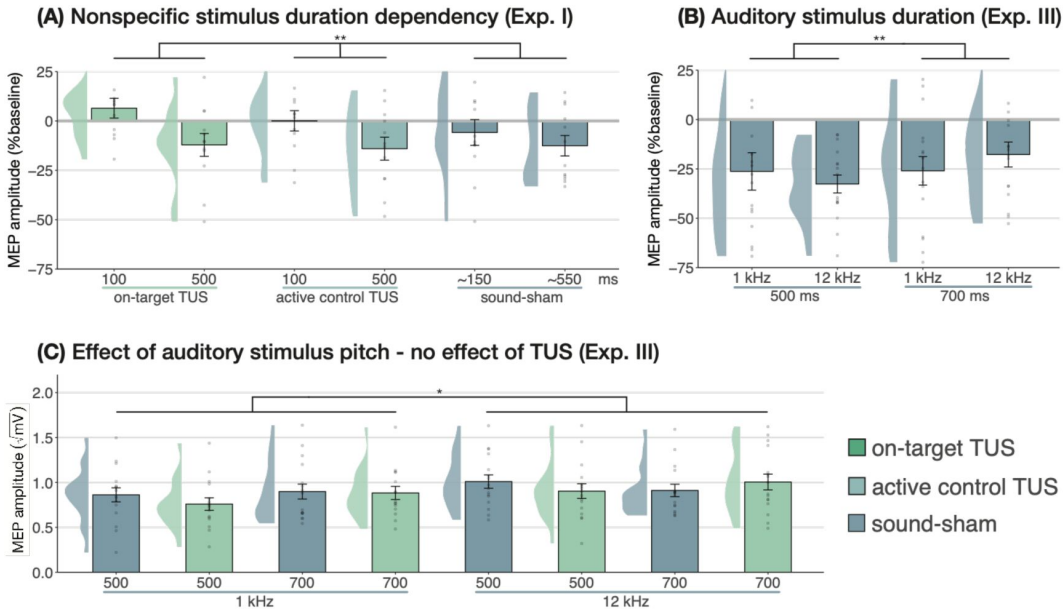

Fig. 6

| Sound-driven effects on corticospinal excitability.

(A) Longer (auditory) stimulus durations resulted in lower MEP amplitudes, regardless of TUS administration, indicating a sound-duration-dependency of motor inhibitory outcomes (Exp. I). **(B)** A significant effect of auditory stimulus duration was also observed in Experiment III. **(C)** The pitch of auditory stimuli also affected MEPs, where lower amplitudes were observed following a 1 kHz tone. There was no effect of TUS. Conventions are as in Figs. **3** and **4C**.

'ultrasonic stimulation' and 'auditory stimulus duration' ($F(1,15) = 4.22$, $p = 0.058$, $\eta_p^2 = 0.22$). No trends were observed for the remaining interactions between these three factors (all $\eta_p^2 < 0.06$, $p > 0.3$). Taken together, these results do not provide evidence for direct ultrasonic neuromodulation but support the influence of auditory stimulation characteristics on motor cortical excitability.

3.4.2. TUS audibility and confound volume

In Experiment IV, we applied TUS to an inactive target – the white matter ventromedial to the left-hemispheric hand motor area – both with and without a continuous auditory masking stimulus. MEP amplitudes did not significantly differ in baseline conditions regardless of whether a continuous sound was being played ($b = 0.03$, $SE = 0.06$, $t(11) = 0.52$, $p = 0.616$), indicating that continuous auditory stimulation alone might not be sufficient to inhibit MEP amplitude.

We additionally applied stimulation at multiple intensities to isolate the effect of auditory confound volume. A linear mixed model with factors 'masking' (no mask/masked) and 'intensity' (4.34/8.69/10.52 Wcm^{-2}) with a random intercept and slope for each factor revealed a significant interaction ($F(2,4038) = 3.43$, $p = 0.033$, $\eta_p^2 = 2{\cdot}10^{-3}$) and an accompanying effect of 'masking' with lesser MEP attenuation when stimulation was masked ($F(1,11) = 11.84$, $p = 0.005$, $\eta_p^2 = 0.52$; **Fig. 7**). Follow-up comparisons revealed significantly less attenuation for masked stimulation at 4.34 W/cm^2 intensity ($F(1,11) = 13.02$, $p = 0.004$, $\eta_p^2 = 0.55$), and a nonsignificant trend for the higher intensities (8.69 W/cm^2: $F(1,11) = 3.87$, $p = 0.077$, $\eta_p^2 = 0.27$; 10.52 W/cm^2: $F(1,11) = 3.47$, $p = 0.089$, $\eta_p^2 = 0.24$). In direct comparisons to baseline, all conditions resulted in a significant inhibition of MEP amplitude (all $t < -3.36$, all $p < 0.007$), with the exception of continuously masked stimulation at the lowest volume, with an intensity of 4.34 W/cm^2 I_{sppa} ($b = -0.06$, $SE = 0.03$, $t(11) = -2.04$, $p = 0.065$).

The data indicate that continuous masking reduces motor inhibition, likely by minimizing the audibility of TUS, particularly when applied at a lower stimulation intensity (i.e., auditory confound volume). The remaining motor inhibition observed during masked trials likely owes to, albeit decreased, persistent audibility of TUS during masking. Indeed, MEP attenuation in the masked conditions descriptively scale with participant reports of audibility. This points towards a role of auditory confound volume in motor inhibition (***Supplementary*** Fig. 8). Nevertheless, one could instead argue that evidence for direct neuromodulation is seen here. This is unlikely for a number of reasons. First, white matter contains a lesser degree of mechanosensitive ion channel expression and there is evidence that neuromodulation of these tracts may occur primarily in the thermal domain (Guo et al., 2022; Sorum et al., 2021). Second, Experiment IV lacks sufficient inferential power in the absence of an additional control and must therefore be interpreted in tandem with Experiments I-III. These experiments revealed no evidence for direct neuromodulation using equivalent or higher stimulation intensities and directly targeting grey matter while also using multiple control conditions. Therefore, we propose that persistent motor inhibition during masked trials owes to continued, though reduced, audibility of the confound (***Supplementary*** Fig. 8). However, future work including an additional control (site) is required to definitively disentangle these alternatives.

3.4.3. Preparatory cueing of TMS

We find that MEP attenuation results from auditory stimulation rather than direct neuromodulation. Two putative mechanisms through which sound cuing may drive motor inhibition have been proposed, positing either that explicit cueing of TMS timing results in compensatory processes that drive MEP reduction (Capozio et al., 2021; Tran et al., 2021), or suggesting the evocation of a startle response that leads to global inhibition (Fisher et al., 2004; Furubayashi et al., 2000; Ilic et al., 2011; Kohn et al., 2004; Wessel & Aron, 2013). Critically, we can dissociate between these theories by exploring the temporal dynamics of MEP attenuation. One would expect a startle response to habituate over time, where MEP attenuation would be reduced during startling initial trials, followed by a normalization throughout the course of the

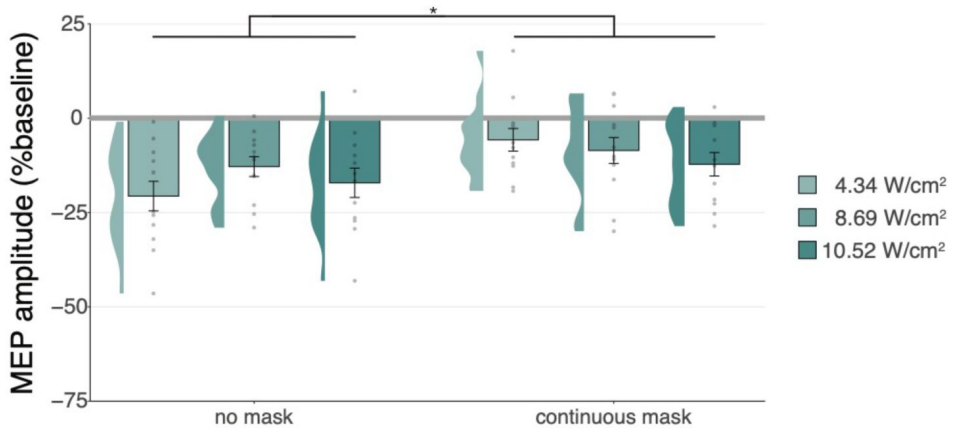

Fig. 7

| Sound-driven effects on corticospinal excitability.

Less MEP attenuation was measured during continuous masking, particularly for lower stimulation intensities (i.e., auditory confound volumes), pointing towards a role of TUS audibility in MEP attenuation.

experiment. Alternatively, if temporally contingent sound-cueing of TMS drives inhibition, MEP amplitudes should decrease over time as the relative timing of TUS and TMS is being learned, followed by a stabilization at a decreased MEP amplitude once this relationship has been learned.

In Experiments I and II, linear mixed models with 'trial number' as a predictor show significant changes in MEP amplitude throughout the experiment, pointing to a learning effect. Specifically, in Experiment I, a significant reduction in MEP amplitude was observed across the first 10 trials where a 500 ms stimulus was delivered (b = –0.04, SE = 0.01, $t(11)$ = –2.88, p = 0.015), following by a stabilization in subsequent blocks (b = $–2{\cdot}10^{-4}$, SE = $3{\cdot}10^{-4}$, $t(11)$ = –0.54, p = 0.601). This same pattern was observed in Experiment II, with a significant reduction across the first 20 trials (b = –0.01, SE = $3{\cdot}10^{-3}$, $t(26)$ = – 4.08, p = $4{\cdot}10^{-4}$), followed by stabilization (b = $6{\cdot}10^{-5}$, SE = $1{\cdot}10^{-4}$, $t(26)$ = 0.46, p = 0.650; **Fig. 8**). The data suggest that the relative timing of TUS and TMS is learned across initial trials, followed by a stabilization at a decreased MEP amplitude once this relationship has been learned. These results could reflect auditory cueing of TMS that leads to a compensatory expectation-based reduction of motor excitability.

Discussion

In this study, we show the considerable impact of auditory confounds during audibly pulsed TUS in humans. We employed improved control conditions compared to prior work across four experiments, one preregistered, at three independent institutions. Here, we disentangle direct neuromodulatory and indirect auditory contributions during ultrasonic neuromodulation of corticospinal excitability. While we corroborated motor inhibitory effects of online TUS (Fomenko et al., 2020; Legon, Bansal, et al., 2018; Xia et al., 2021), we demonstrated that this inhibition also occurs with stimulation of a control region or presentation of a sound alone, suggesting that the auditory confound rather than direct ultrasonic neuromodulation drives inhibition. Further, no direct neuromodulatory effects on overall excitability were observed, regardless of stimulation timing, intensity, or masking. However, we note that an exploratory investigation of temporal dynamics indicated ultrasound might introduce noise to the neural system. Importantly, we found convincing evidence that characteristics of auditory stimuli do globally affect motor excitability, where auditory cueing of TMS pulse timing can affect measures of corticospinal excitability. This highlights the importance of explicit cueing in TMS experimental design. Most importantly, our results call for a reevaluation of earlier findings following audible TUS, and highlight the importance of suitable controls in experimental design (Bergmann & Hartwigsen, 2021; Siebner et al., 2022).

