Published in "Cortex 49(10): 2728–2734, 2014"
which should be cited to refer to this work.

Contributions of pitch and bandwidth to sound-induced enhancement of visual cortex excitability in humans

Lucas Spierer [a,b,1], *Aurelie L. Manuel* [a,1], *Domenica Bueti* [a] *and Micah M. Murray* [a,c,d,*]

[a] *Neuropsychology and Neurorehabilitation Service, Department of Clinical Neurosciences, University Hospital Center and University of Lausanne, Switzerland*
[b] *Neurology Unit, Department of Medicine, Faculty of Sciences, University of Fribourg, Fribourg, Switzerland*
[c] *Department of Radiology, University Hospital Center and University of Lausanne, Switzerland*
[d] *The EEG Brain Mapping Core, Center for Biomedical Imaging (CIBM), Lausanne, Switzerland*

Multisensory interactions have been documented within low-level, even primary, cortices and at early post-stimulus latencies. These effects are in turn linked to behavioral and perceptual modulations. In humans, visual cortex excitability, as measured by transcranial magnetic stimulation (TMS) induced phosphenes, can be reliably enhanced by the co-presentation of sounds. This enhancement occurs at pre-perceptual stages and is selective for different types of complex sounds. However, the source(s) of auditory inputs effectuating these excitability changes in primary visual cortex remain disputed. The present study sought to determine if direct connections between low-level auditory cortices and primary visual cortex are mediating these kinds of effects by varying the pitch and bandwidth of the sounds co-presented with single-pulse TMS over the occipital pole. Our results from 10 healthy young adults indicate that both the central frequency and bandwidth of a sound independently affect the excitability of visual cortex during processing stages as early as 30 msec post-sound onset. Such findings are consistent with direct connections mediating early-latency, low-level multisensory interactions within visual cortices.

1. Introduction

Responses to auditory and visual stimuli have been shown to interact in humans at early stages post-stimulus onset (i.e., within the initial 100 msec; Giard and Peronnet, 1999; Molholm et al., 2002; Cappe et al., 2010; Raij et al., 2010) and within a network of regions including primary auditory as well as primary visual cortices (Martuzzi et al., 2007; Cappe et al., 2010; Raij et al., 2010). Moreover, there have been some demonstrations of the behavioral relevance of such early-latency and low-level multisensory interactions in terms of being linked to reaction time speed, perceptual outcome, or discrimination abilities (e.g., Romei et al., 2007, 2009; Van der Burg et al., 2011; Cappe et al., 2012; Murray et al., 2012).

Whereas support for the latency and locus of these effects is reasonably convincing, establishing the extent to which early-latency effects within primary visual cortex are the consequence of either direct projections from primary or near-primary auditory cortex and/or inputs from higher-level association cortices (e.g., the superior temporal sulcus and/or parietal structures) has been less forthcoming and was our focus

* *Corresponding author*. Centre Hospitalier Universitaire Vaudois, Radiologie, CIBM, Bugnon 46, BH08.078, 1011 Lausanne, Switzerland. E-mail address: Micah.murray@chuv.ch (M.M. Murray).
[1] Equal contributions.

here. To address this question, the tactic in the present study was to vary low-level acoustic features using a within-subject factorial design so as to draw inference regarding the putative source(s) of auditory inputs that are effectuating modulations in visual cortex excitability as indexed by TMS-induced phosphene perception. Specifically, we manipulated the bandwidth and center frequency (pitch) of sounds. This design was predicated on observations in non-human primates that the sharpness of tuning of neurons to frequency and bandwidth progressively decreases from core to belt and to parabelt auditory cortices (e.g., Kosaki et al., 1997; Rauschecker and Tian, 2004; Lakatos et al., 2005; Petkov et al., 2006; Hackett, 2011). Any differential efficacy of either or both of these features in modulating visual cortex excitability (viz. phosphene induction) would therefore be taken as an indication of the extent to which low-level auditory cortices contribute to (and perhaps mediate) such effects.

Anatomical studies in non-human primates have identified monosynaptic projections to primary visual cortex from both primary auditory cortex as well as the superior temporal polysensory region (Falchier et al., 2002; Rockland and Ojima, 2003; Clavagnier et al., 2004; Cappe and Barone, 2005; reviewed in Falchier et al., 2012), making it feasible for direct information transfer between primary cortices (in addition to established indirect, poly-synaptic pathways). Corresponding anatomical data in humans are currently unavailable, though diffusion-based imaging has recently provided evidence for fiber tracts between the superior temporal gyrus and the calcarine sulcus (i.e., low-level auditory regions and primary visual cortex, respectively) (Beer et al., 2011). Additional efforts have been made to apply dynamic causal modeling and effective connectivity to functional magnetic resonance imaging data so as to infer relevant pathways (Lewis and Noppeney, 2010; Noesselt et al., 2010; Powers et al., 2012; Werner and Noppeney, 2010). Despite such evidence, to our knowledge no data have been published associating specific anatomic pathways and early-latency multisensory effects within primary visual cortex.

