City Research Online

City St George's, University of London

Citation: Fassnidge, C., Ball, D., Kazaz, Z., Knudsen, S., Spicer, A., Tipple, A. & Freeman, E. D. (2019). Hearing through your eyes: neural basis of audiovisual cross-activation, revealed by transcranial alternating current stimulation. Journal of Cognitive Neuroscience, 31(6), pp. 922-935. doi: 10.1162/jocn_a_01395

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/21725/

Link to published version: https://doi.org/10.1162/jocn_a_01395

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Uncorrected manuscript
Accepted for publication in Journal of Cognitive Neuroscience

Hearing through your eyes: neural basis of audiovisual cross-activation, revealed by transcranial alternating current stimulation

Christopher Fassnidge[1,2], Danny Ball[1,3], Zainab Kazaz[1], Synøve Knudsen[1,4], Andrew Spicer[1], Anthony Tipple[1], Elliot Freeman[1]

1. Cognitive Neuroscience Research Unit, Department of Psychology, City University London, London, UK
2. King's College London
3. University College London
4. University of Copenhagen

Corresponding author:

Dr Elliot Freeman
City University London
Northampton Square
London EC1V 0HB
United Kingdom
Phone: +44 7739765561
Email: elliot.freeman@city.ac.uk

Keywords:

Synaesthesia, perception, neurostimulation, audiovisual integration, individual differences, psychophysics

Highlights

- Some people report hearing auditory sensations evoked by visual movement and flashes. We report the first neurostimulation investigation of causal brain mechanisms underlying this visual-evoked auditory response (vEAR).

- Results support distinct inhibitory versus integrative functions of alpha and gamma oscillations, respectively

- Alpha-frequency transcranial Alternating Current Stimulation reveals that vision and audition normally compete, but less so in people who report 'hearing' visual movement and flashes

- Gamma-frequency stimulation to temporal areas disrupts performance of a visual sequence-memory task, specifically in people who benefit from being able to recode visual events as sounds

- Results suggest that vEAR depends on disinhibition of normally-occurring latent inter-cortical connections, not just on abnormally increased cross-connectivity

Abstract

Some people experience auditory sensations when seeing visual flashes or movements. This prevalent synaesthesia-like 'visual-evoked auditory response' (vEAR) could result either from over-exuberant cross-activation between brain areas, and/or reduced inhibition of normally-occurring cross-activation. We have used transcranial alternating current stimulation (tACS) to test these theories.

We applied tACS at 10Hz (alpha-band frequency) or 40Hz (gamma-band), bilaterally either to temporal or occipital sites, while measuring same/different discrimination of paired auditory (A) versus visual (V) 'Morse code' sequences. At debriefing, participants were classified as vEAR or non-vEAR depending on whether they reported 'hearing' the silent flashes.

In non-vEAR participants, temporal 10Hz tACS caused impairment of A performance, which correlated with improved V; conversely under occipital tACS, poorer V performance correlated with improved A. This reciprocal pattern suggests that sensory cortices are normally mutually inhibitory, and that alpha-frequency tACS may bias the balance of competition between them. vEAR participants showed no tACS effects, consistent with reduced inhibition, or enhanced cooperation between modalities. In addition, temporal 40Hz tACS impaired V performance, specifically in individuals who showed a performance advantage for V (relative to A). Gamma-frequency tACS may therefore modulate the ability of these individuals to benefit from recoding flashes into the auditory modality, possibly by disrupting cross-activation of auditory areas by visual stimulation.

Our results support both theories, suggesting that vEAR may depend on disinhibition of normally-occurring sensory cross-activation, which may be expressed more strongly in some individuals. Furthermore, endogenous alpha and gamma-frequency oscillations may function respectively to inhibit or promote this cross-activation.

Introduction

Some people can hear what they see: flashing car indicator lights, animated web-browser adverts, neon shop displays, and people's footsteps may all evoke an auditory sensation (Fassnidge, Cecconi Marcotti, & Freeman, 2017; Guttman, Gilroy, & Blake, 2006; Rothen, Bartl, Franklin, & Ward, 2017; Saenz & Koch, 2008). This 'visually-evoked auditory response' (vEAR) is also referred to as 'hearing-motion synaesthesia' (Saenz & Koch, 2008). Its prevalence has been estimated as at least ~5% up to ~20% (Fassnidge et al., 2017; Fassnidge & Freeman, 2018; Rothen et al., 2017). vEAR seems more prevalent than other more canonical varieties of synaesthesia, for example where music, letters or numbers can evoke perceptions of colour (Simner et al., 2006), and thus insights gained from studying vEAR may be more generally applicable as a model for understanding the normal mechanisms by which information from different modalities is combined. However, since the first report (Saenz & Koch, 2008) there has been very little research on this phenomenon. Behaviourally, the ability to 'hear' flashes can benefit discrimination of sequences of flashes (Fassnidge et al., 2017; Guttman et al., 2006; Saenz & Koch, 2008), possibly because the auditory modality is superior at encoding temporal structure (Glenberg, Mann, Altman, Forman, & Procise, 1989; Guttman et al., 2006). This visual discrimination advantage also correlates with impaired auditory detection in the presence of irrelevant flashes (Fassnidge et al., 2017). While there has only been one recent correlational EEG study (Rothen et al., 2017), to our knowledge no study has yet investigated the causal physiological mechanisms underlying vEAR.

There are two popular theories of synaesthesia that might explain vEAR. Firstly, the cross-activation theory (Hubbard, Brang, & Ramachandran, 2011; Ramachandran & Hubbard, 2001) postulates that synaesthetic percepts result from over-abundant connections between different brain regions resulting in greater cross-activation of one sensory representation by another (Bargary & Mitchell, 2008; Hubbard & Ramachandran, 2005; Tomson et al., 2011). Secondly, the disinhibition theory postulates that there is normal cross-connectivity in synaesthetes which is normally inhibited in non-synaesthetes, but disinhibited in synaesthetes (Grossenbacher & Lovelace, 2001; Neufeld et al., 2012), again resulting in greater cross-activation. This latter theory has support from our finding that vEAR is associated with a range of other

diverse perceptual traits, such as earworms, tinnitus, and auditory-evoked phosphenes (Fassnidge & Freeman, 2018). Such generalised phenomena might be more dependent on systemic variables determining cortical excitability or inhibition, rather than on just idiosyncratic patterns of neuroconnectivity (Fassnidge & Freeman, 2018).

In the present study, we used transcranial alternating current stimulation (tACS) to test experimentally whether visual and auditory cortices interact with each other differently in those who experience vEAR versus those who do not. Previous studies show that tACS can entrain endogenous oscillations which may play a role in coordinating coherent patterns of neural activity around the cortex (Antal & Paulus, 2013). The effects of tACS may be complex and non-linear (Herrmann, Rach, Neuling, & Strüber, 2013), thus it is difficult to predict whether tACS should increase or decrease performance in the present case or what specific mechanisms it might affect. However, we can at least make two distinct *a priori* predictions about the effects of tACS in people with versus without vEAR, based on the following logic: if a given group of participants express more of a given process, and tACS modulates that process, then the effects of tACS on their performance should be greater than in a group that expresses less of this same process.

Following this logic, if vEAR participants express *more* cross-activation than non-vEAR, and tACS modulates this cross-activation in some way, then tACS might affect performance of vEAR participants *more* than non-vEAR. For instance, if visual stimulation cross-activates auditory cortex more in vEAR (Hubbard et al., 2011; Ramachandran & Hubbard, 2001), then tACS to temporal areas might modulate visual performance more, compared to non-vEAR participants for whom there is less such visual-to-auditory cross-activation. Conversely, if vEAR participants express *less* inhibition than non-vEAR (Grossenbacher & Lovelace, 2001; Neufeld et al., 2012), and tACS modulates these inhibitory mechanisms, then tACS might have *less* effect in vEAR than non-vEAR participants.

We compared alpha versus gamma frequency tACS in order to test further hypotheses about the functional roles played by endogenous neural oscillations. Alpha

frequencies are thought to play a specific role in cortical inhibition (Bonnefond & Jensen, 2015; Cooper, Croft, Dominey, Burgess, & Gruzelier, 2003; Fu et al., 2001; Jensen & Mazaheri, 2010; Klimesch, 2012) but may also serve to coordinate integration of audiovisual information (Cecere, Rees, & Romei, 2015; Romei, Gross, & Thut, 2012; Schroeder & Foxe, 2005) and active information processing (Palva & Palva, 2007). Gamma frequencies are thought to be important for active information processing and binding of multisensory information, as well as maintenance of information in working memory and intra-cortical signalling (Başar, Başar-Eroglu, Karakaş, & Schürmann, 2000; Engel, Konig, Kreiter, & Singer, 1991; Ray & Maunsell, 2010; Roux & Uhlhaas, 2014; Senkowski, Schneider, Foxe, & Engel, 2008; Senkowski, Talsma, Herrmann, & Woldorff, 2005; Singer, 1999; Yuval-Greenberg & Deouell, 2007). In keeping with these distinct roles (Fries, Reynolds, Rorie, & Desimone, 2001; Jensen & Mazaheri, 2010; Wöstmann, Vosskuhl, Obleser, & Herrmann, 2018), alpha versus gamma-frequency tACS have respectively been found to have inhibitory versus facilitatory effects on performance, for example in attentional paradigms with parietal stimulation sites (Hopfinger, Parsons, & Fröhlich, 2017; Wöstmann et al., 2018), however gamma stimulation can also disrupt cross-hemispheric integration (Helfrich et al., 2014; Strüber, Rach, Trautmann-Lengsfeld, Engel, & Herrmann, 2014).

The above distinct hypothetical roles of endogenous oscillations lead to frequency-specific predictions for the effects of tACS on measures of visual and auditory task performance, following the logic outlined above. If vEAR depends on *reduced* inhibition, and alpha frequencies play a role in inhibition, we might predict that effects of alpha tACS would be *reduced* in vEAR than non-vEAR participants. Alternatively, if gamma frequencies play a role in multisensory binding or signalling, and if vEAR depends on *increased* cross-activation between vision and audition, then we might predict *increased* effects of gamma tACS in individuals whose performance depend upon such cross-activation.

To test these hypotheses we measured how tACS stimulation affected discriminability of pairs of either visual or auditory ‘Morse-code’ sequences (Saenz & Koch, 2008). We measured auditory and visual sequence discrimination performance while

applying tACS at either 10Hz and 40Hz to either occipital or temporal areas, and compared this with performance of the same task in the same modality but under sham stimulation.