No evidence for direct neuromodulation by TUS

Prior studies have highlighted sound-driven effects of TUS in behavioral and electrophysiological research (Airan & Butts Pauly, 2018; Braun et al., 2020; Guo et al., 2018; Johnstone et al., 2021; Sato et al., 2018). Here, we assessed whether the auditory confound of a conventional 1000 Hz pulsed protocol might underlie motor inhibitory effects, which are among the most robust and replicable human findings (Fomenko et al., 2020; Legon, Bansal, et al., 2018; Xia et al., 2021). While we successfully replicated this inhibitory effect, we found the same inhibition following stimulation of a motor control site (contralateral, active) and stimulation of a white-matter control site (ipsilateral, inactive; **Fig. 3**). This contrasts with a prior TUS-TMS study which found that TUS of the contralateral hand motor area did not change motor cortical excitability (Xia et al., 2021). Indeed, in all direct comparisons between on-target and control stimulation, no differences in excitability were observed, pointing towards a spatially nonspecific effect of TUS. Considering further inhibitory effects following administration of an auditory stimulus alone, the data suggest that online TUS motor inhibition is largely driven by the salient auditory confound,

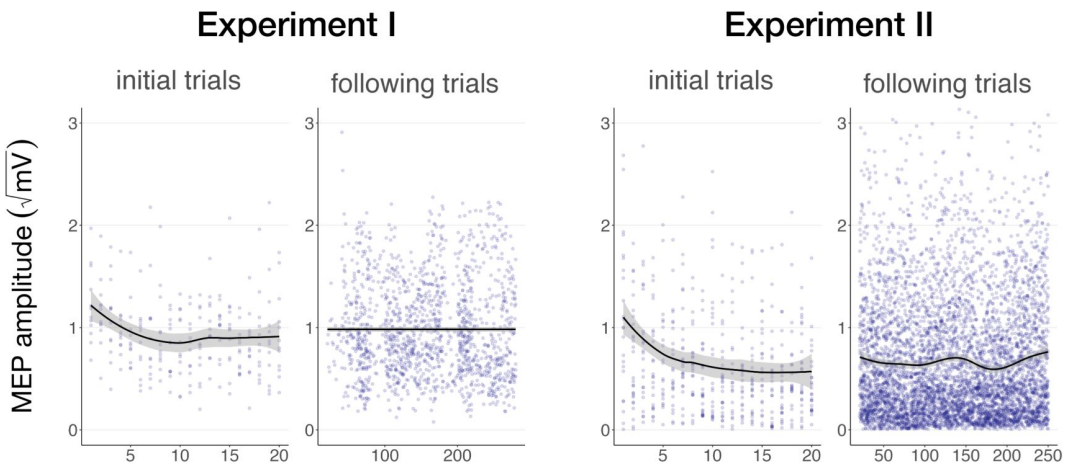

Fig. 8

| Auditory cueing of TMS.

There was a significant reduction in MEP amplitude when participants were first presented with a 500 ms stimulus (initial trials) in Experiment I (left) and Experiment II (right), following by a stabilization of MEP amplitude during the rest of the experiment (following trials), indicating a learning process by which TUS acts as a cue signaling the onset of TMS. The solid line depicts the loess regression fit, and the shaded area represents the 95% confidence interval.

rather than spatially specific and direct neuromodulation. However, an exploratory analysis that tested for effects beyond a global shift in excitation-inhibition balance revealed that TUS might interact with ongoing neural dynamics by introducing dose-dependent noise (**Fig. 4D**).

We found no evidence of a dose-response relationship between TUS intensity (I_{sppa}) and motor inhibition when applying stimulation at a wide range of intensities, nor when testing the relationship between simulated intracranial intensities and changes in excitability (**Fig. 4A-C**). Similarly, administration of a time-locked auditory masking stimulus that effectively reduced TUS detection rates did not provide evidence of direct effects being obscured by audible differences between conditions (**Fig. 5**, **Supplementary** Fig. 5). Taken together, this study presents no evidence for direct and spatially specific TUS inhibition of motor excitability when applying a clearly audible protocol, despite using improved control conditions, higher stimulation intensities, and a larger sample size than prior studies (Fomenko et al., 2020; Legon, Bansal, et al., 2018; Xia et al., 2021). Building on these results, the current challenge is to develop efficacious neuromodulatory protocols with minimal auditory interference. Efforts in this direction are already underway (Mohammadjavadi et al., 2019; Nakajima et al., 2022; Zeng et al., 2022).

Sound-cued motor inhibition

Until now, it was unclear how TUS induced motor inhibition in humans. Here, we show that this inhibition is caused by peripheral auditory stimulation. It is well-known that MEPs are sensitive to both sensory and psychological factors (Duecker et al., 2013). For example, several studies find MEP attenuation following a startling auditory stimulus (Fisher et al., 2004; Furubayashi et al., 2000; Ilic et al., 2011; Kohn et al., 2004; Wessel & Aron, 2013), and have demonstrated the impact of stimulus duration and volume on this inhibition (Furubayashi et al., 2000). It is possible that a similar mechanism is at play for audible TUS protocols. Indeed, we observed modulation of motor cortical excitability dependent upon the characteristics of auditory stimuli, including their duration and timing (**Fig. 6A-B**), their pitch/frequency (**Fig. 6C**), and whether the confound was audible in general, including perceived volume (**Fig. 7**, **Supplementary** Fig. 8).

One possible interpretation of the observed MEP attenuation is that the auditory confound acts as a salient cue to predict the upcoming TMS pulse. Prediction-based attenuation has been reported in both sensory and motor domains (Ford et al., 2007; Tran et al., 2021). For example, MEPs are suppressed when the timing of a TMS pulse can be predicted by a warning cue (Capozio et al., 2021; Tran et al., 2021). In the current experimental setup, participants could also learn the relative timing of the auditory stimulus and the TMS pulse. Indeed, we observe MEP attenuation emerge across initial trials as participants learn when to expect TMS, until a stable (i.e., learned) state is reached (**Fig. 8**). Moreover, no motor inhibition was observed when TUS onset was inaudible or when stimulation timing was potentially too fast to function as a predictive cue (100 ms). Taken together, a parsimonious explanation is expectation-based inhibition of TMS-induced MEPs. This inhibitory response might either reflect inhibition of competing motor programs – a component of motor preparation – or a homeostatic process anticipating the TMS-induced excitation (Capozio et al., 2021; Tran et al., 2021).

Limitations

The precise biomolecular and neurophysiological mechanisms underlying ultrasonic neuromodulation remain under steadily progressing investigation (Weinreb & Moses, 2022; Yoo et al., 2022). A shared interpretation is that mechano-electrophysiological energy transfer is proportional to acoustic radiation force, and thus proportional to stimulation intensity. Accordingly, one could argue that the TUS dose in the present study could have been insufficient to evoke direct neuromodulation. Indeed, despite the applied intensities exceeding prior relevant human work (Fomenko et al., 2020; Legon, Bansal, et al., 2018; Xia et al., 2021) the total

applied neuromodulatory doses are relatively low as compared to, for example, repetitive TUS protocols (rTUS) in animal work (Folloni et al., 2019; Verhagen et al., 2019) or recent human studies (Nakajima et al., 2022).

Alternatively, insufficient neural recruitment could be attributed to stimulation parameters other than intensity. If so, the absence of direct neuromodulation across these experiments might not generalize to parameters outside the tested set. For example, while we replicated and extended prior work targeting the hand motor area at ~30 mm from the scalp (Fomenko et al., 2020; Legon, Bansal, et al., 2018; Xia et al., 2021), other studies have suggested that the optimal stimulation depth to engage the hand motor area may be more superficial (Osada et al., 2022; Siebner et al., 2022).

One might further argue that the TMS hotspot provides insufficient anatomical precision to appropriately target the underlying hand muscle representation with TUS. The motor hotspot may not precisely overly the cortical representation of the assessed muscle due to the increased coil-cortex distance introduced by the TUS transducer. This distance, and the larger TMS coils required to evoke consistent MEPs, results in a broad electric field that is substantially larger than the TUS beam width (e.g., 6 mm for 250 kHz; Fomenko et al., 2020; Legon, Bansal, et al., 2018). Thus, it is possible that a transducer aligned with the center of the TMS coil may not be adequate. Nevertheless, we note that previous work utilizing a similar targeting approach has effectively induced changes in corticospinal motor excitability (Zeng et al., 2022). We also note that our stimulation depth and targeting procedures were comparable to all prior TUS-TMS studies, and that our simulations confirmed targeting (**Fig. 3A**, **Supplementary** Fig. 4). In summary, our main finding that the auditory confound drove motor inhibition in the present study, and likely had an impact in previous studies, holds true.

Considerations and future directions

Crucially, our results do not provide evidence that TUS is globally ineffective at inducing neuromodulation. While the present study and prior research highlight the confounding role of indirect auditory stimulation during pulsed TUS, there remains strong evidence for the efficacy of ultrasonic stimulation in animal work when auditory confounds are accounted for (Mohammadjavadi et al., 2019), or in controlled in-vitro systems such as an isolated retina, brain slices, or neuronal cultures in which the auditory confound carries no influence (Menz et al., 2013; Tyler et al., 2018).

It follows that where an auditory confound could be expected, appropriate control conditions are critical. These controls could involve stimulating a control region, and/or including a matched sound-only sham. In parallel, or perhaps alternatively, the impact of this confound can be mitigated in several ways. First, we recommend that the influence of auditory components be considered in transducer design and selection. Second, masking the auditory confound can help blind participants to experimental conditions. Titrating auditory mask quality per participant to account for intra– and inter-individual differences in subjective perception of the auditory confound would be beneficial. Here, the method chosen for mask delivery must be considered. While bone-conducting headphones align with the bone conduction mechanism of the auditory confound, they might not deliver sound as clearly as in-ear headphones or speakers. Nevertheless, the latter two rely on air-conducted sound. Notably, in-ear headphones could even amplify the perceived volume of the confound by obstructing the ear canal. Importantly, even when using masking stimuli, auditory stimulation could still influence cognitive task performance, among other measures. Alternative approaches could circumvent auditory confounds by testing deaf subjects, or perhaps more practically by ramping the ultrasonic pulse to minimize or even eliminate the auditory confound. This approach still requires validation and will only be relevant for protocols with pulses of sufficient duration. Here, one can expect that the experimental control required to account for auditory confounds might also hold for alternative peripheral effects, such as somatosensory confounds. Longer pulse durations are common in offline rTUS paradigms (Zeng

et al., 2022), with more opportunity for inaudible pulse shaping and the added benefit of separating the time of stimulation from that of measurement. However, appropriate control conditions remain central to make inferences on interventional specificity.