TMS has contributed to these efforts by allowing for more causal inference on the role of specific brain regions at specific latencies in multisensory interactions (Bolognini and Maravita, 2011). For example, several laboratories have shown that the excitability of primary visual cortex, as indexed by phosphene induction[2], is enhanced by the co-presentation of a sound (Romei et al., 2007, 2009; Bolognini et al., 2010; Leo et al., 2011) or a touch (Ramos-Estebanez et al., 2007). In an effort to reveal likely sources of auditory inputs into human primary visual cortex, the authors of these studies identified variations in the efficacy of different sound features (in combination with the latency of observed effects) to modulate visual cortex excitability. Romei et al. (2007) furthermore showed that TMS over the occipital pole over the 60–90 msec post-sound onset period had opposing effects on the simple detection of auditory and visual stimuli (facilitation and slowing, respectively). In fact, the facilitation of simple detection obtained by combining occipital TMS with external auditory stimuli was as great as and correlated with the facilitation of reaction times observed when presenting participants with external auditory–visual stimuli. It has additionally been demonstrated that not all sounds are equally effective in modulating visual cortex excitability. Romei et al. (2009) showed that structured looming sounds selectively and pre-perceptually enhanced visual cortex excitability, and Bolognini et al. (2010) provide evidence for maximal enhancement of visual cortex excitability when the sounds were co-localized at the position of the induced phosphenes.

2. Methods

2.1. Participants

Ten healthy volunteers participated in the study (five women, one left-handed, mean age = 23.1 years, range 20–28 years). All participants reported normal hearing and had normal or corrected-to-normal vision. The study was approved by the Ethics Committee of the Faculty of Biology and Medicine at the University Hospital Center and University of Lausanne. All participants provided written informed consent.

2.2. Stimuli

The stimuli were 300 msec tones and bandpass-filtered noise bursts (22 kHz digitization, 16 bits, 10 msec linear rise/fall time). These sounds were generated according to a 2 × 2 design with factors of center frequency [250 Hz (low) and 6000 Hz (high)] and bandwidth [1 Hz (narrow) and 460 Hz range (broad)]. This resulted in four conditions: 250 Hz (Low/Narrow, LN condition); 6000 Hz (High/Narrow, HN); 20–480 Hz (Low/Broad, LB); and 5770–6230 Hz (High/Broad, HB). These auditory stimuli were presented through two loudspeakers located on each side of the computer monitor at a level judged comfortable by the participant. Because all data were analyzed according to a within-subject design, differences in the intensity of sound presentation across participants cannot influence the statistical outcome. The two center frequencies were chosen to be perceived with comparable loudness according to the revised ISO 226:2003 equal-loudness-level contours standard between 50 and 90 dB SPL.

2.3. TMS apparatus and determination of phosphene threshold (PT)

A 70 mm figure of eight coil (maximum field strength, 2.2 T) and a Magstim Rapid2 Transcranial Magnetic Stimulator were

[2] Phosphenes are the perceived sensation of flashes of light in the absence of visual stimulation following occipital TMS. Phosphenes elicited in low-level visual areas (V1/V2) are generally perceived as brief, static sensations along the horizontal meridian or in the lower quadrant of the hemifield contralateral to the stimulated hemisphere. They are thought to be generated by activation current that is induced by the magnetic field of the TMS pulse (e.g., Allen et al., 2007; Moliadze et al., 2003). When phosphenes are identified and defined, they remain stable within the same participant, thereby providing a reliable measure of visual cortical excitability. The minimum intensity of occipital TMS required to elicit phosphenes (i.e., phosphene threshold or PT) has been routinely used to provide a measure of this excitability (e.g., Pascual-Leone and Walsh, 2001). In studies of cross-modal effects on visual cortex excitability, the PT was first determined for each participant and then stimulator intensity was set at levels below PT. The frequency of phosphenes reported at stimulator intensities below PT was taken as a baseline, with any increases thereupon by non-visual stimuli taken as evidence for cross-modal influences on visual excitability.