Our first experiment compared the effects of alpha-frequency tACS on participants who reported that they could hear the flashes, versus those who did not. We included classical musicians, after having obtained pilot data that suggested they were more likely to experience vEAR, possibly due to their training and auditory imagery abilities. The second experiment compared the effect of gamma versus alpha-frequency tACS on performance in groups defined, not by subjective reports, but by an implicit objective measure of any latent ability to recode visual stimuli into the auditory modality. We quantified this as the ratio between visual and auditory discriminability (V:A). While auditory sequence discrimination is typically much better than visual (Saenz & Koch, 2008), in some participants a latent ability to recode visual stimuli into the auditory modality might provide a specific advantage for visual sequence discrimination (Glenberg et al., 1989; Guttman et al., 2006; Saenz & Koch, 2008), thus giving them a higher V:A ratio.

Our results provide the first insight into the neural mechanisms underlying vEAR. They also offer tentative support for a general disinhibition account of vEAR and of traits such as musicianship and other synaesthesias, which we also found to be strongly associated.

Methods

Participants

In Experiment 1 (10Hz tACS) we tested 36 paid participants who were naïve to the purpose of the study. These included 20 participants recruited from the student population and the local community (age range 18-31 years, mean 23.1, SD 3.74, 7 male. There were also 16 classical musicians from the London Royal College of Music (age range 18-55, mean 24.44, SD 9.92, 9 male). Experiment 2 (40Hz tACS) had 16 naïve participants (age range 18-28 years, mean 20.13, SD 2.64, 1 male), who received course credits or payment for taking part. All participants had normal or corrected vision and reported normal hearing. All procedures were carried out with written informed consent following completion of a safety checklist, and were approved by the local Psychology ethics committee in accordance with the Declaration of Helsinki, under the condition that a first-aid trained experimenter was always present during sessions.

Apparatus and stimuli

The experimental procedure was conducted using an Apple Mac Mini connected to a 17" Sony HMD-A420 cathode ray tube (CRT) display. Auditory stimuli were presented through two Labtec PC speakers both positioned next to each other directly in front of and below the centre of the monitor. Screen resolution was 800x600 pixels with a 120 Hz refresh rate and a viewing distance was set at approximately 57cm (controlled using a chin rest). A small white fixation point marked the centre of the display. Responses were entered using the arrow keys on a standard computer keyboard. Experimental procedures and stimuli were programmed using Psychtoolbox for Matlab.

Figure 1

Visual stimuli consisted of single white circular discs of luminance 81cdm^{-2}, presented centrally on a black background. Disk diameter was 3° of visual angle. Auditory stimuli were sine wave tones of maximum 91dBA sound pressure level, frequency 440Hz. 'Short' and 'Long' events were presented for periods of either 75ms or 300ms

respectively, during which stimulus amplitude immediately decayed linearly from maximum to zero amplitude.

Stimulation was delivered using a battery-driven DC-Stimulator Plus (NeuroConn. Illmenau, Germany) through two 5x5cm conductive-rubber electrodes housed in sponge pads that had been saturated with saline. Electrodes were secured using rubber straps and placed over either the occipital pole (O1, O2) or the temporal lobe (T3, T4), following the international 10-20 system (**Figure 1**).

Design

We employed a 2x2x2x2 design with three within-subjects factors: rhythmic sequence (Visual flashes vs. Auditory beeps), stimulation site (Occipital pole vs. Temporal lobe) and stimulation condition (tACS vs. Sham). Double-blinding was used for the stimulation condition. The fourth between-subjects factor distinguished between people who reported hearing sounds evoked by the flashes (Experiment 1), or those who showed a relative benefit for visual sequence discrimination (both experiments).

Procedure

Participants performed a rhythmic sequence discrimination task (Fassnidge et al., 2017; Saenz & Koch, 2008), over two sessions on separate days. In each session, participants performed the sequence discrimination task under both sham and stimulation conditions. Order of stimulation versus sham was fully counterbalanced across sessions. Order of montage was also independently counterbalanced between sessions.

In the stimulation condition, tACS was applied for fifteen minutes during task performance. The sham condition used identical stimulation parameters except that duration of stimulation was 30 seconds. This was included to replicate any initial sensations experienced in the stimulation session. The current had a sinusoidal waveform with a frequency of 10Hz and intensity of 1000μA. When switched on, the stimulation ramped up to full intensity over 2.5 seconds, then ramped back down at the end of the stimulation sequence, over another 2.5 seconds. Impedance was kept below 10 kΩ for all sessions by dripping saline onto the pads when required.

In the first experimental session, participants read an information sheet and completed a safety screening questionnaire before providing informed consent. This questionnaire screened for contraindications including medical history of epilepsy, brain injury, frequent headaches and anxiety or panic attacks. The experimental task was explained and demonstrated, and the participant was given a practice block. The relevant stimulation location for that session was then located using standard 10-20 procedures. Stimulation site depended on the randomised counterbalancing that had been established beforehand. Electrodes were then applied to the relevant area and secured by a rubber strap. To ensure double-blinding a second experimenter programmed the stimulator prior to each task repetition, delivering either sham or stimulation according to a secret code in the counterbalancing table, but did not interact with the participant or the first experimenter, and stayed outside the lab during data collection.

In each repetition of the sequence discrimination task, there were 50 auditory and 50 visual trials. These were presented in random order in 5 blocks of 20 trials, interleaved with mandatory fixed breaks of 30 seconds. On each trial, two successive rhythmic sequences of stimuli were presented. In half of the trials the events were all visual, and in the other half all auditory (**Figure 1**). The modality of each trial was randomly permuted between trials. Each sequence was composed of eight stimulus events, each of which was presented for a short (75ms) or a long (300ms) duration. Order of events in each sequence was randomised with the constraint that there were four to five transitions between short and long events or vice versa. The interval between events was 100ms, and the interval between the first and second sequence was 500ms. On half of the trials, the two sequences were identical, and on the other half they differed. In ‘different’ trials, the first two events and the last event were always identical between pairs, while the order of the remaining events was randomly permuted. Immediately following the second sequence, participants were required to indicate whether they thought the two sequences were same or different by pressing either the left or right arrow key on a computer keyboard, respectively. No error feedback was given. The response initiated the next trial.

After the first task the participant was given a minimum five-minute break, during which

the second experimenter returned to re-programme the stimulator. Participants then performed a second iteration of the sequence discrimination task. A second experimental session followed no less than 24 hours later, in which the above procedures were repeated for the other stimulation location.

At the end of the both sessions, participants completed a debrief questionnaire asking about any unpleasant effects of the stimulation such as headaches, pain or discomfort, or anxiety. We did not ask participants whether they could distinguish experimental from sham conditions, for this might have created biases on subsequent testing sessions. However as an alternative to this, we interviewed a subset of 27 participants in Experiment 1 about milder experiences such as cutaneous tingling and also flickering phosphenes. Flickering phosphenes have previously been associated with alpha frequency tACS (Kanai, Chaieb, Antal, Walsh, & Paulus, 2008), but not gamma frequency, thus we only asked this question for Experiment 1. We reasoned that participants who could detect such sensations might have had noticeably different experiences in sham and experimental conditions, thus potentially compromising blinding; in contrast those who did not detect any sensations in either session would be must less likely to notice a difference between conditions.

All participants of Experiment 1 were also administered a short structured interview in which they were asked a series of questions about visually-evoked sounds and synaesthetic experiences following questions (Fassnidge et al, 2017).

- *Were you aware of using the flashes as if they were sounds, e.g. "flash, flash-flash" = "beep, beep-beep"*
- *Did you actually hear faint sounds when you saw flashes?*
- *In everyday life, are you ever aware of hearing sounds when you see flashing lights or movement, e.g. shop displays, car indicators, or people walking?*
- *Do you ever experience colours associated with letters, or with music, or tastes or smells associated with sounds?*
- *Have you ever been diagnosed as a synaesthete or do you suspect you might be one?*

Results

Experiment 1

Out of 36 participants, 42% answered 'yes' to the question about hearing flashes, while 66% confirmed that they were aware of using the flashes as if they were sounds. 20% confirmed prior awareness of hearing visual movements, which is similar to the proportion responding positively to the same question in Fassnidge and Freeman's (2018) large-scale internet survey. Responses to the last two questions of the final questionnaire were redundant: only the participants who answered 'yes' to the last synaesthesia question (33%) answered 'yes' to the previous question about specific sensory associations (28%). For simplicity, we consider below only responses to the synaesthesia question. Fisher's exact tests revealed that frequency of reporting hearing flashes was significantly higher in musicians (75%) than non-musicians (12%, slightly lower than estimates of ~20% from Fassnidge et al, 2017) [odds ratio=22.00, $p<0.0001$]. Those who reported hearing flashes also tended to report synaesthetic experiences (56% in musicians vs 12%; but note that self-assessed prevalence typically exceeds estimates based on formal synaesthesia tests; Simner et al., 2006) [odds ratio=24.00, $p<0.0001$]. Musicians also tended to report synaesthesia more frequently [odds ratio=9.43, $p=0.0043$]. **Figure 2** shows these proportions as stacked bar charts. Responses to the remaining questions (about using flashes as if they were sounds, and experiences of visually-evoked sounds in everyday life) showed no significant relationships with musicianship or synaesthesia.

Figure 2

Sensitivity measures (d') for participants' same/different discrimination judgements were first calculated following standard psychophysical methods (Green & Swets, 1966). Three original participants were replaced by new participants because their d' scores for at least one of the auditory conditions was zero. To check whether there were any consistent carry-over effects of tACS on performance in the sham condition when following stimulation, we analysed d' for just the sham conditions in an ANOVA for each of the two electrode montages, grouping data by the order of sham condition.