Conclusion

Transcranial ultrasonic stimulation is rapidly gaining traction as a promising neuromodulatory technique in humans. For TUS to reach its full potential we must identify robust and effective stimulation protocols. Here, we demonstrate that one of the most reliable findings in the human literature – online motor cortical inhibition during a 1000 Hz pulsed protocol – primarily stems from an auditory confound rather than direct neuromodulation. Instead of driving overall inhibition, we found preliminary evidence that TUS might introduce noise to ongoing neural dynamics. Future research must carefully account for peripheral confounding factors to isolate direct neuromodulatory effects of TUS, thereby enabling the swift and successful implementation of this technology in both research and clinical settings.

Data availability

Data and code for the current study can be accessed by reviewers here. An explanation of data access during the review processes can be found on this page.

Declaration of competing interest

Umair Hassan is the head of software development at sync2brain GmbH, where the bossdevice used in Experiment IV was developed. All other authors declare that no competing interests exist.

Acknowledgements

Experiments I & II were supported by the Dutch Research Council (NWO), awarding VIDI fellowships to L.V. (18919) and H.E.M.d.O. (452-17-016), and the Topsector programme Holland High Tech (HiTMaT-38H3). We thank Brittany van Beek for assistance with data acquisition and administration, Sarmad Peymaei for contributions to technical piloting, Norbert Hermesdorf for developing the 3D-printed TUS-TMS clamp, and Gerard van Oijen and Pascal de Water for their technical support. Experiment III was funded by the Canadian Institutes of Health Research (FDN 154292, ENG 173742) and the Natural Science and Engineering Research Council of Canada (RGPIN-2020-04176, RTI-2020-0024). Experiment IV and T.N. were supported by a grant from the Boehringer Ingelheim Foundation to T.O.B.

Author ORCIDs

CRediT authorship contribution statement

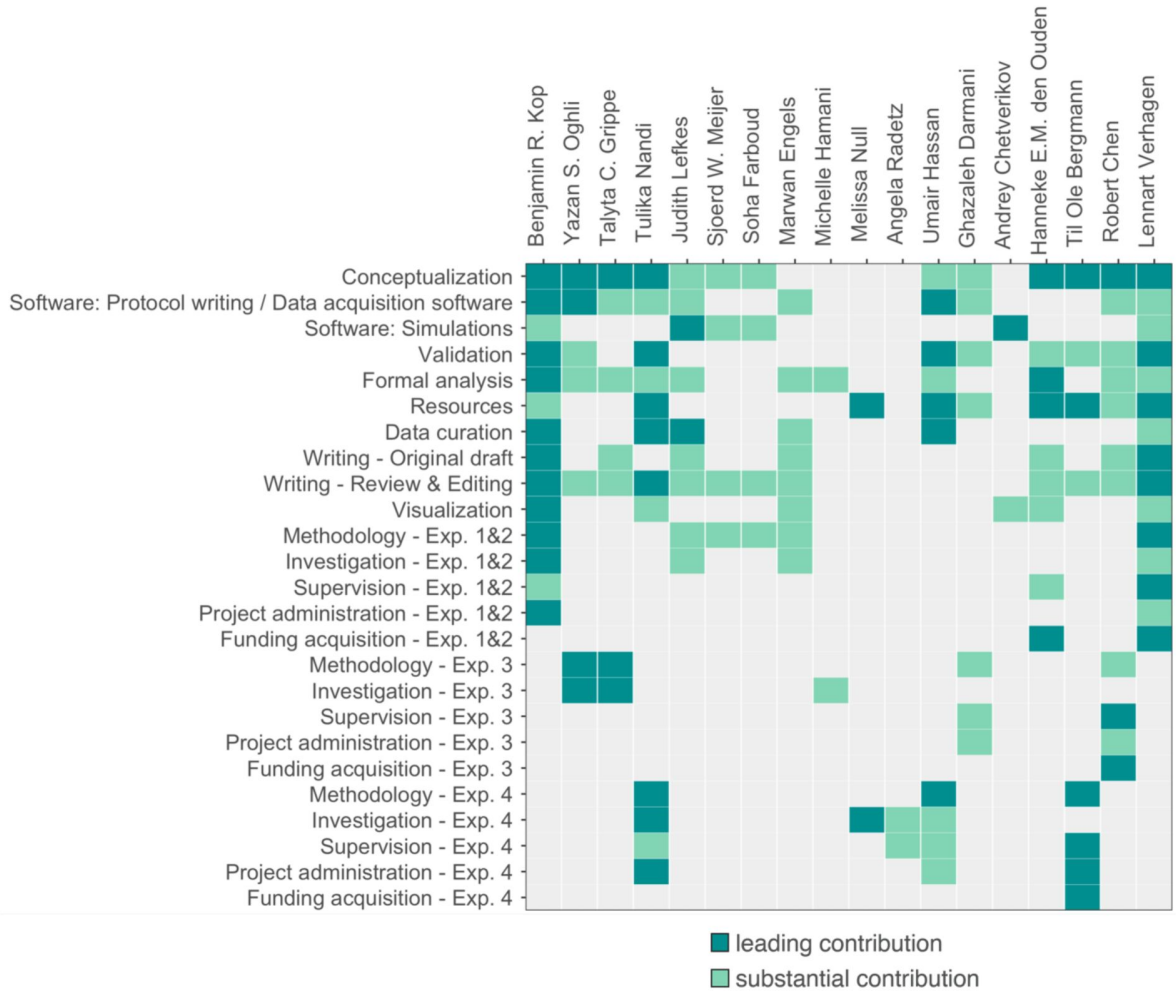

Fig. 6

| Contribution diagram.

This figure depicts the involvement of each author using the CRediT taxonomy (Brand et al., 2015) and categorizes their contributions according to three levels represented by color: 'none' (gray), 'substantial contribution' (light green), 'leading contribution' (dark green).

Author	ORCID
Benjamin R. Kop	https://orcid.org/0000-0001-7817-5845
Yazan Shamli Oghli	https://orcid.org/0000-0002-9206-6714
Talyta C. Grippe	https://orcid.org/0000-0003-3126-8002
Tulika Nandi	https://orcid.org/0000-0001-8161-3770
Judith Lefkes	https://orcid.org/0000-0003-0813-4049
Sjoerd W. Meijer	https://orcid.org/0000-0003-0698-4525
Soha Farboud	https://orcid.org/0000-0003-2774-3907
Marwan Engels	https://orcid.org/0000-0002-2326-3730
Michelle Hamani	*NA*
Melissa Null	https://orcid.org/0000-0002-4147-7214
Angela Radetz	https://orcid.org/0000-0003-2622-3221
Umair Hassan	https://orcid.org/0000-0001-8245-0061
Ghazaleh Darmani	https://orcid.org/0000-0002-4073-1911
Andrey Chetverikov	https://orcid.org/0000-0003-2767-6310
Hanneke E.M. den Ouden	https://orcid.org/0000-0001-7039-5130
Til Ole Bergmann	https://orcid.org/0000-0002-2282-6618
Robert Chen	https://orcid.org/0000-0002-8371-8629
Lennart Verhagen	https://orcid.org/0000-0003-3207-7929

Appendix A. Supplementary information

Supplementary tables and figures can be found in Appendix A.

References

1. Ai L., Mueller J. K., Grant A., Eryaman Y., Legon W (2016) **Transcranial focused ultrasound for BOLD fMRI signal modulation in humans** *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* :1758–1761 https://doi.org/10.1109/EMBC.2016.7591057

2. Airan R. D., Butts Pauly K (2018) **Hearing out Ultrasound Neuromodulation** *Neuron* **98**:875–877 https://doi.org/10.1016/j.neuron.2018.05.031

3. Barr D. J., Levy R., Scheepers C., Tily H. J (2013) **Random effects structure for confirmatory hypothesis testing: Keep it maximal** *Journal of Memory and Language* **68**:255–278 https://doi.org/10.1016/j.jml.2012.11.001

4. Bates D., Mächler M., Bolker B., Walker S (2015) **Fitting Linear Mixed-Effects Models Using lme4** *Journal of Statistical Software* **67** https://doi.org/10.18637/jss.v067.i01

5. Bergmann T. O., Hartwigsen G (2021) **Inferring Causality from Noninvasive Brain Stimulation in Cognitive Neuroscience** *Journal of Cognitive Neuroscience* **33**:195–225 https://doi.org/10.1162/jocn_a_01591

6. Bestmann S., Walsh V (2017) **Transcranial electrical stimulation** *Current Biology* **27**:R1258–R1262 https://doi.org/10.1016/j.cub.2017.11.001

7. Brand A., Allen L., Altman M., Hlava M., Scott J (2015) **Beyond authorship: Attribution, contribution, collaboration, and credit** *Learned Publishing* **28**:151–155 https://doi.org/10.1087/20150211

8. Braun V., Blackmore J., Cleveland R. O., Butler C. R (2020) **Transcranial ultrasound stimulation in humans is associated with an auditory confound that can be effectively masked** *Brain Stimulation* **13**:1527–1534 https://doi.org/10.1016/j.brs.2020.08.014

9. Butler C. R., Rhodes E., Blackmore J., Cheng X., Peach R. L., Veldsman M., Sheerin F., Cleveland R. O (2022) **Transcranial ultrasound stimulation to human middle temporal complex improves visual motion detection and modulates electrophysiological responses. Brain Stimulation: Basic** *Translational, and Clinical Research in Neuromodulation* **15**:1236–1245 https://doi.org/10.1016/j.brs.2022.08.022

10. Cain J. A., Visagan S., Johnson M. A., Crone J., Blades R., Spivak N. M., Shattuck D. W., Monti M. M (2021) **Real time and delayed effects of subcortical low intensity focused ultrasound** *Scientific Reports* **11** https://doi.org/10.1038/s41598-021-85504-y

11. Capozio A., Chakrabarty S., Astill S (2021) **The Effect of Sound and Stimulus Expectation on Transcranial Magnetic Stimulation-Elicited Motor Evoked Potentials** *Brain Topography* **34**:720–730 https://doi.org/10.1007/s10548-021-00867-9

12. Conde V., Tomasevic L., Akopian I., Stanek K., Saturnino G. B., Thielscher A., Bergmann T. O., Siebner H. R (2019) **The non-transcranial TMS-evoked potential is an inherent source of ambiguity in TMS-EEG studies** *NeuroImage* **185**:300–312 https://doi.org/10.1016/j.neuroimage.2018.10.052