used (Magstim Company, Spring Gardens, UK). PT was determined with the following procedure (see also Romei et al., 2007, 2009). Each participant wore a bathing cap to allow for marking of the site at which phosphenes could be elicited and to ensure stimulation of the same site across experimental blocks. The lights were turned off, and participants sat comfortably in a Brainsight Gen3 TMS chair with their chin and forehead supported (http://www.rogue-research.com). Participants kept their eyes open throughout the procedure to determine PT (though were allowed to blink). Stimulator output was initially set at 50% of maximal output. We then positioned the TMS coil approximately 3 cm above the inion with the handle pointing upwards. Single-pulse TMS was then applied at this site and participants were asked to report phosphene if a phosphene was perceived. If a phosphene was not reported, then stimulator intensity was increased 2% and the procedure was repeated. If a phosphene was reported, then 10 trials at that stimulator intensity were completed. If phosphenes were reported on more than five of these 10 trials then stimulator intensity was reduced, and the procedure was repeated. If phosphenes were reported on five or less trials then TMS intensity was again increased until TMS elicited phosphenes on exactly five out of 10 trials. If this site proved ineffective in eliciting phosphenes after stimulator intensity was increased to 70% of maximal output, then the coil was moved leftward by approximately 5 mm and the above procedure repeated with stimulator output initially reduced to 50%. If this position was likewise unsuccessful in identifying PT then the coil was moved leftward an additional 5 mm. If this second leftward position was unsuccessful, then the coil was moved approximately 5 mm to the right of midline, and the procedure repeated. Two participants in addition to those reported in this study were evaluated, but were excluded because they never reported perceiving phosphenes. On average, the PT was of 48.9 ± 1.6% (mean ± s.e.m.) of maximum stimulator output. The coil position at which PT was determined as well as the features of the reported phosphenes (i.e., their shape, size, and location) varied slightly across participants, but were constant for each participant across the experimental blocks. For the experimental blocks the single-pulse TMS was applied at 80% of the individually adjusted PT.

2.4. *Procedure and task*

Participants were seated in Brainsight Gen3 TMS chair in a sound-attenuated booth in front of a 19″ LCD screen and instructed to report when they perceived a phosphene by pressing a response button with their right index finger. All trials consisted in the presentation of one of the four sounds paired with the delivery of a single TMS pulse centered over the occipital pole at a delay of 30, 90, or 150 msec post-sound onset. Then, a response window opened and closed as soon as a response was recorded. In case of no response, the window closed after 4000 msec. The inter-trial interval (i.e., the interval between the closure of the response window and onset of the next trial) was varied pseudo-randomly from 2000 to 3000 msec to avoid anticipation of stimulus onset. The text "Phosphene?" and a fixation cross were presented written in white on a black background during the response window and the inter-trial interval, respectively. Each participant completed five blocks, including four repetitions of each experimental condition and eight randomly intermixed trials involving TMS stimulation in the absence of any sound to establish a baseline measure of visual cortex excitability. Each block thus consisted in a total of 56 trials (four repetitions × four sound conditions × three TMS delays + eight baseline control trials). After the completion of each block, a rest period was provided to participants to maintain high concentration and minimize fatigue. Stimulus presentation, TMS pulse delivery, and behavioral response collection were controlled by E-prime (E-Prime 1.1; Psychology Software Tools, Pittsburgh, PA).

3. Results

The percentage of trials when phosphenes were reported in the absence of sounds was taken as a baseline of visual cortex excitability. The mean (±s.e.m.) percentage was 38.5 ± 5.6%, confirming that the selected stimulator intensity was on average below the phosphene induction threshold throughout the duration of the experiment.

The percentage of trials when phosphenes were reported in the presence of different sounds and at different delays following sound presentation was submitted to a three-way repeated measures analysis of variance (rmANOVA) using center frequency (250 Hz *vs* 6000 Hz), bandwidth (1 Hz *vs* 460 Hz), and delay (30, 90, and 150 msec post-sound onset) as within-subject factors. There were main effects of center frequency [$F_{(1,9)} = 8.277$, $p = .018$, $\eta_p^2 = .479$], bandwidth [$F_{(1,9)} = 7.276$, $p = .024$, $\eta_p^2 = .447$], and delay [$F_{(2,8)} = 5.633$, $p = .030$, $\eta_p^2 = .585$]. Post-hoc contrasts for these main effects are reported below. None of the interactions met the .05 significance criterion (all p's $> .45$). The main effect of center frequency followed from generally higher reports of phosphenes following presentation of sounds with 6000 Hz center frequency than 250 Hz center frequency (57.5% *vs* 44.4%, respectively). The main effect of bandwidth followed from generally higher reports of phosphenes following presentation of narrowband versus broadband sounds (54.4% *vs* 47.5%, respectively). The main effect of delay followed from a general decrease in the reports of phosphenes with greater delays post-sound onset (54.3%, 50.3%, and 48.4%, respectively).