Sham order had no significant main effect and did not interact significantly with Modality (Visual vs. Auditory) or order.

d' for participants in the 10Hz stimulation experiment was analysed in an ANOVA including Site (Occipital versus Temporal), Modality (Visual vs Auditory) and Stimulation (Sham vs. tACS) as repeated measures, and Yes/No responses to the 'Hearing Flashes' debriefing question as a grouping variable. Performance was significantly better overall in participants who reported 'Hearing Flashes' [$F(1,34)=12.34$, $p=.001$, $\eta^2_{partial}=.27$], and auditory (A) d' was significantly higher than visual (V) performance on average [$F(1,35)=134.90$, $p<.00001$, $\eta^2_p=.79$]. Means (and SE) for the different groups and conditions are as follows: 'Yes': V 2.39 (0.17); A 4.12 (0.33), N=15; 'No: V 1.60, (0.19); A 3.05 (0.81), N = 21. There was no significant main effect of tACS versus Sham, or for Site. However, there was a significant interaction between Site, Stimulation and Modality [$F(1,34)=6.43$, $p=.016$, $\eta^2_p=.16$]. Furthermore, these variables interacted significantly with 'Hearing Flashes' [$F(1,34)=7.00$, $p=.012$, $\eta^2_p=.17$]. Similar analyses showed no significant main effects or interactions with grouping variables based on the other questions about using flashes as if they were sounds, or everyday experience of visually-evoked sounds. Given the observed effect sizes and sample sizes, estimated statistical power for detecting an interaction in the present ANOVA design was high (>.9, calculated using G*Power 3.1 software, Faul, Erdfelder, Lang, & Buchner, 2007).

To quantify the effects of tACS, we subtracted Sham from tACS d' scores for each stimulation site separately, so that negative scores represent a decrement in performance. The results are shown in **Figure 3**, with asterisk and bracket annotations indicating significance of Tukey multiple comparisons at $p<.05$. The pattern of results appears to reveal reciprocal effects on task performance of tACS applied at different sites, particularly in non-vEAR participants. For example, A performance was significantly impaired by temporal tACS relative to sham and to occipital stimulation, but there was a non-significant trend for A performance to slightly improve with occipital stimulation. V performance was also significantly poorer with occipital stimulation compared to A, which again showed a non-significant trend to improved discriminability. In contrast, there were no significant deviations from sham

performance in vEAR participants.

An additional analysis assessed the performance of a subsample of 27 participants whose experiences of flicker and tingling sensations were probed in more detail during debriefing. Eight reported flickering phosphenes (sometimes described as occurring in peripheral vision, with blueish colour) and four reported tingling in one or both sessions; the remaining participants reported no such experiences in either session. Participants reporting sensations might arguably have experienced differences between sham versus experimental condition, leading to potential biases and placebo effects; in contrast, those who experienced no sensations would have had no such cues enabling them to distinguish between conditions. We first tested for an association between the frequency of reporting 'Stimulation Sensations' and reporting 'Hearing Flashes' (14 out of 27). If associated, then the critical effects shown in Figure 3 might be explained away as a placebo effect. However there was no significant association [Fisher's exact test: odds ratio=0.25, p=.13]. Bayesian analysis of contingency using JASP software (Jamil et al., 2017) resulted in Bayes Factor of 2.61 (based on a joint multinomial sampling scheme), showing only weak evidence for an association, compared to the null hypothesis of no association.

Table 1

We also tested in an ANOVA whether the subtracted difference between stimulation and sham d's depended on 'Stimulation Sensations' (including in the model 'Hearing Flashes', stimulation Site and Modality as additional independent variables). There were no significant group differences between participants reporting flicker or tingling and those reporting no sensations in either session (N=15) [$F(1,23)=.03$, ns], and no significant interactions of these groups with stimulation site and/or audiovisual task or 'Hearing Flashes' groups [$F=<0.7$]. In contrast, 'Hearing Flashes' still interacted significantly with both Site and Modality [$F(1,23)=9.173$, $p<.009$] in this subsample of participants, showing the same kind of reciprocal pattern seen in Figure 3 (note that this figure shows results for the full sample, not just for the 27 participants analysed here). Complementary results were revealed by a Bayesian ANOVA using JASP software with default priors (Rouder, Morey, Speckman, & Province, 2012). This

analysis specifically compared a two-factor model of the Site x Modality interaction, with the same model but including an interaction with either or both the 'Stimulation Sensations' and 'Hearing Flashes' grouping variables (Table 1). All other main effects and their combinations were considered under the null model. The model with the strongest evidence included the Site x Modality x Hearing Flashes interaction (Bayes Factor 9.23). We divided Bayes Factors for models including the three-factor interaction by the Bayes Factor for the two-factor model, to assess to what extent inclusion of the additional between-subjects grouping factors in the interaction provided a better explanation of our observations than the simple two-factor model alone. We found that the Site x Modality x Stimulation Sensations interaction model had a Bayes Factor of 0.76 relative to the Site x Modality model, thus the latter simpler model was 1.32 times more likely to explain our observations. In contrast, the Site x Modality x Hearing Flashes interaction was 20.42 times more likely than the Site x Modality model alone, and 23.04 times more likely than Site x Modality x Stimulation Sensations. It was also 17.41 times more likely than a more complex interaction model including all four factors. From these analyses, we can therefore conclude that it is highly unlikely that performance was affected by awareness of stimulation sensations or any resulting placebo effects, and much more likely that they are explained by differences in the tendency to experience visually-evoked auditory sensations.

Figure 3

In case the trends seen in Figure 3 had been diluted by participants who overall experienced weaker effects of tACS than others, we correlated individual scores for each task, under occipital stimulation against temporal stimulation (see **Figure 4**). This analysis found significant negative correlations specifically in non-vEAR participants (left graph), confirming that greater decrements in performance related to one stimulation site coincided reciprocally with greater improvement at the other site. In particular, impairment of A performance under temporal stimulation significantly correlated with improved performance of the same task under occipital stimulation [$r(19) = -.69$, $p<.001$]. A similar significant negative correlation was observed for V performance [$r(19) = -.79$, $p<.001$], where greater impairments under occipital

stimulation coincided with greater improvements under temporal stimulation. No such trends were observed for vEAR participants (right graph).

Figure 4

Experiment 2

We tested a separate group of 16 non-vEAR participants with 40Hz tACS and compared their performance with the non-vEAR participants in Experiment 1, excluding the musicians (N=17). This sample size was sufficient to achieve high power (>.9) for detecting an interaction effect in the present mixed ANOVA design, given effect sizes observed in the previous study (calculated using G*Power software). We restricted the sample to non-vEAR in order to test the assumption, made by the disinhibition hypothesis, that there is normal crosstalk between modalities even in people who do not experience synaesthetic phenomena. To assess this crosstalk, we used an implicit measure of the ratio of visual to auditory performance under sham stimulation (V:A). Higher values (e.g. closer to 1) indicate that visual performance is almost as good as auditory, and thus possibly relying more on visual-to-auditory recoding.

Participants were first grouped according to whether their individual V:A ratio was below or above the median across all participants (0.47, SD .25). 8 participants in the 10Hz stimulation group were assigned to the higher-scoring V:A group, and 7 participants in the 40Hz stimulation group. d' values were then analysed in a 4-way ANOVA with Modality and Site as repeated measures, and with frequency (10Hz vs. 40Hz) and V:A group as between-subjects variables. There was a significant interaction of V:A with Modality [$F(1,29)=10.16$, $p=.003$, $\eta^2_p=0.26$], and between V:A, Modality and Frequency [$F(1,29)=13.76$, $p<0.001$, $\eta^2_p=0.32$]. The full pattern of mean results is shown in **Figure 5** (with bracket and asterisk annotations indicating Tukey multiple comparisons significant at $p<.05$). In the group for whom V performance was more similar to A (labelled 'V~A', on the right of the figure), 40Hz stimulation specifically impaired V performance when applied to the Temporal sites [$t(6) = 3.08$, $p<0.022$, Cohen's D = 0.50; see asterisk].

Figure 5

Further analysis of the 40Hz data using ANCOVA with V:A as a continuous variable revealed a significant interaction between V:A with Modality [$F(1,14)=10.29$, $p=.0063$, $\eta^2_p=0.42$]. In the 10Hz data, the corresponding interaction was not significant [$F(1,15)=0.25$, ns]. The 40Hz interactions are visualised in **Figure 6** as scatterplots of performance (relative to sham) against V:A scores. Both occipital and temporal 40Hz tACS impaired V performance more in participants with higher V:A scores [Occipital: $\rho(14)=-0.52$, $p=0.039$; temporal: $\rho(14)=-0.53$, $p=0.035$] (see left two scatterplots in **Figure 6**). There were no other significant correlations. These results suggest that the more an individual relies on their audition to benefit visual task performance, the more their performance can be disrupted by gamma-frequency tACS. The comparable effects of tACS over both stimulation sites suggests that the stimulation may interfere with the auditory recoding of visual information, and also possibly also the transfer of information to auditory cortex from occipital areas.

Figure 6

Discussion

We have performed the first transcranial electrical brain stimulation experiment to investigate the neural basis for the little-known visually-evoked auditory response (vEAR), which may be a frequently occurring type of synaesthesia. Our results support our *a priori* predictions based on the hypothesis that vEAR depends on disinhibition of normally-occurring latent cross-connectivity between visual and auditory cortices. Given prior evidence of the high prevalence of vEAR and present evidence of associations with other synaesthesias and musicianship, these findings may have generalizable implications for understanding mechanisms underlying individual differences in synaesthesia and in normal interactions between modalities, as well as furthering understanding of the functional roles of cortical oscillations and the effects of brain stimulation.

While one popular theory of synaesthesia depends on greater cross-connectivity between cortical areas (Hubbard et al., 2011; Ramachandran & Hubbard, 2001), an alternative explanation is that the cross-connectivity is normal, but just disinhibited in synaesthesia (Grossenbacher & Lovelace, 2001; Neufeld et al., 2012). Our results support both of these hypotheses but depending on the frequency of stimulation and task measures used.

In support of the disinhibition hypothesis, we had reasoned that if people with vEAR had *less* inhibition, and if 10Hz tACS modulates such inhibition, then 10Hz tACS should affect them *less* than people with vEAR. This prediction was confirmed in Experiment 1, where we found less modulation of performance from 10Hz tACS in vEAR than in non-vEAR participants. Curiously, we found clear symmetrical reciprocal effects of stimulation at different sites in non-vEAR participants, where improvements in performance related to one stimulation site correlated with decrements related to stimulation at the other site. For example, the more visual performance improved under temporal stimulation the more it was disrupted by occipital stimulation; conversely the more auditory performance was impaired by temporal stimulation, the more it benefited from occipital stimulation. This antagonistic pattern is suggestive of a mutual inhibition between vision and audition in non-vEAR participants, where stimulation serves to bias the competition in favour of the non-stimulated modality.