13. Darmani G. *et al.* (2022) **Non-invasive transcranial ultrasound stimulation for neuromodulation** *Clinical Neurophysiology* **135**:51–73 https://doi.org/10.1016/j.clinph.2021.12.010

14. Duecker F., de Graaf T. A., Jacobs C., Sack A. T. (2013) **Time– and Task-Dependent Non-Neural Effects of Real and Sham TMS** *PLoS ONE* **8** https://doi.org/10.1371/journal.pone.0073813

15. Fisher R. J., Sharott A, K□hn A. A., Brown P. (2004) **Effects of combined cortical and acoustic stimuli on muscle activity** *Experimental Brain Research* **157**:1–9 https://doi.org/10.1007/s00221-003-1809-6

16. Folloni D., Verhagen L., Mars R. B., Fouragnan E., Constans C., Aubry J.-F., Rushworth M. F. S., Sallet J (2019) **Manipulation of Subcortical and Deep Cortical Activity in the Primate Brain Using Transcranial Focused Ultrasound Stimulation** *Neuron* **101**:1109–1116 https://doi.org/10.1016/j.neuron.2019.01.019

17. Fomenko A. *et al.* (2020) **Systematic examination of low-intensity ultrasound parameters on human motor cortex excitability and behavior** *eLife* **9** https://doi.org/10.7554/eLife.54497

18. Ford J. M., Gray M., Faustman W. O., Roach B. J., Mathalon D. H (2007) **Dissecting corollary discharge dysfunction in schizophrenia** *Psychophysiology* **44**:522–529 https://doi.org/10.1111/j.1469-8986.2007.00533.x

19. Furubayashi T. *et al.* (2000) **The human hand motor area is transiently suppressed by an unexpected auditory stimulus** *Clinical Neurophysiology* **111**:178–183 https://doi.org/10.1016/S1388-2457(99)00200-X

20. Gavrilov L.R., Tsirulnikov E.M (2012) **Focused ultrasound as a tool to input sensory information to humans (Review)** *Acoustical Physics* **58** https://doi.org/10.1134/S1063771012010083

21. Guo H., Hamilton M., Offutt S. J., Gloeckner C. D., Li T., Kim Y., Legon W., Alford J. K., Lim H. H (2018) **Ultrasound Produces Extensive Brain Activation via a Cochlear Pathway** *Neuron* **98**:1020–1030 https://doi.org/10.1016/j.neuron.2018.04.036

22. Guo H., Offutt S. J., Hamilton Ii M., Kim Y., Gloeckner C. D., Zachs D. P., Alford J. K., Lim H. H (2022) **Ultrasound does not activate but can inhibit in vivo mammalian nerves across a wide range of parameters** *Scientific Reports* **12** https://doi.org/10.1038/s41598-022-05226-7

23. Ilic T. V., Pötter-Nerger M., Holler I., Siebner H. R., Ilic N. V., Deuschl G., Volkmann J (2011) **Startle Stimuli Exert Opposite Effects on Human Cortical and Spinal Motor System Excitability in Leg Muscles. Physiological Research** *S* :101–S106 https://doi.org/10.33549/physiolres.932182

24. Johnstone A., Nandi T., Martin E., Bestmann S., Stagg C., Treeby B (2021) **A range of pulses commonly used for human transcranial ultrasound stimulation are clearly audible** *Brain Stimulation* **14**:1353–1355 https://doi.org/10.1016/j.brs.2021.08.015

25. K□hn Andrea A., Sharott A., Trottenberg T., Kupsch A., Brown P. (2004) **Motor cortex inhibition induced by acoustic stimulation** *Experimental Brain Research* **158** https://doi.org/10.1007/s00221-004-1883-4

26. Lee W., Kim H.-C., Jung Y., Chung Y. A., Song I.-U., Lee J.-H., Yoo S.-S (2016) **Transcranial focused ultrasound stimulation of human primary visual cortex** *Scientific Reports* **6** https://doi.org/10.1038/srep34026

27. Legon W., Ai L., Bansal P., Mueller J. K (2018) **Neuromodulation with single-element transcranial focused ultrasound in human thalamus** *Human Brain Mapping* **39**:1995–2006 https://doi.org/10.1002/hbm.23981

28. Legon W., Bansal P., Tyshynsky R., Ai L., Mueller J. K (2018) **Transcranial focused ultrasound neuromodulation of the human primary motor cortex** *Scientific Reports* **8** https://doi.org/10.1038/s41598-018-28320-1

29. Liu C., Yu K., Niu X., He B (2021) **Transcranial Focused Ultrasound Enhances Sensory Discrimination Capability through Somatosensory Cortical Excitation** *Ultrasound in Medicine & Biology* **47**:1356–1366 https://doi.org/10.1016/j.ultrasmedbio.2021.01.025

30. Menz M. D., Oralkan O., Khuri-Yakub P. T., Baccus S. A (2013) **Precise neural stimulation in the retina using focused ultrasound** *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **33**:4550–4560 https://doi.org/10.1523/JNEUROSCI.3521-12.2013

31. Mohammadjavadi M., Ye P. P., Xia A., Brown J., Popelka G., Pauly K. B (2019) **Elimination of peripheral auditory pathway activation does not affect motor responses from ultrasound neuromodulation** *Brain Stimulation* **12**:901–910 https://doi.org/10.1016/j.brs.2019.03.005

32. Murphy K. R. *et al.* (2022) **A tool for monitoring cell type–specific focused ultrasound neuromodulation and control of chronic epilepsy** *Proceedings of the National Academy of Sciences* **119** https://doi.org/10.1073/pnas.2206828119

33. Nakajima K., Osada T., Ogawa A., Tanaka M., Oka S., Kamagata K., Aoki S., Oshima Y., Tanaka S., Konishi S (2022) **A causal role of anterior prefrontal-putamen circuit for response inhibition revealed by transcranial ultrasound stimulation in humans** *Cell Reports* **40** https://doi.org/10.1016/j.celrep.2022.111197

34. Osada T., Nakajima K., Ogawa A., Oka S., Kamagata K., Aoki S., Oshima Y., Tanaka S., Konishi S (2022) **Distributions of cortical depth of the index finger region in the M1: A representative depth parameter for transcranial ultrasound stimulation** *Brain Stimulation* **15**:1348–1350 https://doi.org/10.1016/j.brs.2022.09.012

35. Polanía R., Nitsche M. A., Ruff C. C (2018) **Studying and modifying brain function with non-invasive brain stimulation** *Nature Neuroscience* **21**:174–187 https://doi.org/10.1038/s41593-017-0054-4

36. Sato T., Shapiro M. G., Tsao D. Y (2018) **Ultrasonic Neuromodulation Causes Widespread Cortical Activation via an Indirect Auditory Mechanism** *Neuron* **98**:1031–1041 https://doi.org/10.1016/j.neuron.2018.05.009

37. Siebner H. R. *et al.* (2022) **Transcranial magnetic stimulation of the brain: What is stimulated? – A consensus and critical position paper** *Clinical Neurophysiology* **140**:59–97 https://doi.org/10.1016/j.clinph.2022.04.022

38. Silvanto J., Muggleton N. G., Cowey A., Walsh V (2007) **Neural adaptation reveals state-dependent effects of transcranial magnetic stimulation** *European Journal of Neuroscience* **25**:1874–1881 https://doi.org/10.1111/j.1460-9568.2007.05440.x

39. Sorum B., Rietmeijer R. A., Gopakumar K., Adesnik H., Brohawn S. G (2021) **Ultrasound activates mechanosensitive TRAAK K + channels through the lipid membrane** *Proceedings of the National Academy of Sciences* **118** https://doi.org/10.1073/pnas.2006980118

40. Tran D. M. D., McNair N. A., Harris J. A., Livesey E. J (2021) **Expected TMS excites the motor system less effectively than unexpected stimulation** *NeuroImage* **226** https://doi.org/10.1016/j.neuroimage.2020.117541

41. Tyler W. J., Lani S. W., Hwang G. M (2018) **Ultrasonic modulation of neural circuit activity** *Current Opinion in Neurobiology* **50**:222–231 https://doi.org/10.1016/j.conb.2018.04.011

42. Tyler W. J., Tufail Y., Finsterwald M., Tauchmann M. L., Olson E. J., Majestic C (2008) **Remote Excitation of Neuronal Circuits Using Low-Intensity, Low-Frequency Ultrasound** *PLoS ONE* **3** https://doi.org/10.1371/journal.pone.0003511

43. Verhagen L. *et al.* (2019) **Offline impact of transcranial focused ultrasound on cortical activation in primates** *eLife* **8** https://doi.org/10.7554/eLife.40541

44. Weinreb E., Moses E (2022) **Mechanistic insights into ultrasonic neurostimulation of disconnected neurons using single short pulses** *Brain Stimulation* **15**:769–779 https://doi.org/10.1016/j.brs.2022.05.004

45. Wessel J. R., Aron A. R (2013) **Unexpected Events Induce Motor Slowing via a Brain Mechanism for Action-Stopping with Global Suppressive Effects** *Journal of Neuroscience* **33**:18481–18491 https://doi.org/10.1523/JNEUROSCI.3456-13.2013

46. Xia X., Fomenko A., Nankoo J.-F., Zeng K., Wang Y., Zhang J., Lozano A. M., Chen R (2021) **Time course of the effects of low-intensity transcranial ultrasound on the excitability of ipsilateral and contralateral human primary motor cortex** *NeuroImage* **243** https://doi.org/10.1016/j.neuroimage.2021.118557

47. Yoo S., Mittelstein D. R., Hurt R. C., Lacroix J., Shapiro M. G (2022) **Focused ultrasound excites cortical neurons via mechanosensitive calcium accumulation and ion channel amplification** *Nature Communications* **13** https://doi.org/10.1038/s41467-022-28040-1

48. Zeng K., Darmani G., Fomenko A., Xia X., Tran S., Nankoo J., Shamli Oghli Y., Wang Y., Lozano A. M., Chen R (2022) **Induction of Human Motor Cortex Plasticity by Theta Burst Transcranial Ultrasound Stimulation** *Annals of Neurology, ana* **26294** https://doi.org/10.1002/ana.26294

Article and author information

Benjamin R. Kop
Donders Institute for Brain, Cognition, and Behaviour; Radboud University Nijmegen, the Netherlands
For correspondence: benjamin.kop@donders.ru.nl
ORCID iD: 0000-0001-7817-5845

Yazan Shamli Oghli
Krembil Research Institute, University Health Network; University of Toronto, Canada
ORCID iD: 0000-0002-9206-6714

Talyta C. Grippe
Krembil Research Institute, University Health Network; University of Toronto, Canada
ORCID iD: 0000-0003-3126-8002

Tulika Nandi
Neuroimaging Center; Johannes-Gutenberg University Medical Center Mainz, Germany
ORCID iD: 0000-0001-8161-3770