Given these three main effects in the absence of any interactions and in order to statistically determine whether phosphene induction was increased relative to the above-defined baseline levels, a series of follow-up rmANOVAs were performed. Post-hoc *t*-tests (two-tailed) were corrected for multiple comparisons using the Holm–Bonferroni method (Holm, 1979). First, we tested the data as a function of center frequency, collapsing across bandwidths and delays, and included the TMS-only baseline as an additional condition in a one-way rmANOVA with three levels (TMS-only, 250 Hz and 6000 Hz). This analysis resulted in a main effect of condition [$F_{(2,8)} = 5.232$, $p = .035$, $\eta_p^2 = .567$]. Sounds with 6000 Hz center frequency increased phosphene perception significantly above baseline levels [$t_{(9)} = 3.398$, $p < .008$] as well as levels following presentation of 250 Hz sounds [$t_{(9)} = 2.877$, $p < .02$], whereas sounds with 250 Hz center frequency did not increase phosphene perception above baseline levels [$t_{(9)} = 1.666$,

$p > .12$] (Fig. 1a). Next, we tested the data as a function of bandwidth, collapsing across center frequency and delays, and again included the TMS-only baseline as an additional condition as above. This analysis resulted in a main effect of condition [$F_{(2,8)} = 5.857$, $p = .027$, $\eta_p^2 = .594$]. Narrowband sounds increased phosphene perception significantly above baseline levels [$t_{(9)} = 3.471$, $p < .007$] and above levels observed for broadband sounds [$t_{(9)} = 2.697$, $p < .025$]. Additionally, broadband sounds enhanced phosphene perception above baseline levels [$t_{(9)} = 2.265$, $p < .050$] (Fig. 1b). Lastly, we tested the data as a function of delay, collapsing across center frequency and bandwidth, and again included the TMS-only baseline as an additional condition as above. This analysis resulted in a main effect of condition [$F_{(3,7)} = 5.649$, $p = .028$, $\eta_p^2 = .708$]. TMS delivered 30 msec or 90 msec after sound onset significantly increased phosphene perception above baseline levels [$t_{(9)} = 3.394$, $p < .008$ and $t_{(9)} = 3.204$, $p < .011$, respectively], whereas TMS delivered 150 msec after sound onset did not [$t_{(9)} = 2.231$, $p > .050$] (Fig. 1c). Additionally, TMS delivered 30 msec after sound onset significantly increased phosphene perception above levels observed when TMS was delivered 150 msec after sound onset [$t_{(9)} = 3.524$, $p < .007$]. No other post-hoc contrasts were significant.

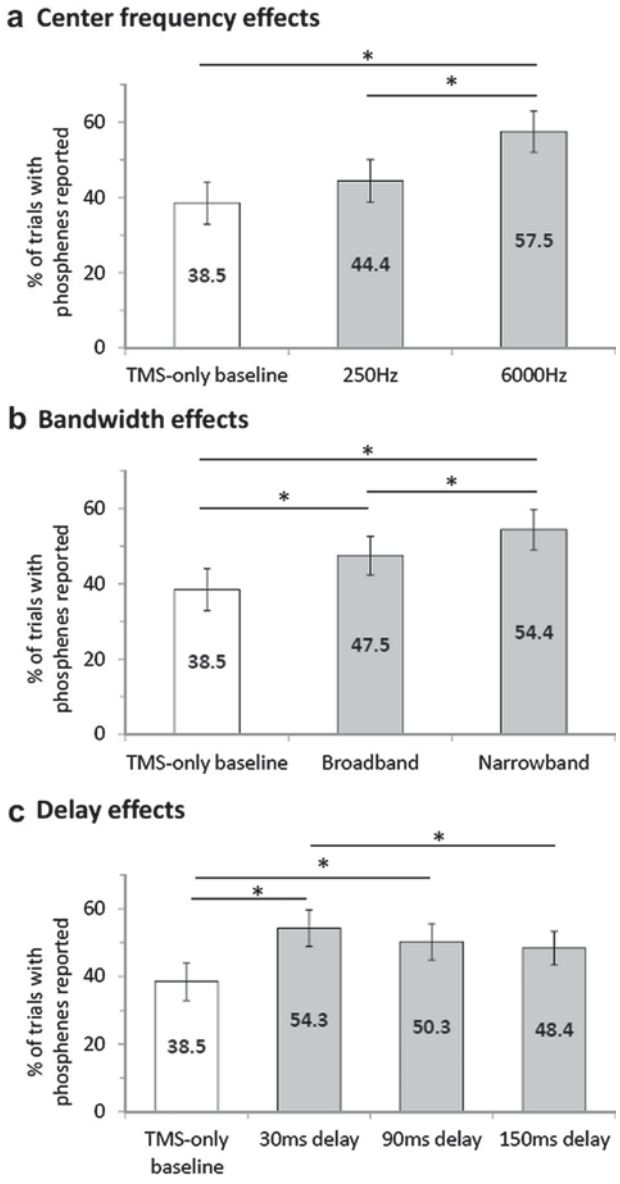

Fig. 1 – Sound-induced modulation of visual cortex excitability. In all panels the y-axis shows the percentage of trials when a phosphene was reported, including the TMS-only baseline condition. Mean (s.e.m. indicated) values across participants are displayed. An asterisk indicates a significant pair-wise difference after correction for multiple comparisons (see Results for details). Panel a displays the results for the main effect of center frequency. Panel b displays the results for the main effect of sound bandwidth. Panel c displays the results for the main effect of delay between sound presentation and TMS stimulation.