This is consistent with fMRI evidence that stimulation in one modality suppresses the BOLD response to an unstimulated modality (Haxby et al., 1994; Johnson & Zatorre, 2005; Langner et al., 2011; Laurienti et al., 2002; Shulman et al., 1997). vEAR participants appeared on average to be immune to this push-pull effect of alpha tACS. This pattern could suggest that there is reduced competition between modalities in both directions, or conversely a general increase in cooperation between vision and audition. Such cooperation could explain why performance in vEAR participants was better both for the visual and the auditory task, as we have previously found (Fassnidge et al., 2017). It also seems consistent with the subjective experience of 'hearing flashes' in vEAR participants, who might experience a nominally unimodal visual stimulus as effectively audiovisual, enabling them to use vision and audition together to perform the discrimination task. However the symmetrical pattern of results also implies that people who are susceptible to visually-evoked auditory sensations might also be susceptible to auditory-evoked visual sensations. Although we did not measure it here, our previous study (Fassnidge et al., 2018) indeed found that reports of visually-evoked auditory sensations and ratings of auditory vividness of silent movies were significantly associated with reports of the converse phenomenon of flashes triggered by sudden sounds, experienced most typically while lying in the dark (Boroojerdi et al., 2000; Jacobs, Karpik, Bozian, & Gøthgen, 1981; Lessell & Cohen, 1979).

These findings all support the theory that some kinds of synaesthesia are characterised by a reduction of physiological disinhibition, which may result in unmasking of connections that are normally existent in most brains (Grossenbacher & Lovelace, 2001). There is previous evidence that visual and auditory brain areas have reciprocal connections (Clavagnier, Falchier, & Kennedy, 2004; Falchier et al., 2010; Gleiss & Kayser, 2008; Schroeder, Lakatos, Kajikawa, Partan, & Puce, 2008), and that different modalities can inhibit each other (Iurilli et al., 2012; Kayser & Remedios, 2012; Mattingley, Driver, Beschin, & Robertson, 1997; Ohshiro, Angelaki, & DeAngelis, 2011; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010; Ward & Mattingley, 2006). Such inhibition may be carried by alpha-frequency oscillations (Bonnefond & Jensen, 2015; Cooper et al., 2003; Fu et al., 2001; Jensen & Mazaheri, 2010; Klimesch, Sauseng, & Hanslmayr, 2007), which might be modulated or

enhanced by tACS (Antal & Paulus, 2013; Herrmann et al., 2013; Hopfinger et al., 2017; Wöstmann et al., 2018).

As illustrated in **Figure 7**, we tentatively propose one way in which, by interfering with such a mutually-inhibitory mechanism, 10Hz tACS might bias the balance of competition between auditory and visual areas. In vEAR individuals, tACS might have little effect because there is weaker prior inhibition to interfere with, and consequently less competition (or more cooperation) between modalities. However, in non-vEAR participants tACS to auditory areas might enhance auditory inhibition, and/or disrupt inhibitory competition of auditory with visual areas, in both cases biasing competition in favour of visual performance, which would consequently improve. Disinhibited visual areas might then send back more competitive inhibition to auditory areas, further reducing auditory performance. Stimulation to visual areas might have the opposite effect of biasing competition in favour of auditory areas, resulting in improved auditory performance and reduced visual performance. Note that although we have schematically indicated inhibition between modalities directly, this inhibition might also come via feedback from higher multimodal areas such as superior temporal sulcus (Grossenbacher & Lovelace, 2001; Neufeld et al., 2012). Also it should be noted that to our knowledge such an effect of tACS specifically on biased competition between modalities has not been observed before, thus this model would need to be substantiated in future research.

Figure 7

While the data support an association between vEAR and reduced crossmodal inhibition, a question remains about the causal role of inhibition. On the one hand, it is possible that the apparent reduction of inhibitory interactions shown in our vEAR participants is the result of using an advantageous attentional strategy of attending to both the sight and the evoked sound of the unimodal flash sequences, while non-vEAR participants would normally adopt a strategy of attending selectively to just the stimulated modality. Alpha tACS could interfere with the ability to suppress the irrelevant modality. This could also explain why musicians tend to have vEAR, if their musical training promotes joint attention to both the sound of music and the

coordinated movements of the conductor and other musicians. There is evidence that musicians do indeed place greater weight on visual movement cues when assessing the quality of a performance (Tsay, 2013). However, an alternative account is that systemic individual differences in the expression of inhibitory mechanisms play a causal role in the experience of vEAR. If so, such individual differences might cause a wider variety of effects on perception than just audiovisual synaesthesia. This is supported by the evident association of vEAR not only with musicianship and general synaesthesia (as found here) but also with tinnitus and musical imagery (Fassnidge & Freeman, 2018). Further evidence from brain stimulation studies suggests that experimental manipulation of cortical excitability (which might result from reduced inhibition) can enhance or induce a variety of cross-sensory phenomena including colour-grapheme and mirror-touch synaesthesia and auditory-modulated visual phosphenes (Bolognini, Miniussi, Gallo, & Vallar, 2013; Bolognini, Senna, Maravita, Pascual-Leone, & Merabet, 2010; Schroeder et al., 2008; Terhune, Tai, Cowey, Popescu, & Cohen Kadosh, 2011). These diverse phenomena are not necessarily dependent on cross-modal attention nor associated with behavioral advantages. Indeed, advantages in visual sequence discrimination are correlated with poorer auditory signal detection in the presence of irrelevant visual stimulation (Fassnidge et al., 2017). Taken all together with the present results, this broad pattern of association suggests that there may be a common causal role for disinhibition in accounting for individual differences in a variety of cross-modal and unimodal phenomena.

This disinhibition theory jointly assumes that vEAR is not only characterised by a reduction of inhibition between vision and audition, but also that such inhibition is necessarily mediated by normally-occurring connections between these areas, whether direct or indirect. This might predict a greater effect of tACS on individuals who have greater connectivity, but Experiment 1 found less effect of tACS in vEAR individuals. This might be because, firstly, an analysis based just on subjective awareness of ‘hearing flashes’ might not be sensitive to the full range of individual differences in latent cross-connectivity. For example, we previously found that some participants suffered impaired auditory signal detection in the presence of an irrelevant visual flash, consistent with a subliminal cross-activation of auditory areas from vision, but this was independent of their awareness of ‘hearing’ these flashes (Fassnidge et

al., 2017). Secondly, if there are latent differences in cross-connectivity, this might not have been affected by 10Hz tACS if that specifically plays a role in modulating inhibition.

Experiment 2 addressed these issues by using a 40Hz tACS combined with an implicit measure of the ratio of visual to auditory task performance, which might increase the more participants are able to recode flashes into the auditory modality (Saenz & Koch, 2008). 40Hz tACS might play a role in disrupting information processing, multisensory binding, or allocation of attentional resources (Başar et al., 2000; Engel et al., 1991; Roux & Uhlhaas, 2014; Senkowski et al., 2008, 2005; Singer, 1999; Yuval-Greenberg & Deouell, 2007). We specifically predicted that if 40Hz tACS performs such functions, then participants who recode visual events into the auditory modality should find it harder to do this during stimulation to temporal areas, because this would interrupt the integration of visual inputs within auditory areas. Our results confirm this prediction: temporal (as well as occipital) 40Hz tACS affected visual sequence discrimination, particularly in those who showed higher V:A ratios. We suggest that in the present study, 40Hz tACS might be more effective at specifically interrupting the signalling between cross-activating brain areas, particularly in high V:A individuals whose latent ability to recode visual events into the auditory modality normally gives them a visual advantage. Given that 40Hz oscillations may also play a role in intracortical signalling (Ray & Maunsell, 2010), it is also possible that 40Hz tACS might also have affected the processing of visual signals once they arrived in temporal areas, without necessarily affecting the inter-cortical transmission of those signals. Such an intra-cortical mechanism might therefore be expected to affect auditory processing too, however present evidence for this specific intra-cortical mechanism is weaker (e.g. see right scatterplot in Figure 5).

In summary, our experimental results provide support for both of the theories outlined previously. The results of Experiment 2 support the *a priori* prediction, based on the hypothesis of greater cross-connectivity between cortical areas (Hubbard et al., 2011; Ramachandran & Hubbard, 2001), that vEAR individuals would be affected more, rather than less by tACS. In contrast, the results of Experiment 1 support the alternative *a priori* prediction, based on disinhibited cross-activation in synaesthesia

(Grossenbacher & Lovelace, 2001; Neufeld et al., 2012), finding that vEAR individuals are affected less by tACS.

This study had some methodological issues which are important to consider. One potentially critical issue is that it was not possible to ensure perfect blinding, given that tACS at low frequencies such as in the alpha to beta range are known to evoke flickering phosphenes in peripheral vision (Kanai et al., 2008). We consider it unlikely that peripheral flickering could have systematically interfered with sequence discrimination of flashes which had high contrast and relatively low temporal frequency, and which were presented at fixation. It also seems implausible that such response biases alone could have conspired to create the detailed pattern of reciprocal effects shown in Figure 3, and an apparent dependence on 'Hearing Flashes'. Nevertheless, we undertook additional analyses to assess whether reported stimulation sensations could have affected performance. Firstly, there was no significant association between 'Hearing Flashes' and reported stimulation sensations, thus stimulation sensations cannot fully explain away the effect of hearing flashes on performance. Secondly, results from classical and Bayesian analyses showed no evidence that stimulation sensations affected the critical cross-over interaction pattern in Figure 3; in contrast, the evidence was much stronger for the role of 'Hearing Flashes' in determining participants' response to stimulation. We therefore conclude that the specific pattern of results is more likely to have depended on Hearing Flashes than on placebo effects due to subjective differences between stimulation conditions. We did not attempt to replicate this analysis with gamma stimulation because this does not tend to evoke phosphenes (Kanai et al., 2008), thus there was less chance of compromised blinding.

Endogenous alpha frequency can vary between individuals from about 8 to 12Hz (Cecere et al., 2015; Zaehle, Rach, & Herrmann, 2010), and one might therefore wonder whether the null effect of tACS stimulation in vEAR participants could reflect a systematic mismatch between endogenous and exogenous stimulation frequency. While we cannot assess this here, we note that mismatches as much as ±2Hz are still effective at influencing perception and performance (Cecere et al., 2015), thus it seems unlikely that such mismatches would be sufficient to entirely eliminate tACS

effects, unless they were substantially greater than 2Hz.