Judith Lefkes
Donders Institute for Brain, Cognition, and Behaviour; Radboud University Nijmegen, the Netherlands
ORCID iD: 0000-0003-0813-4049

Sjoerd W. Meijer
Donders Institute for Brain, Cognition, and Behaviour; Radboud University Nijmegen, the Netherlands
ORCID iD: 0000-0003-0698-4525

Soha Farboud
Donders Institute for Brain, Cognition, and Behaviour; Radboud University Nijmegen, the Netherlands
ORCID iD: 0000-0003-2774-3907

Marwan Engels
Donders Institute for Brain, Cognition, and Behaviour; Radboud University Nijmegen, the Netherlands
ORCID iD: 0000-0002-2326-3730

Michelle Hamani
Krembil Research Institute, University Health Network; University of Toronto, Canada

Melissa Null
Neuroimaging Center; Johannes-Gutenberg University Medical Center Mainz, Germany
ORCID iD: 0000-0002-4147-7214

Angela Radetz
Neuroimaging Center; Johannes-Gutenberg University Medical Center Mainz, Germany
ORCID iD: 0000-0003-2622-3221

Umair Hassan
Neuroimaging Center; Johannes-Gutenberg University Medical Center Mainz, Germany
ORCID iD: 0000-0001-8245-0061

Ghazaleh Darmani
Krembil Research Institute, University Health Network; University of Toronto, Canada
ORCID iD: 0000-0002-4073-1911

Andrey Chetverikov
Donders Institute for Brain, Cognition, and Behaviour; Radboud University Nijmegen, the Netherlands, Department of Psychosocial Science, Faculty of Psychology, University of Bergen, Bergen, Norway
ORCID iD: 0000-0003-2767-6310

Hanneke E.M. den Ouden
Donders Institute for Brain, Cognition, and Behaviour; Radboud University Nijmegen, the Netherlands
ORCID iD: 0000-0001-7039-5130

Til Ole Bergmann
Neuroimaging Center; Johannes-Gutenberg University Medical Center Mainz, Germany, Leibniz Institute for Resilience Research Mainz, Germany
ORCID iD: 0000-0002-2282-6618

Robert Chen
Krembil Research Institute, University Health Network; University of Toronto, Canada
ORCID iD: 0000-0002-8371-8629

Lennart Verhagen
Donders Institute for Brain, Cognition, and Behaviour; Radboud University Nijmegen, the Netherlands
ORCID iD: 0000-0003-3207-7929

Copyright

Editors

Reviewing Editor
Jonas Obleser
University of Lübeck, Lübeck, Germany

Senior Editor
Jonathan Roiser
University College London, London, United Kingdom

Reviewer #1 (Public Review):

Summary: The authors have used transcranial magnetic stimulation (TMS) and motor evoked potentials (MEPs) to determine whether the peripheral auditory confound arising from TUS can drive motor inhibition on its own. They gathered data from three international centers in four experiments testing:

- Experiment 1 (n = 11), two different TUS durations and intensities under sound masking or without.
- Experiment 2 (n = 27) replicates Exp 1 with different intensities and a fixed TUS duration of

500ms.
- Experiment 3 (n = 16) studies the effect of various auditory stimuli testing different duration and pitches while applying TUS in an active site, on-target or no TUS.
- Experiment 4 (n = 12) uses an inactive control site to reproduce the sound without effective neuromodulation, while manipulating the volume of the auditory confound at different US intensities with and without continuous sound masking.

Strengths: This study comes from three very strong groups in noninvasive brain stimulation with long experience in neuromodulation, multimodal and electrophysiological recordings. Although complex to understand due to slightly different methodologies across centers, this study provides quantitative evidence relating to the potential auditory confound in online TUS. The results are in line with reductions seen in motor-evoked responses during online 1kHz TUS, and remarkable efforts were made to isolate peripheral confounds from actual neuromodulation factors, highlighting the confounding effect of sound itself.

Weaknesses: However, there are some points that need attention. In my view, the most important are:

1. Despite the main conclusion of the authors stating that there is no dose-response effect of TUS on corticospinal inhibition, the point estimates for change in MEP and Ipssa indicate a more complex picture. The present data and analyses cannot rule out that there is a dose-response function which cannot be fully attributed to difference in sound (since the relationship in inversed, lower intracranial Isppa leads to higher MEP decrease). These results suggest that dose-response function needs to be further studied in future studies.

2. Other methods to test or mask the auditory confound are possible (e.g., smoothed ramped US wave) which could substantially solve part of the sound issue in future studies or experiments in deaf animals etc.

https://doi.org/10.7554/eLife.88762.2.sa1

Reviewer #2 (Public Review):

Summary:

This study aims to test auditory confounds during transcranial ultrasound stimulation (TUS) protocols that rely on audible frequencies. In several experiments, the authors show that a commonly observed suppression of motor-evoked potentials (MEP) during TUS can be explained by acoustic stimulation. For instance, not only target TUS, but also stimulation of a control site and acoustic stimulation led to suppressed MEP.

The authors have convincingly addressed all of my comments and provided useful additional details. I believe that this is a strong study that will impact the field. Thanks also for making the sound stimuli open-source.

https://doi.org/10.7554/eLife.88762.2.sa0

Author Response

We are grateful for the insightful suggestions and comments provided by the reviewers. Your constructive feedback has been valuable, and we are thankful for the opportunity to address each point.

We appreciate both reviewers' recognition of our devotion to rigorous methodology and experimental control in this study, as evidenced by the comments: "remarkable efforts were

made to isolate peripheral confounds", "a clear strength of the study is the multitude of control conditions … that makes results very convincing", and "thorough design of the study". Indeed, we hope to have provided more than solid, but compelling evidence for sound-driven motor inhibitory effects of online TUS. We hope that this will be reflected in the assessment. Our conclusions are supported by multiple experiments across multiple institutions using exemplary experimental control including (in)active controls and multiple sound-sham conditions. This contrasts with the sole use of flip-over sham or no-stimulation conditions used in the majority of work to date. Indeed, the current study communicates that substantiated inferences on the efficacy of ultrasonic neuromodulation cannot be made under insufficient experimental control.

In response to the reviewers' comments, we have substantially changed our manuscript. Specifically, we have open-sourced the auditory masking stimuli and specified them in better detail in the text, we have improved the figures to reflect the data more closely, we have clarified the intracranial doseresponse relationship, we have elaborated in the introduction, and we have further discussed the possibility of direct neuromodulation. We hope that you agree these changes have helped to substantially improve the manuscript.

> ***Public reviews***
>
> *1.1) Despite the main conclusion of the authors stating that there is no dose-response effects of TUS on corticospinal inhibition, both the comparison of Isppa and MEP decrease for Exp 1 and 2, and the linear regression between MEP decrease (relative to baseline) and the estimated Isppa are significant, arguing the opposite, that there is a dose-response function which cannot be fully attributed to difference in sound (since the relationship in inversed, lower intracranial Isppa leads to higher MEP decrease). These results suggest that doseresponse function needs to be further studied in future studies.*

We thank the reviewer for bringing up this point. While we are convinced our study provides no evidence for a direct neuromodulatory dose-response relationship, we have realized that the manuscript could benefit from improved clarity on this point.

A dose-response relationship between TUS intensity and motor cortical excitability was assessed by manipulating free-water Isppa (Figure 4C). Here, no significant effect of free-water stimulation intensity was observed for Experiment I or II, thus providing no evidence for a dose-response relationship (Section 3.2). To aid in clarity, 'N.S.' has been added to Figure 4C in the revised manuscript.

However, it is likely that the efficacy of TUS would depend on realized intracranial intensity, which we estimated with 3D simulations for on-target stimulation. These simulations resulted in an estimated intracranial intensity for each applied free-water intensity (i.e., 6.35 and 19.06 W/cm2), for each participant. We then tested whether inter-individual differences in intracranial intensity during on-target TUS affected MEP amplitude. We have realized that the original visualization used to display these data and its explanation was unintuitive. Therefore, we have completely revised Supplementary Figure 6. Because of the substantial length of this section, we have not copied it here. Please see the Supplementary material for the implemented improvements.

In brief, we now show MEP amplitudes on the y-axis, rather than expressing values a %change. This plot depicts how individuals with higher intracranial intensities during ontarget TUS exhibit higher MEP amplitudes. However, this same relationship is observed for active control and sound-sham conditions. If there were a direct neuromodulatory doseresponse relationship of TUS, this would be reflected as the difference between on-target and control conditions changing as the estimated intracranial intensity increases. This was not the case. Further, the fact that the difference between on-target stimulation and baseline changes across intracranial intensities is notable, but this occurs to an equal degree in the

control conditions. Therefore, these data cannot be interpreted as evidence for a doseresponse relationship.

We hope the changes in Supplementary Figure 6 will make it clear that there is no evidence for direct intracranial dose-response effects.

> *1.2) Other methods to test or mask the auditory confound are possible (e.g., smoothed ramped US wave) which could substantially solve part of the sound issue in future studies or experiments in deaf animals etc…*

We agree with the reviewer's statement. We aimed to replicate the findings of online motor cortical inhibition reported in prior work using a 1000 Hz square wave modulation frequency. While ramping can effectively reduce the auditory confound, as noted in the discussion, this is not feasible for the short pulse durations (0.1-0.3 ms) employed in the current study (Johnstone et al., 2021). We have further clarified this point in the methods section of the revised manuscript as follows:

"While ramping the pulses can in principle mitigate the auditory confound (Johnstone et al., 2021; Mohammadjavadi et al., 2019), doing so for such short pulse durations (<= 0.3 ms) is not effective. Therefore, we used a rectangular pulse shape to match prior work."

Mitigation of the auditory confound by testing deaf subjects is a valid approach, and has now been added to the revised manuscript in the discussion as follows:

"Alternative approaches could circumvent auditory confounds by testing deaf subjects, or perhaps more practically by ramping the ultrasonic pulse to minimize or even eliminate the auditory confound."

> *1.3) Dose-response function is an extremely important feature for a brain stimulation technique. It was assessed in Exp II by computing the relationship between the estimated intracranial intensities and the modulation of corticospinal excitability (Fig. 3b, 3c). It is not clear why data from Experiment I could not be integrated in a global intracranial dose-response function to explore wider ranges of intracranial intensities and MEP variability.*

We chose not to combine data from Experiment 1 in a global intracranial dose-response function because TUS was applied at different fundamental frequencies and focal depths (Experiment I: 500 kHz, 35 mm; Experiment II: 250 kHz, 28 mm). We have now explicitly communicated this under Supplementary Figure 6:

"It was not appropriate to combine data from Experiments I and II given the different fundamental frequencies and stimulation depths applied… we ran simple linear models for Experiment II, which had a sufficient sample size (n = 27) to assess inter-individual variability."