4. Discussion

This study provides evidence that both the center frequency (pitch) and bandwidth of sounds independently impact the excitability of visual cortex when presented in combination with a subthreshold TMS pulse over the occipital pole. Specifically, 6000 Hz sounds enhanced visual cortex excitability beyond threshold levels, whereas 250 Hz sounds did not, and narrowband sounds enhanced visual cortex excitability beyond threshold levels as well as beyond levels observed with broadband sounds, which were likewise more effective than TMS-alone (Fig. 1a and b). These acoustic features had their maximal effect when the TMS pulse followed sound onset by 30 msec, although effects above baseline were also observed at a delay of 90 msec, but not 150 msec (Fig. 1c). The acoustic and temporal specificity we observed provides a collective pattern that speaks in favor of direct projections from low-level auditory cortices as the principal mediators of cross-modal enhancements in visual cortex excitability.

Our observation that only higher frequency pitch and narrow bandwidth sounds were effective in enhancing visual cortex excitability would suggest that the auditory signal that effectuates the enhancement of visual cortex excitability is relatively un-processed or minimally processed. One possible explanation for the main effect of pitch that we observed can be based on an extrapolation of the anatomic result in non-human primates that it is only more caudal portions of low-level auditory cortices that directly project to primary visual areas (Falchier et al., 2002). If such projections in humans are likewise restricted to more caudal portions, then recent tonotopic mapping would suggest such portions to be more responsive to higher than to lower frequency pitches (Da Costa et al., 2011). The neurophysiologic properties of auditory neurons monosynaptically projecting to primary visual cortex have yet to be determined and at this stage can only be extrapolated based on similar anatomic locations with studies focusing on response properties of neurons within a specific auditory region. The abovementioned anatomic studies (as well as those of Rockland and Ojima, 2003) place the source(s) of monosynaptic auditory inputs into primary visual cortex within caudal portions of low-level auditory regions. Auditory response properties of single neurons have been

well-characterized in core, belt, and parabelt regions in non-human primates (e.g., Rauschecker and Tian, 2004; Lakatos et al., 2005). These studies generally agree that central frequency tuning as well as bandwidth tuning broadens with progression from core to belt and to parabelt regions. For example, belt regions of macaque auditory cortex have been shown to respond more intensively to broadband than to narrowband sounds, whereas more intense responses to narrowband sounds were observed within core regions (Rauschecker and Tian, 2004). The extent to which humans and macaque monkeys exhibit homologous anatomic and neurophysiologic substrates of multisensory integration remains to be fully detailed and will undoubtedly benefit from additional research. In the context of the present study, had projections from belt or other higher-order regions been mediating our effects then a strong prediction would have been for greater enhancement of visual cortex excitability when the TMS pulse was paired with a broadband sound. Instead, the opposite was observed, which supports core regions as the more likely source.

The timing of the present effects likewise provides some constraints on the putative sources of auditory inputs. Our effects were maximal when the auditory stimulus onset preceded the TMS pulse by 30 msec, remained above TMS-only baseline levels when the temporal separation was 90 msec, and did not significantly differ from baseline levels with a temporal separation of 150 msec (Fig. 1c). Response onset within primary auditory cortex in humans has been documented at ~15 msec (Liégeois-Chauvel et al., 1994) with propagation to adjacent regions within ~3 msec (Brugge et al., 2003). In light of these figures and assuming a conduction time to V1 of ~10–12 msec, maximal effects could be expected with a delay of 30 msec between sound onset and TMS delivery (see also Raij et al., 2010).