Average sequence discrimination performance tended to be higher in vEAR participants, as we have found previously (Fassnidge et al., 2017). It is therefore possible that 10Hz tACS stimulation might have affected performance less in the vEAR group if their performance was already on average at a high level. However any ceiling effects might be expected to reduce improvements more than impairments from tACS, while both were reduced to a similar extent in vEAR participants on average; furthermore, this seems unlikely to explain the results of Experiment 2, which depended on groups distinguished by relative performance between auditory and visual conditions, rather than by absolute performance levels.

It is possible that the interval between sham and stimulation conditions may not have been long enough to completely rule out after-effects. However we found no differences in the pattern of results depending on session order; in any case any order effects would be most likely to only weaken apparent stimulation effects, and could not explain away differences between stimulation sites which were tested on separate days.

Our results seem consistent with our *a priori* predictions which specifically assumed modulation of auditory areas. However inference is limited by uncertainty over the precise brain areas which were stimulated by tACS. Particularly in contrast to the occipital electrode montage, our temporal montage may have caused more current flow through the centre of the head, thus the effects we observed with temporal tACS may reflect modulation of medial and subcortical structures, as well as cortical areas outside auditory cortex, in addition to auditory areas. Interpretations about the functions of stimulated areas must therefore be taken with caution. Future studies using more sophisticated ‘high-definition’ tACS and current-flow modelling will be needed to assess whether the effects observed depend specifically on stimulation to auditory and visual cortices.

Finally, it might be wondered whether a single question about whether flashes were actually heard as sounds was sufficient to attribute vEAR abilities to participants. Although no standard diagnostic of this ability exists as yet, we have previously found

that this question is predictive of objective sequence discrimination performance (Fassnidge et al., 2017), while the question about whether flashes were used in the discrimination task as if they were sounds (i.e. not necessarily actually heard) was not predictive. The present results showed a similar dissociation, where the cross-over interaction effects of stimulation depended significantly on reports of hearing flashes (Figure 3), but not significantly on reports of merely 'using' flashes as sounds. This suggests that our questions could validly distinguish people who have a genuine subjective experience of actual auditory sensations (which was predictive of tACS effects), from those who do not but who nevertheless try to use flashes *as if* they were sounds (which was not predictive). The present evidence suggests that our results may reflect specific physiological differences associated specifically with visually-evoked auditory sensations, but this remains to be substantiated using direct physiological measures (e.g. EEG or fMRI).

In conclusion, our results suggest that there are normally-occurring individual variations in the level of cross-activation between visual and auditory cortices, which can be disrupted by 40Hz tACS of auditory areas; however the subjective awareness of 'hearing flashes' may additionally depend on reduced inhibition between visual and auditory cortices, resulting in cooperative rather than competitive interactions between vision and audition, as evidenced by a relative immunity to effects of 10Hz tACS. Although different stimulation frequencies and tasks result in these apparently opposing effects, the pattern of results is altogether consistent with both the cross-connectivity and disinhibition theories of synaesthesia: there are normally-occurring connections between cortical areas which can vary in their richness between individuals, and the disinhibition of these connections is correlated with conscious awareness of vEAR. In addition, our results support distinct functional roles of alpha versus gamma oscillations, in respectively biasing the balance of inhibitory competition between modalities, versus interfering with signal integration. Altogether, the apparent prevalence of vEAR, the availability of behavioural measures, and also the accessibility of auditory and visual cortices to brain stimulation, all make vEAR a uniquely convenient and potentially powerful platform for experimentally testing the neural basis of both synaesthesia and normal multisensory integration, and also for gaining further insights into the causal roles of cortical oscillations of different

frequencies in normal perception.

Acknowledgements
This work was supported by grants from British Academy and Leverhulme (SG151380 and IAF-2017-006 to EF). We thank Marinella Cappelletti for comments on the manuscript, and James Yearsley for advice on Bayesian analysis.

References

Antal, A., & Paulus, W. (2013). Transcranial alternating current stimulation (tACS). *Frontiers in Human Neuroscience*, *7*(June), 317. https://doi.org/10.3389/fnhum.2013.00317

Bargary, G., & Mitchell, K. J. (2008). Synaesthesia and cortical connectivity. *Trends in Neurosciences*. https://doi.org/10.1016/j.tins.2008.03.007

Başar, E., Başar-Eroglu, C., Karakaş, S., & Schürmann, M. (2000). Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International Journal of Psychophysiology*, *39*(2–3), 241–248. https://doi.org/10.1016/S0167-8760(00)00145-8

Bolognini, N., Miniussi, C., Gallo, S., & Vallar, G. (2013). Induction of mirror-touch synaesthesia by increasing somatosensory cortical excitability. *Current Biology*, *23*(10), R436–R437. https://doi.org/10.1016/j.cub.2013.03.036

Bolognini, N., Senna, I., Maravita, A., Pascual-Leone, A., & Merabet, L. B. (2010). Auditory enhancement of visual phosphene perception: the effect of temporal and spatial factors and of stimulus intensity. *Neuroscience Letters*, *477*(3), 109–114. https://doi.org/10.1016/j.neulet.2010.04.044

Bonnefond, M., & Jensen, O. (2015). Gamma activity coupled to alpha phase as a mechanism for top-down controlled gating. *PLoS ONE*, *10*(6), 20–22. https://doi.org/10.1371/journal.pone.0128667

Boroojerdi, B., Bushara, K. O., Corwell, B., Immisch, I., Battaglia, F., Muellbacher, W., & Cohen, L. G. (2000). Enhanced excitability of the human visual cortex induced by short-term light deprivation. *Cereb Cortex*, *10*(5), 529–534. https://doi.org/10.1093/cercor/10.5.529

Cecere, R., Rees, G., & Romei, V. (2015). Individual differences in alpha frequency drive crossmodal illusory perception. *Current Biology*, *25*(2), 231–235. https://doi.org/10.1016/j.cub.2014.11.034

Clavagnier, S., Falchier, A., & Kennedy, H. (2004). Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. *Cognitive, Affective & Behavioral Neuroscience*, *4*(2), 117–126. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15460918

Cooper, N. R., Croft, R. J., Dominey, S. J. J., Burgess, A. P., & Gruzelier, J. H. (2003). Paradox lost? Exploring the role of alpha oscillations during externally vs. internally directed attention and the implications for idling and inhibition hypotheses. *International Journal of Psychophysiology*, *47*(1), 65–74. https://doi.org/10.1016/S0167-8760(02)00107-1

Engel, A., Konig, P., Kreiter, A., & Singer, W. (1991). Interhemispheric synchronization of oscillatory neuronal responses in cat visual cortex. *Science*,

252(5009), 1177–1179. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17797912

Falchier, A., Schroeder, C. E., Hackett, T. A., Lakatos, P., Nascimento-Silva, S., Ulbert, I., … Smiley, J. F. (2010). Projection from visual areas V2 and prostriata to caudal auditory cortex in the monkey. *Cerebral Cortex*, *20*(7), 1529–1538. https://doi.org/10.1093/cercor/bhp213

Fassnidge, C. J., Cecconi Marcotti, C., & Freeman, E. D. (2017). A deafening flash ! Visual interference of auditory signal detection. *Consciousness and Cognition*, *49*, 15–24. https://doi.org/10.1016/j.concog.2016.12.009

Fassnidge, C. J., & Freeman, E. D. (2018). Sounds from seeing silent motion: Who hears them, and what looks loudest? *Cortex*, *103*, 130–141. https://doi.org/10.1016/j.cortex.2018.02.019

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fries, P., Reynolds, J. H., Rorie, a E., & Desimone, R. (2001). Modulation of oscillatory neural synchronization by selective visual attention. *Science*, *291*(February), 1560–1563. https://doi.org/10.1126/science.1055465

Fu, K. M. G., Foxe, J. J., Murray, M. M., Higgins, B. A., Javitt, D. C., & Schroeder, C. E. (2001). Attention-dependent suppression of distracter visual input can be cross-modally cued as indexed by anticipatory parieto-occipital alpha-band oscillations. *Cognitive Brain Research*, *12*(1), 145–152. https://doi.org/10.1016/S0926-6410(01)00034-9

Gleiss, S., & Kayser, C. (2008). Acoustic Noise Improves Visual Perception and Modulates Occipital Oscillatory States, 1–13. https://doi.org/10.1162/jocn

Glenberg, A. M., Mann, S., Altman, L., Forman, T., & Procise, S. (1989). Modality effects in the coding and reproduction of rhythms. *Memory & Cognition*, *17*(4), 373–383. https://doi.org/10.3758/BF03202611

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* New York: Wiley.

Grossenbacher, P. G., & Lovelace, C. T. (2001). Mechanisms of synesthesia: cognitive and physiological constraints. *Trends in Cognitive Sciences*, *5*(1), 36–41. https://doi.org/10.1016/S1364-6613(00)01571-0

Guttman, S. E., Gilroy, L. A., & Blake, R. (2006). Hearing what the eye sees. *Psychological Science*, *16*(3), 228–235. https://doi.org/10.1111/j.0956-7976.2005.00808.x.Hearing

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C.

L. (1994). The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *J Neurosci*, *14*(11 Pt 1), 6336–6353. https://doi.org/7965040

Helfrich, R. F., Knepper, H., Nolte, G., Strüber, D., Rach, S., Herrmann, C. S., … Engel, A. K. (2014). Selective Modulation of Interhemispheric Functional Connectivity by HD-tACS Shapes Perception. *PLoS Biology*, *12*(12), e1002031. https://doi.org/10.1371/journal.pbio.1002031

Herrmann, C. S., Rach, S., Neuling, T., & Strüber, D. (2013). Transcranial alternating current stimulation: a review of the underlying mechanisms and modulation of cognitive processes. *Frontiers in Human Neuroscience*, *7*(June), 279. https://doi.org/10.3389/fnhum.2013.00279

Hopfinger, J. B., Parsons, J., & Fröhlich, F. (2017). Differential effects of 10-Hz and 40-Hz transcranial alternating current stimulation (tACS) on endogenous versus exogenous attention. *Cognitive Neuroscience*, *8*(2), 102–111. https://doi.org/10.1080/17588928.2016.1194261

Hubbard, E. M., Brang, D., & Ramachandran, V. S. (2011). The cross-activation theory at 10. *Journal of Neuropsychology*, *5*(2), 152–177. https://doi.org/10.1111/j.1748-6653.2011.02014.x

Hubbard, E. M., & Ramachandran, V. S. (2005). Neurocognitive mechanisms of synesthesia. *Neuron*, *48*(3), 509–520. https://doi.org/10.1016/j.neuron.2005.10.012

Iurilli, G., Ghezzi, D., Olcese, U., Lassi, G., Nazzaro, C., Tonini, R., … Medini, P. (2012). Sound-driven synaptic inhibition in primary visual cortex. *Neuron*, *73*(4), 814–828. https://doi.org/10.1016/j.neuron.2011.12.026

Jacobs, L., Karpik, A., Bozian, D., & Gøthgen, S. (1981). Auditory-visual synesthesia sound-induced photisms. *Archives of Neurology*, *38*(4), 211–216.