> *1.4) Furthermore, the dose response function as computed with the MEP change relative to baseline shows a significant effect (6.35W/cm2) or a trend (19.06 W/cm2) for a positive linear relationship. This comparison cannot disentangle the auditory confound from the pure neuromodulatory effect but given the direction of the relationship (lower Isppa associated with larger neuromodulatory effect), it is unlikely that it is driven by sound. This relationship is absent for the Active control condition or the Sound Sham condition, more or less matched for peripheral confound. This needs to be further discussed.*

Please refer to point 1.1

> *1.5) The clear auditory confound arises from TUS pulsing at audible frequencies, which can be highly subject to inter-individual differences. Did the authors individually titrate the auditory mask to account for this intra- and inter-individual variability in auditory perception?*

In Experiments I-III, the auditory mask was identical between participants. In Experiment IV, the auditory mask volume and signal-to-noise ratio were adjusted per participant. In the discussion we recommend individualized mask titration. However, we do note that masking successfully blinded participants in Experiment II, despite using uniform masking stimuli (Supplementary Figure 5).

> *1.6) How different is the masking quality when using bone-conducting headphones (e.g., Exp. 1) compared to in-ear headphones (e.g., Exp. 2)?*

In our experience, bone conducting headphones produce a less clear, fuzzier, sound than in-ear headphones. However, in-ear headphones block the ear canal and likely result in the auditory confound being perceived as louder. We have included this information in the discussion of the revised manuscript:

“Titrating auditory mask quality per participant to account for intra- and inter-individual differences in subjective perception of the auditory confound would be beneficial. Here, the method chosen for mask delivery must be considered. While bone-conducting headphones align with the bone conduction mechanism of the auditory confound, they might not deliver sound as clearly as in-ear headphones or speakers. Nevertheless, the latter two rely on airconducted sound. Notably, in-ear headphones could even amplify the perceived volume of the confound by obstructing the ear canal.”

> *1.7) I was not able to find any report on the blinding efficacy of Exp. 1. Do the authors have some data on this?*

We do not have blinding data available for Experiment I. Following Experiment I, we decided it would be useful to include such an assessment in Experiment II.

> *1.8) Was the possibility to use smoothed ramped US wave form ever tested as a control condition in this set of studies, to eventually reduce audibility? For such fast PRF, for fast PRF, the slope would still need to be steep to stimulate the same power (AUC), it might not be as efficient.*

We indeed tested smoothing (ramping) the waveform. There was no perceptible impact on the auditory confound volume. Indeed, prior research has also indicated that ramping over

such short pulse durations is not effective (Johnstone et al., 2021). Taken together, we chose to continue with a square wave modulation as in prior TUS-TMS studies. We have updated the methods section of the manuscript with the following:

“While ramping the pulses can in principle mitigate the auditory confound (Johnstone et al., 2021; Mohammadjavadi et al., 2019), doing so for such short pulse durations (<= 0.3 ms) is not effective. Therefore, we used a rectangular pulse shape to match prior work.”

Importantly, our research shows that auditory co-stimulation can confound effects on motor excitability, and this likely occurred in multiple seminal TUS studies. While some preliminary work has been done on the efficacy of ramping in humans, future work is needed to determine what ramp shapes and lengths are optimal for reducing the auditory confound.

> *1.9) There are other models or experiments that need to be discussed in order to clearly disassociate the TUS effect from the auditory confound effect, for instance, testing deaf animal models or participants, or experiments with multi-region recordings (to rule out the effects of the dense structural connectivity between the auditory cortex and the motor cortex).*

The suggestion to consider multi-region recording in future experiments is important. Indeed, the effects of the auditory confound are expected to vary between brain regions. In the primary motor cortex, we observe a learned inhibition, which is perhaps supported by dense structural connectivity with the auditory system. In contrast, in perceptual areas such as the occipital cortex, one might expect tuned attentional effects in response to the auditory cue. We suggest that it is likely that the impact of the auditory confound also operates on a more global network level. It is reasonable to propose that, in a cognitive task for example, the confound will affect task performance and related brain activity, ostensibly regardless of the extent of direct structural connectivity between the auditory cortex and the (stimulated) region of interest.

Regarding the testing of deaf subjects, this has been included in the revised discussion as follows:

“Alternative approaches could circumvent auditory confounds by testing deaf subjects, or perhaps more practically by ramping the ultrasonic pulse to minimize or even eliminate the auditory confound.”

> *1.10) The concept of stochastic resonance is interesting but traditionally refers to a mechanism whereby a particular level of noise actually enhances the response of non-linear systems to weak sensory signals. Whether it applies to the motor system when probed with suprathreshold TMS intensities is unclear. Furthermore, whether higher intensities induce higher levels of noise is not straightforward neither considering the massive amount of work coming from other NIBS studies in particular. Noise effects are indeed a function of noise intensity, but exhibit an inverted U-shape dose-response relationship (Potok et al., 2021, eNeuro). In general SR is rather induced with low stimulation intensities in particular in perceptual domain (see Yamasaki et al., 2022, Neuropsychologia). In the same order of ideas, did the authors compare inter-trials variability across the different conditions?*

We thank the reviewer for these insightful remarks. Indeed, stochastic resonance is a concept first formalized in the sensory domain. Recently, the same principles have been shown to apply in other domains as well. For example, transcranial electric noise (tRNS) exhibits similar stochastic resonance principles as sensory noise (Van Der Groen & Wenderoth, 2016). Indeed, tRNS has been applied to many cortical targets, including the motor system. In the current manuscript, we raise the question of whether TUS might engage with neuronal activity following principles similar to tRNS. One prediction of this framework would be that TUS might not modulate excitation/inhibition balance overall, but instead exhibit an inverted U-shape dose-dependent relationship with stochastic noise. Please note, we do not use the ‘suprathreshold TMS intensity’ to quantify whether noise could bring a sub-threshold input across the detection threshold, nor whether it could bring a sub-threshold output across the motor threshold. Instead, we use the MEP read-out to estimate the temporally varying excitability itself. We argue that MEP autocorrelation captures the mixture of temporal noise and temporal structure in corticospinal excitability. Building on the non-linear response of neuronal populations, low stochastic noise might strengthen weakly present excitability patterns, while high stochastic noise might override pre-existing excitability. It is therefore not the overall MEP amplitude, but the MEP timeseries that is of interest to us. Here, we observe a non-linear dose-dependent relationship, matching the predicted inverted U-shape.

Importantly, we did not intend to assume stochastic resonance principles in the motor domain as a given. We have now clarified in the revised manuscript that we propose a putative framework and regard this as an open question:

“Indeed, human TUS studies have often failed to show a global change in behavioral performance, instead finding TUS effects primarily around the perception threshold where noise might drive stochastic resonance (Butler et al., 2022; Legon et al., 2018). Whether the precise principles of stochastic resonance generalize from the perceptual domain to the current study is an open question, but it is known that neural noise can be introduced by brain stimulation (Van Der Groen & Wenderoth, 2016). It is likely that this noise is statedependent and might not exceed the dynamic range of the intra-subject variability (Silvanto et al., 2007). Therefore, in an exploratory analysis, we exploited the natural structure in corticospinal excitability that exhibits as a strong temporal autocorrelation in MEP amplitude.”

Following the above reasoning, we felt it critical to estimate noise in the timeseries, operationalized as a t-1 autocorrelation, rather than capture inter-trial variability that ignores the timeseries history and requires data aggregation thereby reducing statistical power. Importantly, we would expect the latter index to capture global variability, putatively masking the temporal relationships which we were aiming to test. The reviewer raises an interesting option, inviting us to wonder if inter-trial variability might be sensitive enough, nonetheless. To this end, we compared inter-trial variability as suggested. This was achieved by first calculating the inter-trial variability for each condition, and then running a three-way repeated measures ANOVA on these values with the independent variables matching our autocorrelation analyses, namely, procedure (on-target/active control)*intensity (6.35/19.06)*masking (no mask/masked). This analysis did not reveal any significant interactions or main effects.

Author response table 1.

Effect	DFn	DFd	F	p
intensity	1	26	2.01136951	0.1679998
procedure	1	26	1.51481729	0.2294241
masking	1	26	0.55249790	0.4639630
intensity:procedure	1	26	0.05069027	0.8236267
intensity:masking	1	26	0.17466900	0.6794268
procedure:masking	1	26	0.27552687	0.6040957
intensity:procedure:masking	1	26	1.96443059	0.1728654

> *1.11) State-dependency/Autocorrelations: These values were extracted from Exp2 which has baseline trials. Can the authors provide autocorrelation values at baseline, with and without auditory mask? Can the authors comment on the difference between the autocorrelation profiles of the active TUS condition at 6.35W/cm2 or at 19.06W/cm2. They should somehow be similar to my understanding. Besides, the finding that TUS induces noise only when sound is present and at lower intensities is not well discussed.*

In the revised manuscript, we have now included baseline in the figure (Figure 4D). Regarding baseline with and without a mask, we must clarify that baseline involves only TMS (no mask), and sham involves TMS + masking stimulus (masked).

The dose-dependent relationship of TUS intensity with autocorrelation is critical. One possible observation would have been that TUS at both intensities decreased autocorrelation, with higher intensities evoking a greater reduction. Here, we would have concluded that TUS introduced noise in a linear fashion.

However, we observed that lower-intensity TUS in fact strengthened pre-existing temporal patterns in excitability (higher autocorrelation), while during higher-intensity TUS these

patterns were overridden (lower autocorrelation). This non-linear relationship is not unexpected, given the non-linear responses of neurons.

If this non-linear dependency is driven by TUS, one could expect it to be present during conditions both with and without auditory masking. However, the preparatory inhibition effect of TUS likely depends on the salience of the cue, that is, the auditory confound. In trials without auditory masking, the salience of the confound in highly dependent on (transmitted) intensity, with higher intensities being perceived as louder. In contrast, when trials are masked, the difference in cue salience between lower and higher intensity stimulation in minimized. Therefore, we would expect for any nuanced dose-dependent direct TUS effect to be best detectable when the difference in dose-dependent auditory confound perception is minimized via masking. Indeed, the dose-dependent effect of TUS on autocorrelation is most prominent when the auditory confound is masked.

“In sum, these preliminary exploratory analyses could point towards TUS introducing temporally specific neural noise to ongoing neural dynamics in a dose-dependent manner, rather than simply shifting the overall excitation-inhibition balance. One possible explanation for the discrepancy between trials with and without auditory masking is the difference in auditory confound perception, where without masking the confound’s volume differs between intensities, while with masking this difference is minimized. Future studies might consider designing experiments such that temporal dynamics of ultrasonic neuromodulation can be captured more robustly, allowing for quantification of possible state-dependent or nondirectional perturbation effects of stimulation.”