Prior TMS studies leave unresolved the sources of auditory inputs that alter visual cortex excitability. On the one hand, Romei et al. (2009) provided evidence that looming sounds enhance visual cortex excitability beyond baseline levels as well as levels observed with other types of sounds at latencies prior to when subjects could reliably discriminate looming from stationary sounds. Differential excitability following from looming versus either stationary or receding sounds first appeared when the TMS pulse was delivered 80 msec after the sound. Control experiments carried out by these authors ruled out explanations in terms of attention/arousal or as being due to the intensity or amplitude envelope. Moreover, they provide evidence that enhancement levels are dependent upon the use of structured (i.e., tonal) stimuli rather than noise bursts, though it should be noted that sounds of all varieties led to enhancement beyond baseline levels (cf. Fig. 4 in Romei et al., 2009). While these data do not unequivocally localize the source of auditory inputs mediating the enhancement of visual cortex excitability, they nonetheless speak in favor of sources that are sensitive to low-level acoustic features and preferentially responsive to structured sounds versus broadband noise bursts; attributes consistent with neural sensitivity within low-level auditory regions (e.g., Rauschecker and Tian, 2004). The present study furthers our understanding of this issue by showing there to be independent contributions of pitch and bandwidth during time windows that overlap with those described by Romei et al. (2007, 2009, 2012) and also by showing that there are acoustic features that fail to enhance visual cortex excitability beyond baseline levels. That is, some sounds were ineffective despite their equivalent perceived loudness, thereby providing one level of evidence against an account of our results in terms of selective attention to the auditory modality. Such an effect would indeed have been predicted to lead to a general enhancement by all sounds irrespective of pitch/bandwidth, or enhanced arousal with higher pitch or broadband sounds.

On the other hand, Bolognini et al. (2010) investigated the potential source(s) of auditory inputs impacting visual cortex excitability by varying the spatial co-registration between sounds (a 20 msec white noise burst) and the perceived location of induced phosphenes. Their dependent measure, in contrast to that used here, was always the difference between the percentages of reported phosphenes when co-presented with sounds versus when TMS was applied alone. They compared these modulations as a function of the spatial alignment between sounds and phosphenes as well as the delay between sound presentation and TMS delivery. The analysis of the data in this manner led the authors to conclude that auditory influences on visual cortex excitability were restricted to situations where the sound was co-localized with the location of peripheral (but not central) phosphenes. Bolognini et al. (2010) considered these results as evidence in favor of a direct-projection mechanism. This interpretation was based on anatomic data from non-human primates showing that monosynaptic projections between low-level auditory cortex and primary visual cortex preferentially, but not exclusively, terminate in peripheral visual field representations (cf. Table 1 in Falchier et al., 2002). More recent findings in humans based on diffusion tensor imaging would instead suggest that fiber tracts from Heschl's gyrus terminate in the occipital pole where the (para)foveal visual field would be represented (Beer et al., 2011). However, in the absence of functional mapping of their seed regions it is difficult to attribute these fiber tracts to specific auditory regions or tonotopic representations, though there is now functional data to link Heschl's gyrus to core auditory regions (Da Costa et al., 2011). More generally, the cumulative data from human electroencephalography, magnetoencephalography, functional magnetic resonance imaging, transcranial magnetic stimulation, and diffusion tensor imaging would support there being early-latency, low-level, and behaviorally-relevant auditory–visual multisensory interactions. These effects occur with central-presented stimuli and involve central visual field representations in humans (Murray et al., 2012).

At first sight, the findings of Bolognini et al. (2010) would therefore appear in sharp contrast with the present results and prior findings examining auditory influences on centrally-perceived phosphenes (Romei et al., 2007, 2009, 2012). However, whether or not a given condition enhanced visual cortical excitability beyond baseline levels was not assessed or discussed. Inspection of their data (cf. Fig. 1 in Bolognini et al., 2010) would instead suggest that enhancement of visual cortex excitability beyond baseline levels was indeed observed both when sounds and phosphenes were co-localized to central positions as well as when sounds were not co-localized with the location of phosphenes but instead were

presented to the opposite hemispace. That is, in their Experiments 1 and 3 it seems to be the case that there was general enhancement of visual cortex excitability by sounds, irrespective of (i) spatial co-localization with the phosphene (when perceived), (ii) peripheral versus central phosphene/sound presentation, and (iii) delay between sound presentation and TMS delivery. Consideration of their data in this manner, albeit based on visual inspection rather than formal statistical analyses, would therefore suggest that spatial features (absolute position or co-localization) are not the main determinant of cross-modal modulation of visual cortex excitability and that such cross-modal modulation occurs for centrally presented sounds and centrally-perceived phosphenes. This pattern is highly consistent with the present results as well as those of Romei et al. (2007, 2009, 2012). Such being said, their finding that co-localized peripheral sounds resulted in further enhancements of phosphene perception is robust and warrants more detailed study to determine its neurobiological basis and whether such effects rely on mechanisms distinct from the abovementioned general effects.