Jamil, T., Ly, A., Morey, R. D., Love, J., Marsman, M., & Wagenmakers, E. J. (2017). Default “Gunel and Dickey” Bayes factors for contingency tables. *Behavior Research Methods*, *49*(2), 638–652. https://doi.org/10.3758/s13428-016-0739-8

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, *4*(November), 186. https://doi.org/10.3389/fnhum.2010.00186

Johnson, J. A., & Zatorre, R. J. (2005). Attention to simultaneous unrelated auditory and visual events: Behavioral and neural correlates. *Cerebral Cortex*, *15*(10), 1609–1620. https://doi.org/10.1093/cercor/bhi039

Kanai, R., Chaieb, L., Antal, A., Walsh, V., & Paulus, W. (2008). Frequency-dependent electrical stimulation of the visual cortex. *Current Biology : CB*, *18*(23), 1839–1843. https://doi.org/10.1016/j.cub.2008.10.027

Kayser, C., & Remedios, R. (2012). Suppressive Competition: How Sounds May Cheat Sight. *Neuron*, *73*(4), 627–629. https://doi.org/10.1016/j.neuron.2012.02.001

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606–617. https://doi.org/10.1016/j.tics.2012.10.007

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Langner, R., Kellermann, T., Boers, F., Sturm, W., Willmes, K., & Eickhoff, S. B. (2011). Modality-specific perceptual expectations selectively modulate baseline activity in auditory, somatosensory, and visual cortices. *Cerebral Cortex*, *21*(12), 2850–2862. https://doi.org/10.1093/cercor/bhr083

Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y.-F., Field, A. S., & Stein, B. E. (2002). Deactivation of Sensory-Specific Cortex by Cross-Modal Stimuli. *Journal of Cognitive Neuroscience*, *14*(3), 420–429. https://doi.org/10.1162/089892902317361930

Lessell, S., & Cohen, M. M. (1979). Phosphenes induced by sound. *Neurology*, *29*(11), 1524.

Mattingley, J. B., Driver, J., Beschin, N., & Robertson, I. H. (1997). Attentional competition between modalities: Extinction between touch and vision after right hemisphere damage. *Neuropsychologia*, *35*(6), 867–880. https://doi.org/10.1016/S0028-3932(97)00008-0

Neufeld, J., Sinke, C., Zedler, M., Dillo, W., Emrich, H. M., Bleich, S., & Szycik, G. R. (2012). Disinhibited feedback as a cause of synesthesia: evidence from a functional connectivity study on auditory-visual synesthetes. *Neuropsychologia*, *50*(7), 1471–1477. https://doi.org/10.1016/j.neuropsychologia.2012.02.032

Ohshiro, T., Angelaki, D. E., & DeAngelis, G. C. (2011). A normalization model of multisensory integration. *Nature Neuroscience*, *14*(6), 775–782. https://doi.org/10.1038/nn.2815

Palva, S., & Palva, J. M. (2007). New vistas for α-frequency band oscillations. *Trends in Neurosciences*, *30*(4), 150–158. https://doi.org/10.1016/j.tins.2007.02.001

Ramachandran, V. S., & Hubbard, E. M. (2001). Psychophysical investigations into the neural basis of synaesthesia. *Proceedings of the Royal Society B: Biological Sciences*, *268*(1470), 979–983. https://doi.org/10.1098/rspb.2000.1576

Ray, S., & Maunsell, J. H. R. (2010). Differences in Gamma Frequencies across Visual Cortex Restrict Their Possible Use in Computation. *Neuron*, *67*(5), 885–896. https://doi.org/10.1016/J.NEURON.2010.08.004

Romei, V., Gross, J., & Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception. *Current Biology*, *22*(9), 807–813. https://doi.org/10.1016/j.cub.2012.03.025

Rothen, N., Bartl, G. J., Franklin, A., & Ward, J. (2017). Electrophysiological Correlates and Psychoacoustic Characteristics of Hearing-Motion Synaesthesia. *Neuropsychologia*, (in press). Retrieved from http://www.sciencedirect.com/science/article/pii/S0028393217303251

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Roux, F., & Uhlhaas, P. J. (2014). Working memory and neural oscillations: Alpha-gamma versus theta-gamma codes for distinct WM information? *Trends in Cognitive Sciences*, *18*(1), 16–25. https://doi.org/10.1016/j.tics.2013.10.010

Saenz, M., & Koch, C. (2008). The sound of change: visually-induced auditory synesthesia. *Current Biology*, *18*(15), 650–651. Retrieved from http://www.sciencedirect.com/science/article/pii/S0960982208007343

Schroeder, C. E., & Foxe, J. (2005). Multisensory contributions to low-level, "unisensory" processing. *Current Opinion in Neurobiology*, *15*(4), 454–458. https://doi.org/10.1016/j.conb.2005.06.008

Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., & Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, *12*(3), 106–113. https://doi.org/10.1016/j.tics.2008.01.002

Senkowski, D., Schneider, T. R., Foxe, J. J., & Engel, A. K. (2008). Crossmodal binding through neural coherence: implications for multisensory processing. *Trends in Neurosciences*, *31*(8), 401–409. https://doi.org/10.1016/j.tins.2008.05.002

Senkowski, D., Talsma, D., Herrmann, C. S., & Woldorff, M. G. (2005). Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. *Experimental Brain Research*, *166*(3–4), 411–426. https://doi.org/10.1007/s00221-005-2381-z

Shulman, G. L., Corbetta, M., Buckner, R. L., Raichle, M. E., Fiez, J. a, Miezin, F. M., & Petersen, S. E. (1997). Top-down modulation of early sensory cortex. *Cerebral Cortex (New York, N.Y. : 1991)*, *7*(3), 193–206. https://doi.org/10.1093/cercor/7.3.193

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., … Ward, J. (2006). Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception*, *35*(8), 1024–1033. Retrieved from http://people.brunel.ac.uk/~hsstnns/reprints/Simner_at_al_2006_Prevalence.pdf

Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of

relations? *Neuron*, *24*(1), 49–65. https://doi.org/10.1016/B978-012370880-9.00287-5

Strüber, D., Rach, S., Trautmann-Lengsfeld, S. A., Engel, A. K., & Herrmann, C. S. (2014). Antiphasic 40 Hz oscillatory current stimulation affects bistable motion perception. *Brain Topography*, *27*(1), 158–171. https://doi.org/10.1007/s10548-013-0294-x

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, *14*(9), 400–410. https://doi.org/10.1016/j.tics.2010.06.008

Terhune, D. B., Tai, S., Cowey, A., Popescu, T., & Cohen Kadosh, R. (2011). Enhanced cortical excitability in grapheme-color synesthesia and its modulation. *Current Biology*, *21*(23), 2006–2009. https://doi.org/10.1016/j.cub.2011.10.032

Tomson, S. N., Avidan, N., Lee, K., Sarma, A. K., Tushe, R., Milewicz, D. M., … Eagleman, D. M. (2011). The genetics of colored sequence synesthesia: suggestive evidence of linkage to 16q and genetic heterogeneity for the condition. *Behav Brain Res*, *223*(1), 48–52. https://doi.org/10.1016/j.bbr.2011.03.071

Tsay, C.-J. (2013). Sight over sound in the judgment of music performance. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(36), 14580–14585. https://doi.org/10.1073/pnas.1221454110

Ward, J., & Mattingley, J. B. (2006). Synaesthesia: An overview of contemporary findings and controversies. *Cortex*, *42*(2), 129–136. https://doi.org/10.1016/S0010-9452(08)70336-8

Wöstmann, M., Vosskuhl, J., Obleser, J., & Herrmann, C. S. (2018, July 1). Opposite effects of lateralised transcranial alpha versus gamma stimulation on auditory spatial attention. *Brain Stimulation*, pp. 752–758. https://doi.org/10.1016/j.brs.2018.04.006

Yuval-Greenberg, S., & Deouell, L. Y. (2007). What you see is not (always) what you hear: induced gamma band responses reflect cross-modal interactions in familiar object recognition. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(5), 1090–1096. https://doi.org/10.1523/JNEUROSCI.4828-06.2007

Zaehle, T., Rach, S., & Herrmann, C. S. (2010). Transcranial Alternating Current Stimulation Enhances Individual Alpha Activity in Human EEG. *PLoS ONE*, *5*(11), e13766. https://doi.org/10.1371/journal.pone.0013766

Figure Legends

Figure 1

a) Sequence discrimination paradigm: (top) visual flash stimuli and (middle) auditory waveforms, and two example paired sequences (bottom); b) TACS electrode montage for Temporal (light grey) and Occipital (dark grey) stimulation sites.

Figure 2

Stacked bar charts illustrating relative proportions of participants in different groups.

Figure 3

Effects of tACS on d' sensitivity (relative to sham stimulation) to differences in visual (V) or auditory (A) sequences, relative to sham, or two occipital versus temporal electrode montages (dark blue and light yellow respectively). Positive values indicate an improvement due to tACS. Error bars show standard error; asterisks and horizontal brackets show significant differences ($p<.05$, Tukey comparisons). Left graph represents participants who did not report visually-evoked auditory sensations (N=21) evoked by the flashes in the visual sequence discrimination condition. Right graph is for participants (mostly musicians) who did report visually-evoked auditory sensations (N=15).

Figure 4

Scatterplot of effects of tACS electrode montage on d' sensitivity (relative to sham stimulation). Separate colours for Visual (dark blue) and Auditory (light yellow) sequence discrimination tasks. Separate graphs for participants who did not report hearing flashes (left) and who did (right). Pearson correlation statistics are shown in legend.