> *1.12) Statistical considerations. Data from Figure 2 are considered in two-by-two comparisons. Why not reporting the ANOVA results testing the main effect of TUS/Auditory conditions as done for Figure 3. Statistical tables of the LMM should be reported.*

Full-factorial analyses and main effects for TUS/Auditory conditions are discussed from Section 3.2 onwards. These are the same data supporting Figure 2 (now Figure 3). We would like to note that the main purpose of Figure 2 is to demonstrate to the reader that motor inhibition was observed, thus providing evidence that we replicated motor inhibitory effects of prior studies. A secondary purpose is to visually represent the absence of direct and spatially specific neuromodulation. However, the appropriate analyses to demonstrate this are reported in following sections, from Section 3.2 onwards, and we are concerned that mentioning these analyses earlier will negatively impact comprehensibility.

Statistical tables of the LMMs are provided within the open-sourced data and code reported at the end of the paper, embedded within the output which is accessible as a pdf (i.e., analysis/analysis.pdf).

> *1.13) Startle effects: The authors dissociate two mechanisms through which sound cuing can drive motor inhibition, namely some compensatory expectation-based processes or the evocation of a startle response. I find the dissociation somehow artificial. Indeed, it is known that the amplitude of the acoustic startle response habituates to repetitive stimulation. Therefore, sensitization can well explain the stabilization of the MEP amplitude observed after a few trials.*

Thank you for bringing this to our attention. Indeed, an acoustic startle response would habituate over repetitive stimulation. A startle response would result in MEP amplitude being significantly altered in early trials. As the participant would habituate to the stimulus, the startle response would decrease. MEP amplitude would then return to baseline levels. However, this is not the pattern we observe. An alternative possibility is that participants learn the temporal contingency between the stimulus and TMS. Here, compensatory

expectation-based change in MEP amplitude would be observed. In this scenario, there would be no change in MEP amplitude during early trials because the stimulus has not yet become informative of the TMS pulse timing. However, as participants learn how to predict TMS timing by the stimulus, MEP amplitude would decrease. This is also the pattern we observe in our data. We have clarified these alternatives in the revised manuscript as follows:

"Two putative mechanisms through which sound cuing may drive motor inhibition have been proposed, positing either that explicit cueing of TMS timing results in compensatory processes that drive MEP reduction (Capozio et al., 2021; Tran et al., 2021), or suggesting the evocation of a startle response that leads to global inhibition (Fisher et al., 2004; Furubayashi et al., 2000; Ilic et al., 2011; Kohn et al., 2004; Wessel & Aron, 2013). Critically, we can dissociate between these theories by exploring the temporal dynamics of MEP attenuation. One would expect a startle response to habituate over time, where MEP amplitude would be reduced during startling initial trials, followed by a normalization back to baseline throughout the course of the experiment as participants habituate to the starling stimulus. Alternatively, if temporally contingent sound-cueing of TMS drives inhibition, MEP amplitudes should decrease over time as the relative timing of TUS and TMS is being learned, followed by a stabilization at a decreased MEP amplitude once this relationship has been learned."

> *1.14) Can the authors further motivate the drastic change in intensities between Exp1 and 2? Is it due to the 250-500 carrier difference? It this coming from the loss power at 500kHz?*

The change in intensities between Experiments I and II was not an intentional experimental manipulation. Following completion of data acquisition, our TUS system received a firmware update that differentially corrected the 250 kHz and 500 kHz stimulation intensities. In this manuscript, we report the actual free-water intensities applied during our experiments.

> *1.15) Exp 3: Did 4 separate blocks of TUS-TMS and normalized for different TMS intensities used with respect to baseline. But how different was it. Why adjusting and then re adjusting intensities?*

The TMS intensities required to evoke a 1 mV MEP under the four sound-sham conditions significantly differed from the intensities required for baseline. In the revised appendix, we have now included a figure depicting the TMS intensities for these conditions, as well as statistical tests demonstrating each condition required a significantly higher TMS intensity than baseline.

TMS intensities were re-adjusted to avoid floor effects when assessing the efficacy of ontarget TUS. Sound-sham conditions themselves attenuate MEP amplitude. This is also evident from the higher TMS intensities required to evoke a 1 mV MEP under these conditions. If direct neuromodulation by TUS would have further decreased MEP amplitude, the concern was that effects might not be detectible within such a small range of MEP amplitudes.

> *1.16) In Exp 4, TUS targeted the ventromedial WM tract. Since direct electrical stimulation on white matter pathways within the frontal lobe can modulate motor output probably through dense communication along specific white matter pathways (e.g., Vigano et al., 2022, Brain), how did the authors ensure that this condition is really ineffective? Furthermore, the stimulation might have covered a lot more than just white matter. Acoustic and thermal simulations would be helpful here as well.*

Thank you for pointing out this possibility. Ultrasonic and electrical stimulation have quite distinct mechanisms of action. Therefore, it is challenging to directly compare these two approaches. There is a small amount of evidence that ultrasonic neuromodulation of white

matter tracts is possible. However, the efficacy of white matter modulation is likely much lower, given the substantially lesser degree of mechanosensitive ion channel expression in white matter as opposed to gray matter (Sorum et al., 2020, PNAS). Further, recent work has indicated that ultrasonic neuromodulation of myelinated axonal bundles occurs within the thermal domain (Guo et al., 2022, SciRep), which is not possible with the intensities administered in the current study. Nevertheless, based on Experiment IV in isolation, it cannot be definitively excluded that there TUS induced direct neuromodulatory effects in addition to confounding auditory effects. However, Experiment IV does not possess sufficient inferential power on its own and must be interpreted in tandem with Experiments I-III. Taken together with those findings, it is unlikely that a veridical neuromodulation effect is seen here, given the equivalent or lower stimulation intensities, the substantially deeper stimulation site, and the absence of an additional control condition in Experiment IV. This likelihood is further decreased by the fact that inhibitory effects under masking descriptively scale with the audibility of TUS.

Off-target effects such as unintended co-stimulation of gray matter when targeting white matter is always an important factor to consider. Unfortunately, individualized simulations for Experiment IV are not available. However, the same type of transducer and fundamental frequency was used as in Experiment II, for which we do have simulations. Given the size of the focus and the very low in-situ intensities extending beyond the main focal point, it is incredibly unlikely that effective stimulation was administered outside white matter in a meaningful number of participants. Nevertheless, the reviewer is correct that this can only be directly confirmed with simulations, which remain infeasible due to both technical and practical constraints. We have included the following in the revised manuscript:

“The remaining motor inhibition observed during masked trials likely owes to, albeit decreased, persistent audibility of TUS during masking. Indeed, MEP attenuation in the masked conditions descriptively scale with participant reports of audibility. This points towards a role of auditory confound volume in motor inhibition (Supplementary Fig. 8). Nevertheless, one could instead argue that evidence for direct neuromodulation is seen here. This unlikely for a number of reasons. First, white matter contains a lesser degree of mechanosensitive ion channel expression and there is evidence that neuromodulation of these tracts may occur primarily in the thermal domain (Guo et al., 2022; Sorum et al., 2021). Second, Experiment IV lacks sufficient inferential power in the absence of an additional control and must therefore be interpreted in tandem with Experiments I-III. These experiments revealed no evidence for direct neuromodulation using equivalent or higher stimulation intensities and directly targeting grey matter while also using multiple control conditions. Therefore, we propose that persistent motor inhibition during masked trials owes to continued, though reduced, audibility of the confound (Supplementary Fig. 8). However, future work including an additional control (site) is required to definitively disentangle these alternatives.”

> *1.17) Still for Exp 4. the rational for the 100% MSO or 120% or rMT is not clear, especially with respect to Exp 1 and 2. Equipment is similar as well as raw MEPs amplitudes, therefore the different EMG gain might have artificially increased TMS intensities. Could it have impacted the measured neuromodulatory effects?*

Experiment IV was conducted independently at a different institute than Experiments I-II. In contrast to Experiments I-II, a gel pad was used to couple TUS to the participant’s head. The increased TMS-to-cortex distance introduced by the gel pad necessitates higher TMS intensities to compensate for the increased offset. In fact, in 9/12 participants, the intended intensity at 120% rMT exceeded the maximum stimulator output. In those cases, we defaulted to the maximum stimulator output (i.e., 100% MSO). We have clarified in the revised supplementary material as follows:

"We aimed to use 120% rMT (n =3). However, if this intensity surpassed 100% MSO, we opted for 100% MSO instead (n = 9). The mean %MSO was 94.5 ± 10.5%. The TMS intensities required in this experiment were higher than those required in Experiment I-II using the same TMS coil, though still within approximately one standard deviation. This is likely due to the use of a gel pad, which introduces more distance between the TMS coil and the scalp, thus requiring a higher TMS intensity to evoke the same motor activity."

Regarding the EMG gain, this did not affect TMS intensities and did not impact the measured neuromodulatory effects. The EMG gain at acquisition is always considered during signal digitization and further analyses.

> *1.18) Exp. 4. It would be interesting to provide the changes in MEP amplitudes for those subjects who rated "inaudible" in the self-rating compared to the others. That's an important part of the interpretation: inaudible conditions lead to inhibition, so there is an effect. The auditory confound is not additive to the TUS effect.*

Previously, we only provided participant's ratings of audibility, and showed that conditions that were rated as inaudible more often showed less inhibition, descriptively indicating that inaudible stimulation does not lead to inhibition. This interpretation is in line with our conclusion that the TUS auditory confound acts as a cue signaling the upcoming TMS pulse, thus leading to preparatory inhibition.

We have now included an additional plot and discussion in Supplementary Figure 8 (Subjective Report of TUS Audibility). Here, we show the change in MEP amplitude from baseline for the three continuously masked TUS intensities as in the main manuscript, but now split by participant rating of audibility. Descriptively, less audible sounds result in no marked change or a smaller change in MEP amplitude. This supports our conclusion that direct neuromodulation is not being observed here. When participants were unsure whether they could hear TUS, or when they did hear TUS, more inhibition was observed. However, this is still to a lesser degree than unmasked stimulation which was nearly always audible, and likely also more salient. This also supports our conclusion that these results indicate a role of cue salience rather than direct neuromodulation. Regarding masked conditions where participants were uncertain whether they heard TUS, the sound was likely sufficient to act as a cue, albeit potentially subliminally. After all, preparatory inhibition is not a conscious action undertaken by the participant either. We would also like to note that participants reported perceived audibility after each block, not after each trial, so selfreported audibility was not a fine-grained measurement. The data from Experiment IV suggest that the volume of the cue has an impact on motor inhibition. Taken together with the points mentioned in 1.16, it is not possible to conclude there is evidence for direct neuromodulation in Experiment IV.

> *1.19) I suggest to re-order sub panels of the main figures to fit with the chronologic order of appearance in the text. (e.g Figure 1 with A) Ultrasonic parameters, B) 3D-printed clamp, C) Sound-TMS coupling, D) Experimental condition).*

We have restructured the figures in the manuscript to provide more clarity and to have greater alignment with the eLife format.