The magnitude of the enhancements in visual cortex excitability is highly consistent with prior findings. Here, effects were on the order of ~10–20% versus baseline levels (cf. Fig. 1). This is similar to what was observed by Romei et al. (2007, 2009). The increase from baseline in Romei et al. (2007) was ~15–20% (see their Fig. 4; i.e., phosphenes were reported on TMS-only trials roughly 30% of the time and increased to a maximum of roughly 50%). In Romei et al. (2009) only looming sounds led to the values doubling those observed at baseline. The other sounds (which were stationary or receding) led to increases again on the order of 15–20% (see Fig. 2 in Romei et al., 2009). In Bolognini et al. (2010) increases, when present, were likewise on the order of 15–20%, with the exception of the one condition and delay in Experiment 1 that led to a near-doubling.

Several domains were not specifically investigated here, but nonetheless warrant continued study. For one, the present study used a limited sample of two pitches and two bandwidths. A fuller stimulus set would be necessary to derive tuning curves for auditory influences on visual cortex excitability. Likewise, a fuller stimulus set may prove more effective in revealing different latencies of auditory inputs to visual cortices. Secondly, there is mounting evidence that visual excitability (and cross-modal influences upon such) is state-dependent such that the phase of ongoing oscillations at the time of TMS delivery can play a central role in modulating cortical excitability and can be reset by preceding sounds (Romei et al., 2012). Thirdly, it will be important to examine inter-individual variations in tonotopic representations and their consequences on cross-modal modulation of visual cortex excitability. A fourth, but by no means exhaustive domain, would be to capitalize upon these and related findings to optimize parameters of sensory substitution devices in visually-impaired individuals (e.g., Amedi et al., 2007). In conclusion, the present study provides evidence in support of there being direct projections from low-level auditory cortex to primary visual cortex that can impact the excitability of visual neurons and in turn perception/behavior.

Author contributions

All co-authors contributed to the experimental design. A.M. and D.B. acquired the data. L.S., A.M., and M.M.M. analyzed the data. M.M.M. and L.S. wrote the paper with input from all authors.

Conflict of interest

The authors declare that they have no competing financial interests.

Acknowledgments

This work has been supported by the Swiss National Science Foundation (Grant 310030B-133136 to MMM). We thank Troy Hackett and Arnaud Falchier for helpful discussions during the preparation of this manuscript. We are extremely grateful to Dr. Bolognini for sharing some of her raw data from her 2010 publication.

REFERENCES

Allen EA, Pasley BN, Duong T, and Freeman RD. Transcranial magnetic stimulation elicits coupled neural and hemodynamic consequences. *Science*, 317: 1918–1921, 2007.

Amedi A, Stern WM, Camprodon JA, Bermpohl F, Merabet L, Rotman S, et al. Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nature Neuroscience*, 10(6): 687–689, 2007.

Beer AL, Plank I, and Greenlee MW. Diffusion tensor imaging shows white matter tracts between human auditory and visual cortex. *Experimental Brain Research*, 213(2–3): 299–308, 2011.

Bolognini N and Maravita A. Uncovering multisensory processing through non-invasive brain stimulation. *Frontiers in Psychology*, 2: 46, 2011.

Bolognini N, Senna I, Maravita A, Pascual-Leone A, and Merabet LB. Auditory enhancement of visual phosphene perception: The effect of temporal and spatial factors and of stimulus intensity. *Neuroscience Letters*, 477(3): 109–114, 2010.

Brugge JF, Volkov IO, Garell PC, Reale RA, and Howard 3rd MA. Functional connections between auditory cortex on Heschl's gyrus and on the lateral superior temporal gyrus in humans. *Journal of Neurophysiology*, 90: 3750–3763, 2003.

Cappe C and Barone P. Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *European Journal of Neuroscience*, 22: 2886–2902, 2005.

Cappe C, Thut G, Romei V, and Murray MM. Auditory-visual multisensory interactions in humans: Timing, topography, directionality, and sources. *Journal of Neuroscience*, 30(38): 12572–12580, 2010.

Cappe C, Thelen A, Romei V, Thut G, and Murray MM. Looming signals reveal synergistic principles of multisensory integration. *Journal of Neuroscience*, 32(4): 1171–1182, 2012.

Clavagnier S, Falchier A, and Kennedy H. Long-distance feedback projections to area V1: Implications for multisensory integration, spatial awareness, and visual consciousness. *Cognitive, Affective, and Behavioral Neuroscience*, 4(2): 117–126, 2004.

Da Costa S, van der Zwaag W, Marques JP, Frackowiak RS, Clarke S, and Saenz M. Human primary auditory cortex follows the shape of Heschl's gyrus. *Journal of Neuroscience*, 31(40): 14067–14075, 2011.

Falchier A, Clavagnier S, Barone P, and Kennedy H. Anatomical evidence of multimodal integration in primate striate cortex. *Journal of Neuroscience*, 22: 5749–5759, 2002.