Figure 5

Comparison of d' for two frequencies of stimulation (rows) relative to sham, for occipital and temporal electrode montages (dark blue and light yellow respectively), for two groups defined by V:A (ratio of Visual to Auditory sequence discriminability columns). Left column: V poorer than A; Right column: V more similar to A. Error bars show within-subjects standard errors; asterisks and horizontal brackets show significant differences ($p<.05$, Tukey comparisons).

Figure 6

Scatterplots of effects of 40Hz tACS on d' (Y axes, relative to sham) and different tasks and montages (columns), each plotted against V:A ratios for individual non-vEAR participants (X axes). Lines of best fit are shown with correlation statistics. Filled and open datapoints distinguish between the V<A and V~A groups shown in Figure 5.

Figure 7

Schematic illustration of the theoretical effects of 10Hz tACS on inhibitory signalling via alpha frequency oscillations (wave patterns), biasing the balance of competition between auditory and visual areas. a) in non-vEAR subjects, tACS to auditory areas might reduce inhibition of visual areas, which become more excitable and send back inhibition to auditory areas (b) in vEAR subjects, tACS might have less effect because there is less prior inhibition between areas. In vEAR, auditory and visual areas effectively cooperate rather than compete.

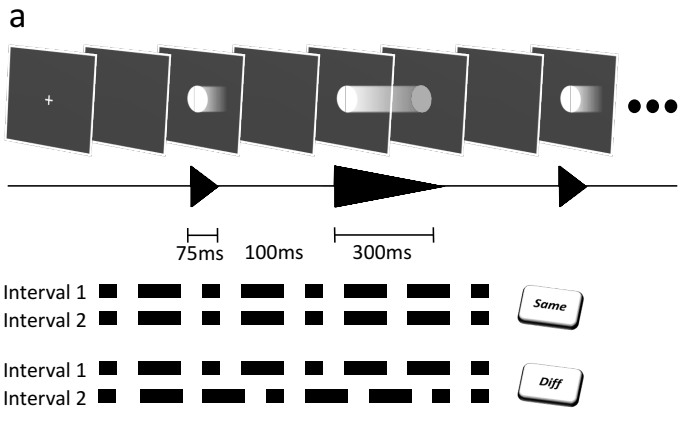
Table 1

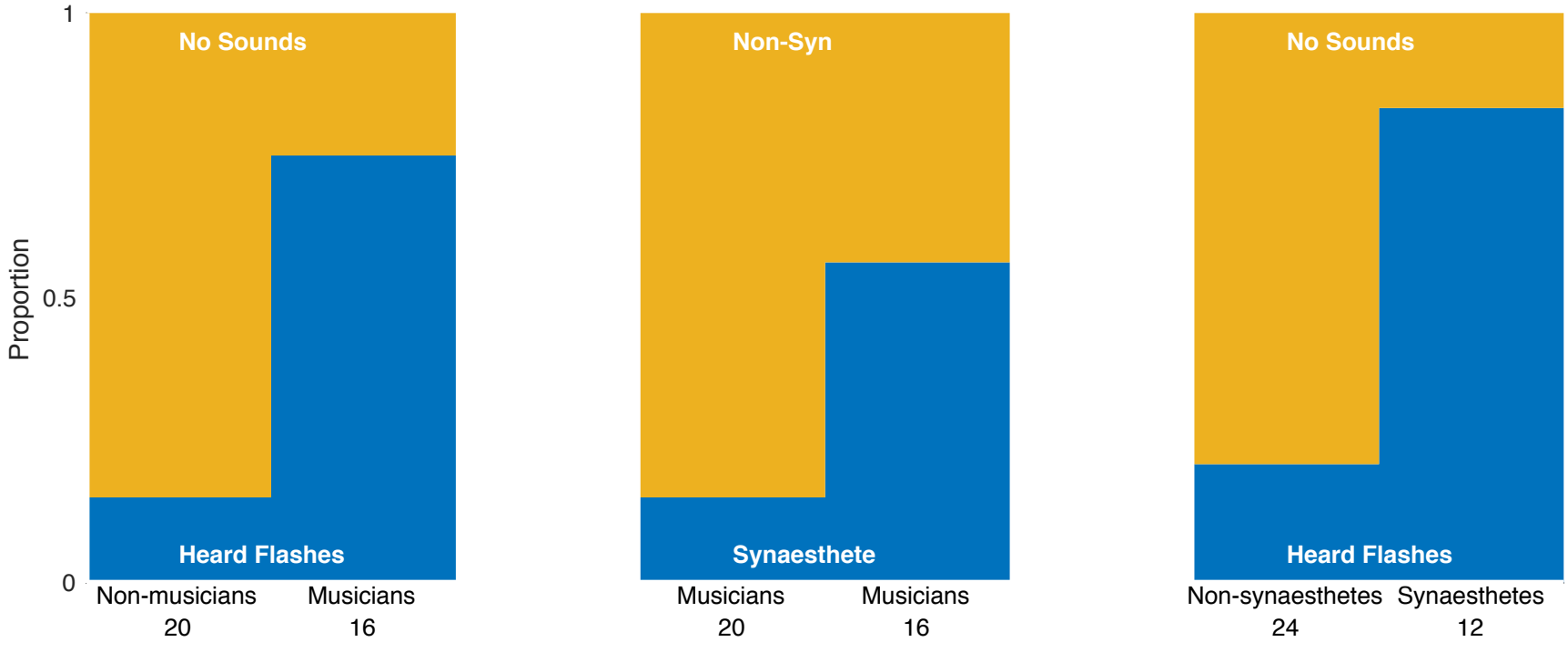
Bayesian ANOVA: Model Comparison

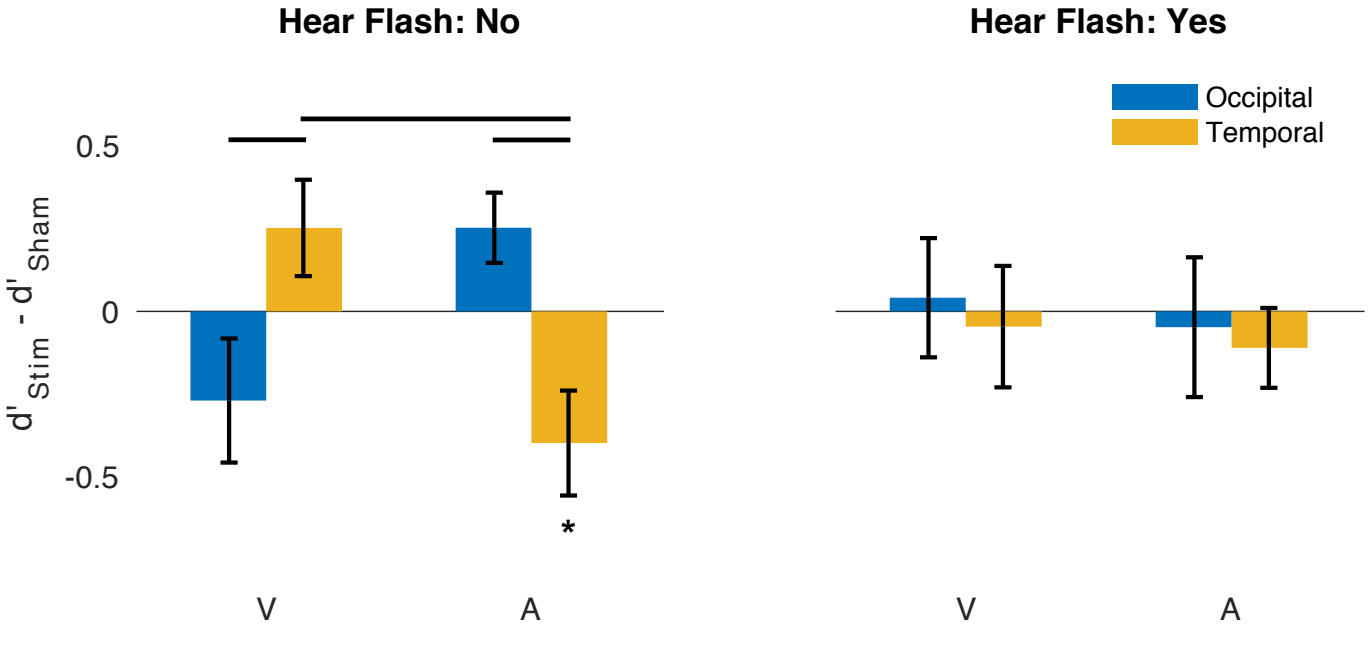
Models	P(Model)	P(Model | data)	BF_{Model}	BF_{10}	error %
Null model	0.17	0.037	0.19	1.000	
OT ✻ VA	0.17	0.083	0.45	2.24	10.76
OT ✻ VA ✻ StimulationSensations (+ OT ✻ VA)	0.17	0.028	0.15	0.76	13.30
OT ✻ VA ✻ HeardFlashes (+ OT ✻ VA)	0.17	0.65	9.23	17.54	55.58
OT ✻ VA ✻ StimulationSensations ✻ HeardFlashes (+ OT ✻ VA + OT ✻ VA ✻ StimulationSensations + OT ✻ VA ✻ HeardFlashes)	0.17	0.10	0.53	2.59	15.090