> *2.1) Although auditory confounds during TUS have been demonstrated before, the thorough design of the study will lead to a strong impact in the field.*

We thank the reviewer for recognition of the impact of our work. They highlight that auditory confounds during TUS have been demonstrated previously. Indeed, our work builds upon a larger research line on auditory confounds. The current study extends on the

confound's presence by quantifying its impact on motor cortical excitability, but perhaps more importantly by invalidating the most robust and previously replicable findings in humans. Further, this study provides a way forward for the field, highlighting the necessity of (in)active control conditions and tightly matched sham conditions for appropriate inferences in future work. We have amended the abstract to better reflect these points:

"Primarily, this study highlights the substantial shortcomings in accounting for the auditory confound in prior TUS-TMS work where only a flip-over sham control was used. The field must critically reevaluate previous findings given the demonstrated impact of peripheral confounds. Further, rigorous experimental design via (in)active control conditions is required to make substantiated claims in future TUS studies."

> *2.2) A few minor [weaknesses] are that (1) the overview of previous related work, and how frequent audible TUS protocols are in the field, could be a bit clearer/more detailed*

We have expanded on previous related work in the revised manuscript:

"Indeed, there is longstanding knowledge of the auditory confound accompanying pulsed TUS (Gavrilov & Tsirulnikov, 2012). However, this confound has only recently garnered attention, prompted by a pair of rodent studies demonstrating indirect auditory activation induced by TUS (Guo et al., 2022; Sato et al., 2018). Similar effects have been observed in humans, where exclusively auditory effects were captured with EEG measures (Braun et al., 2020). These findings are particularly impactful given that nearly all TUS studies employ pulsed protocols, from which the pervasive auditory confound emerges (Johnstone et al., 2021)."

> *2.3) The acoustic control stimulus can be described in more detail*

We have elaborated upon the masking stimulus for each experiment in the revised manuscript as follows:

Experiment I: "In addition, we also included a sound-only sham condition that resembled the auditory confound. Specifically, we generated a 1000 Hz square wave tone with 0.3 ms long pulses using MATLAB. We then added white noise at a signal-to-noise ratio of 14:1. This stimulus was administered to the participant via bone-conducting headphones."

Experiment II: "In this experiment, the same 1000 Hz square wave auditory stimulus was used for sound-only sham and auditory masking conditions. This stimulus was administered to the participant over in-ear headphones."

Experiment III: "Auditory stimuli were either 500 or 700 ms in duration, the latter beginning 100 ms prior to TUS (Supplementary Fig. 3.3). Both durations were presented at two pitches. Using a signal generator (Agilent 33220A, Keysight Technologies), a 12 kHz sine wave tone was administered over speakers positioned to the left of the participant as in Fomenko and colleagues (2020). Additionally, a 1 kHz square wave tone with 0.5 ms long pulses was administered as in Experiments I, II, IV, and prior research (Braun et al., 2020) over noisecancelling earbuds."

Experiment IV: "We additionally applied stimulation both with and without a continuous auditory masking stimulus that sounded similar to the auditory confound. The stimulus consisted of a 1 kHz square wave with 0.3 ms long pulses. This stimulus was presented through wired bone-conducting headphones (LBYSK Wired Bone Conduction Headphones). The volume and signal-to-noise ratio of the masking stimulus were increased until the participant could no longer hear TUS, or until the volume became uncomfortable."

In the revised manuscript we have also open-sourced the audio files used in Experiments I, II, and IV, as well as a recording of the output of the signal generator for Experiment III:

"Auditory stimuli used for sound-sham and/or masking for each experiment are accessible here: https://doi.org/10.5281/zenodo.8374148."

> *2.4) The finding that remaining motor inhibition is observed during acoustically masked trials deserves further discussion.*

We agree. Please refer to points 1.16 and 1.18.

> *2.5) In several places, the authors state to have "improved" control conditions, yet remain somewhat vague on the kind of controls previous work has used (apart from one paragraph where a similar control site is described). It would be useful to include more details on this specific difference to previous work.*

In the revised manuscript, we have clarified the control condition used in prior studies as follows:

Abstract:

"Primarily, this study highlights the substantial shortcomings in accounting for the auditory confound in prior TUS-TMS work where only a flip-over sham control was used."

Introduction:

"To this end, we substantially improved upon prior TUS-TMS studies implementing solely flip-over sham by including both (in)active control and multiple sound-sham conditions."

Methods:

"We introduced controls that improve upon the sole use of flip-over sham conditions used in prior work. First, we applied active control TUS to the right-hemispheric face motor area, allowing for the assessment of spatially specific effects while also better mimicking ontarget peripheral confounds. In addition, we also included a sound-only sham condition that closely resembled the auditory confound."

> *2.6) I also wondered how common TUS protocols are that rely on audible frequencies. If they are common, why do the authors think this confound is still relatively unexplored (this is a question out of curiosity). More details on these points might make the paper a bit more accessible to TUS-inexperienced readers.*

Regarding the prevalence of the auditory confound, please refer to point 2.2.

Peripheral confounds associated with brain stimulation can have a strong impact on outcome measures, often even overshadowing the intended primary effects. This is well known from electromagnetic stimulation. For example, the click of a TMS pulse can strongly modulate reaction times (Duecker et al., 2013, PlosOne) with effect sizes far beyond that of direct neuromodulation. Unfortunately, this consideration has not yet fully been embraced by the ultrasonic neuromodulation community. This is despite long known auditory effects of TUS (Gavrilov & Tsirulnikov, 2012, Acoustical Physics). It was not until the auditory confound was shown to impact brain activity by Guo et al., and Sato et al., (2018, Neuron) that the field began to attend to this phenomenon. Mohammadjavadi et al., (2019, BrainStim) then showed that neuromodulation persisted even in deaf mice, and importantly, also demonstrated that ramping ultrasound pulses could reduce the auditory brainstem response (ABR). Braun and colleagues (2020, BrainStim) were the first bring attention to the auditory confound in humans, while also discussing masking stimuli. This was followed by a study from Johnstone and colleagues (2021, BrainStim) who did preliminary work assessing both masking and ramping in humans. Recently, Liang et al., (2023) proposed a new form of masking colourfully

titled the ‘auditory Mondrian’. Further research into the peripheral confounds associated with TUS is on the way.

However, we agree that the confound remains relatively unexplored, particularly given the substantial impact it can have, as demonstrated in this paper. What is currently lacking is an assessment of the reproducibility of previous work that did not sufficiently consider the auditory confound. The current study constitutes a strong first step to addressing this issue, and indeed shows that results are not reproducible when using control conditions that are superior to flip-over sham, like (in)active control conditions and tightly matched soundsham conditions. This is particularly important given the fundamental nature of this research line, where TUS-TMS studies have played a central role in informing choices for stimulation protocols in subsequent research.

We would speculate that, with TUS opening new frontiers for neuroscientific research, there comes a rush of enthusiasm wherein laying the groundwork for a solid foundation in the field can sometimes be overlooked. Therefore, we hope that this work sends a strong message to the field regarding how strong of an impact peripheral confounds can have, also in prior work. Indeed, at the current stage of the field, we see no justification not to include proper experimental control moving forward. Only when we can dissociate peripheral effects from direct neuromodulatory effects can our enthusiasm for the potential of TUS be warranted.

> *2.7) Results, Fig. 2: Why did the authors not directly contrast target TUS and control conditions?*

Please refer to point 1.1.

> *2.8) The authors observe no dose-response effects of TUS. Does increasing TUS intensity also increase an increase in TUS-produced sounds? If so, should this not also lead to doseresponse effects?*

We thank the reviewer for this insightful question. Yes, increasing TUS intensity results in an increased volume of the auditory confound. Under certain circumstances this could lead to ‘dose-response’ effects. In the manuscript, we propose that the auditory confounds acts as a cue for the upcoming TMS pulse, thus resulting in MEP attenuation once the cue is informative (i.e., when TMS timing can be predicted by the auditory confound). In this scenario, volume can be taken as the salience of the cue. When the auditory confound is sufficiently salient, it should cue the upcoming TMS pulse and thus result in a reduction of MEP amplitude.

If we take Experiment II as an example (Figure 3B), the 19.06 W/cm2 stimulation would be louder than the 6.35 W/cm2 intensity. However, as both intensities are audible, they both cue the upcoming TMS pulse. One could speculate that the very slight (nonsignificant) further decrease for 19.06 W/cm2 stimulation could owe to a more salient cueing.

One might notice that MEP attenuation is less strong in Experiment I, even though higher intensities were applied. Directly contrasting intensities from Experiments I and II was not feasible due to differences in transducers and experimental design. From the perspective of sound cueing of the upcoming TMS pulse, the auditory confound cue was less informative in Experiment I than Experiment II, because TUS stimulus durations of both 100 and 500 ms were administered, rather than solely 500 ms durations. This could explain why descriptively less MEP attenuation was observed in Experiment I, where cueing was less consistent.

Perhaps more convincing evidence of a sound-based ‘dose-response’ effect comes from Experiment IV (Figure 4B). Here, we propose that continuous masking reduced the salience of the auditory confound (cue), and thus, less MEP attenuation was be observed. Indeed, we see less MEP change for masked stimulation. For the lowest administered volume during masked

stimulation, there was no change in MEP amplitude from baseline. For higher volumes, however, there was a significant inhibition of MEP amplitude, though it was still less attenuation than unmasked stimulation. These results indicate a 'doseresponse' effect of volume. When the volume (intensity) of the auditory confound was low enough, it was inaudible over the continuous mask (also as reported by participants), and thus it did not act as a cue for the upcoming TMS pulse, therefore not resulting in motor inhibition. When the volume (intensity) was higher, less participants reported not being able to hear the stimulation, so the cue was to a given extent more salient, and in line with the cueing hypothesis more inhibition was observed.

In summary, because the volume of the auditory confound scales with the intensity of TUS, there may be dose-response effects of the auditory confound volume. Along the border of (in)audibility of the confound, as in masked trials of Experiment IV, we may observe dose-response effects. However, at clearly audible intensities (e.g., Experiment I & II), the size of such an effect would likely be small, as both volumes are sufficiently audible to act as a cue for the upcoming TMS pulse leading to preparatory inhibition.

> *2.9) I wonder if the authors could say a bit more on the acoustic control stimulus. Some sound examples would be useful. The authors control for audibility, but does the control sound resemble the one produced by TUS?*

Please refer to point 2.3.

> *2.10) The authors' claim that the remaining motor inhibition observed during masked trials is due to persistent audibility of TUS relies "only" on participants' descriptions. I think this deserves a bit more discussion. Could this be evidence that there is a TUS effect in addition to the sound effect?*

Please refer to points 1.16 and 1.18.