Falchier A, Cappe C, Barone P, and Schroeder CE. Sensory convergence in low-level cortices. In Stein BE (Ed), *The New Handbook of Multisensory Processes*. Cambridge, USA: MIT Press, 2012: 67–79.

Giard MH and Peronnet F. Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, 11: 473–490, 1999.

Hackett TA. Information flow in the auditory cortical network. *Hearing Research*, 271(1–2): 133–146, 2011.

Holm S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2): 65–70, 1979.

Kosaki H, Hashikawa T, He J, and Jones EG. Tonotopic organization of auditory cortical fields delineated by parvalbumin immunoreactivity in macaque monkeys. *Journal of Comparative Neurology*, 386(2): 304–316, 1997.

Lakatos P, Pincze Z, Fu KM, Javitt DC, Karmos G, and Schroeder CE. Timing of pure tone and noise-evoked responses in macaque auditory cortex. *NeuroReport*, 16(9): 933–937, 2005.

Leo F, Romei V, Freeman E, Ladavas E, and Driver J. Looming sounds enhance orientation sensitivity for visual stimuli on the same side as such sounds. *Experimental Brain Research*, 213(2–3): 193–201, 2011.

Lewis R and Noppeney U. Audiovisual synchrony improves motion discrimination via enhanced connectivity between early visual and auditory areas. *Journal of Neuroscience*, 30(37): 12329–12339, 2010.

Liégeois-Chauvel C, Musolino A, Badier JM, Marquis P, and Chauvel P. Evoked potentials recorded from the auditory cortex in man: Evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology*, 92: 204–214, 1994.

Martuzzi R, Murray MM, Michel CM, Maeder PP, Thiran JP, Clarke S, et al. Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cerebral Cortex*, 17(7): 1672–1679, 2007.

Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, and Foxe JJ. Multisensory auditory-visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Brain Research. Cognitive Brain Research*, 14: 115–128, 2002.

Moliadze V, Zhao Y, Eysel U, and Funke K. Effect of transcranial magnetic stimulation on single-unit activity in the cat primary visual cortex. *Journal of Physiology*, 553: 665–679, 2003.

Murray MM, Cappe C, Romei V, Martuzzi R, and Thut G. Auditory-visual multisensory interactions in human primary cortices: Synthesis and controversies. In Stein BE (Ed), *The New Handbook of Multisensory Processes*. Cambridge, USA: MIT Press, 2012: 223–238.

Noesselt T, Tyll S, Boehler CN, Budinger E, Heinze HJ, and Driver J. Sound-induced enhancement of low-intensity vision: Multisensory influences on human sensory-specific cortices and thalamic bodies relate to perceptual enhancement of visual detection sensitivity. *Journal of Neuroscience*, 30(41): 13609–13623, 2010.

Pascual-Leone A and Walsh V. Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science*, 292: 510–512, 2001.

Petkov CI, Kayser C, Augath M, and Logothetis NK. Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biology*, 4(7): e215, 2006.

Powers 3rd AR, Hevey MA, and Wallace MT. Neural correlates of multisensory perceptual learning. *Journal of Neuroscience*, 32(18): 6263–6274, 2012.

Raij T, Ahveninen J, Lin FH, Witzel T, Jääskeläinen IP, Letham B, et al. Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *European Journal of Neuroscience*, 31: 1772–1782, 2010.

Ramos-Estebanez C, Merabet LB, Machii K, Fregni F, Thut G, Wagner TA, et al. Visual phosphene perception modulated by subthreshold crossmodal sensory stimulation. *Journal of Neuroscience*, 27: 4178–4181, 2007.

Rauschecker JP and Tian B. Processing of band-passed noise in the lateral auditory belt cortex of the rhesus monkey. *Journal of Neurophysiology*, 91(6): 2578–2589, 2004.

Rockland KS and Ojima H. Multisensory convergence in calcarine visual areas in macaque monkey. *International Journal of Psychophysiology*, 50: 19–26, 2003.

Romei V, Murray MM, Merabet L, and Thut G. Occipital TMS has opposing effects on visual and auditory stimulus detection: Implications for multisensory interactions. *Journal of Neuroscience*, 27: 11465–11472, 2007.

Romei V, Murray MM, Cappe C, and Thut G. Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Current Biology*, 19(21): 1799–1805, 2009.

Romei V, Gross J, and Thut G. Sounds reset rhythms of visual cortex and corresponding human visual perception. *Current Biology*, 22(9): 807–813, 2012.

Van der Burg E, Talsma D, Olivers CN, Hickey C, and Theeuwes J. Early multisensory interactions affect the competition among multiple visual objects. *NeuroImage*, 55: 1208–1218, 2011.

Werner S and Noppeney U. Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. *Journal of Neuroscience*, 30(7): 2662–2675, 2010.