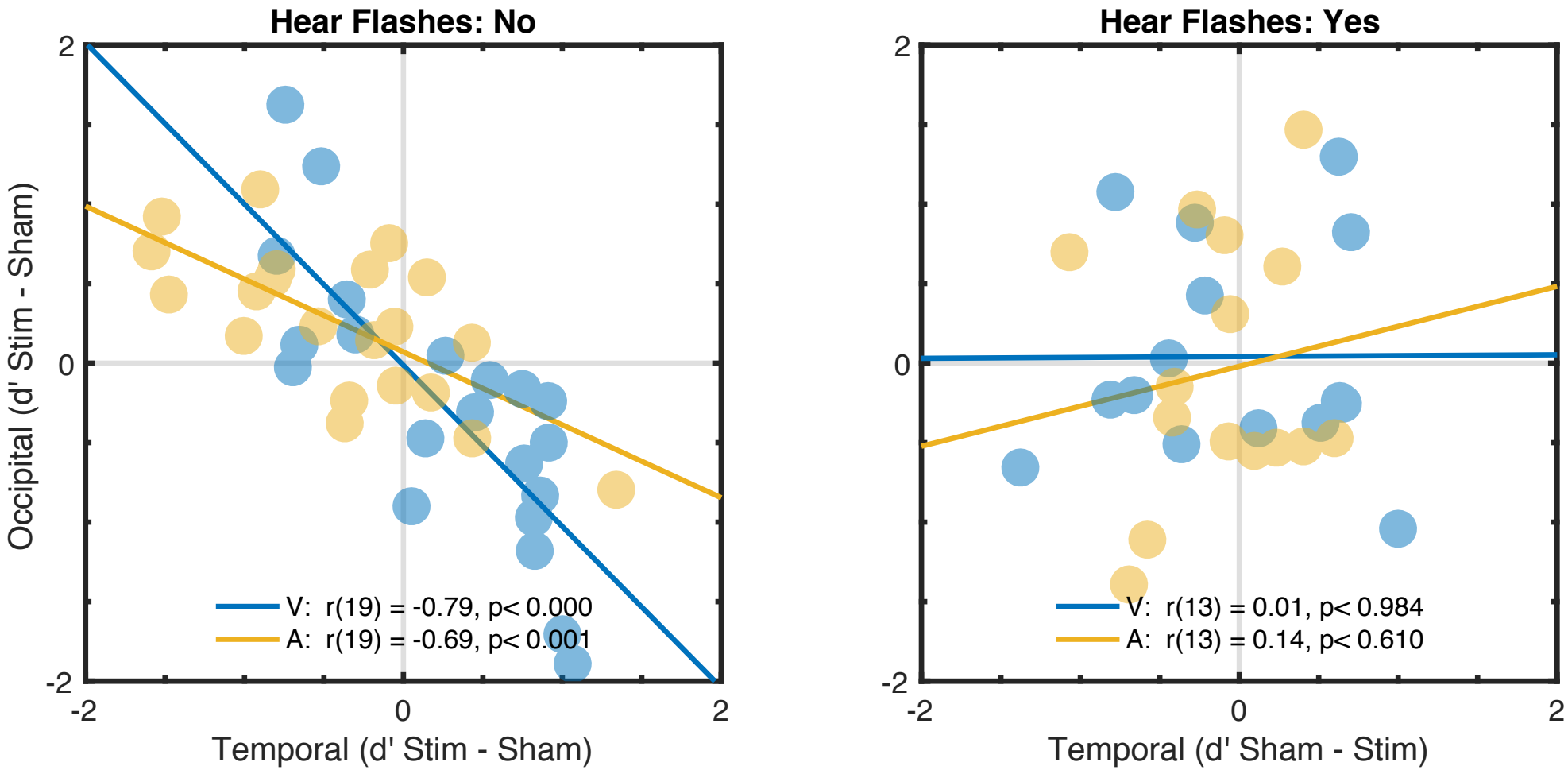
Table 1

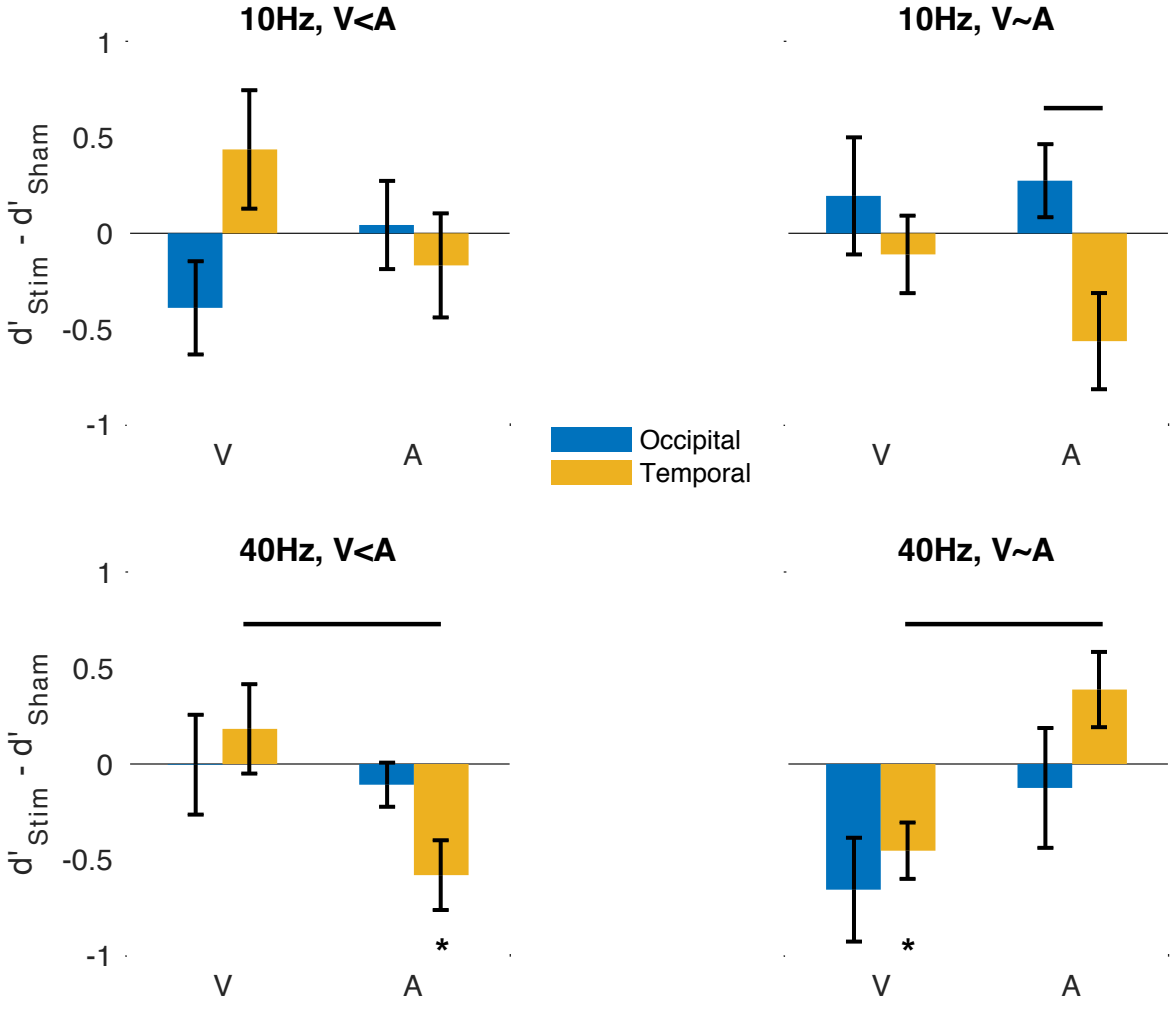
Results of Bayesian ANOVA comparing two-way interaction between Occipital vs Temporal stimulation sites (OT) and visual vs auditory task (VA), with models that also include grouping variables for 'Hearing Flashes' and 'Stimulation Sensations'. All main effects and interactions of no interest are represented under the null model. See main text for details.

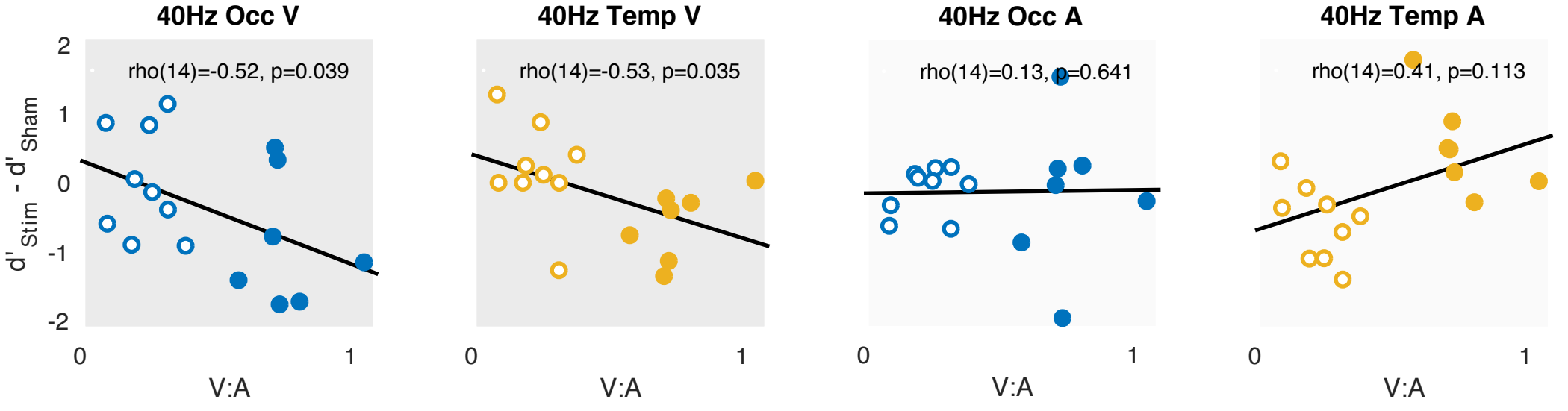

a
75ms
100ms
300ms
Interval 1
Interval 2
Same
Interval 1
Interval 2
Diff
b
NASION
INION
10%
10%
10%
Nasion
Inion


Proportion
1
0.5
0
No Sounds
Heard Flashes
Non-musicians
20
Musicians
16
Non-Syn
Synaesthete
Musicians
20
Musicians
16
No Sounds
Heard Flashes
Non-synaesthetes
24
Synaesthetes
12


Hear Flash: No
Hear Flash: Yes
Occipital
Temporal
0.5
0
-0.5
d'Stim - d'Sham
*
V
A
V
A


Hear Flashes: No
Occipital (d' Stim - Sham)
Temporal (d' Stim - Sham)
V: r(19) = -0.79, p< 0.000
A: r(19) = -0.69, p< 0.001
2
0
-2
-2
0
2
Hear Flashes: Yes
Temporal (d' Sham - Stim)
V: r(13) = 0.01, p< 0.984
A: r(13) = 0.14, p< 0.610


10Hz, V<A
10Hz, V~A
40Hz, V<A
40Hz, V~A
d'Stim - d'Sham
1
0.5
0
-0.5
-1
V
A
Occipital
Temporal
*


40Hz Occ V
rho(14)=-0.52, p=0.039
40Hz Temp V
rho(14)=-0.53, p=0.035
40Hz Occ A
rho(14)=0.13, p=0.641
40Hz Temp A
rho(14)=0.41, p=0.113
d' Stim - d' Sham
2
1
0
-1
-2
0
1
V:A